ECONSTOR
Make Your Publications Visible.

A Service of

ZBW Leibniz-Informationszentrum Wirtschaft
Leibniz Information Centre for Economics

Campbell, Douglas et al.

Working Paper

Robustness Report: "Going to a Better School: Effects and Behavioral Responses", by Cristian Pop-Eleches and Miguel Urquiola (2013)

I4R Discussion Paper Series, No. 133

Provided in Cooperation with:
The Institute for Replication (I4R)

Suggested Citation: Campbell, Douglas et al. (2024) : Robustness Report: "Going to a Better School: Effects and Behavioral Responses", by Cristian Pop-Eleches and Miguel Urquiola (2013), I4R Discussion Paper Series, No. 133, Institute for Replication (I4R), s.l.

This Version is available at:
https://hdl.handle.net/10419/298014

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

WWW.ECONSTOR.EU

Mitglied der Leibniz-Gemeinschaft

INSTITUTE for REPLICATION

No. 133
I4R DISCUSSION PAPER SERIES

Robustness Report: “Going to a Better School: Effects and Behavioral Responses”, by Cristian Pop-Eleches and Miguel Urquiola (2013)

Douglas Campbell
Abel Brodeur
Magnus Johannesson
Joseph Kopecky
Lester Lusher
Nikita Tsoy

June 2024

INSTITUTE for REPLICATION

I4R DISCUSSION PAPER SERIES

I4R DP No. 133

Robustness Report: "Going to a Better School: Effects and Behavioral Responses", by Cristian Pop-Eleches and Miguel Urquiola (2013)

Douglas Campbell[1], Abel Brodeur[2], Magnus Johannesson[3], Joseph Kopecky[4], Lester Lusher[5], Nikita Tsoy[6]

[1]New Economic School, Moscow/Russia
[2]University of Ottawa/Canada and IZA – Institute of Labor Economics, Bonn/Germany
[3]Stockholm School of Economics, Stockholm/Sweden
[4]Trinity College Dubin/Ireland
[5]University of Pittsburgh/USA
[6]Institute for Computer Science, Artificial Intelligence and Technology (INSAIT), and Sofia University, Sofia/Bulgaria

JUNE 2024

Any opinions in this paper are those of the author(s) and not those of the Institute for Replication (I4R). Research published in this series may include views on policy, but I4R takes no institutional policy positions.

I4R Discussion Papers are research papers of the Institute for Replication which are widely circulated to promote replications and meta-scientific work in the social sciences. Provided in cooperation with EconStor, a service of the ZBW – Leibniz Information Centre for Economics, and RWI – Leibniz Institute for Economic Research, I4R Discussion Papers are among others listed in RePEc (see IDEAS, EconPapers). Complete list of all I4R DPs - downloadable for free at the I4R website.

I4R Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

Editors

Abel Brodeur
University of Ottawa

Anna Dreber
Stockholm School of Economics

Jörg Ankel-Peters
RWI – Leibniz Institute for Economic Research

E-Mail: joerg.peters@rwi-essen.de
RWI – Leibniz Institute for Economic Research
Hohenzollernstraße 1-3
45128 Essen/Germany
www.i4replication.org

ISSN: 2752-1931

Robustness Report: "Going to a Better School: Effects and Behavioral Responses", by Cristian Pop-Eleches and Miguel Urquiola (2013)*

Douglas Campbell† Abel Brodeur‡ Magnus Johannesson§
Joseph Kopecky¶ Lester Lusher‖ Nikita Tsoy**

May 12, 2024

Abstract

Pop-Eleches and Urquiola (2013) apply a regression discontinuity to the Romanian secondary school system, and notably find that (a) students who go to a better school get higher scores on an exam used for university admission, (b) parents of students who get into a better school help their kids less with homework, and (c) kids who go to a slightly better school report more negative interactions with peers. We first reproduce all regression tables in Pop-Eleches and Urquiola (2013), and then test for robustness by unstacking the data, multi-way clustering, altering the cutoffs, altering control variables, and conducting influential analysis. Overall, we find the results for finding (a), (b), and (c) are robust in 100%, 42%, and 60% of the robustness checks we ran, and the t/z scores were on average 93%, 69%, and 92% as large as the original study.

JEL Classification: I21, I28, J13
Keywords: Education, Peer Effects, Economics of Education

*This robustness report was written as part of a meta-robustness project entitled "The robustness reproducibility of the American Economic Review". See Campbell, Brodeur, Dreber, Johannesson, Kopecky, Lusher and Tsoy (2024) for more information.

†Corresponding author. New Economic School, E-mail: dolcampb@gmail.com
‡University of Ottawa and IZA. E-mail: abrodeur@uottawa.ca
§Department of Economics, Stockholm School of Economics, E-mail: magnus.johannesson@hhs.se
¶Department of Economics, Trinity College Dublin, E-mail: jkopecky@tcd.ie
‖Department of Economics, University of Pittsburgh, E-mail: lesterlusher@pitt.edu
**INSAIT, Sofia University, E-mail: nikita.tsoy@insait.ai

1 Introduction

Pop-Eleches and Urquiola (2013) – hereafter, PU – study the impact of going to a better school on exam scores, and on student and parental behavior. Applying a stacked regression discontinuity to the Romanian secondary school system, they find that (a) students who go to a better school get higher scores on an exam used for university admission, (b) parents of students who get into a better school help their kids less with homework, and (c) kids who go to a slightly better school report more negative interactions with peers, among other interesting findings.[1]

We first replicate all the regression tables in PU exactly using the data and code provided by the authors on the AER website. We then run several robustness checks on what we view as the key identified result. These include unstacking the data, multi-way clustering, altering the cutoffs, altering control variables, and conducting influential analysis by dropping observations with large dfbetas.[2]

We find that the first result, that students who score slightly above the cutoff for the high school entrance exam go on to score better on a college entrance exam than students who score slightly below the cutoff, was extremely robust. We ran 20 robustness checks, and all had a coefficient in the same direction, and were statistically significant at 5%. In addition, t-values on average were 93% as large as the original paper. Frequently, in our robustness checks, we estimated a larger coefficient and/or t-value than the original study. And the smallest t-value we estimated came when we rather severely shrank the sample.

The second result – that parents of students who scored just above the cutoff to go to a better school helped their kids less with homework – was robust with a coefficient in the same direction and statistically significant at the 5% level in just 42% of cases. That said, the results were often borderline significant, as 79% of

[1] Note that the original paper features many interesting results. These were three of the most salient in our view, and so we focus this robustness report on these findings.

[2] Note that these were a bit challenging to implement as the original data set did not have any variables defined: we cross-checked results with the paper to infer what each variable was. This leaves open the possibility that we may have incorrectly defined a variable.

robustness checks were significant at 10%, and 89% were significant at 15%. Varying the window around the cutoff slightly in either direction tended to erode statistical significance, but no matter which cutoff one uses, all results are qualitatively similar in our view. This result did prove sensitive to influential analysis, however, as the results do weaken when outliers are removed using dfbeta statistics.

The third result, that kids who go to a slightly better school report more negative interactions with peers, was mostly robust. When we shrank the window size around the cutoff exam scores by 40%, the results did weaken, but this can be expected given the smaller sample size. While only 60% of robustness tests were significant at 5%, 73% were significant at 10%, and the average t-value in our robustness checks was 92% as large as the original paper. The results also strengthened with influential analysis.

Overall, it makes sense that the second and third results were less robust than the first, given that they used survey data and a much smaller sample size (the first result used 1.25 million observations vs. around 6,500 for the latter two). Our influential analysis of the second result is a reason for concern, but otherwise, our results appear qualitatively similar to the original paper. Our results do suggest that one might like to have larger sample sizes to verify the latter two results.

2 Replication

2.1 Regression model

All three of the key results we focus on have the same empirical strategy. The idea is to compare students who score just above a high school entrance exam cutoff compared to otherwise similar students who score just below.

$$y_{iz}^{k} = \alpha 1\{t_i - \tilde{t}_z \geq 0\} + \eta(t_i - \tilde{t}_z) + \psi(t_i - \tilde{t}_z)1\{t_i - \tilde{t}_z \geq 0\} + w_z + v_i \quad (1)$$

For student i, cutoff j, and result k, and where the data is stacked so that

student i can be in the data more than one time relative to various cutoffs in the same locality. The variable t is exam scores, and $\tilde{t}_z$ is the exam score cutoff to get into school z.

The three results we will focus on, in order, are (a) scores on an exam used for university admission, (b) an indicator variable for whether parents of students who get into a better school help their kids with homework often, and (c) whether kids report more negative interactions with peers using an index value.[3] In all three regressions we focus on, PU use a cutoff of one point – thus all scores within one point of a cutoff are included in the regression, and the key variable of interest will be a dummy variable for being above the cutoff.

2.2 Effect on Baccalaureate Grade

2.2.1 Computational Reproduction

We first completed computational reproduction for all the regression tables in this paper, and, using the author-provided data and code, were able to reproduce the coefficients from the original result exactly.[4] In this note, we simply show the reproduced results for the findings we focus on. For the impact of going to a better school on the baccalaureate grade (used for college admissions), we chose Table 4, Panel B, column (1) as the benchmark. We reproduce this result in our Table 1, column (1). The original paper estimated a coefficient of 0.18 with a t-value of 7.90 on the indicator variable for having a grade above a cutoff.

[3]This is an index created from a questionnaire on whether: "in the last month, their peers have: (i) been mean to them, (ii) hit them, (iii) taken their things without asking, or (iv) made them feel marginalized. The possible responses for each of these items ranged from zero (did not happen at all in the past month) to five (happened daily)" according to PU.

[4]Note that this study was not pre-registered. We implemented the computational replication before designing the robustness plan, and our study gives full leeway to the replicators to add more robustness checks after coming in contact with the data and even after beginning to run robustness checks. All robustness checks run on the paper's key results are recorded here.

2.2.2 Robustness Results: Baccalaureate Grade

First, we experimented with dropping with the interactive control (the third term) in Equation 1. The negative coefficient on this variable in Table 1 column (1) was a concern, and overall, the relation between entrance exam scores and baccalaureate scores is concave up. When we drop this control in column (2), however, the estimated impact of going to a better school increases. When we control for raw transition score and transition score squared (instead of the normalized score), the results are little changed. In column (4), we add back in the interactive control from the original, and the coefficient and significance level are similar to the benchmark. In column (5), we use all data and create two indicator variables: one for being within one point in either direction of a school cutoff, and the second is being within one point above a cutoff. Once again, we estimate a coefficient of .022, a larger coefficient and a much larger t-value (18.95) than the original paper. We conclude this result is robust to these changes in how we control for transition scores.

Table 1: Impact on Baccalaureate Grade: Robustness

	(1) b/se/t/p	(2) b/se/t/p	(3) b/se/t/p	(4) b/se/t/p	(5) b/se/t/p
1{Grade ≥ Cutoff}	0.0181	0.0224	0.0219	0.0173	
	(0.0023)	(0.0023)	(0.0023)	(0.0023)	
	[7.90]	[9.87]	[9.68]	[7.55]	
	{0.000}	{0.000}	{0.000}	{0.000}	
Transition score (normalized)	0.802	0.766			
	(0.0040)	(0.0028)			
	[199.3]	[270.2]			
	{0.000}	{0.000}			
Above cutoff*score	-0.0585			-0.0634	-0.0393
	(0.0044)			(0.0044)	(0.0025)
	[-13.4]			[-14.4]	[-15.7]
	{0.000}			{0.000}	{0.000}
Transition score			-0.137	-0.102	0.331
			(0.037)	(0.037)	(0.029)
			[-3.65]	[-2.77]	[11.6]
			{0.000}	{0.006}	{0.000}
Transition score squared			0.0577	0.0580	0.0285
			(0.0023)	(0.0024)	(0.0018)
			[24.6]	[24.7]	[16.2]
			{0.000}	{0.000}	{0.000}
1{1 ≥ Grade ≥ Cutoff}					0.0221
					(0.0012)
					[18.9]
					{0.000}
1{Cutoff-1 ≥ Grade ≥ Cutoff+1}					-0.0364
					(0.0016)
					[-22.6]
					{0.000}
Observations	1256035	1256035	1256035	1256035	2546208

Notes: This table first reproduces Table 4, Panel B, column (1) of PU in the first column. All regressions include cutoff fixed effects and are clustered at the individual level. Columns (1) through (4) include only observations within one point of a school's grade cutoff. Column (5) includes all data. The dependent variable is the grade on the baccalaureate.

Next, we experimented with altering the cutoffs.[5] First, we widen the cutoff from one to 1.2 in column (1) of Table 2. The coefficient increases from .018 to .020. In column (2), we reduce the cutoff to 0.8, and this time the coefficient falls slightly to .016, although is still highly significant. Narrowing the cutoff window to 0.4 results in a coefficient on scoring above the cutoff of .018. Lastly, we tried greatly shrinking the cutoff to 0.05. Here, the estimated coefficient on scoring above the cutoff increases to 0.027 and is still significant at 5%, even though the number

[5]Note that the authors, to their credit, also displayed results for multiple cutoffs for each result.

of observations used fell from 1.25 million to 59,000. Given that we restricted the number of observations to this extent, it makes sense that the t-value fell to 2.10. In column (5), we again use all data and create two dummy variables: one for being within 0.1 points of a cutoff in either direction and one for being within 0.1 points above a cutoff. Once again we get an estimate of .022, and highly significant with a t-value of 6.25. We conclude this result is robust to changes in the particular cutoff used.

Table 2: Impact on Baccalaureate Grade: Altering the Cutoff

	(1)	(2)	(3)	(4)	(5)
	b/se/t/p	b/se/t/p	b/se/t/p	b/se/t/p	b/se/t/p
1{Grade ≥ Cutoff}	0.0198	0.0157	0.0180	0.0267	
	(0.0021)	(0.0025)	(0.0034)	(0.013)	
	[9.30]	[6.38]	[5.26]	[2.10]	
	{0.000}	{0.000}	{0.000}	{0.036}	
Transition score (normalized)	0.793	0.809	0.775	0.683	
	(0.0034)	(0.0050)	(0.011)	(0.31)	
	[230.3]	[163.1]	[67.6]	[2.23]	
	{0.000}	{0.000}	{0.000}	{0.026}	
Above cutoff*score	-0.0438	-0.0715	-0.0365	-0.132	-0.0192
	(0.0033)	(0.0060)	(0.015)	(0.43)	(0.0018)
	[-13.5]	[-11.9]	[-2.47]	[-0.31]	[-10.8]
	{0.000}	{0.000}	{0.013}	{0.758}	{0.000}
1{.1 ≥ Grade ≥ Cutoff}					0.0218
					(0.0035)
					[6.25]
					{0.000}
1{Cutoff-.1 ≥ Grade ≥ Cutoff+.1}					-0.0118
					(0.0026)
					[-4.48]
					{0.000}
Transition score					0.336
					(0.029)
					[11.8]
					{0.000}
Transition score squared					0.0276
					(0.0018)
					[15.7]
					{0.000}
Observations	1461418	1024576	528822	59523	2546208
Point Cutoff	1.20	0.80	0.40	0.05	None

Notes: All regressions include cutoff fixed effects and are clustered at the individual level. Columns (1) through (4) include only observations within one point of a school's grade cutoff. Column (5) includes all data. The dependent variable is the grade on the baccalaureate.

Next, we examine clustering schemes and changing the way the data is stacked.

One worry might be that the author's stacking method – whereby one student's score appears in the data multiple times relative to different cutoffs – is unconventional and might inflate t-values artificially. The authors address this in part by clustering at the student level. First, we try "unstacking" the data – we restrict each student to one observation relative to the nearest cutoff. When we do this in Table 3 column (1), we get a coefficient of 0.028, and still highly significant. In column (2), we multiway cluster by cutoff and by town, and the t-value increases.[6] In column (3), we use the original benchmark regression, and cluster by cutoff and student, and do find that this matters somewhat. We get a t-value of 4.12 vs. 7.90 in the benchmark. When we three-way cluster by student, town, and cutoff, the t-value falls to 3.94. Regardless, this result is still highly significant at 1%. We conclude this result is robust to various clustering schemes.

Table 3: Unstacking and Multiway Clustering

	(1)	(2)	(3)	(4)
	b/se/t/p	b/se/t/p	b/se/t/p	b/se/t/p
1{Grade ≥ Cutoff}	0.0281	0.0281	0.0181	0.0181
	(0.0060)	(0.0059)	(0.0044)	(0.0046)
	[4.66]	[4.74]	[4.12]	[3.94]
	{0.000}	{0.000}	{0.000}	{0.000}
Trans. score (normalized)	0.753	0.753	0.802	0.802
	(0.035)	(0.037)	(0.0091)	(0.018)
	[21.5]	[20.2]	[87.7]	[44.9]
	{0.000}	{0.000}	{0.000}	{0.000}
Above cutoff*score	-0.0277	-0.0277	-0.0585	-0.0585
	(0.035)	(0.038)	(0.012)	(0.015)
	[-0.78]	[-0.72]	[-5.09]	[-3.99]
	{0.434}	{0.470}	{0.000}	{0.000}
Observations	199620	199620	1256035	1256035
Additional Clustering	No	+Town	+Student	+Student & Town

Notes: Columns (1) and (2) "unstack" the data by limiting each student to one observation (the closest absolute cutoff). All columns cluster by cutoff, Column (2) additionally clusters by town, column (3) additionally clusters by student, and column (4) additionally clusters by both student and town. The dependent variable is the grade on the baccalaureate.

Next, we experimented with letting the controls for transition score vary by year

[6]To be more precise, by "town" here, we mean the first five digits of the cutoff fixed effect variable "uazY" in the data. This appears to be the cutoff fixed effect identifier minus year and cutoff number information.

and by number of within-town cutoffs, and with various levels of clustering. The results are in Table 4 and are robust.

Table 4: Baccalaureate Grade: Additional Robustness

	(1) b/se/t/p	(2) b/se/t/p	(3) b/se/t/p	(4) b/se/t/p
1{Transition grade ≥ cutoff}	0.0174	0.0181	0.0157	0.0154
	(0.0023)	(0.0023)	(0.0045)	(0.0046)
	[7.58]	[7.90]	[3.46]	[3.37]
	{0.000}	{0.000}	{0.001}	{0.001}
Observations	1256035	1256035	1024576	1024576
Clustering	One-Way	One-Way	Three-Way	Three-Way
Added Controls	Year	Schools	None	Year
Point Cutoff	1.00	1.00	0.80	0.80

Notes: All regressions include cutoff fixed effects. The dependent variable is the grade on the baccalaureate.

Lastly, we conduct influential analysis in Table 5. In column (1), we compute dfbeta statistics for each observation using our benchmark regression, and then purge outliers with a dfbeta statistic with an absolute value higher than the (standard) cutoff of $2/\sqrt{N} = 2/\sqrt{1256053} = .00178$.[7] The coefficient on being above the cutoff increases to .023. In column (2), we test the version using a cutoff of 0.8, which earlier had a slightly smaller coefficient than the benchmark. When we purge outliers for this version here, the coefficient is unchanged at .016.

Table 5: Influential Analysis

	(1)	(2)
1{Transition grade ≥ cutoff}	0.023	0.16
	(0.0019)	(0.0043)
	[12.10]	[37.42]
	{0.000}	{0.000}
Observations	1172203	972004
Cutoff	1.00	0.80

Notes: Column (1) is the benchmark regression with outliers dropped. Column (2) is the same regression but with a cutoff of 0.8 and multiway clustered by cutoff and student. Each regression includes cutoff fixed effects. The dependent variable is the grade on the baccalaureate.

[7] We use the dfbeta post-regression command in Stata.

2.3 Second Result: Parents Help Child With Homework

The next outcome variable we look at is whether parents report helping children with homework often.[8] This is survey data and a relatively small subsample of the original data. We reproduce the benchmark results from Table 7, panel C, column (7) in our Table 14, column (1). We confirm a coefficient of -.043, a t-value of -2.26, and a p-value of .024. Next, we first try alternative methods of clustering, first by two-way clustering at the cutoff level in column (2), and additionally the town level in column (3). The errors are hardly changed or even fall. In column (4) we drop the interactive term, and the results strengthen. Lastly, in column (5), we add in several demographic control variables, and here, the standard errors get slightly larger, but the coefficient is essentially unchanged and the p-value is still below .05. These demographic controls include dummy variables for ethnicity (Romanian, Hungarian, and Gypsies), and whether the mother is married and employed.

Table 6: Impact on Parents Helping with Homework

	(1)	(2)	(3)	(4)	(5)
	b/se/t/p	b/se/t/p	b/se/t/p	b/se/t/p	b/se/t/p
1{Grade ≥ Cutoff}	-0.0427	-0.0427	-0.0427	-0.0506	-0.0419
	(0.019)	(0.019)	(0.015)	(0.018)	(0.020)
	[-2.26]	[-2.24]	[-2.76]	[-2.79]	[-2.05]
	{0.024}	{0.026}	{0.008}	{0.006}	{0.042}
Transition score (normalized)	-0.00574	-0.00574	-0.00574	0.0404	-0.0228
	(0.033)	(0.041)	(0.037)	(0.020)	(0.040)
	[-0.17]	[-0.14]	[-0.15]	[2.00]	[-0.56]
	{0.863}	{0.889}	{0.878}	{0.048}	{0.574}
Above cutoff*score	0.0644	0.0644	0.0644		0.0796
	(0.039)	(0.046)	(0.046)		(0.045)
	[1.66]	[1.39]	[1.41]		[1.76]
	{0.097}	{0.165}	{0.164}		{0.081}
Observations	6487	6487	6487	6487	5791
Clustering	One-Way	Two-Way	Three-Way	Two-Way	Two-Way
Added controls?	None	None	None	None	+Demogr.

Notes: This table first reproduces Table 7, Panel C, column (7) of PU in the first column. All regressions include cutoff fixed effects. Column (1) is clustered at the student level, columns (2), (4), and (5) are additionally clustered at the cutoff level, and column (3) is additionally clustered at the cutoff and town levels. The dependent variable is an indicator for whether parents help children with their homework regularly.

[8]Note, we use PU's definition, but these children are high school students.

Next, we vary the cutoffs in Table 7. The benchmark included all observations for student*school cutoff pairs where the students were within one point of the cutoff. We alternatively try cutoffs of 1.4, 1.2, 0.8, and 0.6. Of these, we find that varying the cutoff in either direction weakens the results slightly. Moving the cutoff up to 1.2, for example, shrinks the coefficient from -.042 to -.0283, with a t-value of 1.59, and p-value of 1.11 – no longer statistically significant at conventional thresholds. The same happens when we shrink the window. That said, in each case, the results are not substantially different from the original benchmark, in our view. In addition, when we drop the (perhaps dubious) interaction term in column (5), the results strengthen somewhat. Our subjective view is that these results call for more evidence on this question (and perhaps larger sample sizes), but the results here are not too worrying.

Table 7: Parental Help: Varying the Cutoffs

	(1)	(2)	(3)	(4)	(5)
	b/se/t/p	b/se/t/p	b/se/t/p	b/se/t/p	b/se/t/p
1{Grade ≥ Cutoff}	-0.0332	-0.0283	-0.0298	-0.0198	-0.0369
	(0.017)	(0.018)	(0.021)	(0.024)	(0.020)
	[-1.97]	[-1.59]	[-1.42]	[-0.83]	[-1.80]
	{0.049}	{0.111}	{0.156}	{0.404}	{0.071}
Transition score (normalized)	-0.0318	-0.0452	-0.0437	-0.0564	0.0169
	(0.026)	(0.030)	(0.043)	(0.059)	(0.024)
	[-1.22]	[-1.53]	[-1.02]	[-0.95]	[0.70]
	{0.222}	{0.126}	{0.307}	{0.342}	{0.487}
Above cutoff*score	0.0846	0.0943	0.0919	0.0596	
	(0.029)	(0.033)	(0.051)	(0.072)	
	[2.96]	[2.84]	[1.81]	[0.83]	
	{0.003}	{0.004}	{0.070}	{0.408}	
Observations	8286	7460	5323	4155	5323
Cutoff	1.4	1.2	0.8	0.6	0.8

Notes: All regressions include cutoff fixed effects. The dependent variable is an indicator for whether parents help children with their homework regularly.

Next, we alter the control variables, controlling for the raw transition score in column (1) of Table 8, and adding in a square term as well in column (2). In columns (3) and (4), all the data is used and dummy variables for being within a point of the cutoff and a point above the cutoff are included. The key result is mostly robust,

although the coefficients are just a bit more than half of the original in the latter two regressions. They are still statistically significant at 10%.

Table 8: Parental Help: Alternative Functional Forms for Controls

	(1) b/se/t/p	(2) b/se/t/p	(3) b/se/t/p	(4) b/se/t/p
1{Grade ≥ Cutoff}	-0.0506	-0.0491		
	(0.018)	(0.018)		
	[-2.78]	[-2.69]		
	{0.005}	{0.007}		
Transition score	0.0404	-0.0819	-0.0891	0.0244
	(0.017)	(0.15)	(0.071)	(0.074)
	[2.32]	[-0.53]	[-1.25]	[0.33]
	{0.021}	{0.594}	{0.210}	{0.743}
Transition score squared		0.00763	0.00603	-0.00172
		(0.0096)	(0.0043)	(0.0045)
		[0.79]	[1.41]	[-0.39]
		{0.429}	{0.157}	{0.699}
1{1 ≥ Grade ≥ Cutoff}			-0.0220	-0.0233
			(0.012)	(0.013)
			[-1.79]	[-1.81]
			{0.074}	{0.070}
1{Cutoff+1 ≥ Grade ≥ Cutoff-1}			-0.00356	0.00237
			(0.017)	(0.018)
			[-0.21]	[0.13]
			{0.837}	{0.895}
Observations	6487	6487	11723	10582

Notes: All regressions include cutoff fixed effects. The dependent variable is an indicator for whether parents help children with their homework regularly. Column (4) additionally includes several demographic controls, excluded for space.

As the dependent variable is an indicator variable for whether parents often help their children with homework, it is also natural to fit a non-linear model. In Table 9 we first run a logit regression on the original specification, and then once again vary the cutoffs in columns (2) and (3). In column (4), we use all the data and include additional dummy variables for being within a point of the cutoff (omitted for space), and for being within one point above the cutoff (our key variable of interest). In this case, all regressions are at least significant at 10%. We attempted to run a version of this with demographic controls (full and partial) but our regression failed to converge.

Table 9: Parental Help: Robustness Using Logit

	(1) b/se/t/p	(2) b/se/t/p	(3) b/se/t/p	(4) b/se/t/p
1{Grade ≥ Cutoff}	-0.309	-0.225	-0.213	
	(0.13)	(0.14)	(0.12)	
	[-2.34]	[-1.66]	[-1.73]	
	{0.019}	{0.097}	{0.084}	
1{1 ≥ Grade ≥ Cutoff}				-0.160
				(0.086)
				[-1.86]
				{0.063}
Observations	6220	5073	7264	11611
Cutoff	1.00	0.80	1.20	None

Notes: All regressions are estimated using logit, and include cutoff fixed effects. The dependent variable is an indicator for whether parents help children with their homework regularly. Other controls excluded for space.

Lastly, we conduct influential analysis in Table 10. In column (1), we compute dfbeta statistics for each observation using our benchmark regression, and then purge outliers with a dfbeta statistic with an absolute value higher than the (standard) cutoff of $2/\sqrt{N} = 2/\sqrt{6487} = .025$.[9] The coefficient on being above the cutoff is now just -.01, and not statistically significant. In column (2), we test the version using a cutoff of 1.2, and when we purge outliers for this version here, the coefficient flips to 0.09 and is statistically significant at 1%. These results are worrying, and thus we conclude that this result is not robust to influential analysis.

Table 10: Parental Help: Influential Analysis

	(1) b/se/t/p	(2) b/se/t/p
1{Transition grade ≥ $cutoff$}	-0.0102	0.0901
	(0.011)	(0.018)
	[-0.89]	[5.03]
	{0.373}	{0.000}
Observations	6055	6880
Cutoff	1.00	1.20

Notes: All regressions include cutoff fixed effects. The dependent variable is an indicator for whether parents help children with their homework regularly.

[9] We use the dfbeta post-regression command in Stata.

3 Impact on Students' Negative Interactions

The third result we test robustness for is an indicator variable for whether a student experiences negative interactions at school. We run all the same regressions for whether parents help with homework, only with a different outcome variable. Table 11 column (1) is the benchmark, and an exact reproduction of PU Table 8, Panel C, column (5). We confirm a coefficient of .045, a t-value of 2.4, and a p-value of .017. Next, we use alternative methods of clustering, first by two-way clustering at the cutoff level in column (2), and additionally at the town level in column (3). The errors are little changed. In column (4) we drop the interactive term, and the results weaken slightly. Lastly, in column (5), we add in several demographic control variables, and here, the standard errors get slightly larger, but the coefficient is essentially unchanged and the p-value is still below .05.

Table 11: Impact on Negative Interactions: Robustness

	(1)	(2)	(3)	(4)	(5)
	b/se/t/p	b/se/t/p	b/se/t/p	b/se/t/p	b/se/t/p
1{Grade ≥ Cutoff}	0.0454	0.0454	0.0454	0.0396	0.0406
	(0.019)	(0.018)	(0.020)	(0.017)	(0.020)
	[2.40]	[2.52]	[2.28]	[2.28]	[2.02]
	{0.017}	{0.013}	{0.027}	{0.024}	{0.045}
Transition score (normalized)	-0.0909	-0.0909	-0.0909	-0.0567	-0.108
	(0.039)	(0.042)	(0.042)	(0.018)	(0.045)
	[-2.33]	[-2.19]	[-2.16]	[-3.19]	[-2.39]
	{0.020}	{0.030}	{0.035}	{0.002}	{0.018}
Above cutoff*score	0.0477	0.0477	0.0477		0.0744
	(0.042)	(0.047)	(0.045)		(0.049)
	[1.12]	[1.02]	[1.06]		[1.52]
	{0.261}	{0.308}	{0.292}		{0.130}
Observations	6500	6499	6499	6499	5794
Clustering	One-Way	Two-Way	Three-Way	Two-Way	Two-Way
Added controls?	None	None	None	None	+Demogr.

Notes: T-values in parentheses. This table first reproduces Table 8, Panel C, column (5) of PU in the first column. All regressions include cutoff fixed effects. Column (1) is clustered at the student level, columns (2), (4), and (5) are additionally clustered at the cutoff level, and column (3) is additionally clustered at the cutoff and town levels. The dependent variable is an indicator variable for whether a student experiences negative interactions at school.

Next, we vary the cutoffs in Table 7. The benchmark included all observations for

student*school cutoff pairs where the students were within one point of the cutoff. We alternatively try cutoffs of 1.4, 1.2, 0.8, and 0.6. Of these, we find that the results are robust to increasing the cutoff, but do weaken slightly when we reduce the cutoff. Trimming the cutoff to 0.6, for example, shrinks the coefficient from .045 to .019, with a t-value of just .82, and a p-value of .41 – no longer statistically significant at conventional thresholds. That said, the sample size also shrinks to just 4,161 observations, so it can be expected that significance would shrink. For that reason, we do not read too much into these results. We also drop the interaction term for a cutoff of 0.8 in column (5) and find the results are little changed.

Table 12: Negative Interactions: Varying the Cutoffs

	(1)	(2)	(3)	(4)	(5)
	b/se/t/p	b/se/t/p	b/se/t/p	b/se/t/p	b/se/t/p
1{Grade ≥ Cutoff}	0.0362	0.0387	0.0346	0.0188	0.0309
	(0.017)	(0.018)	(0.020)	(0.023)	(0.020)
	[2.13]	[2.15]	[1.70]	[0.82]	[1.57]
	{0.033}	{0.031}	{0.089}	{0.413}	{0.116}
Transition score (normalized)	-0.0912	-0.0965	-0.0723	-0.0394	-0.0412
	(0.030)	(0.034)	(0.043)	(0.060)	(0.023)
	[-3.04]	[-2.81]	[-1.68]	[-0.66]	[-1.79]
	{0.002}	{0.005}	{0.092}	{0.510}	{0.073}
Above cutoff*score	0.0720	0.0757	0.0472	0.0593	
	(0.031)	(0.037)	(0.049)	(0.071)	
	[2.29]	[2.07]	[0.96]	[0.84]	
	{0.022}	{0.039}	{0.338}	{0.403}	
Observations	8299	7475	5329	4161	5329
Cutoff	1.4	1.2	0.8	0.6	0.8

Notes: All regressions include cutoff fixed effects. The dependent variable is an indicator for whether a student experiences negative interactions at school.

Next, we alter the control variables, controlling for the raw transition score in column (1) of Table 8, and adding in a square term as well in column (2). In columns (3) and (4), all the data is used and dummy variables for being within a point of the cutoff and a point above the cutoff are included. The key result is robust in the first two columns, but the coefficients decline in the latter two columns. When we include a full panoply of demographic controls, and use the entire data with fixed effects, the results do appear to weaken.

Table 13: Parental Help: Varying the Cutoffs

	(1) b/se/t/p	(2) b/se/t/p	(3) b/se/t/p	(4) b/se/t/p
1{Grade ≥ Cutoff}	0.0396	0.0369		
	(0.018)	(0.018)		
	[2.25]	[2.06]		
	{0.024}	{0.039}		
Transition score	-0.0567	0.160	0.0123	0.0236
	(0.017)	(0.16)	(0.071)	(0.077)
	[-3.29]	[1.03]	[0.17]	[0.31]
	{0.001}	{0.305}	{0.862}	{0.760}
Transition score squared		-0.0135	-0.00295	-0.00376
		(0.0095)	(0.0042)	(0.0046)
		[-1.42]	[-0.70]	[-0.82]
		{0.157}	{0.481}	{0.413}
1{1 ≥ Grade ≥ Cutoff}			0.0191	0.0138
			(0.012)	(0.013)
			[1.54]	[1.05]
			{0.123}	{0.293}
1{Cutoff+1 ≥ Grade ≥ Cutoff-1}			-0.0438	-0.0443
			(0.017)	(0.018)
			[-2.51]	[-2.40]
			{0.012}	{0.016}
Observations	6499	6499	11745	10595

Notes: All regressions include cutoff fixed effects. The dependent variable is an indicator for whether a student experiences negative interactions at school Column (4) additionally includes several demographic controls, excluded for space.

Lastly, we conduct influential analysis in Table 10. In column (1), we compute dfbeta statistics for each observation using our benchmark regression, and then purge outliers with a dfbeta statistic with an absolute value higher than the (standard) cutoff of $2/\sqrt{N} = 2/\sqrt{6500} = .025$.[10] The results are robust. In column (2), we test the version using a cutoff of 1.2, and when we purge outliers for this version here, the results are statistically significant at 1%. Thus we conclude this result is robust to this particular influential analysis.

[10]We use the dfbeta post-regression command in Stata.

Table 14: Parental Help: Influential Analysis

	(1) b/se/t/p	(2) b/se/t/p
1{Transition grade cutoff}	0.0331	0.0268
	(0.0077)	(0.0065)
	[4.32]	[4.12]
	{0.000}	{0.000}
Observations	6273	7211
Cutoff	1.00	1.20

Notes: All regressions include cutoff fixed effects. The dependent variable is an indicator for whether a student experiences negative interactions at school.

4 Conclusion

Overall, we find that the first result (baccalaureate scores) is impressively robust, and the other two results were a bit less robust, which is understandable given the extreme difference in sample size (1.25 million vs. 6,500). The exogenous impact of going to a better school does appear to positively impact baccalaureate scores, and this result was robust with a coefficient in the same direction and statistically significant at 5% in all 20 robustness checks we ran and with t-values 92% of the original. Often we estimated slightly larger coefficients than the original study.

The finding that the parents of students who score slightly above a cutoff to go to a better school help less with their homework is somewhat less robust, with just 40% of checks statistically significant at 5% and with a coefficient in the same direction as the original study, but 79% of the checks were significant at 10%, and some that were not significant cut the sample and thus may have been expected to reduce t-values. The one concern was that influential analysis was problematic.

Lastly, the finding that students who scored above the cutoff reported more negative interactions with peers was robust in 60% of regressions. Smaller sample sizes vs. the original, and particularly when we narrowed the cutoff window, were a primary contributor to higher p-values. There was one problematic control when we

included all data and demographic controls, but otherwise, this result was mostly robust. The average t-value was 92% as large as in the original study.

References

Campbell, D., Brodeur, A., Dreber, A., Johannesson, M., Kopecky, J., Lusher, L. and Tsoy, N.: 2024, The robustness reproducibility of the american economic review, *Technical report*.

Pop-Eleches, C. and Urquiola, M.: 2013, Going to a better school: Effects and behavioral responses, *American Economic Review* **103**(4), 1289–1324.