

# Appendix — For Online Publication

## A. APPENDIX FIGURES AND TABLES

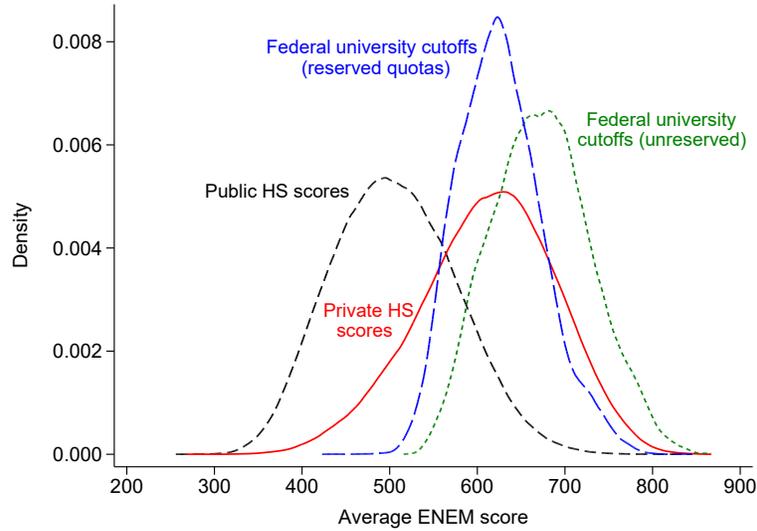
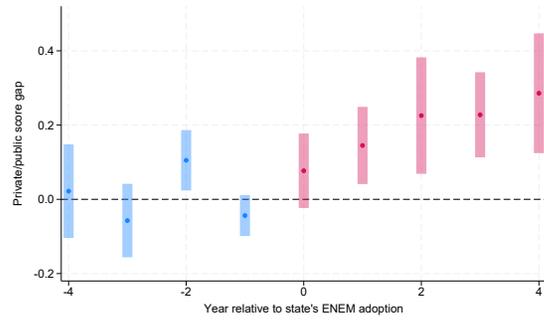


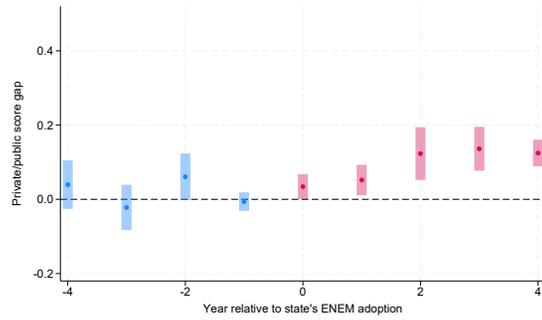
FIGURE A1. Distributions of average ENEM scores and federal university admission cutoffs

*Notes:* This figure displays the distribution of ENEM test scores and federal university cutoff scores. The solid red line shows the distribution of average ENEM scores for private school students in the analysis sample who took the ENEM exam in cohorts prior to each state’s ENEM adoption year (i.e., cohorts with  $\text{HighStakes}_{st} = 0$ ). The dashed black line represents the corresponding distribution for public school students. In both distributions, average ENEM scores are calculated as the average score across math, language arts, natural science, and social science.

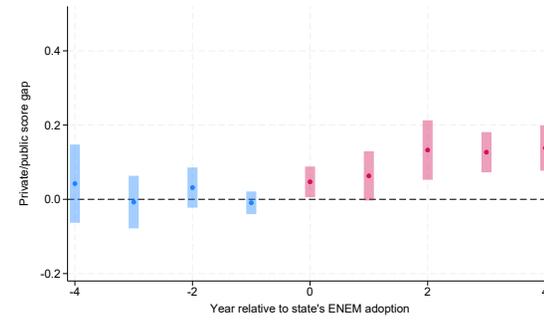
The green short-dashed line shows the distribution of cutoff scores for unreserved admissions to all federal university programs in 2016. The blue long-dashed line shows the corresponding distribution for reserved quotas at federal university programs, which include quotas for public high school students, low-SES students, and/or underrepresented minority students.



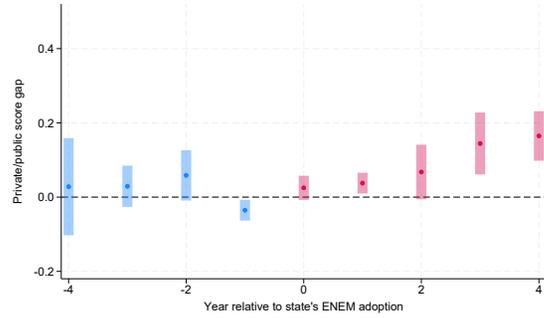
Panel A. Math



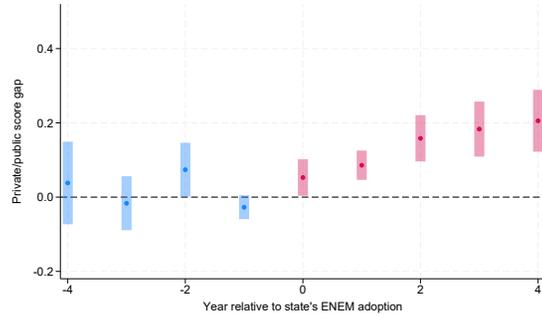
Panel B. Language arts



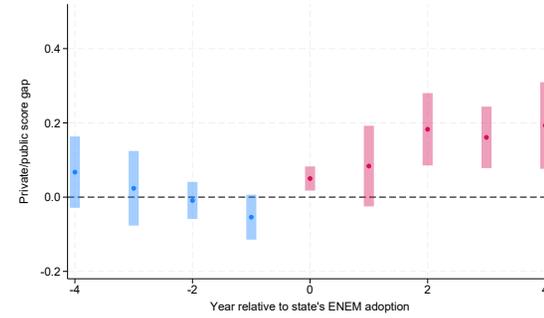
Panel C. Social sciences



Panel D. Natural sciences



Panel E. Average (core subjects)



Panel F. Writing

FIGURE A2. Callaway and Sant’Anna (2021) event studies for the impact of ENEM adoption on private/public test score gaps

*Notes:* This figure plots event study estimates of the impact of ENEM adoption on private/public test score gaps using the method of Callaway and Sant’Anna (2021). We estimate the regressions using data from 2009–2017 ENEM exam takers in our analysis sample (column C of Table 1). We collapse the data at the state  $\times$  year ( $st$ ) level and define the dependent variables to be the *difference* in mean ENEM subject scores (in SD units) between private and public schools in each  $st$  cell. “Average (core subjects)” is the average score across math, language arts, natural science, and social science.

We estimate the event study coefficients with Callaway and Sant’Anna (2021)’s `csdid` package for Stata using the binary treatment variable,  $HighStakes_{st}$ , and not-yet-treated states as the control group. We then plot the results using the authors’ `csdid_plot` command. Coefficients in the post-treatment period (red circles) represent the change in the private/public test score gap between treated and not-yet-treated states in event time  $t$  relative to one year before the treated group’s ENEM adoption year (event time  $-1$ ). Coefficients in the pre-treatment period (red circles) represent the change in the private/public test score gap between treated and not-yet-treated states in event year  $t$  relative to event year  $t - 1$ . Light-colored rectangles depict 95% confidence intervals using standard errors clustered at the state level.

TABLE A1. Summary of Brazilian high school and college systems

(A)	(B)	(C)	(D)	(E)	(F)	(G)
Institution type	# of schools	Prop. of schools	# of students	Prop. of students	# students per school	Attended a private HS
<b>Panel A. High school seniors in 2009</b>						
Federal high schools	100	0.004	9,772	0.005	98	1.000
State high schools	16,583	0.702	1,823,524	0.849	110	0.000
Municipal high schools	373	0.016	23,156	0.011	62	0.000
Private high schools	6,567	0.278	290,366	0.135	44	1.000
All high schools	23,623	1.000	2,146,818	1.000	91	0.140
<b>Panel B. New college enrollees in 2009</b>						
Federal universities	59	0.025	225,112	0.108	3,815	0.471
State universities	40	0.017	119,489	0.057	2,987	0.370
Municipal universities	9	0.004	22,453	0.011	2,495	0.319
Private universities	225	0.094	1,018,698	0.489	4,528	0.458
Public technical colleges	168	0.070	55,609	0.027	331	0.259
Private technical colleges	1,888	0.790	640,021	0.307	339	0.331
All colleges	2,389	1.000	2,081,382	1.000	871	0.401

*Notes:* This table presents summary statistics on the Brazilian high school and college systems. Panel A presents statistics for students who were high school seniors in 2009 using data from the National Primary and Secondary School Census (*Censo Escolar*). Panel B presents statistics for students who were new college enrollees in 2009 using data from the National Higher Education Census (*Censo da Educação Superior*).

Column (A) categorizes high schools by ownership (federal, state, municipal, or private), and colleges by both ownership and institution type (university or technical college). “University” includes both *Universidade* and *Centro Universitário* institutions. “Technical colleges” include *Faculdade*, *Instituto Federal de Educação Ciência e Tecnologia*, and *Centro Federal de Educação Tecnológica* institutions. Columns (B) and (C) show the number and proportion of schools in each category. Columns (D) and (E) show the number and proportion of students who attended schools in each category. Column (F) shows the number of students per school (column D divided by column B). Column (G) shows the proportion of students at each school type who attended a private high school. Throughout the paper, we include the small number of federal high schools in the group of private high schools since both tend to enroll wealthier and high-achieving students.

TABLE A2. Proportion of federal university enrollees admitted using the ENEM by state and year

(A)	(B)	(C)	(D)	(E)	(F)	(G)	(H)	(I)	(J)	(K)	(L)	(M)
Proportion admitted using ENEM by exam year ( $\text{ProportionENEM}_{st}$ )												
State	# 2009 enrollees	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017
Pernambuco (PE)	7,375	0.00	<b>0.98</b>	<b>0.93</b>	<b>0.94</b>	<b>0.93</b>	<b>0.94</b>	<b>0.92</b>	<b>0.91</b>	<b>0.92</b>	<b>0.92</b>	<b>0.95</b>
Amazonas (AM)	2,821	0.00	0.00	<b>0.32</b>	<b>0.36</b>	<b>0.38</b>	<b>0.42</b>	<b>0.43</b>	<b>0.42</b>	<b>0.50</b>	<b>0.47</b>	<b>0.47</b>
Espirito Santo (ES)	3,302	0.00	0.00	<b>0.99</b>	<b>0.92</b>	<b>1.00</b>	<b>0.90</b>	<b>0.92</b>	<b>0.89</b>	<b>0.87</b>	<b>0.94</b>	<b>0.92</b>
Maranhão (MA)	2,359	0.00	0.00	<b>0.98</b>	<b>0.88</b>	<b>0.95</b>	<b>0.94</b>	<b>0.91</b>	<b>0.87</b>	<b>0.89</b>	<b>0.78</b>	<b>0.92</b>
Mato Grosso (MT)	3,582	0.00	0.00	<b>0.64</b>	<b>0.93</b>	<b>0.90</b>	<b>0.86</b>	<b>0.83</b>	<b>0.94</b>	<b>0.95</b>	<b>0.89</b>	<b>0.97</b>
Paraná (PR)	6,820	0.00	0.00	<b>0.53</b>	<b>0.51</b>	<b>0.51</b>	<b>0.47</b>	<b>0.58</b>	<b>0.60</b>	<b>0.63</b>	<b>0.57</b>	<b>0.57</b>
Rio Grande do Sul (RS)	12,723	0.00	0.00	<b>0.42</b>	<b>0.37</b>	<b>0.42</b>	<b>0.41</b>	<b>0.39</b>	<b>0.41</b>	<b>0.62</b>	<b>0.63</b>	<b>0.62</b>
São Paulo (SP)	5,774	0.00	0.00	<b>0.87</b>	<b>0.95</b>	<b>0.89</b>	<b>0.80</b>	<b>0.74</b>	<b>0.70</b>	<b>0.74</b>	<b>0.77</b>	<b>0.71</b>
Ceará (CE)	3,704	0.00	0.00	0.00	<b>0.85</b>	<b>0.85</b>	<b>0.82</b>	<b>0.85</b>	<b>0.83</b>	<b>0.87</b>	<b>0.77</b>	<b>0.81</b>
Mato Grosso do Sul (MS)	2,710	0.00	0.00	0.00	<b>0.66</b>	<b>0.63</b>	<b>0.54</b>	<b>0.65</b>	<b>0.72</b>	<b>0.71</b>	<b>0.62</b>	<b>0.56</b>
Paraíba (PB)	7,160	0.00	0.00	0.02	<b>0.40</b>	<b>0.45</b>	<b>0.55</b>	<b>0.94</b>	<b>0.90</b>	<b>0.92</b>	<b>0.93</b>	<b>0.90</b>
Acre (AC)	955	0.00	0.00	0.00	0.00	<b>0.96</b>	<b>0.76</b>	<b>0.94</b>	<b>0.89</b>	<b>0.82</b>	<b>0.86</b>	<b>0.83</b>
Alagoas (AL)	3,008	0.00	0.00	0.00	0.00	<b>0.90</b>	<b>0.83</b>	<b>0.87</b>	<b>0.88</b>	<b>0.94</b>	<b>0.95</b>	<b>0.93</b>
Piauí (PI)	3,592	0.01	0.00	0.55	0.29	<b>0.82</b>	<b>1.00</b>	<b>0.71</b>	<b>0.97</b>	<b>0.96</b>	<b>0.58</b>	<b>0.98</b>
Rio Grande do Norte (RN)	6,443	0.00	0.00	0.34	0.15	0.06	<b>0.61</b>	<b>0.70</b>	<b>0.82</b>	<b>0.82</b>	<b>0.79</b>	<b>0.82</b>
Rondônia (RO)	1,106	0.00	0.00	0.08	0.02	0.31	<b>0.77</b>	<b>0.69</b>	<b>0.79</b>	<b>0.81</b>	<b>0.95</b>	<b>0.96</b>
Bahia (BA)	6,251	0.00	0.00	0.16	0.16	0.31	0.36	<b>0.90</b>	<b>0.83</b>	<b>0.82</b>	<b>0.81</b>	<b>0.77</b>
Distrito Federal (DF)	5,296	0.00	0.00	0.00	0.01	0.10	0.04	<b>0.23</b>	<b>0.24</b>	<b>0.20</b>	<b>0.21</b>	<b>0.21</b>
Minas Gerais (MG)	20,918	0.00	0.08	0.19	0.20	0.29	0.36	<b>0.72</b>	<b>0.74</b>	<b>0.75</b>	<b>0.72</b>	<b>0.73</b>
Pará (PA)	2,937	0.00	0.00	0.00	0.00	0.03	0.09	<b>0.89</b>	<b>0.82</b>	<b>0.78</b>	<b>0.76</b>	<b>0.81</b>
Rio de Janeiro (RJ)	16,871	0.02	0.02	0.22	0.37	0.49	0.45	<b>0.77</b>	<b>0.77</b>	<b>0.78</b>	<b>0.50</b>	<b>0.48</b>
Roraima (RR)	842	0.00	0.00	0.00	0.07	0.11	0.17	<b>0.48</b>	<b>0.46</b>	<b>0.38</b>	<b>0.41</b>	<b>0.39</b>
Goiás (GO)	3,636	0.00	0.00	0.00	0.01	0.14	0.19	0.36	<b>0.85</b>	<b>0.83</b>	<b>0.77</b>	<b>0.84</b>
Tocantins (TO)	1,602	0.00	0.00	0.14	0.15	0.20	0.08	0.18	<b>0.83</b>	<b>0.88</b>	<b>0.78</b>	<b>0.84</b>
Santa Catarina (SC)	4,632	0.00	0.00	0.15	0.14	0.15	0.12	0.12	0.15	<b>0.35</b>	<b>0.34</b>	<b>0.34</b>
Amapá (AP)	309	0.00	0.00	0.00	0.00	0.49	0.43	0.44	0.40	0.38	<b>0.98</b>	<b>0.61</b>
Sergipe (SE)	3,309	0.00	0.00	0.01	0.02	0.03	0.02	0.01	0.01	0.06	0.02	0.02
All states	140,037	0.00	0.07	0.30	0.37	0.45	0.48	0.66	0.69	0.72	0.67	0.68

Notes: This table shows the proportion of federal university enrollees who were admitted using the ENEM by state and year. Column (A) lists the 27 states of Brazil. Column (B) shows the number of new federal university enrollees in each state in the 2009 calendar year. Columns (C)–(M) show the proportion of new federal university enrollees who were admitted using the ENEM based on the year students took the ENEM exam (the calendar year prior to the enrollment year). The sample for these statistics consists of new enrollees in bachelor’s programs at federal universities, using data from the Brazilian Higher Education Census.

The figures in columns (C)–(M) are the values we use for the continuous treatment variable,  $\text{ProportionENEM}_{st}$ , where  $s$  denotes states and  $t$  denotes years. Bolded figures represent state  $\times$  years that we classify as high stakes using the binary treatment variable,  $\text{HighStakes}_{st}$ . See Section 3.3 for details on the definition of  $\text{ProportionENEM}_{st}$  and  $\text{HighStakes}_{st}$ .

TABLE A3. Characteristics of federal universities by state’s ENEM adoption year

(A)	(B)	(C)	(D)	(E)	(F)	(G)	(H)	(I)	(J)	(K)	(L)
	Year of state’s ENEM adoption ( $\tau_s^*$ )										
Characteristic	2008	2009	2010	2011	2012	2013	2014	2015	2016	Never	All
<b>Panel A. University characteristics</b>											
# states	1	7	3	3	2	6	2	1	1	1	27
# federal universities	3	15	6	3	3	24	2	1	1	1	59
# 2009 enrollees	7,375	37,381	13,574	7,555	7,549	53,115	5,238	4,632	309	3,309	140,037
Mean university size	2,458	2,492	2,262	2,518	2,516	2,213	2,619	4,632	309	3,309	2,373
Mean cutoff score (2016)	670	682	669	656	657	711	660	707	670	655	688
<b>Panel B. Characteristics of 2009 enrollees</b>											
Age at enrollment	23.64	24.53	23.80	24.31	24.17	24.15	23.60	24.38	26.81	24.76	24.21
Female	0.49	0.46	0.48	0.50	0.45	0.49	0.48	0.44	0.51	0.48	0.48
White	0.42	0.67	0.33	0.46	0.40	0.58	0.46	0.85	0.40	0.30	0.56
Black	0.07	0.07	0.40	0.08	0.07	0.09	0.08	0.04	0.15	0.10	0.12
Brown	0.46	0.19	0.24	0.44	0.51	0.31	0.43	0.09	0.41	0.57	0.29
Private high school	0.40	0.43	0.60	0.26	0.74	0.52	0.58	0.79	0.87	0.62	0.52

*Notes:* This table shows characteristics of federal universities and their student bodies by their state’s ENEM adoption year. Columns (B)–(K) categorize the federal universities by the year in which their state adopted the ENEM exam,  $\tau_s^*$ , as defined in Section 3.3. Column (L) includes all federal universities. The ENEM adoption years for each state are:

- 2008: Pernambuco.
- 2009: Amazonas, Espirito Santo, Maranhão, Mato Grosso, Paraná, Rio Grande do Sul, São Paulo.
- 2010: Ceará, Mato Grosso do Sul, Paraíba.
- 2011: Acre, Alagoas, Piauí.
- 2012: Rio Grande do Norte, Rondônia.
- 2013: Bahia, Distrito Federal, Minas Gerais, Pará, Rio de Janeiro, Roraima.
- 2014: Goiás, Tocantins.
- 2015: Santa Catarina.
- 2016: Amapá.
- Never: Sergipe.

Data on enrollment size and student characteristics are from the Brazilian Higher Education Census. In Panel A, the number of universities, the number of enrollees, and the mean university size are defined using new 2009 enrollees in bachelor’s programs at federal universities. In Panel B, some demographic variables are missing in the 2009 census year, so we compute student characteristics using students who enrolled in 2009 but appear in any census year between 2009 and 2018.

The mean cutoff score (2016) is from a public data request from the centralized admission platform SISU (*Sistema de Seleção Unificada*). These averages correspond to non-reserved quotas for bachelor’s degree programs at federal universities in the year 2016. The cutoff scores are typically weighted averages of ENEM scores in up to five subjects (math, language arts, natural science, social science, and writing) and are presented in ENEM scale score units.

We obtained the SISU data in March 2020 from public data request number #23480008751201932.

TABLE A4. Visualization of stacked dataset for event studies and robustness tests

(A)	(B)	(C)	(D)	(E)	(F)	(G)	(H)	(I)	(J)	(K)	(L)	(M)	(N)
ENEM adoption yr		Number of test takers in sample (in thousands) by exam year											
Treated group	Control group	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	Total
2008	2009	114	123										237
2008	2010	29	30	31									90
2008	2011	11	12	13	14								50
2008	2012	13	15	15	16	18							78
2008	2013	68	78	70	81	88	89						474
2008	2014	18	21	20	21	25	25	25					154
2008	2015	12	13	14	15	16	17	17	17				120
2008	2016	7	8	9	9	11	11	11	11	10			86
2008	Never	8	9	10	10	12	12	12	11	11	11	10	115
2009	2010	131	139	121									391
2009	2011	113	121	104	120								456
2009	2012	115	124	105	122	128							594
2009	2013	<b>170</b>	<b>186</b>	<b>161</b>	<b>187</b>	<b>198</b>	<b>198</b>						1,100
2009	2014	120	130	110	126	135	134	136					890
2009	2015	114	122	104	121	126	126	127	129				969
2009	2016	109	116	99	114	121	120	121	123	124			1,048
2009	Never	110	117	100	115	121	121	122	124	125	127	117	1,299
2010	2011	28	28	28	41								125
2010	2012	31	31	29	43	46							180
2010	2013	86	94	84	108	115	118						606
2010	2014	35	37	34	48	52	55	56					318
2010	2015	29	29	28	42	44	46	48	47				314
2010	2016	24	24	23	36	38	40	42	41	40			309
2010	Never	25	25	24	37	39	41	43	42	41	40	36	393
2011	2012	12	13	12	14	16							67
2011	2013	67	76	67	79	85	86						460
2011	2014	16	19	17	19	22	22	22					138
2011	2015	10	11	11	13	14	14	14	14				101
2011	2016	5	6	6	7	8	8	8	8	8			63
2011	Never	6	7	7	8	9	9	9	9	8	9	8	89
2012	2013	70	79	69	81	87	88						474
2012	2014	19	22	19	21	24	25	24					154
2012	2015	13	14	13	15	16	16	16	16				119
2012	2016	8	9	7	9	10	10	10	10	10			84
2012	Never	9	10	8	10	11	11	11	11	11	11	10	112
2013	2014	74	85	74	86	94	95	97					605
2013	2015	68	77	68	80	85	87	89	90				644
2013	2016	63	72	63	74	80	81	83	84	83			682
2013	Never	64	73	63	75	80	82	84	84	84	85	80	855
2014	2015	17	20	17	20	22	23	23	23				167
2014	2016	12	15	12	14	17	17	17	17	18			140
2014	Never	13	16	13	15	17	18	18	18	19	19	18	184
2015	2016	6	7	6	8	8	9	9	9	9			72
2015	Never	7	8	7	9	9	10	10	10	10	11	7	97
2016	Never	2	3	2	3	3	4	4	4	4	4	4	37

Fig 3 & Tab 4, Col D sample				813	1,016	1,054	1,014	740	535	345	181	162	5,859
Tab 4, Col E sample	170	186	161	187	198	198							1,100
Tab 4, Col F sample	2,079	2,274	1,898	2,087	2,052	1,864	1,311	950	615	318	289		15,738

Notes: This table presents the stacked dataset for our event studies (Figure 3) and robustness tests (Table 4). Columns (A) and (B) show the pairwise combinations of ENEM adoption years,  $\tau_s^*$  and  $\tau_{s'}^*$ . Columns (C) through (N) indicate the number of observations (in thousands) or each pair by exam year. The boxed cells show the sample for both Figure 3 and column (D) of Table 4, defined by the pairwise treatment effects that we can estimate using 2009–2017 exam takers. The bold cells show the sample for column (E) of Table 4, which includes 2007–2012 exam takers in the 2009 vs. 2013 pair of ENEM adoptions years. The sample for column (F) of Table 4 includes all cells in this table. The bottom three rows summarize the totals for each sample by exam year.

TABLE A5. The impact of ENEM adoption on private/public test score gaps using alternative samples

	(A)	(B)	(C)	(D)	(E)
Dependent variable	Main sample	Appear any year	Appear pre-ENEM	Participation pre-ENEM	Participation all years
Math score	0.158* (0.079)	0.097 (0.058)	0.112* (0.060)	0.158** (0.070)	0.127* (0.069)
Language arts score	0.076*** (0.026)	0.048*** (0.017)	0.054*** (0.013)	0.088*** (0.028)	0.072*** (0.024)
Natural science score	0.065* (0.034)	0.047 (0.032)	0.046 (0.030)	0.102** (0.036)	0.080** (0.035)
Social science score	0.081*** (0.023)	0.066** (0.026)	0.069*** (0.025)	0.119** (0.044)	0.103** (0.043)
Average score (core subjects)	0.110** (0.040)	0.074** (0.035)	0.081** (0.032)	0.135*** (0.043)	0.110** (0.040)
Writing score	0.102* (0.058)	0.124** (0.048)	0.129** (0.049)	0.037 (0.105)	0.006 (0.119)
<i>N</i> (# exam takers)	2,512,214	10,991,098	6,774,892	856,165	718,748
Appear in INEP:	All years	Any year	At least 2005–2008	At least 2005–2008	All years
Participation rate:				≥ 50% in 2005–2008	≥ 50% in all years

*Notes:* This table examines the robustness of our estimates of the effects of ENEM adoption on the private/public school test score gaps to using alternative samples.

The dependent variables are ENEM subject scores in SD units. “Average score (core subjects)” is the average score across math, language arts, natural science, and social science.

Each column shows  $\beta^{\text{gap}}$  coefficients from equation (2) estimated on different samples of high schools. Column (A) reproduces the estimates in the main sample, which consists of high schools that appeared in the INEP report in each year from 2005 to 2015. Column (B) relaxes the selection criterion and includes high schools that appeared at least once in the INEP report from 2005 to 2015. Column (C) includes high schools that appeared in the INEP report every year in the pre-ENEM period, i.e., from 2005 to 2008. In addition to the requirements in column (C), column (D) requires a participation rate of over 50% every year in the pre-ENEM period, meaning that at least 50% of the seniors in the high school took the ENEM in those years. Finally, column (E) includes high schools listed in the INEP report every year from 2005 to 2015 and maintained a participation rate of over 50% in all years.

Note that in certain years, INEP required a minimum participation rate for high schools to be included in the report. As a result, columns (A) to (C) inherently satisfy this participation requirement. The criteria in columns (D) and (E) impose additional participation requirements beyond the inherent INEP requirements. Details on the INEP requirement on participation rate can be found in Appendix C.3.

Parentheses contain standard errors clustered at the state level. \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .

TABLE A6. The impact of ENEM adoption on demographic test score gaps

	(A)	(B)	(C)	(D)	(E)	(F)
Dependent variable	Private/ public school	White/ non-white	Male/ female	Mother college/ no college	Father college/ no college	High/low income
Math score	0.158* (0.079)	0.069* (0.038)	-0.017 (0.016)	0.111* (0.058)	0.082 (0.060)	0.125* (0.067)
Language arts score	0.076*** (0.026)	0.027 (0.021)	-0.007 (0.012)	0.047 (0.034)	0.011 (0.034)	0.050* (0.025)
Natural science score	0.065* (0.034)	0.041* (0.023)	-0.019* (0.011)	0.039 (0.031)	-0.001 (0.033)	0.055** (0.025)
Social science score	0.081*** (0.023)	0.060*** (0.020)	-0.017 (0.011)	0.067** (0.030)	0.022 (0.028)	0.073*** (0.024)
Average score (core subjects)	0.110** (0.040)	0.057** (0.024)	-0.017 (0.012)	0.076* (0.041)	0.033 (0.039)	0.088** (0.038)
Writing score	0.102* (0.058)	0.005 (0.033)	-0.014 (0.014)	0.052 (0.039)	0.066 (0.045)	0.061* (0.032)
<i>N</i> (# exam takers)	2,512,214	2,387,052	2,512,214	2,489,743	2,489,191	2,487,270
Low-stakes average score gap	1.229	0.477	0.271	0.882	1.027	0.892

*Notes:* This table shows estimates of the impact of ENEM adoption on test score gaps between students from different demographic groups, as listed in the column headers.

The table displays  $\beta^{\text{gap}}$  coefficients from versions of equation (2) estimated at the individual level. The dependent variables are ENEM subject scores in SD units. “Average score (core subjects)” is the average score across math, language arts, natural science, and social science. We estimate the regressions using data from 2009–2017 ENEM exam takers in our analysis sample (column C of Table 1). Some columns have slightly smaller sample sizes due to missing values of demographic variables.

Column (A) reproduces our benchmark  $\beta^{\text{gap}}$  coefficients from column (E) of Table 3. Columns (B)–(F) display  $\beta^{\text{gap}}$  coefficients from a specification of equation (2) that replaces the  $\text{Private}_h$  dummy with an indicator for being in the more advantaged demographic group. Column (B) defines “non-white” as students who self-identify as Black, Brown, or Indigenous. Columns (D) and (E) define “college” as having a college or post-graduate degree. Column (F) defines a student as “high-income” if his/her reported family income was greater than or equal to twice the minimum wage in the exam year.

The bottom row shows the gap in the average score (core subjects) in exam cohorts prior to each state’s ENEM adoption year (i.e., cohorts with  $\text{HighStakes}_{st} = 0$ ). Parentheses contain standard errors clustered at the state level.

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

TABLE A7. Clustered and wild  $t$  bootstrap  $p$  values for effects of ENEM adoption on test scores

Dependent variable	(A)	(B)	(C)	(D)
	DD coefficients			
	All schools	Private schools	Public schools	Private/public gap
Math score	0.022 (0.690) [0.781]	0.143 (0.021) [0.090]	-0.015 (0.829) [0.857]	0.158 (0.056) [0.096]
Language arts score	0.035 (0.334) [0.438]	0.068 (0.002) [0.016]	-0.008 (0.820) [0.881]	0.076 (0.007) [0.002]
Natural science score	0.026 (0.520) [0.613]	0.062 (0.052) [0.264]	-0.003 (0.951) [0.885]	0.065 (0.071) [0.134]
Social science score	0.019 (0.581) [0.677]	0.056 (0.064) [0.356]	-0.024 (0.402) [0.450]	0.081 (0.002) [0.070]
Average score (core subjects)	0.029 (0.497) [0.633]	0.095 (0.013) [0.082]	-0.014 (0.752) [0.799]	0.110 (0.012) [0.036]
Writing score	0.049 (0.168) [0.128]	0.165 (0.031) [0.052]	0.063 (0.065) [0.118]	0.102 (0.090) [0.138]

*Notes:* This table examines the robustness of our estimates of the effects of ENEM adoption on test scores to the inference method. The point estimates reproduce the regression coefficients shown in columns (B) through (E) of Table 3. Parentheses contain  $p$  values derived from standard errors clustered at the state level, as in Table 3. Brackets show  $p$  values from a wild  $t$  bootstrap with 1000 replications that imposes the null hypothesis, as recommended by Cameron et al. (2008).

TABLE A8. Robustness to controls for affirmative action adoption

Covariates	(A)	(B)	(C)	(D)
	Benchmark model	Controls for affirmative action adoption		
		Federal universities	Federal & state univ.	All universities
<b>Panel A. Math score</b>				
ProportionENEM <sub>s(h)t</sub> × Private <sub>h</sub>	0.158* (0.079)	0.151* (0.078)	0.157* (0.079)	0.158* (0.080)
ProportionAA <sub>s(h)t</sub> × Private <sub>h</sub>		0.005 (0.088)	0.039 (0.122)	0.029 (0.116)
<b>Panel B. Language arts score</b>				
ProportionENEM <sub>s(h)t</sub> × Private <sub>h</sub>	0.076*** (0.026)	0.080*** (0.023)	0.074*** (0.024)	0.084*** (0.026)
ProportionAA <sub>s(h)t</sub> × Private <sub>h</sub>		-0.109** (0.051)	-0.105 (0.071)	-0.126 (0.076)
<b>Panel C. Natural science score</b>				
ProportionENEM <sub>s(h)t</sub> × Private <sub>h</sub>	0.065* (0.034)	0.062* (0.033)	0.063* (0.034)	0.066* (0.035)
ProportionAA <sub>s(h)t</sub> × Private <sub>h</sub>		-0.024 (0.061)	-0.043 (0.085)	-0.016 (0.072)
<b>Panel D. Social science arts score</b>				
ProportionENEM <sub>s(h)t</sub> × Private <sub>h</sub>	0.081*** (0.023)	0.075*** (0.024)	0.080*** (0.022)	0.076*** (0.025)
ProportionAA <sub>s(h)t</sub> × Private <sub>h</sub>		0.029 (0.063)	0.036 (0.089)	0.091 (0.080)
<b>Panel E. Average score (core subjects)</b>				
ProportionENEM <sub>s(h)t</sub> × Private <sub>h</sub>	0.110** (0.040)	0.106** (0.040)	0.108** (0.040)	0.111** (0.042)
ProportionAA <sub>s(h)t</sub> × Private <sub>h</sub>		-0.029 (0.067)	-0.021 (0.096)	-0.006 (0.085)
<b>Panel F. Writing score</b>				
ProportionENEM <sub>s(h)t</sub> × Private <sub>h</sub>	0.102* (0.058)	0.083 (0.052)	0.103* (0.054)	0.086 (0.054)
ProportionAA <sub>s(h)t</sub> × Private <sub>h</sub>		0.199** (0.078)	0.206* (0.106)	0.285** (0.133)
<i>N</i> (# exam takers)	2,512,214	2,512,214	2,512,214	2,512,214

*Notes:* This table examines the robustness of our results on private/public school test score gaps to controls for affirmative action adoption. The sample includes ENEM test takers in our high school senior sample (column C of Table 1). The dependent variables are the ENEM subject scores listed in the panel titles (in SD units). Column (A) replicates our benchmark results from column (E) of Table 3, which are the  $\beta^{\text{gap}}$  coefficients on the interaction between ProportionENEM<sub>s(h)t</sub> and a dummy for private high schools, Private<sub>h</sub> from equation (2). In columns (B)–(D) we add in a measure of the adoption of affirmative action at the state × year level, ProportionAA<sub>s(h)t</sub>, and its interaction with Private<sub>h</sub>. We compute ProportionAA<sub>s(h)t</sub> as the proportion of all new enrollees in state *s*(*h*) and year *t* who were admitting through reserved quotas using higher education census data. Columns (B)–(D) define ProportionAA<sub>s(h)t</sub> using only federal universities, federal and state universities, and all universities, respectively.

Parenteses contain standard errors clustered at the state level.

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

TABLE A9. The impact of ENEM adoption on the number of slots offered by federal universities

Dependent variable	(A)	(B)	(C)
	Low-stakes mean	DD coefficients	
		Continous treatment	Binary treatment
<b>Panel A. Institution <math>\times</math> program level</b>			
Log # admitted students	4.416	0.046 (0.093)	0.007 (0.045)
Prop. admitted through reserved quotas	0.155	0.032 (0.047)	0.004 (0.028)
$N$ (# institutions $\times$ programs $\times$ years)	10,544	39,027	39,027
<b>Panel B. Institution level</b>			
Log # admitted students	8.693	-0.111 (0.081)	-0.015 (0.032)
Prop. admitted through reserved quotas	0.155	0.037 (0.051)	0.008 (0.031)
$N$ (# institutions $\times$ years)	148	553	553
<b>Panel C. State level</b>			
Log # admitted students	9.527	-0.083 (0.089)	-0.007 (0.035)
Prop. admitted through reserved quotas	0.155	0.031 (0.047)	0.007 (0.030)
$N$ (# states $\times$ years)	71	243	243

*Notes:* This table presents estimates of the impact of ENEM adoption on the number of admission slots offered by federal universities.

The dependent variables are: (1) the logarithm of the number of admitted students (a proxy for the number of slots); and (2) the proportion of students admitted through quotas reserved for low-income or minority students. We compute each variable at three levels: institution  $\times$  program  $\times$  year (Panel A), institution  $\times$  year (Panel B), and state  $\times$  year (Panel C).

Column (A) shows the mean of each variable for admission cohorts prior to each state's ENEM adoption year (i.e., cohorts with  $\text{HighStakes}_{st} = 0$ ). Columns (B) and (C) display  $\beta$  coefficients from a specification of equation (1) estimated at the level corresponding to the one the dependent variable is defined. Column (B) uses the continuous treatment variable,  $\text{ProportionENEM}_{st}$ , while column (C) uses the binary treatment variable,  $\text{HighStakes}_{st}$ .

Regressions for the log number of admitted students use fixed weights over time, based on the number of admitted students in 2009. Regressions for the proportion of students admitted through quotas reserved for low-income or minority students are weighted by the number of admitted students in each respective year.

Parentheses contain standard errors clustered at the state level. \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .

TABLE A10. The impact of ENEM adoption on test score gaps by test prep activity

	(A)	(B)	(C)	(D)
Dependent variable	Private/ public school	Prep/ public school	Prep/ other private school	Took/didn't take a prep course
Math score	0.158* (0.079)	0.202** (0.094)	0.052 (0.052)	0.215** (0.080)
Language arts score	0.076*** (0.026)	0.115** (0.045)	0.040 (0.031)	0.194** (0.083)
Natural science score	0.065* (0.034)	0.134*** (0.043)	0.075** (0.028)	0.185** (0.076)
Social science score	0.081*** (0.023)	0.172*** (0.030)	0.097*** (0.030)	0.212** (0.089)
Average score (core subjects)	0.110** (0.040)	0.180*** (0.051)	0.076** (0.034)	0.232** (0.093)
Writing score	0.102* (0.058)	0.244 (0.153)	0.154 (0.124)	0.028 (0.045)
<i>N</i> (# exam takers)	2,512,214	1,779,119	807,293	500,938
Low-stakes average score gap	1.229	1.147	-0.087	0.371

*Notes:* This table presents estimates of the impact of ENEM adoption on test score gaps between students who did/did not engage in test prep activities.

We define two different measures of students’ test prep activity. Our first measure of test prep activity defines a set of “prep schools” with curricula specifically focused on preparation for college admission exams. We obtained lists of schools that use test-oriented curricula from the websites of four prominent test prep companies: *Sistema Anglo*, *Sistema pH*, *Elite Rede de Ensino*, and *Curso Objetivo*. These lists provided the names and street addresses of each test-oriented school. We then obtained names and addresses for each school in our sample from INEP’s administrative [school catalog](#) data. We geocoded the addresses in both datasets using the `ggmap` package for *R*. Finally, we matched schools whose addresses were within 100 meters of each other and manually checked all matches using school names. We define schools that matched using this procedure as “prep schools” (all of the matched schools are private high schools).

Our second measure of test prep activity is based on a variable from the ENEM questionnaire that indicates whether individuals took an entrance exam preparation course. The question does not distinguish between courses focused on the ENEM exam and those focused on other *vestibular* exams. This question is only available in the 2009–2011 and 2013–2014 cohorts, and there are missing values for many students in these cohorts.

The table displays  $\beta^{\text{gap}}$  coefficients from specifications of equation (2) estimated at the individual level. Column (A) reproduces our benchmark  $\beta^{\text{gap}}$  coefficients from column (E) of Table 3. Columns (B) to (D) present  $\beta^{\text{gap}}$  coefficients from specifications of equation (2) where we replace the  $\text{Private}_h$  dummy with an indicator for prep schools (columns B and C) or with an indicator for individuals who took a prep course (column D). The bottom row shows the gap in the average score (core subjects) for exam cohorts prior to each state’s ENEM adoption year (i.e., cohorts with  $\text{HighStakes}_{st} = 0$ ). The dependent variables are ENEM subject scores in SD units. “Average score (core subjects)” is the average score across math, language arts, natural science, and social science. We estimate the regressions using data from 2009–2017 ENEM exam takers in our analysis sample (column C of Table 1)

Parentheses contain standard errors clustered at the state level. \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .

TABLE A11. Clustered and wild  $t$  bootstrap  $p$  values for effects of ENEM adoption on the informativeness of test scores for longer-run outcomes

Dependent variable	(A)	(B)	(C)	(D)
	Benchmark model DD coefficients		Binary treatment DD coefficients	
	Raw corr.	Within- program	Raw corr.	Within program
Enrolled in any college by 2019	0.036 (0.000) [0.000]		0.022 (0.000) [0.000]	
Finished college within 5 years of ENEM	0.014 (0.091) [0.096]	0.026 (0.001) [0.028]	0.007 (0.116) [0.152]	0.015 (0.010) [0.044]
Earned a college degree by 2019	0.033 (0.037) [0.144]	0.038 (0.001) [0.016]	0.016 (0.179) [0.244]	0.023 (0.005) [0.032]
Appears in RAIS in 2016–2018	0.056 (0.209) [0.501]	0.020 (0.095) [0.238]	0.014 (0.567) [0.675]	0.008 (0.310) [0.352]
Persisted in college for 1 year	0.008 (0.581) [0.563]	0.024 (0.004) [0.014]	0.007 (0.564) [0.595]	0.011 (0.097) [0.092]
Persisted in college for 3 years	0.033 (0.017) [0.056]	0.043 (0.000) [0.000]	0.018 (0.047) [0.064]	0.024 (0.000) [0.008]
Completed program within 5 years	0.044 (0.025) [0.140]	0.035 (0.003) [0.030]	0.022 (0.063) [0.092]	0.018 (0.019) [0.042]
Fraction of college credits completed	0.003 (0.812) [0.785]	0.014 (0.292) [0.382]	−0.011 (0.330) [0.342]	0.013 (0.064) [0.126]
Hourly wage (BRL)	0.046 (0.017) [0.176]	0.027 (0.013) [0.180]	0.027 (0.001) [0.036]	0.012 (0.061) [0.128]
Log hourly wage	−0.029 (0.050) [0.284]	−0.001 (0.900) [0.905]	−0.017 (0.053) [0.136]	−0.001 (0.866) [0.889]

*Notes:* This table examines the robustness of our estimates of the effects of ENEM adoption on the informativeness of test scores for student longer-run outcomes to the inference method. The point estimates reproduce the regression coefficients shown in columns (B) through (E) of Table 5. Parentheses contain  $p$  values derived from standard errors clustered at the state level, as in Table 5. Brackets show  $p$  values from a wild  $t$  bootstrap with 1000 replications that imposes the null hypothesis, as recommended by Cameron et al. (2008).

TABLE A12. Effects of ENEM adoption on the correlation of average ENEM scores and outcomes controlling for demographics

	(A)	(B)	(C)	(D)	(E)
	Control variables (cumulative)				
Dependent variable: Correlation between average ENEM scores and...	No controls	Private HS	+ SES & race	+ Age & gender	HS dummies
<b>Panel A. Outcomes for all exam takers</b>					
Enrolled in any college by 2019	0.036*** (0.006)	0.028*** (0.008)	0.036*** (0.007)	0.031*** (0.011)	0.028*** (0.009)
Finished college within 5 years of ENEM	0.014* (0.008)	0.016** (0.006)	0.012 (0.008)	0.005 (0.011)	0.008 (0.008)
Earned a college degree by 2019	0.033** (0.015)	0.017 (0.013)	0.014 (0.017)	0.009 (0.019)	0.012 (0.014)
Appears in RAIS in 2016–2018	0.056 (0.044)	0.009 (0.024)	0.009 (0.022)	0.009 (0.021)	0.012 (0.019)
<i>N</i> (# exam takers)	1,257,079	1,257,079	1,257,079	1,257,079	1,257,079
<b>Panel B. Outcomes for college enrollees</b>					
Persisted in college for 1 year	0.008 (0.014)	0.013 (0.010)	0.012 (0.009)	0.013 (0.009)	−0.000 (0.010)
Persisted in college for 3 years	0.033** (0.013)	0.023** (0.009)	0.021** (0.008)	0.022** (0.008)	0.006 (0.011)
Completed program within 5 years	0.044** (0.018)	0.026 (0.018)	0.018 (0.020)	0.020 (0.021)	0.008 (0.015)
Fraction of college credits completed	0.004 (0.014)	−0.008 (0.013)	−0.007 (0.013)	−0.006 (0.013)	−0.002 (0.017)
<i>N</i> (# in higher ed.)	959,339	959,339	959,339	959,339	959,339
<b>Panel C. Outcome for individuals in RAIS</b>					
Hourly wage (BRL)	0.047** (0.018)	0.049*** (0.016)	0.040** (0.015)	0.046*** (0.016)	0.053*** (0.016)
Log hourly wage	−0.029** (0.014)	−0.037** (0.015)	−0.026** (0.012)	−0.013 (0.015)	0.007 (0.010)
<i>N</i> (# in RAIS)	326,827	326,827	326,827	326,827	326,703

*Notes:* This table present estimates of the impacts of ENEM adoption on the correlation between average ENEM scores and student outcomes, controlling for demographic characteristics.

The table shows  $\beta$  coefficients from equation (5) estimated on the analysis sample. Column (A) reproduces the coefficients from column (B) of Table 5, where the dependent variable,  $Y_{st}$ , is the raw correlation coefficient between average (core subjects) ENEM scores and the longer-run outcomes of individuals who attended high school in state  $s$  and took the ENEM in year  $t$ . In columns (B)–(E), the dependent variables are state  $\times$  year correlation coefficients between ENEM scores and outcomes after residualizing all variables on demographic controls. These columns incrementally include four sets of demographic controls: a dummy for private high schools (column B); dummies for family income categories, mother’s education, father’s education, and race (column C); a gender dummy and a linear term in age (column D); and high school dummies (column E). Each column includes the new controls plus all controls from the preceding columns.

Parentheses contain standard errors clustered at the state level. \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .

TABLE A13. Effects of ENEM adoption on exam performance by topic area in language arts, natural science, and social science

(A)	(B)	(C)	(D)	(E)	(F)					
						Outcome: Proportion of correct answers				
						Public students		Private/public gap		
Question group	$N_q$	Mean	$\beta$ (SE)	Mean	$\beta^{\text{gap}}$ (SE)					
<b>Panel A. Language arts</b>										
Communication (1–4)	54	0.468	−0.003 (0.009)	0.193	0.019 (0.009)**					
Foreign language (5–8)	80	0.424	−0.009 (0.013)	0.234	−0.005 (0.008)					
Body language (9–11)	29	0.521	−0.018 (0.010)*	0.139	0.016 (0.013)					
Art (12–14)	42	0.425	0.005 (0.012)	0.188	0.019 (0.007)**					
Literary text (15–17)	63	0.342	−0.004 (0.006)	0.157	0.022 (0.006)***					
Linguistics (18–20)	41	0.449	−0.003 (0.015)	0.193	0.020 (0.012)					
Argumentation (21–24)	63	0.450	0.000 (0.005)	0.171	0.013 (0.006)**					
Portuguese (25–27)	39	0.410	−0.011 (0.008)	0.205	0.017 (0.010)					
Social communication (28–30)	33	0.455	0.001 (0.011)	0.186	0.024 (0.007)***					
All coefficients equal ( $p$ value)			0.034	0.021						
<b>Panel B. Natural science</b>										
Human constructions (1–4)	57	0.274	0.002 (0.009)	0.166	0.013 (0.006)**					
Technology (5–7)	34	0.255	0.005 (0.005)	0.102	0.010 (0.004)**					
Environmental conservation (8–12)	58	0.360	0.005 (0.006)	0.178	0.004 (0.005)					
Ecosystems (13–16)	54	0.316	−0.004 (0.008)	0.204	0.018 (0.011)					
Scientific methods (17–19)	49	0.297	−0.006 (0.010)	0.189	0.006 (0.008)					
Physics (20–23)	57	0.269	−0.010 (0.008)	0.158	0.009 (0.004)*					
Chemistry (24–27)	65	0.246	−0.001 (0.002)	0.127	0.006 (0.007)					
Biology (28–30)	31	0.416	−0.012 (0.014)	0.192	0.018 (0.014)					
All coefficients equal ( $p$ value)			0.023	0.045						
<b>Panel C. Social science</b>										
Culture (1–5)	73	0.385	−0.011 (0.006)*	0.184	0.022 (0.007)***					
Geography (6–10)	66	0.353	0.004 (0.006)	0.214	0.006 (0.006)					
Social institutions (11–15)	74	0.376	−0.006 (0.008)	0.173	0.021 (0.005)***					
Technology (16–20)	61	0.376	−0.010 (0.005)*	0.186	0.009 (0.007)					
Citizenship (21–25)	62	0.432	0.002 (0.007)	0.192	0.015 (0.009)					
Society and nature (26–30)	69	0.405	−0.001 (0.008)	0.168	0.013 (0.007)*					
All coefficients equal ( $p$ value)			0.000	0.000						

*Notes:* This table presents estimates of the impacts of ENEM adoption on performance on different topic areas of the language arts (Panel A), natural science (Panel B), and social science (Panel C) tests.

This table is analogous to the math test results in Panel B of Table 6. The sample includes 2009–2017 ENEM exam takers in our analysis sample (column C of Table 1). Regressions are at the high school ( $h$ )  $\times$  year ( $t$ )  $\times$  exam question ( $q$ ) level. The dependent variable is the proportion of correct answers in each  $htq$  cell. Each row displays results from a separate regression for questions in the same topic area, as defined by ENEM test designers.

Column (A) specifies the group of questions contained in each regression. Column (B) indicates the number of questions in each group. Column (C) shows the mean proportion of correct answers for public school students in cohorts prior to each state’s ENEM adoption year (i.e., cohorts with  $\text{HighStakes}_{st} = 0$ ). Column (E) shows the mean private/public gap in the proportion of correct answers in those cohorts. Columns (D) and (F) display the  $\beta$  and  $\beta^{\text{gap}}$  coefficients from equation (2) estimated for each group of questions. The last row of each panel reports  $p$  values from  $F$  tests that the topic area coefficients in columns (D) or (F) are jointly equal.

Parentheses contain standard errors clustered at the state level. \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .

TABLE A14. The relationship between skill informativeness for longer-run outcomes and the impact of ENEM stakes on private/public skill gap

	(A)	(B)	(C)	(D)	(E)	(F)
	Mean informativeness	Change in return to a correct answer from 1 SD increase in $\beta^{\text{gap}}$				
Dependent variable: Informativeness for...	All subjs	Math	Lang. arts	Nat. science	Soc. science	All subjs
<b>Panel A. Outcomes for all exam takers</b>						
Enrolled in any college by 2019	0.101	0.015** (0.007)	0.012** (0.005)	0.000 (0.006)	-0.000 (0.003)	0.006** (0.003)
Finished college within 5 years of ENEM	0.029	0.004* (0.002)	0.005*** (0.002)	0.001 (0.002)	0.000 (0.001)	0.002 (0.001)
Earned a college degree by 2019	0.071	0.011** (0.005)	0.010*** (0.003)	0.002 (0.005)	-0.001 (0.002)	0.004* (0.002)
Appears in RAIS in 2016–2018	-0.016	-0.000 (0.002)	0.008*** (0.002)	0.000 (0.001)	-0.001 (0.001)	0.002* (0.001)
<i>N</i> (# competencies)	120	30	30	30	30	120
<b>Panel B. Outcomes for college enrollees</b>						
Persisted in college for 1 year	0.012	0.002** (0.001)	0.001 (0.001)	0.000 (0.001)	-0.000 (0.000)	0.001* (0.000)
Persisted in college for 3 years	0.047	0.007*** (0.002)	0.003 (0.002)	0.000 (0.002)	-0.001 (0.001)	0.002** (0.001)
Completed program within 5 years	0.023	0.003 (0.002)	0.005*** (0.002)	0.001 (0.002)	0.001 (0.001)	0.002 (0.001)
Fraction of college credits completed	0.048	0.007*** (0.003)	0.006*** (0.002)	0.001 (0.002)	-0.001 (0.001)	0.003** (0.001)
<i>N</i> (# competencies)	120	30	30	30	30	120
<b>Panel C. Outcome for individuals in RAIS</b>						
Hourly wage (BRL)	13.285	3.052*** (0.726)	1.766*** (0.604)	0.039 (0.730)	-0.893 (0.538)	0.831** (0.413)
Log hourly wage	0.153	0.034*** (0.008)	0.021*** (0.007)	0.002 (0.008)	-0.008 (0.006)	0.011** (0.004)
<i>N</i> (# competencies)	120	30	30	30	30	120

*Notes:* This table presents OLS estimates of the relationship between the informativeness of ENEM competencies for longer-run outcomes and the change in the private/public gap in performance on the higher-stakes exam.

We measure the informativeness of each exam competency for a given outcome as the mean difference in a given outcome between individuals who got each question correct and incorrect using 2009–2014 (excluding 2011) ENEM participants in our analysis sample, and then average these differences across all questions in the same competency. For outcomes in Panel C, the sample excludes 2013–2014 ENEM participants. To measure the effect of ENEM stakes on the private/public gap,  $\beta^{\text{gap}}$ , we estimate equation (2) separately for groups of ENEM questions in each competency (as in Table 6, Panel C, column F). These coefficients represent the change in the private/public gap in the proportion of correct answers resulting from an increase in ENEM stakes.

Column (A) shows the mean informativeness averaged across all subjects and competencies. Columns (B)–(F) show OLS coefficients from bivariate regressions of informativeness on the  $\beta^{\text{gap}}$  coefficients, estimated by subject and pooling across all subjects. We normalize the OLS coefficients to represent a 1 SD increase in  $\beta^{\text{gap}}$ .

Parentheses contain robust standard errors. \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .

TABLE A15. The relationship between skill informativeness for longer-run outcomes and the impact of ENEM stakes on private/public skill gap, controlling for question characteristics

Dependent variable: Informativeness for...	(A)	(B)	(C)	(D)
	Mean informativeness	Change in return to a correct answer from 1 SD increase in $\beta^{\text{gap}}$		
		Bivariate regression	Controls for % correct	IRT controls
<b>Panel A. Outcomes for all exam takers</b>				
Enrolled in any college by 2019	0.1011	0.0063** (0.0031)	0.0011 (0.0010)	0.0028 (0.0020)
Finished college within 5 years of ENEM	0.0294	0.0020 (0.0013)	0.0009** (0.0004)	0.0011* (0.0006)
Earned a college degree by 2019	0.0707	0.0045* (0.0023)	0.0026** (0.0011)	0.0029* (0.0015)
Appears in RAIS in 2016–2018	−0.0162	0.0018* (0.0011)	0.0017 (0.0010)	0.0001 (0.0010)
<i>N</i> (# competencies)	120	120	120	120
<b>Panel B. Outcomes for college enrollees</b>				
Persisted in college for 1 year	0.0123	0.0007* (0.0003)	0.0000 (0.0001)	0.0003 (0.0003)
Persisted in college for 3 years	0.0472	0.0023** (0.0011)	0.0006 (0.0004)	0.0014* (0.0008)
Completed program within 5 years	0.0225	0.0017 (0.0012)	0.0012** (0.0005)	0.0013* (0.0008)
Fraction of college credits completed	0.0481	0.0027** (0.0012)	0.0017*** (0.0006)	0.0019** (0.0009)
<i>N</i> (# competencies)	120	120	120	120
<b>Panel C. Outcome for individuals in RAIS</b>				
Hourly wage (BRL)	13.2852	0.8314** (0.4131)	0.8954** (0.4140)	0.7313* (0.4185)
Log hourly wage	0.1534	0.0110** (0.0043)	0.0098** (0.0040)	0.0085** (0.0040)
<i>N</i> (# competencies)	120	120	120	120

*Notes:* This table presents OLS estimates of the relationship between the informativeness of ENEM competencies for longer-run outcomes and the change in the private/public performance gap on the higher-stakes exam, controlling for question characteristics.

This table is analogous to Table A14, but includes specifications that control for question characteristics. Column (B) reproduces the OLS estimates shown in column (F) of Table A14. Column (C) adds a quadratic term for the mean proportion of correct answers in each competency, interacted with subject dummies. Column (D) includes a quadratic term for the mean IRT parameters in each competency, interacted with subject dummies. We normalize the OLS coefficients to represent a 1 SD increase in  $\beta^{\text{gap}}$ .

Parentheses contain robust standard errors. \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .

TABLE A16. Heterogeneity by topics in the *Me Salva!* study guide (math test)

(A) Question group	(B) $N_q$	(C)–(F) Outcome: Proportion of correct answers			
		Public students		Private/public gap	
		Mean	$\beta$ (SE)	Mean	$\beta^{\text{gap}}$ (SE)
<b>Panel A. Study guide words matched to question text</b>					
Not in study guide	262	0.306	−0.007 (0.011)	0.186	0.026 (0.010)**
In study guide	143	0.267	−0.003 (0.008)	0.161	0.021 (0.008)**
cone ( <i>cone</i> )	8	0.471	−0.010 (0.026)	0.210	0.024 (0.030)
cube ( <i>cubo</i> )	8	0.179	−0.000 (0.008)	0.182	0.029 (0.010)***
cylinder ( <i>cilindro</i> )	10	0.374	−0.011 (0.013)	0.157	0.021 (0.022)
directly ( <i>diretamente</i> )	8	0.284	0.009 (0.011)	0.218	0.030 (0.010)***
median ( <i>mediana</i> )	12	0.226	0.009 (0.010)	0.130	0.055 (0.021)**
parallelepiped ( <i>paralelepípedo</i> )	9	0.314	0.001 (0.034)	0.342	0.043 (0.028)
possibilities ( <i>possibilidade</i> )	8	0.275	−0.015 (0.006)**	0.126	0.021 (0.017)
possible ( <i>possíveis</i> )	10	0.310	−0.015 (0.023)	0.168	0.011 (0.022)
prism ( <i>prisma</i> )	8	0.368	−0.002 (0.015)	0.136	0.014 (0.025)
probability ( <i>probabilidade</i> )	25	0.234	−0.010 (0.008)	0.131	0.028 (0.010)**
pyramid ( <i>pirâmide</i> )	10	0.337	−0.035 (0.018)*	0.138	0.042 (0.017)**
rectangle ( <i>retângulo</i> )	10	0.277	0.006 (0.011)	0.192	0.024 (0.014)*
square ( <i>quadrado</i> )	35	0.256	−0.001 (0.005)	0.197	0.009 (0.006)
triangle ( <i>triângulo</i> )	8	0.231	−0.002 (0.006)	0.120	0.038 (0.010)***
Fewer than 8 occurrences	40	0.268	−0.005 (0.007)	0.146	0.016 (0.008)*
In vs. not in study guide ( $p$ value)			0.310	0.157	
<b>Panel B. Study guide concepts matched to solutions</b>					
Not in study guide	231	0.317	−0.005 (0.010)	0.193	0.023 (0.011)**
In study guide	173	0.259	−0.004 (0.009)	0.155	0.025 (0.008)***
Geometric formulas	64	0.261	−0.004 (0.009)	0.174	0.026 (0.007)***
Proportions (“Rule of 3”)	12	0.330	−0.011 (0.018)	0.205	0.058 (0.010)***
Manipulating fractions	58	0.248	−0.004 (0.008)	0.161	0.029 (0.009)***
Radicals	13	0.195	0.009 (0.006)*	0.139	0.024 (0.014)*
Combinatory and statistical analysis	25	0.224	−0.003 (0.004)	0.103	0.024 (0.015)
Probability	36	0.255	−0.010 (0.010)	0.138	0.025 (0.007)***
Trigonometric formulas	8	0.231	0.006 (0.006)	0.104	0.017 (0.007)**
In vs. not in study guide ( $p$ value)			0.791	0.591	

*Notes:* This table presents estimates of the impact of ENEM adoption on student performance on math questions covered in a study guide by the test prep company *Me Salva!*. The sample includes 2009–2017 ENEM exam takers in our analysis sample (column C of Table 1). Regressions are at the high school ( $h$ )  $\times$  year ( $t$ )  $\times$  exam question ( $q$ ) level. The dependent variable is the proportion of correct answers in each  $htq$  cell. Panel A groups question based on whether the question text contains keywords from the *Me Salva!* study guide, restricting to words that appear in eight or more questions. Panel B groups questions based on whether the solution text (prepared by another test prep company, *Descomplica*) requires concepts from the *Me Salva!* study guide. See Appendix C.4 for details on these data sources and our match between *Me Salva!* topics and ENEM questions.

Column (A) specifies the group of questions for each regression. Column (B) indicates the number of questions in each group. Column (C) shows the mean proportion of correct answers for public school students in cohorts prior to each state’s ENEM adoption year (i.e., cohorts with  $\text{HighStakes}_{st} = 0$ ). Column (E) shows the mean private/public gap in the proportion of correct answers in those cohorts. Columns (D) and (F) display the  $\beta$  and  $\beta^{\text{gap}}$  coefficients from equation (2) estimated for each group of questions. The last row of each panel reports  $p$  values from an  $F$  test that the coefficients in the first and second rows are equal.

Parentheses contain standard errors clustered at the state level. \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .

## B. THEORETICAL APPENDIX

**B.1. Impacts of exam stakes on test score gaps and informativeness.** This section presents a simple framework that illustrates how exam stakes can impact inequality in test scores and their informativeness for longer-run outcomes. This framework helps to interpret our main empirical results and shed light on the potential mechanisms.

B.1.1. *Model setup.* We consider a population of exam takers that are characterized by their socioeconomic status (SES) and their abilities. Let  $X_i$  denote an observable measure of individual  $i$ 's SES, such as an indicator for attending a private high school. We let  $a_i$  denote individual  $i$ 's ability to perform well on a college admission exam, which is not directly observable in data. We refer to  $a_i$  as *test ability* to emphasize that it may be distinct from other abilities that help the individual perform well in college and the labor market.

We assume that the stakes of the college admission exam vary across cohorts, and that test takers in high-stakes cohorts engage in more test prep. For simplicity, we assume that individuals are randomly assigned to either a low-stakes or a high-stakes exam cohort, and we let  $H_i$  be a binary indicator for the high-stakes cohort.<sup>31</sup> Individuals have a stronger incentive to perform well in the high-stakes exam cohort, and so we assume that these individuals engage in additional test prep. This additional prep increases their test ability by an amount that we denote by  $e_i$ . We interpret  $e_i$  broadly; for example, it may include studying test prep books, taking preparatory courses, focusing more intently during the school year, or exerting more effort on the exam.

A test taker's preparedness for the exam, as a function of their cohort, is given by:

$$(B1) \quad \theta_i = a_i + H_i e_i.$$

where we refer to  $\theta_i$  as *test skill*. Individuals in low-stakes cohorts have test skill  $\theta_i = a_i$ , while individuals in high-stakes cohorts have test skill  $\theta_i = a_i + e_i$ .

An individual's score on the college admission exam is a noisy measure of their test skill. We let  $T_i$  denote individual  $i$ 's test score, which is observable in the data. We assume test scores are given by:

$$(B2) \quad T_i = \theta_i + \epsilon_i^T,$$

where  $\epsilon_i^T$  is random noise that reflects variation in test performance due to factors like guessing and health on exam day.

We let  $Y_i$  denote an observable measure of college success, such as persisting in college after enrolling or completing a college degree. We model  $Y_i$  as a function of test ability,  $a_i$ ,

---

<sup>31</sup> Our empirical analysis relies on a parallel trends version of this assumption.

test prep,  $e_i$ , and SES,  $X_i$ :

$$(B3) \quad Y_i = \alpha a_i + \beta e_i + \gamma X_i + v_i + \epsilon_i^Y.$$

The parameter  $\alpha$  represents the effect of test ability on college success; it is natural to assume  $\alpha > 0$  since individuals with high test scores tend to perform better in college. Similarly, an individual's SES may help them succeed in college above and beyond their test ability, and so we assume  $\gamma > 0$ . The parameter  $\beta$  allows for the possibility that test prep directly affects college performance, although, as we discuss below, the sign of this parameter is less clear. In addition to these three factors, we allow college success to depend on other abilities that are unrelated to test skill, which we denote by  $v_i$ , and a random noise term,  $\epsilon_i^Y$ .

B.1.2. *Effects of exam stakes on test score gaps.* In Section 4, we find that increasing the stakes of a college admission exam increases test score gaps between high- and low-SES students. If  $X_i$  is a binary indicator for high SES, then this result can be written as:

$$(B4) \quad \mathbb{E}[T_i|X_i = 1, H_i = 1] - \mathbb{E}[T_i|X_i = 0, H_i = 1] > \mathbb{E}[T_i|X_i = 1, H_i = 0] - \mathbb{E}[T_i|X_i = 0, H_i = 0].$$

In other words, the mean difference in test scores,  $T_i$ , between high- and low-SES exam takers is larger in the high-stakes cohort ( $H_i = 1$ ) than in the low-stakes cohort ( $H_i = 0$ ).

Plugging in equations (B1)–(B2) and using the factor that  $H_i$  is independent of  $a_i$ , expression (B4) becomes:

$$(B5) \quad \mathbb{E}[e_i|X_i = 1] - \mathbb{E}[e_i|X_i = 0] > 0.$$

Expression (B5) states that higher stakes exams lead to larger socioeconomic test score gaps increase if high SES students engage in more test prep than low SES students.

B.1.3. *Effects of exam stakes on informativeness.* In Section 5, we find that high-stakes exam scores are more informative for various measures of college success than lower-stakes exam scores. Our measure of informativeness is the correlation between exam scores,  $T_i$ , and a college outcome,  $Y_i$ , or:

$$(B6) \quad \text{Corr}(T_i, Y_i) = \frac{\text{Cov}(T_i, Y_i)}{\sigma_T \sigma_Y},$$

where  $\sigma_T$  and  $\sigma_Y$  are the standard deviations of test scores and outcomes. Ignoring these standard deviations for simplicity, our main result can be written as:

$$(B7) \quad \text{Cov}(T_i, Y_i|H_i = 1) > \text{Cov}(T_i, Y_i|H_i = 0).$$

Using equations (B1)–(B3), expression (B7) simplifies to:

$$(B8) \quad \gamma \text{Cov}(X_i, e_i) + (\alpha + \beta) \text{Cov}(a_i, e_i) + \text{Cov}(v_i, e_i) + \beta \text{var}(e_i) > 0.$$

Expression (B8) shows that there are two broad channels through which additional incentives for test prep could increase the predictive validity of exam scores:

- (1) **Correlational channels:**  $\gamma \text{Cov}(X_i, e_i) + (\alpha + \beta) \text{Cov}(a_i, e_i) + \text{Cov}(v_i, e_i) > 0$ .

First, high-stakes exam scores may be more predictive of college outcomes because they are more correlated with predetermined characteristics that also benefit students in college. These predetermined characteristics could include SES,  $X_i$ , test ability,  $a_i$ , or other unobserved abilities that impact college success,  $v_i$ . If the amount of test prep,  $e_i$ , is sufficiently correlated with any of these characteristics, this would lead test scores to be more informative for college outcomes.

Our results in Section 4 suggest that  $\text{Cov}(X_i, e_i) > 0$ , so all else equal, one would expect a correlation between test prep and SES to increase the informativeness of test scores. For example, wealthy students may have greater access to test prep services, and family wealth also helps students succeed in college.

The signs of the covariances between test prep,  $e_i$ , and the ability terms,  $a_i$  and  $v_i$ , are *a priori* less clear. On the one hand, students who know that they are likely to succeed in college may have the strongest incentives to prep for the higher-stakes exam, which would imply positive covariances between  $e_i$  and  $a_i$  or  $v_i$ . On the other hand, students with low test ability may have the greatest benefits from engaging in test prep, which would imply negative covariances. Thus we do not have a strong prior on the signs of the covariances between test prep and abilities.

- (2) **Skill accumulation channel:**  $\beta \text{var}(e_i) > 0$ .

A second potential mechanism is that test prep may be directly beneficial for college success. This channel is operative if the test skills that individuals accumulate from prepping,  $e_i$ , directly improve their college outcomes,  $Y_i$ . In our framework, this skill accumulation channel exists if  $\beta > 0$ . Brazil’s ENEM exam was designed to be closely aligned with high school and college curricula, so test designers certainly hoped that students would learn useful material through their exam preparation. On the other hand, high-stakes exams are often criticized for creating incentives to learn test-taking strategies that are not useful outside the exam. This perverse incentive could imply that  $\beta = 0$  or even  $\beta < 0$  if test prep crowds out other useful learning.<sup>32</sup>

---

<sup>32</sup> There is another potential channel that is outside the scope of our framework: high-stakes exam scores may be more informative for college outcomes if they increase the “match quality” between individuals and college programs. This channel could arise if test scores,  $T_i$ , have a causal effect on outcomes,  $Y_i$ , through their influence on which college and/or major students attend. In other words, high-stakes exam scores may be more informative for college success if the distribution of these scores leads to better student/college matches

Distinguishing between the correlational and skill accumulation channels is challenging because of unobserved abilities that impact both test scores and college success ( $a_i$  and  $v_i$ ). Figure 6 presents evidence that the increase in informativeness is likely to be partly driven by a correlational channel, as the correlation between test scores and outcomes decreases when we control for demographic characteristics. Yet we also find that the increase in exam stakes caused students to perform better across a wide range of exam skills that are related to high school and college curricula (Table 6). While this result may be due to either skill accumulation or a correlation with unobserved abilities, it shows that higher-stakes exams help colleges identify students who are more academically prepared for their programs.

## B.2. Relationship between correlation coefficients and Blackwell informativeness.

In our empirical analysis, we measure the informativeness of test scores using the correlation coefficient between student outcomes and test scores, which is different from the theoretical concept of Blackwell informativeness used in Frankel and Kartik (2019). This subsection explores the relationship between these two measures of informativeness. Specifically, we prove that with binary states and signals, an increase in a signal’s Blackwell informativeness leads to a higher correlation between states and signal realizations.

Suppose that both the state  $\Theta$  and the signal  $S$  take one of two values, high or low, coded as 1 and 0. We use  $\theta_1$  and  $\theta_0$  interchangeably with  $\Theta = 1$  and  $\Theta = 0$ , and similarly for  $s_1$  and  $s_0$ . The signal realizations are generated according to the following conditional probabilities:

		Signal $S$	
		0	1
State $\Theta$	0	$p_0$	$1 - p_0$
	1	$1 - p_1$	$p_1$

TABLE B1. Joint distribution of signals and states

Table B1 shows  $\Pr(s | \theta)$ , the conditional probability of observing a signal  $s$  given a state  $\theta$ . A signal is completely characterized by  $p_0$  and  $p_1$ . We require that the signal realizations be appropriately ordered:  $p_1 > 1 - p_0$  and  $p_0 > 1 - p_1$ , meaning the high signal realization occurs more frequently when the state is high compared to when the state is low, and vice versa for the low signal. This is likely true in the context of test scores as signals for abilities.

Following Frankel and Kartik (2019), we say that signal  $S$  is (Blackwell) more informative than signal  $S'$  if and only if for any prior, the distribution of posterior beliefs after observing  $S$  is a mean-preserving spread of the distribution of posterior beliefs after observing  $S'$ . In

---

(e.g., on the basis of academic preparation). To distinguish between this channel and the correlational/skill accumulation channels, we follow the standard practice that testing agencies use to measure predictive validity. Specifically, we present results that examine the correlation between  $T_i$  and  $Y_i$  after de-meaning both variables *within* college programs. These results reflect only variation in informativeness among students who enrolled in the same colleges and majors.

the context of binary signals, it is equivalent to saying that the interval between the two posteriors after observing the high and low realizations under  $S$  contains the interval between the two posteriors under  $S'$ .

We are interested in the relationship between Blackwell informativeness of a signal and the linear correlation between the signal realizations and the states. We show that when a signal becomes Blackwell more informative, the linear correlation increases. The following proposition formalizes this statement,

**Proposition 1.** *Consider a binary state  $\Theta$  and two binary signals  $S$  and  $S'$ . If  $S$  is more informative Blackwell than  $S'$ , then*

$$\text{Corr}(\Theta, S) > \text{Corr}(\Theta, S').$$

To prove Proposition 1, we express the correlation in terms of quantities that directly relate to a signal's Blackwell informativeness. We first note that the correlation coefficient can be written as

$$(B9) \quad \text{Corr}(\Theta, S) = \frac{\text{Cov}(\Theta, S)}{\sqrt{\text{Var}(\Theta)\text{Var}(S)}}$$

$$(B10) \quad = \frac{\sum_i (\theta_i - \mathbb{E}(\Theta))(s_i - \mathbb{E}(S))}{\sqrt{\sum_i (\theta_i - \mathbb{E}(\Theta))^2 (s_i - \mathbb{E}(S))^2}}.$$

For any prior  $\mu = \Pr(\theta = 1) = \Pr(\theta_1)$  and any binary signal  $S$ , denote the posterior after observing  $s_0$  as  $\mu_0 = \Pr(\theta_1 | s_0)$  and the posterior after observing  $s_1$  as  $\mu_1 = \Pr(\theta_1 | s_1)$ .

We want to express  $p_0$  and  $p_1$  in terms of  $\mu$ ,  $\mu_0$ , and  $\mu_1$ . Using Bayes' theorem:

$$(B11) \quad \begin{aligned} p_1 &= \Pr(s_1 | \theta_1) \\ &= \frac{\Pr(\theta_1 | s_1) \Pr(s_1)}{\Pr(\theta_1)} \\ &= \frac{\mu_1 (p_1 \mu + (1 - p_0)(1 - \mu))}{\mu}. \end{aligned}$$

Similarly we can rewrite  $p_0$ :

$$(B12) \quad \begin{aligned} p_0 &= \frac{\Pr(\theta_0 | s_0) \Pr(s_0)}{\Pr(\theta_0)} \\ &= \frac{(1 - \mu_0)(p_0(1 - \mu) + (1 - p_1)\mu)}{1 - \mu}. \end{aligned}$$

Rewriting (B11) and (B12), we obtain:

$$\begin{aligned} \mu_0(1 - \mu)p_0 &= (1 - \mu_0)\mu(1 - p_1) \\ (1 - \mu_1)\mu p_1 &= \mu_1(1 - \mu)(1 - p_0), \end{aligned}$$

or equivalently,

$$(B13) \quad p_0 = \frac{1 - \mu_0}{\mu_0} \frac{\mu}{1 - \mu} (1 - p_1)$$

$$(B14) \quad p_1 = \frac{\mu_1}{1 - \mu_1} \frac{1 - \mu}{\mu} (1 - p_0).$$

Substituting (B14) into (B13) and simplifying, we solve for  $p_0$  and  $1 - p_0$ :

$$(B15) \quad p_0 = \frac{1 - \mu_0}{1 - \mu} \frac{\mu_1 - \mu}{\mu_1 - \mu_0}, \quad 1 - p_0 = \frac{1 - \mu_1}{1 - \mu} \frac{\mu - \mu_0}{\mu_1 - \mu_0}.$$

Similarly, we solve for  $p_1$  and  $1 - p_1$ :

$$(B16) \quad p_1 = \frac{\mu_1}{\mu} \frac{\mu - \mu_0}{\mu_1 - \mu_0}, \quad 1 - p_1 = \frac{\mu_0}{\mu} \frac{\mu_1 - \mu}{\mu_1 - \mu_0}.$$

We can also express the variances and covariance in terms of  $\mu$ ,  $\mu_0$ , and  $\mu_1$ . Since  $\Theta$  is a Bernoulli random variable with success probability  $\mu$ , its variance is

$$(B17) \quad \text{Var}(\Theta) = \mu(1 - \mu).$$

Signal  $S$  is also a Bernoulli random variable but with success probability  $\Pr(s_1) = p_1\mu + (1 - p_0)(1 - \mu)$ . Therefore, its variance can be written as

$$(B18) \quad \begin{aligned} \text{Var}(S) &= [p_1\mu + (1 - p_0)(1 - \mu)][1 - p_1\mu - (1 - p_0)(1 - \mu)] \\ &= \frac{(\mu_1 - \mu)(\mu - \mu_0)}{(\mu_1 - \mu_0)^2}. \end{aligned}$$

The covariance between the state and the signal can be written as

$$(B19) \quad \begin{aligned} \text{Cov}(\Theta, S) &= \text{E}[\Theta S] - \text{E}[\Theta]\text{E}[S] \\ &= \Pr(\Theta = 1, S = 1) - \Pr[\Theta = 1] \Pr[S = 1] \\ &= \frac{(\mu_1 - \mu)(\mu - \mu_0)}{\mu_1 - \mu_0}. \end{aligned}$$

Now we are ready to derive the correlation coefficient:

$$(B20) \quad \begin{aligned} \text{Corr}(\Theta, S) &= \frac{\text{Cov}(\Theta, S)}{\sqrt{\text{Var}(\Theta)\text{Var}(S)}} \\ &= \frac{1}{\sqrt{\mu(1 - \mu)}} \sqrt{(\mu_1 - \mu)(\mu - \mu_0)} \end{aligned}$$

Expression B20 shows that the correlation coefficient can be represented in terms of the distance between the posteriors and the prior under the signal.

We can rewrite the correlation coefficient as a function of quantities directly related to Blackwell informativeness. Let  $x \equiv \mu_1 - \mu$  and  $y \equiv \mu - \mu_0$ . Then,

$$\text{Corr}(\Theta, S) = \frac{1}{\sqrt{\mu(1-\mu)}} \sqrt{xy}.$$

For a fixed prior  $\mu$ , if signal  $S$  is Blackwell more informative than signal  $S'$ , by Blackwell's theorem,  $x$  and  $y$  must be weakly greater than  $x'$  and  $y'$  (with at least one strict inequality). As a result,  $\text{Corr}(\Theta, S) > \text{Corr}(\Theta, S')$ . Therefore, we conclude that when a binary signal becomes Blackwell more informative about the state, the correlation between the state and the signal realizations will also increase.

We have proved this claim for the case of binary signals and binary states to illustrate our point. We conjecture that similar results should hold for multiple and continuous states and signals, provided the signal realizations and states are appropriately ordered. However, the proof of this general result is beyond the scope of this paper.

## C. EMPIRICAL APPENDIX

C.1. **Variable definitions.** This section provides details on the variables that we use in our paper. These variables come from the following datasets (see Section 3):

- Data on ENEM exam takers in the years 2007–2017 (INEP, 2019a). This dataset contains student-level test scores and demographic variables collected from a survey that test-takers complete when they enroll to take the exam.
- INEP’s Higher Education Census (*Censo da Educação Superior*) for the years 2010–2019 (INEP, 2022a). This dataset contains information on a student’s enrollment in higher education institutions, enrollment year, graduation year, credits completed in his/her program, and total credits required in the program curriculum.
- Brazil’s employee-employer dataset, the RAIS (*Relação Anual de Informações Sociais*), for the years 2016–2018 (RAIS, 2022). This dataset contains information on an individual’s employment, average monthly earnings, and average monthly contracted hours.
- INEP’s School Census (*Censo Escolar*) for the years 2007–2017 (INEP, 2022b). This dataset contains annual student-level data on enrollment and grade level in all Brazilian primary and secondary schools.

### C.1.1. *Test scores (ENEM data).*

- **Subject scores.** The post-2009 ENEM scores, as reported to the public, are scaled to have a mean of 500 and an SD of 100 in the population of 2009 high school seniors who took the exam. Throughout the paper, we report ENEM scores in SD units relative to this population. For ENEM scores in math, language arts, natural science, and social science, our transformation is:

$$\text{Transformed subject score} = \frac{\text{Raw subject score} - 500}{100}.$$

After the transformation, a score of zero in our paper is equivalent to the performance of the average high school senior who took the ENEM in 2009, and a score of one is 1 SD higher within this population. These transformations preserve the comparability of test scores across cohorts.

The 2007–2008 ENEM reported only a single core component score plus a writing score. To define scores for each subject, we first categorize the multiple choice questions into math, language arts, natural science, and social science, and then compute a separate score for each subject using the IRT parameters.

Since the reference populations differ for the 2007–2008 and 2009–2017 exams, in regressions where we use scores from both periods, we standardize the scores to have mean 0 and SD 1 within each year of our sample.

- **Average score (core subjects).** The post-2009 average scores are calculated by taking the average of the four subject scores, and then standardizing relative to the reference population. Our transformation is:

$$\text{Transformed average score} = \frac{\text{Raw average score} - 500}{86.7},$$

where 86.7 is the SD of the average score in the reference population. After transformation, a score of zero in our paper is equivalent to the performance of the average high school senior who took the ENEM in 2009, and a score of one is 1 SD higher within this population.

For the 2007–2008 ENEM, the average score is defined as the core component score. In the regressions where we use scores from both the 2007–2008 and 2009–2017 exams, we standardize the scores to have mean 0 and SD 1 within each year of our sample.

- **Writing scores.** The post-2009 writing score is also standardized relative to the reference population. Our transformation is:

$$\text{Transformed writing score} = \frac{\text{Raw writing score} - 597}{137},$$

where 137 is the SD of the writing score in the reference population.

For the 2007–2008 ENEM, we use the writing score as included in the data. In regressions where we use scores from both the 2007–2008 and 2009–2017 exams, we standardize the writing scores to have mean 0 and SD 1 within each year of our sample.

### C.1.2. *Exam-taker characteristics (ENEM data).*

- **Private high school.** We define “private high schools” to include both private and federal high schools since their students are comparable in terms of socioeconomic status and achievement. Only 0.5 percent of all Brazilian students attended a federal high school in 2009 (see Appendix Table A1). “Public high schools” include both state and municipality schools.
- **Race.** In Brazil, race is commonly classified in five groups: *branco* (white), *pardo* (brown), *preto* (Black), *amarelo* (yellow), and indigenous. Since Asian and indigenous people represent a small proportion of the population in Brazil (less than three percent in our sample), we use indicators for three major racial groups: *branco* (white), *pardo* (brown), and *preto* (Black). We set the indicator variables to missing if the student declined to declare his/her racial identity (about two percent of the students in our sample).
- **Parental education.** Parental education is measured based on the highest educational completed by each parent, categorized into eight levels ranging from “none” to “post-graduate.” Using these categorical variables, we define two indicator variables, “Mother completed college” and “Father completed college,” which equal one if the respective parent has completed a “college” or “post-graduate” degree.

- **Family income.** Family income is measured in multiples of the minimum wage. We define an indicator variable “Family income > 2x min. wage” (or “High-income”), which equals one if the reported income exceeds twice the minimum wage.

#### C.1.3. *College outcomes (Censo da Educação Superior).*

- **Completed program within 5 years.** An indicator variable that equals one if the student completed their program in INEP’s higher education census within 5 years of their enrollment year. This variable is defined for ENEM exam takers who appear in the higher education census using the first college program they enrolled in.
- **Earned a college degree by 2019.** An indicator variable that equals one if the student is completed any degree from an institution in INEP’s higher education census by 2019. This variable is defined for all ENEM exam takers.
- **Enrolled in any college by 2019.** An indicator variable that equals one if the student is ever observed in the INEP’s higher education census for the years 2010–2019. This variable is defined for all ENEM exam takers.
- **Finished college with 5 years of ENEM.** An indicator variable that equals one if the student is completed any degree from an institution in INEP’s higher education census within 5 years of taking the ENEM exam. This variable is defined for all ENEM exam takers.
- **Fraction of college credits completed.** The number credits completed on the student’s last record in INEP’s higher census census divided by the total credits required to graduate from the program. This variable is defined for ENEM exam takers who appear in the higher education census using the first college program they enrolled in.
- **Persisted in college for 1 year (or 3 years).** An indicator variable that equals one if the student’s last year on record in INEP’s higher education census is greater than or equal to 1 year (or 3 years) after their enrollment year. This variable is defined for ENEM exam takers who appear in the higher education census using the first college program they enrolled in.

#### C.1.4. *Labor market outcomes (RAIS data).*

- **Appears in RAIS in 2016–2018.** An indicator variable that equals one if a student appears in the 2016–2018 RAIS, indicating that the student has engaged in formal employment. This variable is defined for all ENEM exam takers.
- **Hourly wage (BRL).** Average monthly earnings divided by average monthly contracted hours computed using the 2016–2018 RAIS. If the individual is observed with earnings in multiple years in 2016–2018, we take the average across the multiple years. Wages are expressed in terms of nominal Brazilian Reals.
- **Log hourly wage.** The natural log of Hourly wage (BRL), computed as described above.

C.1.5. *High-school enrollment (ENEM data and Censo Escolar).*

- **Log number of exam takers in school.** The natural log of the number of ENEM test-takers per school/cohort as measured in the ENEM data.
- **Log number of high school seniors in school.** The natural log of the number of enrolled senior students per school/cohort as measured in the *Censo Escolar*.
- **Number of schools attended in past 3 years.** The number of distinct high schools that a senior student attended in the three years before taking the ENEM exam as measured in the *Censo Escolar*. This variable is computed by appending the *Censo Escolar* data for the years 2007–2017 and counting the number of different high schools that each senior student attended in the three years preceding their senior year, using a unique student identifier across years.

C.2. **Data and merging.** Our base dataset contains all individuals who took the ENEM exam from 2007 to 2017. This dataset includes student-level and question-level information. The student-level data includes self-reported demographic and socioeconomic status (SES) measures, such as sex, race, high-school type (public/private), parental education, and family income. The question-level data includes student responses to each exam question, the question subject, and the skill tested. From this dataset, we only keep high school seniors with a non-zero, non-missing score on each subject test. These restrictions exclude, for example, individuals who took the exam after graduating from high school or who missed one of the testing days.

To measure long-run outcomes, we combine the 2009–2014 ENEM records with two individual-level administrative datasets using individuals’ national ID numbers (*Cadastro de Pessoas Físicas*).<sup>33</sup> The linkage was conducted in the secured data room at the INEP facilities in Brasília, Brazil. We exclude students with missing national IDs (0.04 percent) and those who took the exam more than three times in our data (0.10 percent). We also exclude the 2011 cohort because of a data issue: the crosswalk variable that INEP created to match individuals across their different datasets is not correctly defined for this cohort.

We measure college outcomes using Brazil’s higher-education census from 2010 to 2019. This dataset offers comprehensive information about all college enrollees, including their university of enrollment, major, the academic year when they enrolled, and their year of graduation. 65.6 percent of high-school seniors taking the ENEM during 2009–2014 appear in the census data.

We measure labor-market outcomes using an administrative employee-employer matched dataset called RAIS (*Relação Anual de Informações Sociais*) from 2016 to 2018. The RAIS

---

<sup>33</sup> We do not include 2015–2017 ENEM takers in our analysis of long-run outcomes because we do not observe many of these outcomes in these cohorts given the timing of our data.

contains data on workers employed in the formal sector. It does not include data on individuals working within the informal sector, those who are self-employed, or individuals who are currently unemployed. This dataset includes worker-level and firm-level information. The worker-level data includes educational attainment, occupation, and earnings. The firm-level data includes total employee count, the industry they operate within, and their geographical location.

**C.3. Sample definition.** This section describes the criteria for high schools to be included in the INEP annual performance reports, which we use to define our main analysis sample.

As noted in the main text, our analysis sample consists of the set of high schools that were included in the annual ENEM performance reports created by INEP in *each* year in 2005–2015. High schools were included in the performance reports if they met two conditions. First, the high school needed to have at least 10 test-takers who declared that they would graduate that year. Before 2011, all individuals taking the ENEM were recorded as test-takers for the calculation of the participation rate, regardless of whether they completed the test or not. Since 2011, to be considered a test-taker, an individual has to complete the four subject tests plus the writing essay, and obtain a non-zero score in all subjects.<sup>34</sup> Second, starting in 2009, the INEP required a minimum *participation rate* to be included in the report. The participation rate is the total number of ENEM test-takers in the high school divided by the number of enrolled students in the final years of high school, based on the records of the High School Census. Between 2009 and 2010, the minimum participation rate to be included in the report was 2 percent. In 2011, INEP increased the minimum participation rate to 50 percent.

At the schools in our sample, the average ENEM participation rate over the 2005–2015 period was 70 percent. Table 2 presents balance tests that show that the number of exam takers in our analysis sample and the characteristics of these exam takers did not change significantly when the stakes of the ENEM increased. Appendix Table A5 shows the robustness of our main results to alternative sample definitions.

**C.4. Categorization of math questions.** This subsection provides details on the categorization of math questions that we use for the heterogeneity analyses in Table 6 and Appendix Table A16.

In Table 6, we categorize math questions into 7 *topic areas* (Panel B) and 30 *competencies* (Panel C) defined by ENEM test designers. These topic areas and competencies are in the

---

<sup>34</sup> Only the following grades are considered in the report: 3rd and 4th grade of regular high school (*ensino médio regular 3a e 4a série*), 3rd and 4th grade of teaching track high school (*ensino médio magistério 3a e 4a série*), non-serialized regular and teaching track high school (*ensino médio não-seriado, regular e magistério*), vocational education and high school for youth and adult education (*educação profissionalizante e ensino médio para educação de jovens e adultos*).

TABLE C1. Topic areas and competencies for ENEM math exam

(A)	(B)	(C)	(D)
Topic area	Competency (and reference number)	Topic area	Competency (and reference number)
Numbers	Recognize numbers (1)	Proportions	Solve problems using proportions (16)
Numbers	Identify numerical patterns (2)	Proportions	Use proportions to construct arguments (17)
Numbers	Solve problems using numbers (3)	Proportions	Evaluate interventions using proportions (18)
Numbers	Use numbers to construct arguments (4)	Algebra	Identify algebraic relationships (19)
Numbers	Evaluate interventions using numbers (5)	Algebra	Interpret Cartesian graphs (20)
Geometry	Project 3D objects into 2D space (6)	Algebra	Solve problems using algebra (21)
Geometry	Identify geometric shapes (7)	Algebra	Use algebra to construct arguments (22)
Geometry	Solve problems using geometry (8)	Algebra	Evaluate interventions using algebra (23)
Geometry	Use geometry to construct arguments (9)	Interpreting data	Make inferences using data in tables/graphs (24)
Measurements	Identify units of measurement (10)	Interpreting data	Solve problems using data in tables/graphs (25)
Measurements	Use scales in everyday situations (11)	Interpreting data	Use tables/graphs to construct arguments (26)
Measurements	Solve problems using magnitudes (12)	Statistics	Calculate statistical quantities from data (27)
Measurements	Use measurements to construct arguments (13)	Statistics	Solve problems using statistics (28)
Measurements	Evaluate interventions using measurements (14)	Statistics	Use statistics to construct arguments (29)
Proportions	Identify proportional relationships (15)	Statistics	Evaluate interventions using statistics (30)

*Notes:* This table shows the 7 topic areas (columns A and C) and 30 competencies (columns B and D) for the ENEM math exam. Labels are translated and shortened by the authors from the descriptions in `Matriz_Referencia_Enem.pdf`, which is included with the microdata.

ITENS\_PROVA\_\*\*\*\*.TXT files of the microdata. The labels for each topic area and competency are defined in `Matriz_Referencia_Enem.pdf`, which is included with the microdata and is also available online at the link in this footnote.<sup>35</sup> Table C1 shows the translated and shortened labels that we use for Table 6. We also present results by topic area for language arts, natural science, and social science in Appendix Table A13.

In Appendix Table A16, we define groups of questions based on whether the questions are related to topics covered in a study guide created by *Me Salva!*, which is a well-known Brazilian test prep company. The study guide is called *The Approved Book: One topic per day to pass the ENEM*.<sup>36</sup> In Panel A of Appendix Table A16, we match keywords from the *Me Salva!* study guide to the text of each question. In Panel B of Appendix Table A16, we match concepts from the *Me Salva!* study guide to solutions to each question that were created by another well-known test prep company called *Descomplica*. *Descomplica* creates solutions to each question from previous ENEM exams and makes them publicly-available online to help students prepare.<sup>37</sup>

We match keywords and concepts in the *Me Salva!* study guide to ENEM questions using text analysis. The *Me Salva!* study guide contains tips for solving questions in seven different

<sup>35</sup> See: [https://download.inep.gov.br/download/enem/matriz\\_referencia.pdf](https://download.inep.gov.br/download/enem/matriz_referencia.pdf) (accessed in June 2023).

<sup>36</sup> The Portuguese title is *O Livro do Aprovado: Um conteúdo por dia para passar no ENEM*. See: <https://cdn.mesalva.com/uploads/medium/attachment/MS2018-livro-do-aprovado.pdf> (accessed in June 2023).

<sup>37</sup> See: <https://descomplica.com.br/gabarito-enem/questoes/?cor=azul> (accessed in June 2023). One question is behind a paywall (Question 145 in the *Azul* book of the 2015 math exam), and thus we exclude it from our analysis in Panel B of Table A16).

content areas, and in each content area, there are keywords that appear in bold in the text. For Panel A of Appendix Table A16, we search the question text for the keywords from the study guide. For Panel B of Appendix Table A16, we search the *Descomplica* solutions for both the keywords and for regular expressions that indicate questions in which the solution is likely to depend on concepts from the study guide. The content areas and search terms that we use are as follows; the keywords that we use in both Panels A and B appear in italics, and the regular expression searches that we use in Panel B appear in plain text:

- **Geometric formulas.** *cilindro, cone, cubo, equilatero, esfera, hexagonal, hexagono, losango, paralelepipedo, piramide, prisma, quadrado, quadrangular, retangulo, trapezio, triangular, triangulo.*
- **Proportions (“Rule of 3”).** *diretamente, grandezas, inversamente, proporcionais, proporção, regra de três.*
- **Manipulating fractions.** Solutions that contain at least two fractions ( $[(0-9)+/[0-9]+)$  and also an equals sign (=).
- **Radicals.** Solutions that contain a square or cube root sign ( $\sqrt{\quad}$  or  $\sqrt[3]{\quad}$ ).
- **Combinatory and statistical analysis.** *arranjo, combinação, mediana, moda, permutação, possibilidades.* Solutions that contain arrangement or combination notation ( $A[0-9],[0-9]$  or  $C[0-9],[0-9]$ ).
- **Probability.** *combinação, possiveis, probabilidade.* Solutions that contain combination notation ( $C[0-9],[0-9]$ ).
- **Trigonometric formulas.** *cos, cossec, cosseno, cotg, sec, sen, seno, tangente, tg.*

In all cases we trim plural, adjective, and masculine/feminine endings to words before matching. In Panel A, we display results for keywords that appear in eight or more questions, and we group all other words into the “Fewer than 8 occurrences” category. In Panel B, we display results separately for each of the seven *Me Salva!* content areas. In both panels, we also show results pooling across all questions that do/don’t match any search term in the study guide.