Online Appendix for
Compulsory Voting Can Increase Political Inequality: Evidence From Brazil

Gabriel Cepaluni
São Paulo State University

F. Daniel Hidalgo
Massachusetts Institute of Technology

January 28, 2016

Appendix

A Laws on Compulsory Voting

1988 Brazilian Constitution

Paragraph 1. Electoral enrollment and voting are:

I. mandatory for persons over eighteen years of age;

II. optional for:

  (a) the illiterate;

  (b) those over seventy years of age;

  (c) those over sixteen and under eighteen years of age.

Law No. 4737, JULY 15, 1965 (Electoral Code)

Article 7. Voters who fail to vote and who do not justify their absences to the electoral judge within thirty (30) days after the election will pay three (3) to ten (10) percent of the minimum wage in the region, imposed by the electoral judge and charged as provided in Article 367.

Paragraph 1. Without a voting proof, a paid fine or a proper justification in the last election, voters may not:

I. To compete in a test for public office ("concurso público") or to take over a public job;

II. To receive wages, payments, salaries or earnings from public jobs two months after the election;

III. To participate in a public bid;

IV. To get public loans from a public bank or from the State (in all its levels: Federal, State, municipality etc.);

V. To obtain a passport or a ID card;

VI. To renew its enrollment in a public education institution or any education institution supervised by the government;

VII. To perform any act for which requires discharge from military service or income tax.

Paragraph 2. The Brazilian born or naturalized, over 18 years, may not take the actions listed in the previous paragraph without proof of being listed unless exempted in Articles 5 and 6, Paragraph 1.

Paragraph 3. It will be cancelled the registration of voters who have not voted for three (3) consecutive elections and who have not payed the fine or have not justified their absences within six (6) months from the date of the last election that they were supposed to vote.

Article 8. The Brazilian born who do not register up to 19 years of age or the naturalized Brazilian that do not enlist until a year after acquiring the Brazilian nationality will pay 3 (three) to ten (10) percent of the value of minimum wage in the region imposed by the electoral judge and charged at the time of the voter registration.

Sole Paragraph. The penalty does not apply to non-registered voters that registered to vote up to one hundredth days before the subsequent election up until their 19 years old birthdays.

Article 9. Those responsible for violations of the Articles 7 and 8 will pay one (1) to three (3) minimum wages prevailing in the electoral zone or disciplinary suspension within thirty (30) days.

Article 10. The electoral judge will provide a document of exemption of lawful sanctions to those who do not vote for a justified reason and not registered under Articles 5 and 6, Paragraph 1.

Article 11. Voters may pay the fine for not voting for the electoral judge in the electoral zone they are in if they are outside of their electoral zones and need a voting discharge document issued by the Electoral Court.

B Survey Details

We hired the survey firm *Instituto Análise* in 2014 to explore whether penalties for individuals that fail to vote might unequally affect citizens of different social status. The firm interviewed 1230 individuals above 18 years old in face-to-face interviews.

Instituto Análise first selects a randomly stratified sample of municipalities. The stratification takes into consideration the geographic location of the municipalities (regions and states), its size and its type (countryside, metropolitan region or capital city). In a second stage the firm selects a randomly stratified sample of a set of census blocks (*setores censitários*) inside of each of these municipalities. In this stage the firm takes into consideration several variables to conduct the stratified randomization, such as median income, literacy and family size. The Brazilian Institute of Geography and Statistics (*Instituto Brasileiro de Geografia e Estatística—IBGE*) provides statistical information on the municipalities and its census blocks to design the sample. Finally, the results of the survey are weighted to represent the Brazilian population. Calculations of standard errors and confidence intervals incorporates the two stage stratified sampling design.

C Estimating the Voting Age Population

For the 17–18 year old sample, we estimate the number of living Brazilian citizens born on each day in our estimation window. To do so using the PNAD data, we model the number of people with a birth date on any given day as a smooth function of day of the year, as well a intercept shift by day of the week. Allowing the number of people born each day to vary smoothly by date accounts for seasonality in births, while including day-of-the-week fixed effects adjust for the fact that fewer births occur on weekends.

To show that this modeling strategy is necessary, in figure A.1 we present the per-day total number of births in our registration data for citizens born in a 1 year window around October 7, 1992. Because these individuals are 19 or 20 in 2012, they are all affected by the compulsory voting law and thus differential registration rates by date should be minimal. As is visible in the data, there are relatively more births in March and April than in October. Moreover, there are substantially fewer births, on average, on Saturdays and Sundays than weekdays.

To estimate the total number of eligible voters with birth dates on any given day, we rely on the National Household Survey (*Pesquisa Nacional por Amostra de Domicílios* or PNAD), a large scale survey administered by the Brazilian census bureau. The target population of the survey is all residents of Brazil and the annual sample size is over 350,000 individuals. Crucially, the PNAD

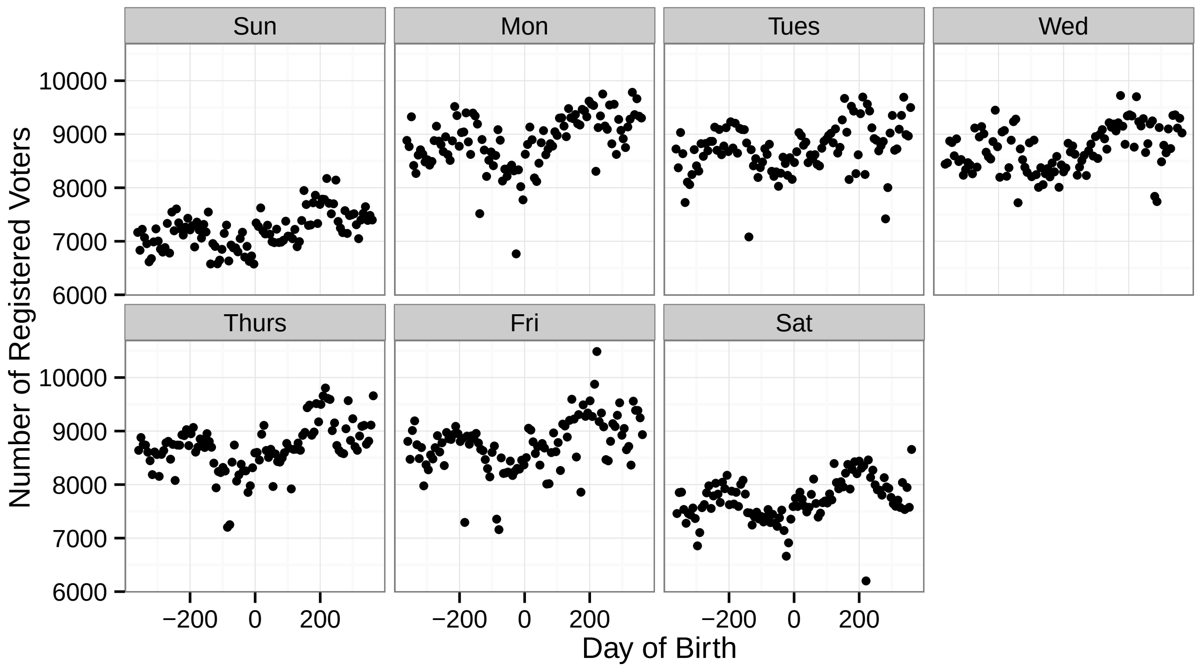

Figure A.1: Per-day total number of births, by weekday, for citizens born in a 1 year window around October 7, 1992.

includes date of birth in their public microdata. Specifically, we sum the sampling weights of all respondents born on each day to generate an estimate of the voting age population with every birth-date in our estimation window.[1] Figure A.2 plots these estimates separately by each day of the week for respondents born in 1994 and 1995.

Figure A.2 shows that the per day estimates are fairly variable, suggesting a considerable amount of sampling error, especially when compared to the registration data plotted in figure A.1. To generate more stable estimates, we assume that the average number of people with birth dates on each day varies smoothly by day, with an allowance for additive shifts by day of the week. To estimate smoothed day-specific population totals, we fit a generalized additive model (GAM) to a one year window around the target date of October 7, 1994 of the following form:

$$\mathbb{E}[y_i] = w_i + f(x_i)$$

where y_i is the total number of births on day i, w_i is a day of the week fixed effect and x_i is

[1]Because we aggregate data from the 2009, 2011, 2012, and 2013 PNAD, we divide each respondent's sampling weight by four to adjust for the multiple waves.

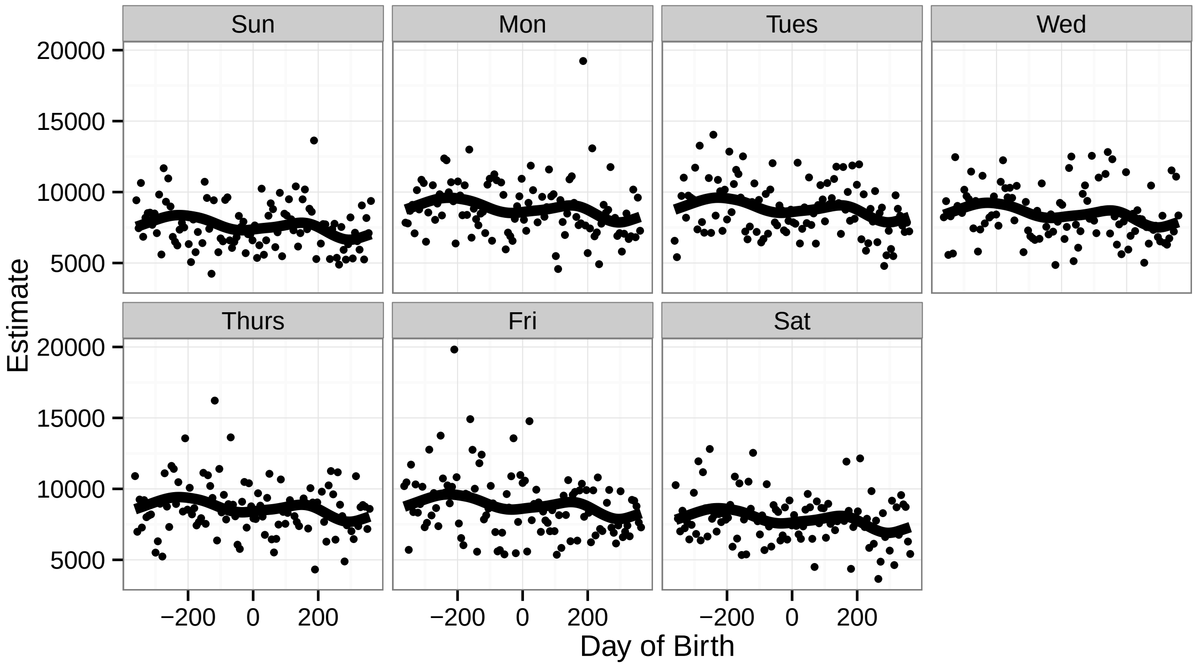

Figure A.2: Estimated number of individuals by birth-date for individuals born in 1994 and 1995. Data has been faceted by day of the week. X axis has been centered at October 7, 1994. Black line is an estimate from a generalized additive model. Points are unsmoothed estimates from the 2009, 2011, 2012, and 2013 PNAD.

the day variable, centered around the threshold. To estimate the model, we use a penalized cubic regression spline with 30 knots evenly spaced across the support of the data. The penalty term is chosen to minimize the generalized cross validation criterion. For details see Wood (2006). The model estimates are represented by the thick black lines in figure A.2.

C.1 Estimating Strata-specific Voting Age Population

Because we report results for several subsamples, when estimating treatment effects for particular populations, we estimate the denominator for the target populations by restricting the PNAD survey samples to the relevant population. When estimating heterogeneous treatment effects by covariate strata, i.e. education, we construct the samples by estimating the GAM separately in each stratum and then aggregate the per-stratum estimates into a combined dataset.

For education, we use self-reported education according to the PNAD survey. While this data is self-reported and thus subject to response biases[2], the education data in the voter registry is also

[2]The PNAD estimates are very similar to census estimates that use much larger samples. For example, the 2010

self-reported. If response biases are similar for the two sources, then self-reported education in the PNAD should be a good proxy for self-reported education in the voter registry. For the full voter registry, the proportion of voters without primary education is 50%. In the PNAD, the proportion of voters without a primary education is 53%. The similarity of these proportions across both datasets suggests that differential response bias is not a large concern.

D Smoothness of the Forcing Variable

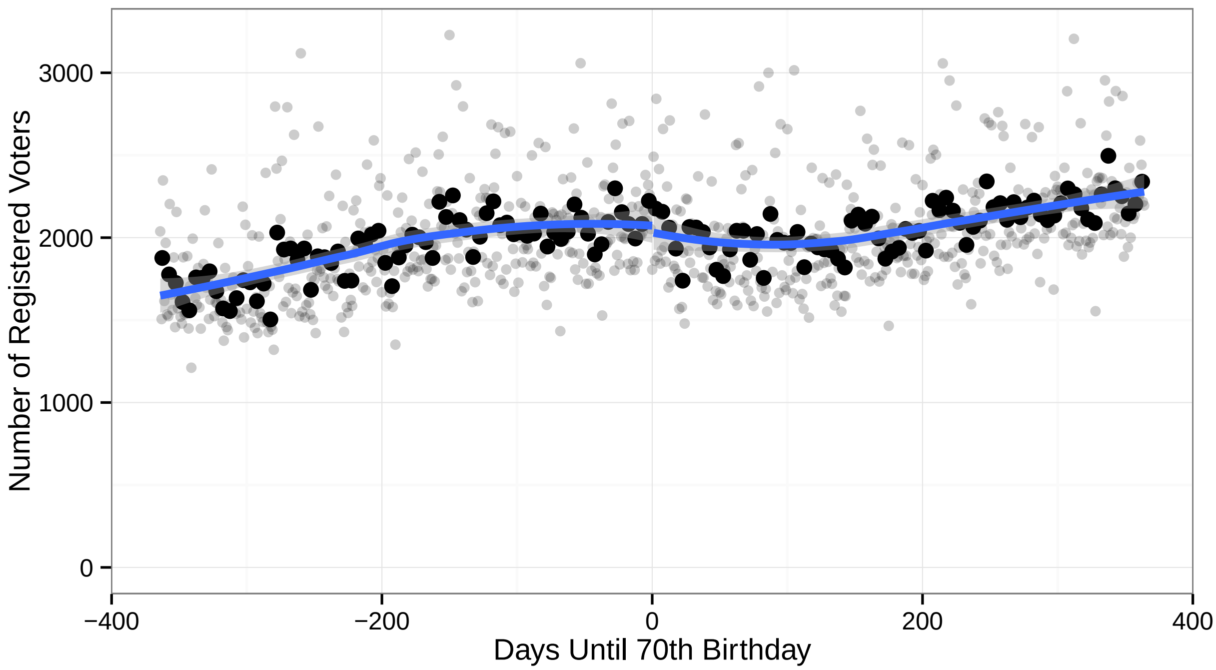

Figure A.3: Number of registered voters by birthdate for the 1942 sample. Dark points represent average turnout in 5 day bins, while gray points represent average turnout in 1 day bins. Blue lines are estimates from a loess model estimated separately around the threshold

As discussed in the main text, for voters born in 1942, we condition on registration status. To show that this does not induce post-treatment bias, we plot the number of registered voters by each birthdate in Figure A.3. As evidenced by the plot, there is no discontinuous jump at the threshold.

census reported that among residents between 20 and 24, 75% had completed primary education. According to the PNAD for the same age group, we obtained the exact same figure. Both however, are based on self-reported education.

E Covariate Balance

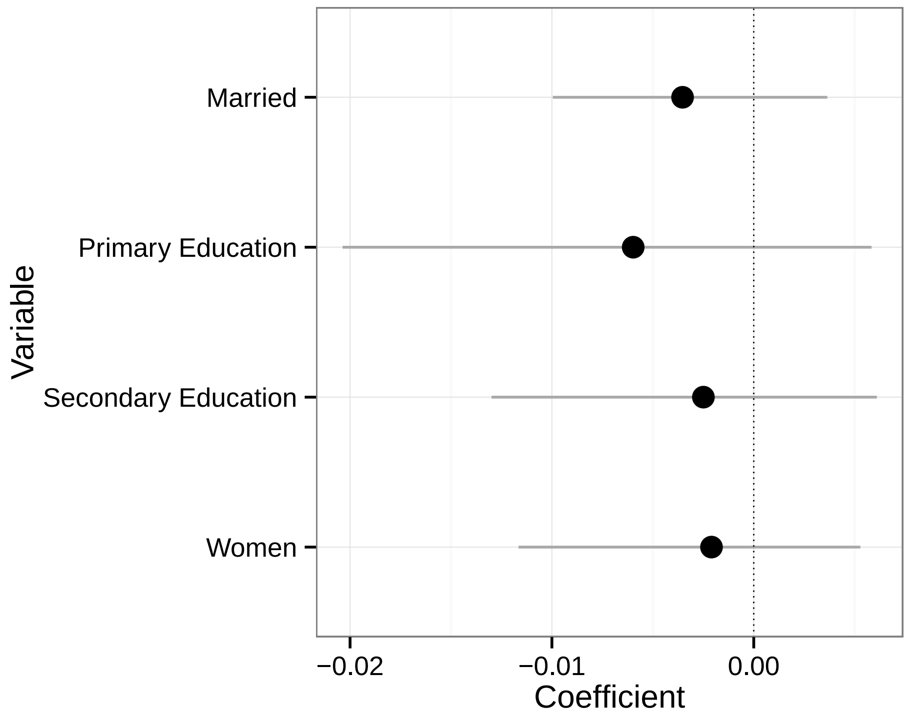

Figure A.4: Covariate Balance for 1942 Sample. Dots represent point estimates from a local linear regression and and lines are 95% robust confidence intervals proposed by Calonico et al. (2014).

While covariates included in the voter registry are few, we checked for smoothness in these pre-treatment variables for voters born in 1942. As evidenced by the point estimates from local linear regressions plotted in Figure A.4, imbalances are small and statistically insignificant.

For voters born in 1994, we check balance on the same variables using the PNAD data, as the survey measures variables of interest for the target population. As evidenced by the point estimates from local linear regressions plotted in Figure A.5, imbalances are small for the education variables.[3] For gender, there is some evidence of imbalance. Conditioning on this variable, however, barely changes the point estimate: the effect of compulsory voting on turnout in the full sample after controlling for gender is 0.119, which is virtually identical to the estimate reported in the main text.

[3]Given that very few 17–18 year olds would be married, we do not check balance on this variable.

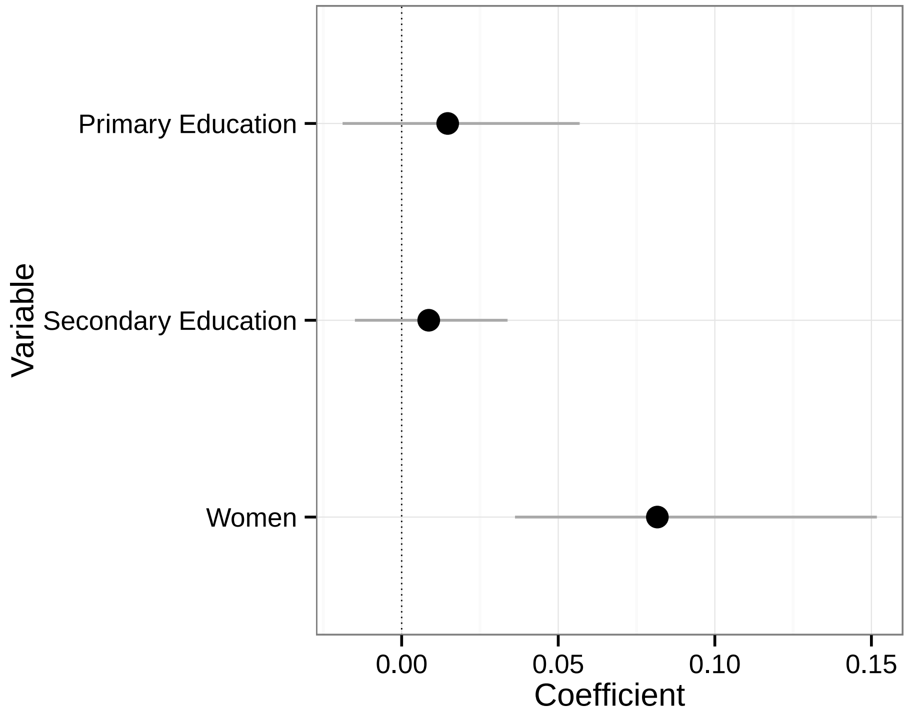

Figure A.5: Covariate Balance for the 1994 Sample. Dots represent point estimates from a local linear regression and lines are 95% robust confidence intervals proposed by Calonico et al. (2014).

F Discontinuity Plot for 17–18 Year Old Sample

Figure A.6 shows discontinuities in turnout by education status for 17 and 18 year olds.

G Alternative Results Under a Local Randomization Assumption

In this paper, we follow the advice of Lee and Card (2008) and aggregate our data to the day-of-birth-level, which is the finest level of resolution of our forcing variable. This approach is conservative because it increases standard errors relative to standard approaches that treat all respondent observations as IID. As argued by Lee and Card (2008), regression discontinuities with *discrete* forcing variables may be inconsistent with the standard theoretical results (Hahn et al., 2001) because the discrete nature of the forcing variable no longer allows for the computation of averages within arbitrarily small neighborhoods of the threshold. Lee and Card (2008) treat this issue as an additional source of specification error when seeking to model the true underlying regression function and they recommend that analysts adjust their uncertainty estimates to reflect

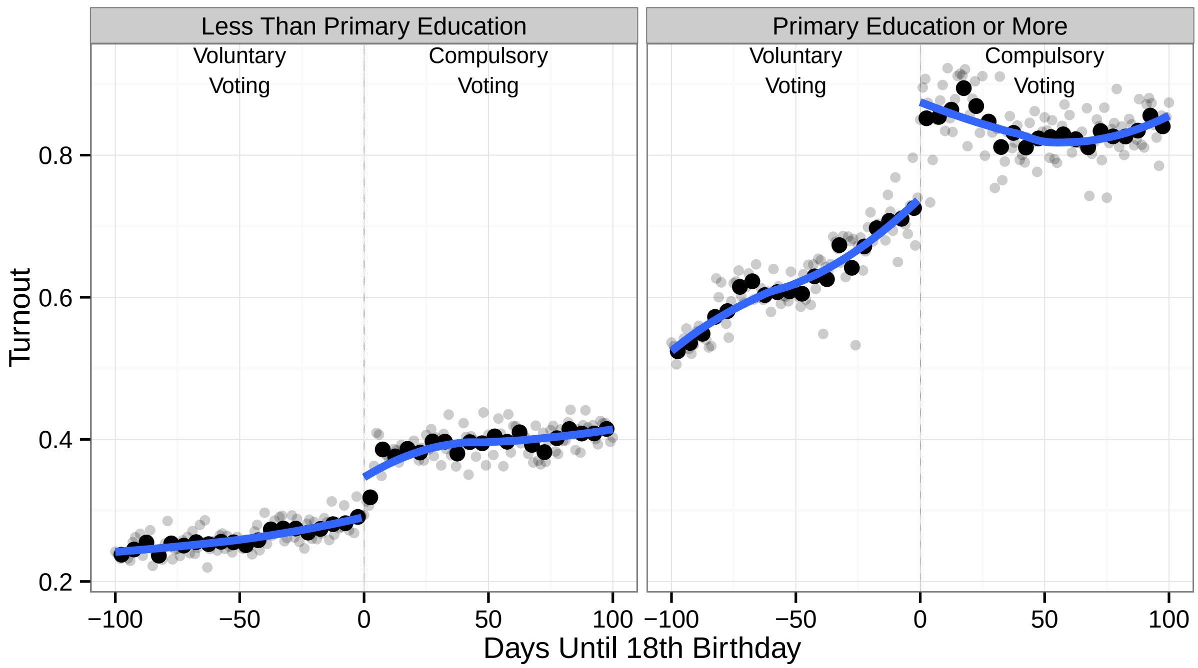

Figure A.6: Effect of Compulsory Voting by Education Status for the 17–18 Year Old Sample. Dark points represent average turnout in 5 day bins, while gray points represent average turnout in 1 day bins. Blue lines are estimates from a loess model.

this potential source of error. In their framework, mis-specification error is modeled as a random shock that occurs at the level of the forcing variable unit (i.e. day) and thus this error is correlated for all units with the same value of the forcing variable. This correlation leads to solutions such as clustering standard errors or aggregating to the level of the forcing variable. We follow their latter suggestion and collapse the data to the finest level of resolution of the forcing variable. Additionally, the discretization from exact time-of-birth to day-of-birth is minor relative to similar papers which use year or quarter of birth as the forcing variable and employ the solution proposed by Lee and Card (2008).[4]

One may sidestep this issue by using an alternative identification assumption: local randomization. In the main text of the paper, we estimate the effect of compulsory voting using a "continuity" approach (Hahn et al., 2001), which assumes that the conditional expectation of potential outcomes is continuous at the threshold. An alternative framework for analyzing regression dis-

[4]In practice, using the aggregated or individual level data makes almost no difference to the size of our standard errors. The standard errors for our aggregated 69–70 year old sample (reported in the main text) are 0.004, 0.005, and 0.005 for the full sample, the more educated sample, and the less educated sample, respectively. For comparison, the standard errors estimated using the individual level data (same bandwidths) are 0.005, 0.006, and 0.006.

	Age 69–70			
	All	Educated	Less Educated	Difference
Estimate	0.031	0.068	0.017	0.051
Std. Error	0.009	0.015	0.011	0.019
95% CI	[0.013, 0.048]	[0.037, 0.098]	[-0.004, 0.038]	[0.014, 0.087]
Bandwidth	2	2	2	
# of Individuals	8,610	2,439	6,171	

Table A.1: Effect of Compulsory Voting on Turnout Under a Local Randomization Assumption. Table reports difference in average turnout between voters born within 2 days after and before the threshold. The "Difference" column shows the estimated difference between the "Educated" and "Less Educated" samples. Standard errors are heteroskedasticity consistent.

continuities is the "local randomization" approach (Cattaneo et al., 2015), which treats assignment to treatment "as if" randomly assigned near the threshold. Under this assumption, one can analyze the data within the chosen bandwidth as a randomized experiment. Given the arbitrary nature of the threshold and the lack of an obvious mechanism that would make voters born immediately before October 7, 1942 systematically different from those born immediately after, we analyze our data under this alternative assumption. We restrict our analysis to 69–70 year-old sample to avoid the differential registration problem in the 17–18 year old sample.

To estimate treatment effects, we use the difference-in-mean estimator on voters born either 2 days before or after the threshold. In other words, our treatment group is composed of voters born on the October 7 and 8, while the control group is composed of those born in the two days preceding that date. First, we check if the number of voters in treatment and control are consistent with random assignment with probability 1/2. A simple binomial test reports a p-value of 0.87, which is consistent with "as if" random assignment. Treatment effects can be found in Table A.1. These estimates are similar to our local linear estimates and the same basic pattern is evident: compulsory voting has much stronger effects on more educated voters than less educated voters.

H Incorporating Sampling Uncertainty for the 17–18 Year-Old Sample

As discussed in the text, we impute the number of 17- and 18-year-old citizens born on each day from the PNAD surveys administered by the Brazilian government. In our main text we report robust standard errors and confidence intervals that ignores the sampling uncertainty from these

	Dependent variable:	
	DV: Turnout	
	(1)	(2)
Compulsory Voting	0.117*	0.074*
	(0.054)	(0.014)
Primary Education		0.436*
		(0.018)
Compulsory Voting x Primary Education		0.069*
		(0.022)
Intercept	0.544*	0.287*
	(0.035)	(0.007)
Bandwidth (Days)	92	92
Number of Individuals	1,775,921	1,775,921
Observations	370	370
Note:	Bootstrapped standard errors in parentheses.*p<0.05	

Table A.2: Results for 17–18 year-old sample incorporating sampling uncertainty from PNAD survey. Table reports coefficients from a local linear regression on turnout rates by day of birth and education status. Coefficients on forcing variable have been omitted. Data weighted by the number of registered voters in each birth-date-education cell.

surveys. To incorporate the sampling uncertainty from the PNAD, we use a two stage procedure where we first resample respondents and primary sampling units within the strata used by the Brazilian census agency and construct turnout-by-day estimation dataset based on this bootstrap sample. We then sample with replacement from this estimation dataset and use this bootstrap sample to compute point estimates. The standard errors we report in Table A.2 are the standard deviation of 1000 bootstrapped point estimates. All conclusions remain unchanged. We used the same bandwidth that was selected for the full 17–18 year old sample in Table 1.

I Robustness to Alternative Bandwidth Choices

For the 17-18 year-old sample, we evaluate robustness of our point estimates to different bandwidth choices. In Figure A.7, we show how point estimates vary for bandwidths ranging from 15 days to 300 days. In all cases, we use the local linear estimator with triangle kernel weighting. Dashed lines are conventional 95% confidence intervals. While point estimates tend to diminish as bandwidth decreases (especially for voters with less than primary education), the gap between

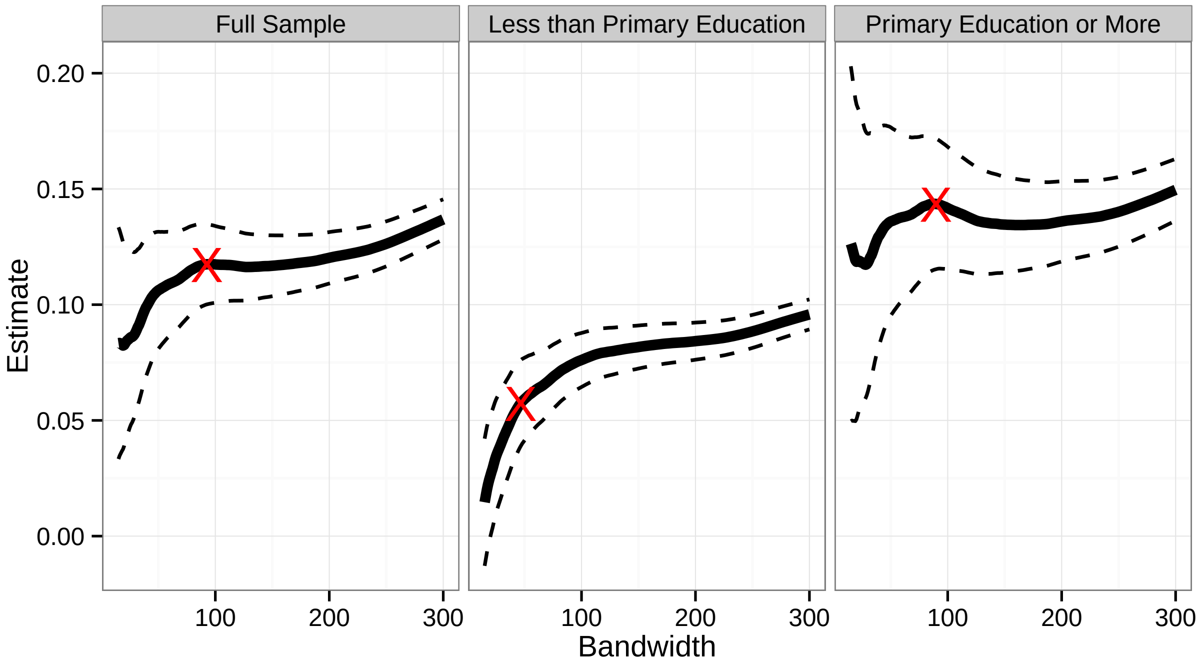

Figure A.7: Robustness to alternative bandwidth choices for the 17–18 year old sample. This plot shows how coefficients vary by using bandwidths ranging from 15 days to 300 days for estimates using the full sample, citizens with primary education or more, or citizens with less than primary education. Points marked with an "X" are bandwidths selected using the Calonico et al. (2014) procedure. Dashed lines are 95% confidence intervals.

estimates for more and less educated voters remains for all bandwidths.

J Placebo Thresholds

As an indirect test of the continuity assumption, we estimate treatment effects at placebo thresholds in both our sample. As shown in Figures A.8 and A.9, estimates at placebo cutoffs are generally close to 0.

K Heterogeneity by All Levels of Education.

In the main text, we aggregate our education variable into having completed primary education or not. Aggregating the data in this way is especially important for the 17–18 year-old sample because voters may go onto receive more education after registering and this would not be reflected in our data. For the 69–70 year-old sample, this is less of a concern because all voters were

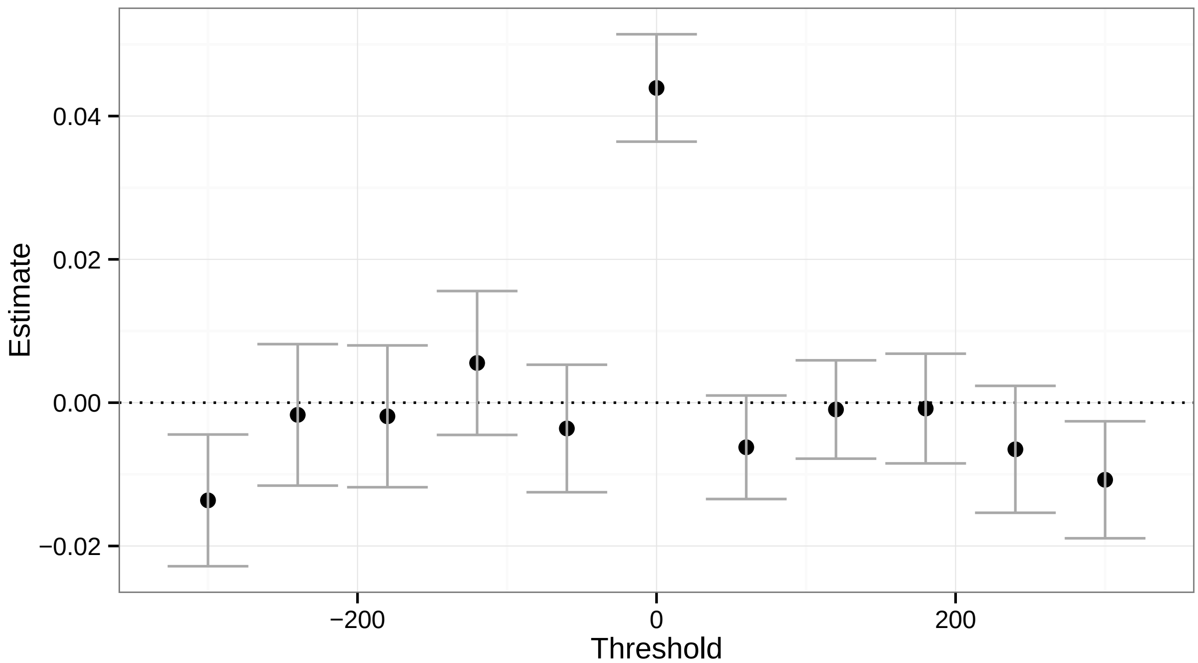

Figure A.8: Placebo Thresholds (69–70 Year Old Sample). Each dot represents a point estimate from a local linear regression using different thresholds of the forcing variable. With the exception of the threshold at 0, all other cutoffs are "placebo" thresholds. Bandwidths selected using the Calonico et al. (2014) procedure.

re-registered in 1986 and thus the data will reflect their education level when they were 43 or 44 (or later) years old. Generally, education data is only updated when a voter re-registers due to a change in residence and they wish to change their assigned polling location.

In Figure A.10, we report our estimates by the most disaggregated education data available. As evident by the plot, effect estimates increase monotonically from "read and write" (i.e. no formal education) through "secondary education incomplete". From "Secondary Education Completed" through "College Completed", however, there is no clear trend. In all cases, however, estimates for "Primary Completed" or above are larger than the estimates for "Primary Incomplete" and "Read and Write". The number of voters who fall into the more educated categories is small relative to those with less education, which make the estimates more variable (as reflected in the wider confidence intervals).

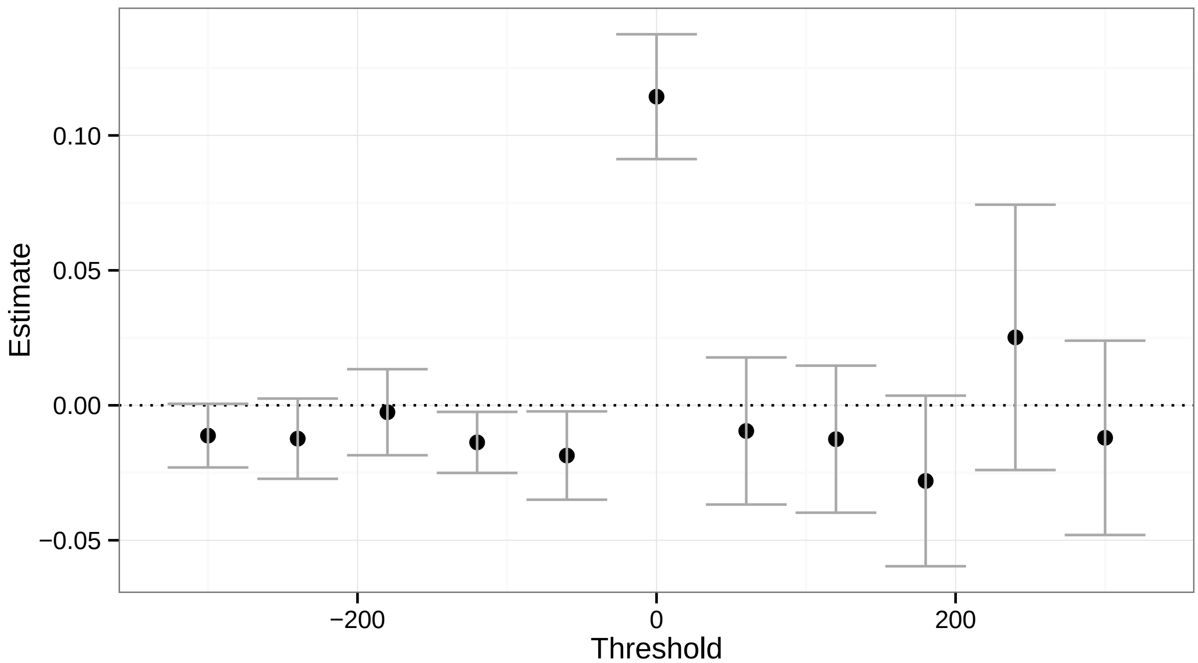

Figure A.9: Placebo Thresholds (17–18 Year Old Sample). Each dot represents a point estimate from a local linear regression using different thresholds of the forcing variable. With the exception of the threshold at 0, all other cutoffs are "placebo" thresholds. Bandwidths selected using the Calonico et al. (2014) procedure.

L Lack of Information as an Alternative Mechanism

An alternative hypothesis that might explain the observed heterogeneity in treatment effects is a correlation between information about the compulsory voting law and education. Some voters might not know that voting is voluntary starting at 70, for example, and continue to vote because they erroneously believe they will be sanctioned if they abstain. If this lack of information is correlated with education, then the larger effect of compulsory voting on the more educated might simply be a function of greater information. To test for this possibility, we asked survey respondents the age at which voting becomes compulsory (i.e. 18) and the age at which it becomes voluntary once again (i.e. 70). If the proportion of respondents answering correctly to these questions is correlated with education, then this alternative explanation becomes more plausible.

In Figure A.11, we plot the proportion of respondents who know the correct ages as a function of educational attainment. In addition to plotting the estimates from the full sample, we also present estimates from voters over 60 and voters under 30.

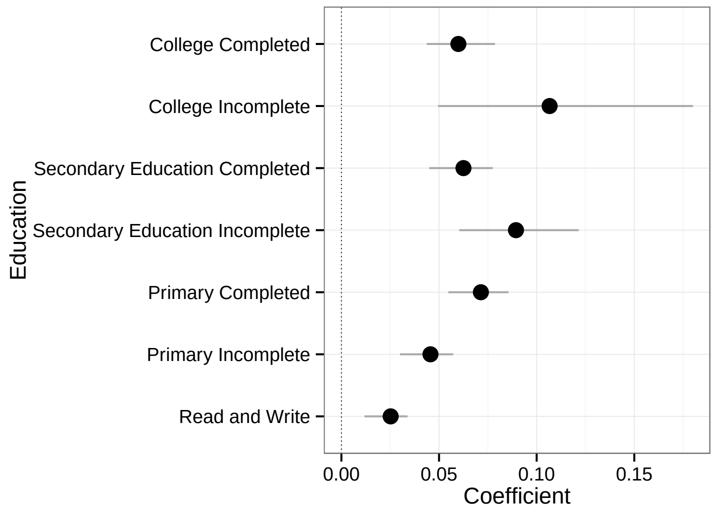

Figure A.10: Effects of Compulsory Voting by All Education Levels. Plot reports coefficients from a triangle kernel weighted local linear regression on turnout rates by day of birth. MSE-optimal bandwidths and robust confidence intervals are estimated using the Calonico et al. (2014) procedure.

In the case of the 18 year old threshold in the full sample, there is a positive correlation between knowing the correct answer and education. This correlation is driven by the difference between voters who completed less than primary education and voters with secondary or higher educational attainment. Knowledge about the threshold at age 70 in the full sample, however, exhibits a *negative*, albeit modest, correlation with education status. While the correlation between education and knowledge is consistent with the alternative explanation in the case of the threshold at 18, knowledge about the 70 year old threshold does not fit the alternative theory. Given the consistent pattern of effect heterogeneity by education across both thresholds, the information hypothesis does not account for our results in the full sample. In the subsamples, standard errors are larger, which makes our inferences less certain. Nevertheless, we do not find consistent differences across education levels and knowledge, which further supports our hypothesis.

References

Calonico, S., M. D. Cattaneo, and R. Titiunik (2014). Robust nonparametric confidence intervals for regression-discontinuity designs. *Econometrica 82*(6), 2295–2326.

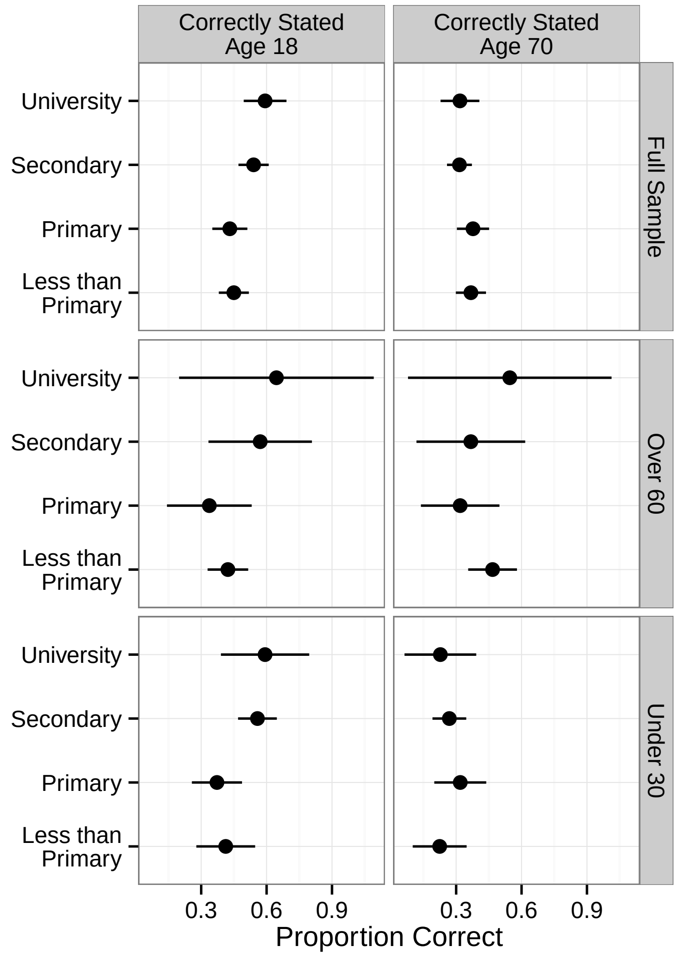

Figure A.11: Proportion Correctly Stating CV Thresholds by Education Status and Age Group. Estimates taken from a survey asking respondents to provide age at which voting is compulsory (left panel) or voluntary (right panel). Lines represent 95% confidence intervals.

Cattaneo, M. D., B. R. Frandsen, and R. Titiunik (2015). Randomization inference in the regression discontinuity design: An application to party advantages in the us senate. *Journal of Causal Inference 3*(1), 1–24.

Hahn, J., P. Todd, and W. Van der Klaauw (2001). Identification and estimation of treatment effects with a regression-discontinuity design. *Econometrica 69*(1), 201–209.

Lee, D. S. and D. Card (2008). Regression discontinuity inference with specification error. *Journal of Econometrics 142*(2), 655–674.

Wood, S. (2006). *Generalized additive models: an introduction with R*. CRC press.