**Appendices (Online)**

[*Appendix A:*  Questionnaire 1](#_Toc63150704)

[Introduction 2](#_Toc63150705)

[Questions of theoretical interest 3](#_Toc63150706)

[Post-survey questions 6](#_Toc63150707)

[*Appendix B:* The RDD Sample 10](#_Toc63150708)

[*Table B1:* Sample Summary Statistics 10](#_Toc63150709)

[*Appendix C:* Robustness Tests 11](#_Toc63150710)

[*Figure C1:* Sensitivity of Main RD Estimates to Bandwidth Choice 11](#_Toc63150711)

[*Figure C2:* Sensitivity of Main RD Estimates to Donut Hole Bandwidth 12](#_Toc63150712)

[*Figure C3:* RD Plot with Pre-treatment Variable: Father’s Education 13](#_Toc63150713)

[*Table C1:* RD Estimate of Effect of University Attendance on Father’s Education 14](#_Toc63150714)

[*Figure C4:* RD Plot with Pre-treatment Variable: Childhood Socio-Economic Status 14](#_Toc63150715)

[*Table C2:* RD Estimate of Effect of University Attendance on Childhood Socio-Economic Status 15](#_Toc63150716)

[*Table C3:* Local Randomization RD Estimates 16](#_Toc63150717)

[*Table C4:* Local Randomization RD Estimates of Effect of University Attendance on Pre-Treatment Variables 16](#_Toc63150718)

[*Appendix D:* World Values Survey (WVS) Analyses 17](#_Toc63150719)

[Measuring social capital using the World Value Survey 17](#_Toc63150720)

[*Figure D1*: University education and social capital in 50 countries with control variables 18](#_Toc63150721)

[*Appendix E:* Power Analyses 19](#_Toc63150722)

[*Appendix F:* Tertiary Educational Attainment 20](#_Toc63150723)

[*Figure F1:* Partial or complete tertiary educational attainment among 15-24 year-olds in 2010 20](#_Toc63150724)

# *Appendix A:* Questionnaire

This is the text of the questionnaire, entitled “Views of Romanian Society and Politics,” completed by survey respondents in Romania. The Romanian translation is available upon request.

The questionnaire is designed to test a number of different outcomes which form the basis of four studies that we conducted, as indicated below. Each was pre-registered separately at E-Gap. Each is now completed and under review. The main hypothesis for each study shows statistically significant results in the expected direction. We mention this to address concerns that we might be “fishing” among separately registered studies – conducting many different analyses and only reporting a subset which produced statistically significant results – which is not the case.

Normally, these four studies would be conducted in a serial fashion. However, in this case there is a finite sampling frame composed of individuals who fall near the cutoff point of our RDD. Once surveyed, it is inconvenient and potentially problematic to re-survey these individuals. Consequently, we needed to include all outcomes in the same survey. Only those questions falling under the rubric of social capital (below) are employed for the current project.

**Study 1: Social Capital**

*V* = voting

*P* = other modes of political participation

*M* = membership

*T* = trust

**Study 2: Cultural liberalism**

*CL* = cultural liberalism (labeled “traditional values” in the pre-registration document)

**Study 3: Ideology**

*I =* ideology (labeled “political liberalism” in the pre-registration document)

**Study 4: Corruption**

*C* = corruption/law-abidingness

## Introduction

This is a short survey of attitudes towards politics and society in Romania. It should take five to ten minutes to complete. It is entirely voluntary and you may exit the survey at any time.

Participants are eligible to win a free smart phone or tablet. The link to qualify for that offer is at the end of the survey.

All information provided in this survey is confidential.

If you have questions about the survey, or about any other aspect of this process, please contact one of the names below.

We appreciate your help!

Researchers:

* John Gerring, Professor, Department of Government, University of Texas at Austin, USA
* Emanuel Coman, Assistant Professor, Department of Political Science, Trinity College, Dublin, Ireland
* Stephen Jessee, Associate Professor, Department of Government, University of Texas at Austin, USA

## Questions of theoretical interest

*[Question order within this section is randomized.]*

1. For each of the following actions indicate whether you think it can be justified in some circumstances.

*10-point scale: 1 = Never justifiable … 5 = Sometimes justifiable … 10 = Always justifiable*

* Claiming government benefits to which you are not entitled *(C)*
* Avoiding a fare on public transport *(C)*
* Stealing property *(C)*
* Cheating on taxes if you have a chance *(C)*
* Someone accepting a bribe in the course of their duties *(C)*
* Homosexuality *(CL)*
* Prostitution *(CL)*
* Abortion *(CL)*
* Divorce *(CL)*
* Sex before marriage *(CL)*
* For a man to beat his wife *(CL)*
* Parents beating children *(CL)*
* For a magistrate or policeman to accept a bribe to close a case *(C)*
* Providing information accessed or obtained as part of your job, to people to whom it may be of interest *(C)*
* To receive money (a police officer) to not fine someone while in traffic *(C)*
* To favor the firm of a friend to obtain funds/ contracts from/with public institutions *(C)*
* For a bureaucrat to accept a present to fasten the issuing of a document *(C)*
* To help relatives or friends avoid sanctions or to be advantaged in solving some problems *(C)*
* To hire a relative in a public institution *(C)*
* For a public employee to receive a present AFTER they have favorably pursued your request, without breaking their work conduct. *(C)*
* To offer money or presents to a doctor to treat you *(C)*
* To give money or presents to a doctor because you are happy with the way they treated you *(C)*
* To give presents to a teacher because you are pleased with the way they educated your children *(C)*

1. To what extent do you agree or disagree with the following statements?

*Fully agree, Partially agree, Neither agree nor disagree, Partially disagree, Fully disagree*

* Most people can be trusted *(T)*
* Political leaders can be trusted *(T)*
* Punishment for corruption is too mild *(C)*
* Corruption is inevitable, and has always been around *(C)*
* Corruption is forgivable if no alternative is available in order to get something done *(C)*
* If I witnessed an act of corruption, I would report it *(C)*

1. How common are the following practices among people you know (friends, family, business associates)?

*Never, Occasionally, Every now and then, Regularly, All the time*

* Giving or receiving bribes in exchange for favors or work conducted, or to get out of a jam *(C)*

1. In the 2019 European Union elections, did you vote? *(V)*

*Yes, No, Don’t remember, Not old enough to qualify*

1. How interested are you in politics? *(P)*

*Not at all interested, Not very interested, Somewhat interested, Very interested*

1. For each of the following organizations, indicate your current involvement.

*Not a member, Inactive member, Active member*

* Church or religious organization *(CL)*
* Sport or recreational organization *(M)*
* Art, music or educational organization *(M)*
* Labor union *(M)*
* Political party *(P)*
* Environmental organization *(M)*
* Professional association *(M)*
* Humanitarian or charitable organization *(M)*
* Consumer organization *(M)*
* Self-help group, mutual aid group *(M)*
* Other organization *(M)*

1. Here are some forms of political action that people can take. For each one, indicate your involvement…

*I would never under any circumstances do this, I might do this, I definitely would do this or have done this*

* Voting *(V)*
* Signing a petition *(P)*
* Joining in boycotts *(P)*
* Attending peaceful demonstrations *(P)*
* Joining strikes *(P)*
* Any other act of protest *(P)*

1. How do you feel about the following statements?

*Strongly agree, Agree, Neither agree nor disagree, Disagree, Strongly disagree*

* When jobs are scarce, men should have more right to a job than women *(CL)*
* When jobs are scarce, employers should give priority to people of this country over immigrants *(CL)*
* If a woman earns more money than her husband, it's almost certain to cause problems *(CL)*
* Having a job is the best way for a woman to be an independent person *(CL)*

1. Apart from weddings and funerals, about how often do you attend religious services these days? *(CL)*

* More than once a week
* Once a week
* Once a month
* Only on special holy days
* Once a year
* Less often
* Never, practically never

1. How important is God in your life? *(CL)*

*10-point scale: 1=“not at all important,” 10 = “very important”*

1. All things considered, would you say that the world is better off, or worse off, because of science and technology? *(CL)*

*10-point scale: 1 = “the world is a lot worse off”… 10 = “the world is a lot better off”*

1. Below are various groups of people. Please indicate any that you would *not* like to have as neighbors…

* Drug addicts *(CL)*
* People of a different race *(CL)*
* People who have AIDS *(CL)*
* Immigrants/foreign workers *(CL)*
* Homosexuals *(CL)*
* People of a different religion *(CL)*
* Unmarried couples living together *(CL)*
* People who speak a different language *(CL)*

1. For each of the following statements, indicate your views.

*strongly agree, agree, neither agree nor disagree, disagree, strongly disagree*

* One of my main goals in life has been to make my parents proud *(CL)*
* When a mother works for pay, the children suffer *(CL)*
* On the whole, men make better political leaders than women do *(CL)*
* A university education is more important for a boy than for a girl *(CL)*
* On the whole, men make better business executives than women do *(CL)*
* Being a housewife is just as fulfilling as working for pay *(CL)*

1. If there were a national election tomorrow, for which party on this list would you vote? *(L)*

* The Save Romania Union
* The Social-Democratic Party
* The National Liberal Party
* Democratic Alliance of Hungarians in Romania
* Alliance of Liberals and Democrats
* People’s Movement Party
* Other

## Post-survey questions

1. What was the last high school that you attended?

* County [prepopulated menu]
* City [prepopulated menu]
* High school [prepopulated menu]

1. Have you finished high school?

* No
* Yes, without Bac
* Yes, with Bac

*[If the latter, then ask…]*

1. In what year did you finish high school?

*[If high school finished in 2019, then ask…]*

1. Which of the following best describes your current situation?

* In the process of applying for admission to university, with intention to start this fall
* Admitted to university but have not started yet
* Enrolled in university and attending
* None of the above

1. Have you ever attended, or do you plan to attend, a vocational school (post-secondary)?

* I am planning to attend
* I am currently attending
* I have attended

1. Have you ever – currently, or in the past – attended university (even for a short period of time)?

* No [skip the following questions on university education]
* Yes

1. Are you currently enrolled in university?

* No
* Yes

1. How many years of university education have you completed?

* Less than one year
* One year
* Two years
* Three years
* Four years
* Five years or more

1. What is the name of the last university you attended, or the university you currently attend or plan to attend?

*pre-populated list*

1. What is, or was, your major course of study for the BA?

*pre-populated list*

1. Please indicate your sex

* Male
* Female

1. Are you currently (answer one of the following)… *(TV: code Married and Widowed as traditional, all others as non-traditional)*

* Married
* Living together as married
* Divorced
* Separated
* Widowed
* Single

1. Have you had any children? *(TV: code as interval scale, with more children indicating more traditional values)*

* No children
* One child
* Two children
* Three children
* Four children
* Five children
* Six children
* Seven children
* Eight or more children

1. Are you employed now or not?

* If yes, about how many hours a week?
* If more than one job: only for the main job (code one answer):
* Yes, has paid employment:
* Full time employee (30 hours a week or more)
* Part time employee (less than 30 hours a week)
* Self employed
* No, no paid employment:
* Retired/pensioned
* Housewife not otherwise employed
* Student
* Unemployed
* Other (write in):\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

1. Would you describe yourself as belonging to the

* Upper class
* Upper middle class
* Lower middle class
* Working class
* Lower class

1. Imagine an income scale on which 1 indicates the lowest income group and 10 the highest income group in your country. Indicate in what group your household is, counting all wages, salaries, pensions and other incomes that come in.

*10-point scale: 1 = Lowest group…10 = Highest group*

1. Please indicate your year of birth

*year*

1. In what country do you currently live?

*pre-populated list*

*[Ask only if answer to previous question is Romania]*

1. In what town or city do you currently live?

*pre-populated list*

1. What is the highest educational level that your father attained?

* No formal education
* Incomplete primary school
* Complete primary school
* Incomplete secondary school: technical/vocational type
* Complete secondary school: technical/vocational type
* Incomplete secondary: university-preparatory type
* Complete secondary: university-preparatory type
* Some university-level education, without degree
* University-level education, with degree
* Master degree
* Doctoral (PhD)

1. Would you describe your family – when you were growing up – as belonging to the

* Upper class
* Upper middle class
* Lower middle class
* Working class
* Lower class

1. If you wish, please tell us what ethnic group you identify with.

*pre-populated list*

# *Appendix B:* The RDD Sample

### *Table B1:* Sample Summary Statistics

|  |  |
| --- | --- |
| Female | 43*%* |
| Single (never married) | 77 |
| One or more children | 6 |
| Employed | 52 |
| Romanian ethnicity | 91 |
| Father’s education level |  |
| Vocational school or less | 11 |
| Some university | 17 |
| Completed university | 48 |
| More than university | 23 |
| Childhood SES |  |
| Lower class | 10 |
| Working class | 34 |
| Lower middle class | 31 |
| Upper middle class | 21 |
| Lower class | 5 |
| Current SES |  |
| Lower class | 10 |
| Working class | 30 |
| Lower middle class | 5 |
| Upper middle class | 23 |
| Lower class | 33 |
| Age |  |
| 18 | 2 |
| 19 | 12 |
| 20 | 27 |
| 21 | 15 |
| 22 | 21 |
| 23 | 18 |
| 24 | 4 |
| 25+ | 2 |
| Descriptive statistics for the sample of Romanians whose *bac* scores fall within our stated bandwidth and who completed our survey. *N*=1,515 | |

# *Appendix C:* Robustness Tests

This appendix presents the results of the various robustness tests mentioned in the paper.

In Figure C1, we explore how varying the bandwidth of our RD analysis affects the estimated treatment effect. Because we surveyed only individuals whose initial BAC scores are within .2 of the passage threshold (i.e. those who scored between 5.8 and 6.2), we conduct these analyses by narrowing the bandwidth sequentially at each distance. In other respects, these analyses replicate the benchmark RD analysis as described in the paper, clustering standard errors by score and reporting robust confidence intervals. As with any such analysis, narrowing the bandwidth decreases the precision of the estimates. Nonetheless, it is informative to survey the point estimates derived from these sub-samples to see if there is any overall pattern.

### *Figure C1:* Sensitivity of Main RD Estimates to Bandwidth Choice

****

*Points show estimated effects based on given bandwidth radius (i.e. using only observations within that radius distance of the threshold of 6) and bars show robust 95% confidence intervals, all estimated using the local polynomial RD method implemented in the rdrobust function using the same arguments as the main analysis in the paper. Note that bandwidth radius of .2 (far right point and bar) is the same as bandwidth used in the main analysis in the paper (i.e. using respondents over the entire range of bac scores in our sample).*

Figure C1 demonstrates that the estimated effect of university attendance on social capital is remarkably consistent across the chosen bandwidths. The only noticeable change is that the estimates become more precise as the bandwidth is widened, as expected (since the sample is expanding).

Next, we conduct a so-called “donut hole” robustness check (e.g. Batarji et al. 2011). This test accomplishes much the same thing as the previous test except that here we *narrow* the bandwidth by removing observations lying at designated radii from the cutoff (6.0). In the unlikely scenario that some students were able to manipulate their scores so as to achieve a barely passing grade on the *bac*, this test eliminates such potentially problematic data points from the analysis. Another benefit of this approach is in assessing how dependent the main results are to extrapolations near the threshold. As previously, we anticipate that dropping observations will decrease the precision of our estimates. Thus, we are more interested in the point estimates for each analysis than the confidence intervals.

### *Figure C2:* Sensitivity of Main RD Estimates to Donut Hole Bandwidth

****

*Points show estimated effects based on given donut hole radius (i.e. dropping observations within that radius distance of the threshold of 6) and bars show robust 95% confidence intervals, all estimated using the local polynomial RD method implemented in the rdrobust function using the same arguments as the main analysis in the paper. Note that donut hole radius of 0 (far left point and bar) is the same as bandwidth used in the main analysis in the paper (i.e. using all respondents in our sample).*

Results displayed in Figure C2 show the estimated treatment effects when observations within a given distance (from 0 to .2) of the threshold are dropped from the sample are extremely stable. There is no evidence of meaningful changes in the estimated effect that might result from students being able to sort themselves around the cutoff.

Figures C3 and C4 show RD plots substituting each of two notable pre-treatment covariates as our dependent variable: father’s education and the socio-economic status (SES) of the family during childhood. Both tap what is likely the biggest potential confounder in the relationship between education and social capital, namely educational resources and parental concern with education during childhood. Both of these measures are based on self-reports from our survey (see questions 16 and 17 in Appendix A above). As such, one must be somewhat wary of errors that are correlated with the treatment. For example, respondents who attend university, and thereby come into regular contact with affluent students, might be inclined to exaggerate their family’s socioeconomic status in order to fit in. Alternatively, they might downgrade their family’s socioeconomic status, which by reference to their affluent acquaintances seems meager. This sort of bias seems less likely with respect to father’s education.

In any case, these two measures, while not perfect, offer an opportunity to assess whether those who narrow pass, or narrowly fail, the BAC are similar in background characteristics that might serve as confounders. For these analyses, we adopt the format of the benchmark RD analysis, as described in the paper.

### *Figure C3:* RD Plot with Pre-treatment Variable: Father’s Education

****

*Points show average level of father’s education among respondents having each unique value of bac score. Vertical line denotes (fuzzy) treatment threshold of 6. Horizontal lines on each pane show averages for all respondents above/below threshold. The size of each point is proportional to the number of observations at that Bac score value.*

### *Table C1:* RD Estimate of Effect of University Attendance on Father’s Education

|  |  |  |  |
| --- | --- | --- | --- |
|  | Estimated effect on Father’s education level | 95% CI | p-value |
| University Attendance | -.05 | [-1.23, .84] | .72 |

*Analysis using rdrobust function from rdrobust R package (version 0.99.9) clustering standard errors by bac score, using all of the data in our sampling bandwidth of 5.8 to 6.2, and otherwise using the function’s default arguments. Confidence interval and p-value are based on robust bias-corrected results. N=1170 (531 obs. below threshold, 639 obs. above threshold)*

### *Figure C4:* RD Plot with Pre-treatment Variable: Childhood Socio-Economic Status

****

*Points show average reported childhood socio-economic status among respondents having each unique value of bac score. Vertical line denotes (fuzzy) treatment threshold of 6. Horizontal lines on each pane show averages for all respondents above/below threshold. The size of each point is proportional to the number of observations at that Bac score value.*

### *Table C2:* RD Estimate of Effect of University Attendance on Childhood Socio-Economic Status

|  |  |  |  |
| --- | --- | --- | --- |
|  | Estimated effect on childhood SES | 95% CI | p-value |
| University attendance | .01 | [-.52, .55] | .96 |

*Analysis using rdrobust function from rdrobust R package (version 0.99.9) clustering standard errors by bac score, using all of the data in our sampling bandwidth of 5.8 to 6.2, and otherwise using the function’s default arguments. Confidence interval and p-value are based on robust bias-corrected results. N=1166 (526 obs. below threshold, 640 obs. above threshold)*

Visual inspection of Figures C3 and C4 do not suggest a jump at the threshold value of 6. Likewise, formal RD analyses presented in Tables C1 and C2 show that the estimated “effects” of university attendance on these pre-treatment variables are small in magnitude and nowhere close to statistically significant. Thus, this robustness check shows no evidence of problematic differences between treated and untreated observations in terms of their values on relevant pre-treatment covariates.

Finally, we conduct RD analyses mirroring the results in the paper, this time using a local randomization (LR) framework (Cattaneo, Titiunik, and Vasquez-Bare 2016) rather than the continuity-based ones in the main paper. The LR approach relies on different assumptions than the continuity-based framework. Loosely speaking, the LR approach assumes that within the chosen bandwidth, the potential outcomes for the dependent variable are unrelated to the score variable. If we obtain similar estimates using the LR approach as in the main analyses in our paper, this might be thought to add credibility to the findings. Table C3 shows estimated treatment effects from local randomization RD analyses for each of the three dependent variables we consider.

In the rightmost column of Table C3, we conduct a local randomization RD analysis using the entire sample (in which *bac* scores range from 5.8 to 6.2). This wider bandwidth contains a much larger number of observations but also a more heterogeneous sample. As such, this analysis will have lower variance, but possibly much larger bias. As such, it should probably be viewed skeptically. (In our view, the continuity-based RD analyses discussed in the paper offer a superior approach.)

Table C3 shows estimated treatment effects from local randomization RD analyses. The far-right column of the table shows estimates using our entire sample (bandwidth of .2, i.e. from 5.8 to 6.2). These estimates are highly significant and substantively large, albeit not as large as those from our main, continuity-based, RD analyses in the paper. It is also comforting that varying the bandwidth used in these local randomization analyses produces little change in the estimated effects (see the middle columns of Table C3). In fact, we can even conduct a local randomization analysis using only the two score variable values closest to the threshold (5.983333 and 6) and we obtain a significant positive treatment effect estimate despite a sample size of just over 100 observations.

### *Table C3:* Local Randomization RD Estimates

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Bandwidth | minimum | .075 | .1 | .125 | .15 | .175 | .2 |
| Estimated treatment effect (p-value) | .61 (.002) | .48 (<.001) | .49 (<.001) | .46 (<.001) | .45 (<.001) | .48 (<.001) | .43 (<.001) |
| *N* (*N*above, *N*below) | 104 (53, 51) | 421 (181, 240) | 612 (285, 327) | 698 (320, 378) | 801 (366, 435) | 1,007 (460, 547) | 1,210 (542, 668) |

*Analysis using rdrandinf function from rdlocrand R package (version 0.7.1), using all of the data in our sampling bandwidth of 5.8 to 6.2, and otherwise using the function’s default arguments. “minimum” bandwidth uses only observations with either the highest passing score below the cutoff (.5983333) or the lowest passing score (6).*

As a final robustness test, we conduct local randomization RD analyses using the same approach described above, this time with each of the family background variables as outcomes. Results for these pretreatment covariate diagnostics for father’s education level and self-reported childhood SES are presented in Table C4.

These results show some evidence of differences in pre-treatment variables, particularly for father’s education. This might be viewed as evidence against the assumptions necessary for local randomization RD analyses, at least within the full bandwidth used in our data (recall the continuity-based RD analyses in the paper showed no evidence of jumps in either of these pre-treatment variables at the cutoff, which might suggest that the continuity-based RD framework is more appropriate for our application).

In the local randomization framework, a common approach is to employ automated window selection procedures using one or more pre-treatment covariates to determine the size of the window around the score variable threshold within which it is plausible that the local randomization holds (i.e. such that there is not evidence of differences in the pre-treatment covariates on average between those above and below the treatment threshold within the selected window). We conduct this analysis using the window selection procedure implemented in the rdlocrand package in R. We specify both father’s education and childhood SES as pre-treatment covariates to serve as balance checks, using the function’s standard arguments.

The window selected by the function is 5.9 to 6.1 (i.e. .1 on either side of our threshold of 6 on the score variable). Conveniently, Table C3 shows the results of estimating a local randomization RD using only observations within this bandwidth, giving a point estimate of .49 (*p*<.001). This estimate is quite similar to those obtained in other windows, with the exception of the minimum window, which uses only a small amount of the available data and therefore is less precise.

### *Table C4:* Local Randomization RD Estimates of Effect of University Attendance on Pre-Treatment Variables

|  |  |  |  |
| --- | --- | --- | --- |
|  | Estimate | 95% CI | p-value |
| Father’s Education | -.21 | [-.58, .06] | .02 |
| Childhood SES | -.09 | [-.34, .06] | .17 |

*Analysis using rdrandinft function from rdlocrandt R package (version 0.7.1) using all of the data in our sampling bandwidth of 5.8 to 6.2. N=1166 (526 obs. below threshold, 640 obs. above threshold)*

# *Appendix D:* World Values Survey (WVS) Analyses

## Measuring social capital using the World Value Survey

To create the measurement of social capital used in the analyses graphed in Figure 1 of the main text, we apply the procedure described on page 20 to a series of questions in WVS wave 6, grouped under four subcategories: (i) voting (V); (ii) trust (T); (iii) membership in organizations (M); (iv) other modes of political participation (P). We used the following questions:

V227: *When national elections take place, do you vote always, usually or never? (V)* (recoded “3”= never; “4”=usually; “5”=always)

V85-V89: *Now I’d like you to look at this card. I’m going to read out some forms of political action that people can take, and I’d like you to tell me, for each one, whether you have done any of these things, whether you might do it or would never under any circumstances do it (recoded "3"- would never do; “4”-might do; “5”-have done).*

* *V85- signing a petition (P)*
* *V86- joining in boycotts (P)*
* *V87- attending peaceful demonstrations (P)*
* *V88- joining strikes (P)*
* *V89- any other acts of protests (P)*

V84*: How interested would you say you are in politics? (recoded “4”-not at all interested; “5”- not very interested; “6”-somewhat interested; “7”-very interested) (P).*

V25-V35: *Now I am going to read off a list of voluntary organizations. For each organization, could you tell me whether you are an active member, an inactive member or not a member of that type of organization (coded “0”-don’t belong; “1”-inactive member; “2”- active member)?*

* *V25- church or religious organization (M)*
* *V26- sport or recreational organization (M)*
* *V27- art, music or educational organization (M)*
* *V28- labor union (M)*
* *V29- political party (P)*
* *V30- environmental organization (M)*
* *V31- professional association (M)*
* *V32- humanitarian or charitable organization (M)*
* *V33- consumer organization (M)*
* *V34- self-help group, mutual aid group (M)*
* *V35- other organization (M)*

The following countries were dropped from the analysis because they had missing observations on various variables: Belarus (missing V85-V89); Taiwan (missing V35 and V89); Ecuador (missing V227); Germany (missing V35); Kuwait (missing V86-V89 and V28-V29); Morocco (missing V89); Qatar (missing V86-V89 and V28-V29); Singapore (missing V85-V89); Spain (missing B89 almost entirely); Uzbekistan (missing V85-V89).

### *Figure D1*: University education and social capital in 50 countries with control variables



OLS estimates, surrounded by 95% confidence intervals, of our index of social capital regressed against a binary indicator of university education. Models include the following covariates: gender, age, socioeconomic status.

# *Appendix E:* Power Analyses

Schochet (2008) proposes an approach for conducting power calculations that involves calculating a design effect for a regression discontinuity analysis as compared to a simple difference of means test in a simple randomized controlled trial (RCT). The design effect under a linear data generating process depends on the distribution of the score variable around the threshold for treatment assignment. Specifically, the design effect is calculated as where is the correlation between the score variable and treatment assignment.

Because we actually have collected data on the population distribution of the score variable, a natural estimate of this design effect can be constructed based on this. These data show a roughly uniform in the bandwidth we plan to sample from, which Schochet shows produces a design effect of 4.

Furthermore, because our RD design is fuzzy rather than sharp, we must also adjust for noncompliance, in particular the difference in treatment probability for those with scores below the threshold and those above, which would be 1 in a RCT. We expect this difference to be in the ballpark of .7 in our application (for example, if only 5% of those with *bac* scores below 6 attended university, while 75% of those with scores of at least 6 attended university, the difference in treatment probabilities would be .75-.05=.7).

Finally, we require an estimate (educated guess) about the likely size of the effect of our treatment on our dependent variable relative to the standard deviation of the dependent variable. Obviously, this will be quite speculative, but a natural starting point is to use estimates of this relationship from observation studies, as in the analyses of the World Values Survey in Appendix D above. Regression a standardized social capital measure on a dummy variable for university attendance produces an estimated regression coefficient of .49 with a standard error of .08. Note that this is slightly different from the coefficient estimate plotted for Romania in Appendix D. This is because the present regression uses a dependent variable that is standardized to have mean zero and standard deviation 1 among respondents from Romania as the dependent variable in our RD analysis will be. In Appendix D, by contrast, the dependent variable is standardized across all respondents in the World Values Survey (not just those in Romania) to better facilitate interpretation and comparison of estimates across multiple countries.

We conduct power analyses under two different assumptions: that the true effect size of interest in our RD is the same as the estimated relationship between university attendance and social capital among Romanian World Values Survey respondents, and that the true effect size is half of this value. These two assumptions would imply that to achieve 80% power at the .05 significance level in our fuzzy RD analysis, we would require either 1,097 or 4,365 respondents, respectively.

There are several reasons these estimated sample sizes are obviously uncertain. We find it likely that a sample of between 2,000 and 3,000 will be sufficient for our purposes. First, analyzing only respondents whose *bac* scores are within a narrow bandwidth, we are likely to decrease the variance of the dependent variable relative. Second, our social capital variable contains more component variables and is thus likely to include lower measurement error than the variable used in the World Values Survey regressions.

# *Appendix F:* Tertiary Educational Attainment

Using data from Barro and Lee (2013), we calculate the percentage of 15-24 year-olds who have completed at least partial tertiary education in 2010. As shown in Figure F1, Romania falls very near to the middle: 79th of 146 countries.

### *Figure F1:* Partial or complete tertiary educational attainment among 15-24 year-olds in 2010

