Online Appendix for

Am I obliged to vote? A regression discontinuity analysis of compulsory voting with ill-informed voters

March 2, 2021

Contents

S1 Online survey sample characteristics **2**

S2 Including or not weekend days and holidays **4**

S3 Regression discontinuity estimates **8**

S4 Continuity-based approach to estimate the effect of CV **10**

S5 Specification checks for the RD analyses **11**

S1 Online survey sample characteristics

The survey was carried out by Netquest, the only online survey firm in Brazil that holds the ISO 26362 certification for online panels. Below we present in Table S1.1 sociodemographic characteristics of our online sample survey and compare them with the Brazilian National Household Sample Survey (*Pesquisa Nacional por Amostra de Domicílios*, PNAD), fielded at about the same time (3rd quarter of 2018). The Brazilian National Household Sample Survey is a quarterly study conducted by the *Instituto Brasileiro de Geografia e Estatística* (IBGE), the Brazilian census agency. As we can see, the online sample matches very closely the sociodemographic characteristics of the Brazilian population. We encounter small differences for education and social class, with our online sample being slightly more educated and better off than the general population. If anything, the population true level of knowledge about the compulsory voting laws in Brazil is lower than what we uncover from our online survey.

Table S1.1: Sociodemographic characteristics of our online sample as compared to those from the Brazilian National Household Sample Survey

	Online survey	Brazilian National Household Survey
Men	49.5	48.3
Women	50.5	51.7
Age	36.7	35.6
High school degree	37.6	23.9
College degree	15.7	12.1
North	7.8	8.6
Northeast	27.9	27.5
Southeast	42.4	42.2
South	14.5	14.3
Midwest	7.4	7.7
White	44.5	43.2
Brown (*pardo*)	41.2	46.9
Black (*preto*)	11.3	8.8
Asian (*amarelo*)	2.2	0.7
Indigenous	0.7	0.4
Social class†		
A	2.6	2.8
B1	7.6	4.6
B2	14.8	16.4
C1	29.5	21.6
C2	17.6	26.1
DE	27.9	28.5

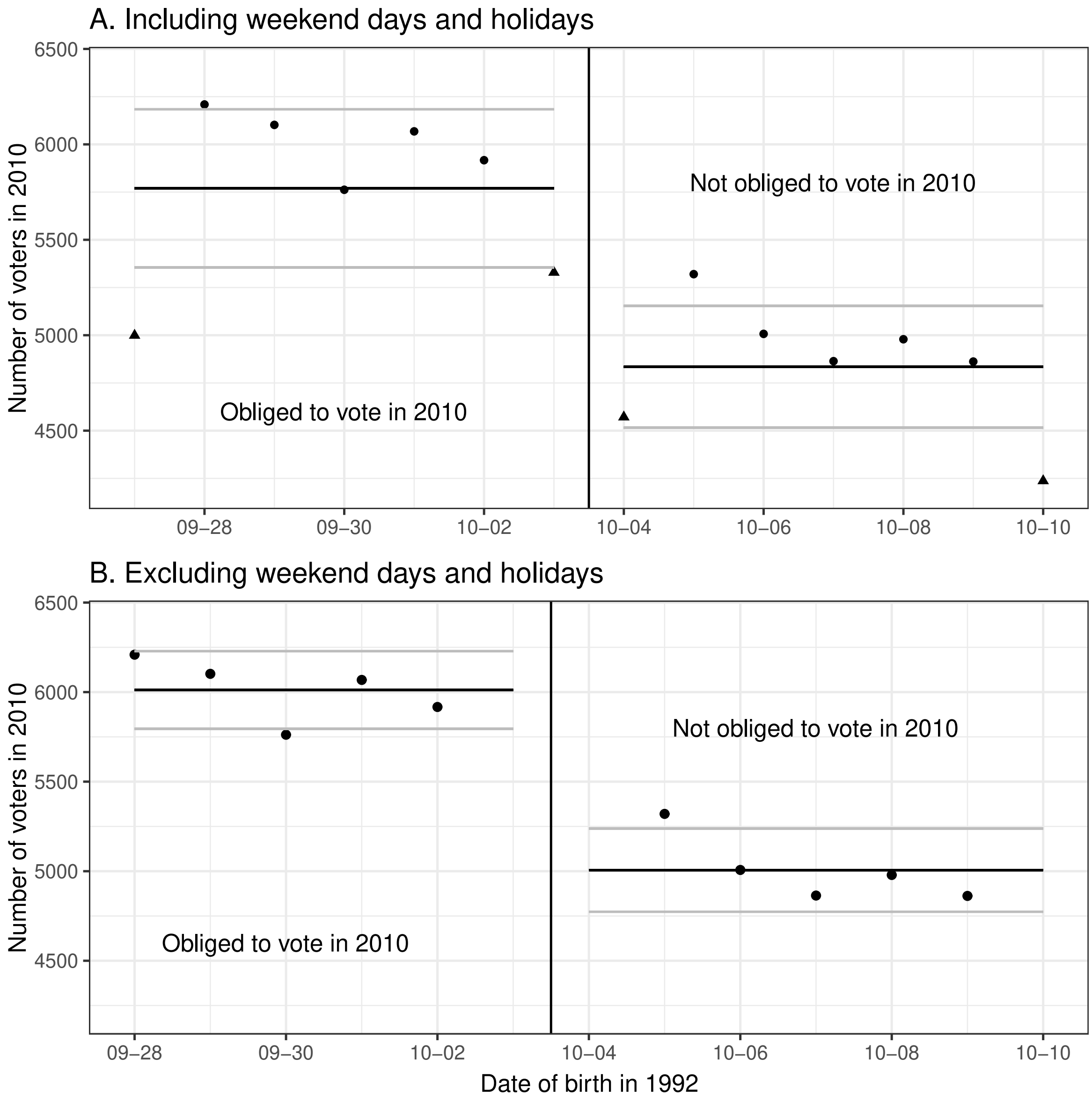
Note: All entries are percentages except for age.
†Data for social class at the national level come from the *Associação brasileira de empressas de pesquisa*. Their classification is based on data also collected from the IBGE (the Brazilian census agency) through the Family Budget Study (*Pesquisa de Orçamento Familiar*).

S2 Including or not weekend days and holidays

There has been a world-wide decrease in births on Sundays, Saturdays, and holidays (Goodman, Nelson and Maciosek 2005). Such decline is due, in part, to increasing elective interventions like induction and scheduled Caesareans (Mancuso et al. 2004). For that reason, we excluded weekend and holiday births (including Christmas and New Year's eves) from our calculations. Below, we illustrate how much lower births are during weekend days and holidays for the 2010 election. This section also presents the effects of compulsory voting on participation when considering all dates. The estimates are very similar to those reported in Figure 2.

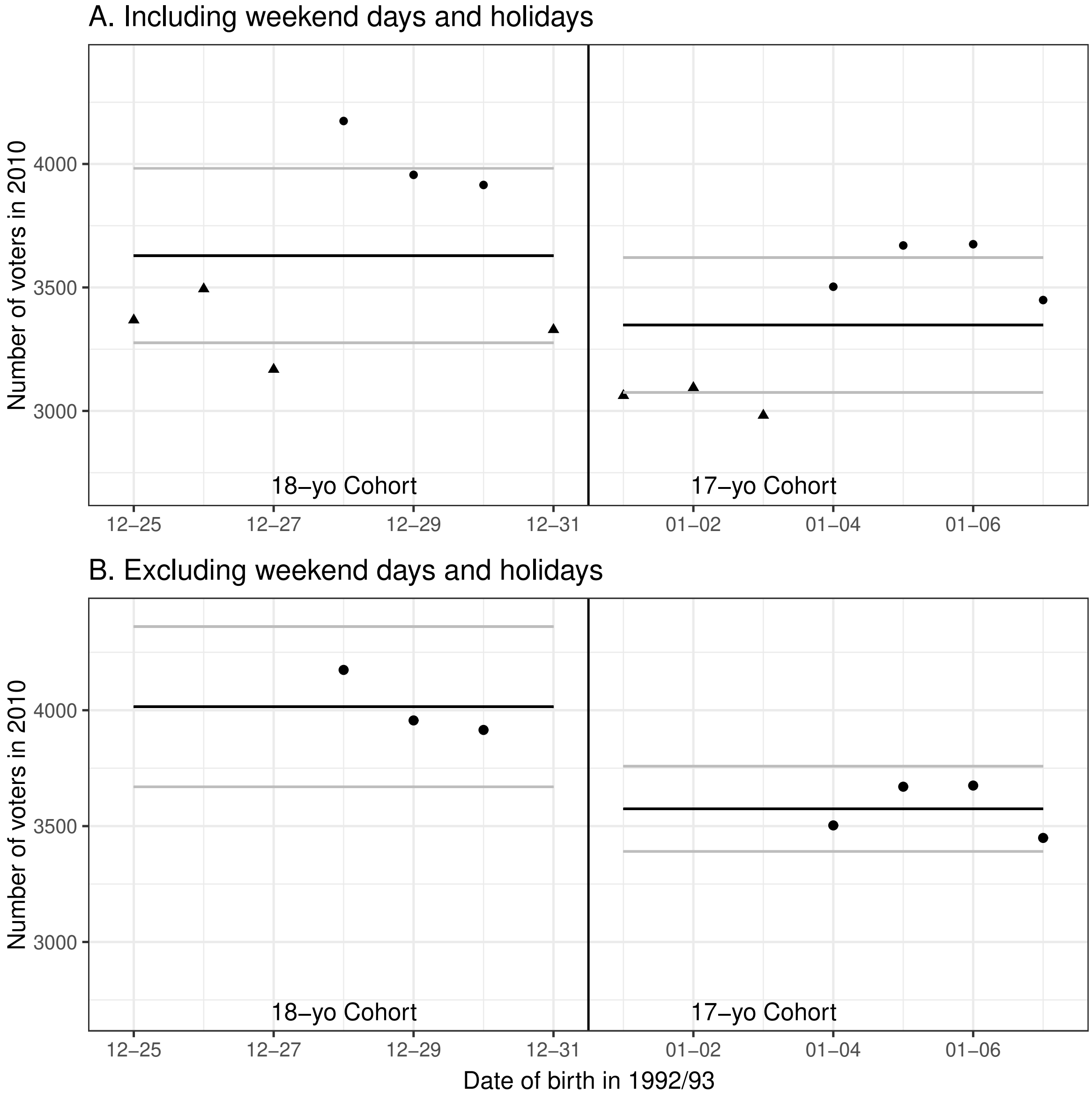
Figures S2.1 and S2.2 below plot the number of voters in 2010 binned by date of birth around the discontinuities of interest (Election Day and end-of-year). Both figures S2.1 and S2.2 have two panels, A and B. Panel A includes weekend and holiday births and panel B excludes them. Weekend and holiday births are indicated by triangles instead of circles in panel A. It is easy to see that the number of births during weekends and holidays is systematically lower among young voters, justifying why these days are excluded from our analyses. In Table S2.2 below, we present the effects of compulsory voting at the two critical moments, considering, this time, all dates. The estimated effects are very close to the ones presented in Figure 2 in the paper, but exhibit (not surprisingly) larger confidence intervals. The estimated effect for 2010 on Election Day was 20.1% [15.6; 23.7], as compared to 19.3% [9.0; 23.9] when considering weekend days and holidays. For the end of the year, the estimate reported in Figure 2 is 12.3% [5.8; 15.5], as compared to 8.4% [-2.7; 20.5] when considering weekend days and holidays. This last estimate fails only short of statistical significance despite its much larger confidence interval (more than twice as large). But, more importantly for present purposes, the substance of our findings remains unchanged whether we include or not weekend days and holidays in our analyses.

Figure S2.1: CV effect at Election Date in the 2010 election, by date of birth



Note: Dots and triangles indicate the number of voters who voted in the first round of the 2010 presidential elections by date of birth (7 days on each side of the cutoff). The dark line indicates the mean number of voters and the grey lines represent the 95% confidence intervals.

Figure S2.2: CV effect at the end of the year in the 2010 election, by date of birth



Note: Dots and triangles indicate the number of voters who voted in the first round of the 2010 presidential elections by date of birth (7 days on each side of the cutoff). The dark line indicates the mean number of voters and the grey lines represent the 95% confidence intervals.

Table S2.2: Compulsory voting effects on participation at Election Day and the end of the year including weekend and holiday births, 2010 Brazilian Presidential Elections

	Election Day	**End-of-year**
Control	4834.3	3347.9
Treatment	5769.4	3629.1
Effect	935.1	281.3
95% CI	[781.0, 1184.6]	[-89.2, 685.4]
Eff. number of obs. [l,r]	[7, 7]	[7, 7]
% increase	19.3	8.4
95% CI	[9.0, 23.9]	[-2.7, 20.5]

Note: Entries are average number of voters as calculated by the local randomization approach using a bandwidth of 7 days.

S3 Regression discontinuity estimates

Tables S3.3 and S3.4 present the details about the regression discontinuity estimates shown in Figure 2 in the paper.

Table S3.3: Compulsory voting effects on participation at Election Day and the end of the year, 2010, 2014 and 2018 Brazilian Presidential Elections

	Election Day	**End-of-year**
2010		
Control	5006.4	3574.3
Treatment	6011.6	4015.0
Effect	1005.2	440.8
95% CI	[781.0, 1184.6]	[206.4, 553.1]
Eff. number of obs. [l,r]	[5, 5]	[4, 3]
% increase	20.1	12.3
95% CI	[15.6, 23.7]	[5.8, 15.5]
2014		
Control	4113.6	2755.3
Treatment	5269.8	3147.0
Effect	1156.2	391.8
95% CI	[632.8, 1679.5]	[157.0, 572.2]
Eff. number of obs. [l,r]	[5, 5]	[4, 3]
% increase	28.1	14.2
95% CI	[15.4, 40.8]	[5.7, 20.8]
2018		
Control	3744.0	2470.0
Treatment	4439.0	2740.5
Effect	695.0	270.5
95% CI	[290.0, 1019.0]	[40.7, 500.3]
Eff. number of obs. [l,r]	[4, 5]	[4, 4]
% increase	18.6	11.0
95% CI	[7.7, 27.2]	[1.6, 20.3]

Note: Entries are average number of voters as calculated by the local randomization approach using a bandwidth of 7 days.

Table S3.4: Compulsory voting effects on participation at Election Day and the end of the year, 2008, 2012 and 2016 Brazilian Municipal Elections

	Election Day	**End-of-year**
2008		
Control	5823.0	4664.5
Treatment	6473.6	5042.7
Effect	650.6	378.2
95% CI	[172.2, 1129.1]	[182.9, 495.7]
Eff. number of obs. [l,r]	[4, 5]	[4, 3]
% increase	11.2	8.1
95% CI	[3.0, 19.4]	[3.9, 10.6]
2012		
Control	5609.5	4406.8
Treatment	6425.2	4605.2
Effect	815.7	198.4
95% CI	[292.4, 1194.5]	[22.0, 374.9]
Eff. number of obs. [l,r]	[4, 5]	[5, 5]
% increase	14.5	4.5
95% CI	[5.2, 21.3]	[0.5, 8.5]
2016		
Control	4750.0	3666.3
Treatment	5438.6	3969.0
Effect	688.6	302.8
95% CI	[437.2, 801.1]	[-133.0, 599.6]
Eff. number of obs. [l,r]	[5, 5]	[4, 3]
% increase	14.5	8.3
95% CI	[9.2, 16.9]	[-3.6, 16.4]

Note: Entries are average number of voters as calculated by the local randomization approach using a bandwidth of 7 days.

S4 Continuity-based approach to estimate the effect of CV

Table S4.5 below also presents the same analyses as those presented in Figure 2 in the paper (also presented in the upper panel of Table S3.3 above) for 2010, but considering this time date of birth as a continuous running variable. Estimates were calculated using the rdrobust R package by Calonico et al. (2018) using the MSE-optimal bandwith selection and a polynomial of order 1. The estimates also come with the recommended robust bias corrected confidence intervals. As can been seen from Table S4.5, the results are very similar to the ones presented in Table S3.3 where we (correctly) treat date of birth as a discrete running variable. Again, the estimates are not exactly the same—and nor should we expect them to be—but the substance of the findings remains unaltered. We are more confident in the results reported in Table S3.3 because the approach correctly treats the running variable—date of birth—as discrete.

Table S4.5: Compulsory voting effects on participation at Election Day and the end of the year treating date of birth as a continuous running variable, 2010 Brazilian Presidential Elections

	Election Day	**End-of-year**
Control	4929.0	3279.0
Treatment	5756.0	3634.4
Effect	827.1	355.4
95% CI	[360.5, 1293.6]	[-12.5, 723.3]
95% robust bias corrected CI	[207.3, 1294.0]	[-24.2, 839.1]
Eff. number of obs. [l,r]	[24, 25]	[25, 26]
% increase	16.8	10.8
95% CI	[4.2, 26.3]	[-0.7, 25.6]

Notes:
By treating the running variable as continuous, we include in our estimations all observations.
The bandwidth is 89 days which corresponds to the number of days from the day after Election Day (October 3) to the New Year, the two cutpoints of interest.

S5 Specification checks for the RD analyses

Below, we present a series of specification checks using the data from the 2010 Brazilian Presidential election.

S5.1: Covariate balance

One common practice in regression discontinuity designs is to check for covariate balance to conclude that the randomization process was "successful" (Lee and Lemieux 2014). Our participation data are fairly limited, but provide nevertheless information about the gender of voters who turned out on Election Day. A balance analysis presented in Table S5.6 shows a slight gender imbalance at the end of the year in 2010, considering the same bandwidth of 7 days. Specifically, we find a smaller proportion of women voters before the end of the year (treatment). The difference in proportions between women and men at the end of the year, although statistically significant, is substantively small at .034. We estimated the effect of compulsory voting at the two discontinuities for men and women separately to examine the extent to which our findings are driven by gender imbalances. The findings, presented in Table S5.7, indicate that compulsory voting exerts, as expected, a positive effect on participation among *both* men and women at *both* Election Day and the end of the year. The effect at the end of the year, however, is stronger among men, as compared to women. This last finding may suggest that women are slightly more informed than men about compulsory voting laws in Brazil, given that they are mostly affected when they should: at Election Day. Men, for their part, are substantively affected at both Election Day and at the end of the year, suggesting a misinterpretation of the law. Interestingly, this imbalance is in line with findings from our survey. Indeed, we find that young women (18-24) are more knowledgeable about compulsory voting than young men. To the question *"All voters who are 18 on Election Day are obliged to vote"*, 94.7% of young women respondents got it right while 91.6% of the young men did so ($p = .10$). But, more importantly, to the question *"All voters who turn 18 this year, including those who turn 18 after Election Day, are obliged to vote"*, 38.0% of young women respondents got it right while only 29.5 of the young men did so ($p = .02$).

Table S5.6: Proportion of women voters on Election Day and at the end of the year, 2010 Brazilian Presidential Elections

	Control	Treatment	Difference (p-value†)
2010			
Election Day	.505	.505	.000 (.918)
End-of-year	.522	.488	.034 (<.001)

†Chi-square two-sample test for equality of proportions with continuity correction.

Table S5.7: Compulsory voting effects on participation at Election Day and the end of the year be gender, 2010 Brazilian Presidential Elections

	Election Day		**End-of-year**	
2010	**Women**	**Men**	**Women**	**Men**
Control	2525.8	2480.6	1865.5	1708.8
Treatment	3035.8	2975.8	1958.7	2056.3
Effect	510.0	495.2	93.2	347.6
95% CI	[421.8, 598.2]	[349.1, 641.3]	[-44.2, 212.1]	[242.6, 387.0]
Bandwidth	7	7	7	7
Eff. number of obs. [l,r]	[5, 5]	[5, 5]	[4, 3]	[4, 3]
% increase	20.2	20.0	5.0	20.3
95% CI	[16.7, 23.7]	[14.1, 25.9]	[-2.4, 11.4]	[14.2, 22.6]

Notes:
Entries are average number of voters, as calculated by the local randomization approach using a bandwidth of 7 days.
The cutoff date of the running variable for Election Day is October 3 and that for the end of the year is December 31.
Weekend and holiday births (including Christmas and New Year's eves) are excluded from the analyses.

S5.2: Placebo tests

We also conducted a series of placebo tests for jumps at points other than our two discontinuities. In Table S5.8, we present the estimated effects of compulsory voting on participation by considering the same cutoffs, but for voters either one year younger or older. Specifically, we examined the effect of turning 17 or 19 on Election Day and at the end of the year to see if similar jumps are observed as the ones estimated for those reaching 18 on those same days in 2010. The analysis consists of four possible placebo tests, two for each of the two cutoffs. All placebo tests report, as expected, no statistically significant effect at both cutoffs of interest. Overall, we are very confident that the effects among voters turning 18 at Election Day and the end of the year are real and can be attributed to compulsory voting.

Table S5.8: Placebo tests at Election Day and the end of the year (± 1 year), 2010 Brazilian Presidential Elections

	Election Day		**End-of-year**	
2010	**17-yo**	**19-yo**	**17-yo**	**19-yo**
Control	2932.4	7498.8	2447.4	7363.5
Treatment	2938.0	7575.8	2499.0	7292.7
Effect	5.6	77.0	51.6	-70.8
95% CI	[-85.9, 120.0]	[-103.1, 343.7]	[-102.6, 160.3]	[-687.8, 397.3]
Eff. number of obs. [l,r]	[5, 5]	[5, 5]	[5, 4]	[4, 3]
% increase	0.2	1.0	2.1	-1.0
95% CI	[-2.9, 4.1]	[-1.4, 4.6]	[-4.2, 6.6]	[-9.3, 5.4]

Notes:
Entries are average number of voters as calculated by the local randomization approach using a bandwidth of 7 days.
Weekend and holiday births are excluded from the analyses.

S5.3: Different bandwidth sizes

Finally, the estimates presented in Figure 2 in the paper (and in Tables S3.3 and S3.4 above) rely on a 7-day bandwidth. This bandwidth was chosen for representing the smallest window one could imagine while balancing the number of weekdays on either side of the cutoff. In Figures S5.3 and S5.4 we show how our estimates are sensitive to alternative bandwidth choices. Specifically, we present estimates of the effect of compulsory voting (in number of voters) at both Election Day and the end of the year in 2010 while varying the bandwidth from 3 to 21 days. We present our results from a minimum bandwidth of 3 days because it represents the smallest number of observations needed to calculate a p-value for the effect of interest. Estimates are presented as triangles and the 7-day bandwidth estimates presented in Table S3.3 are shown as a circles. Estimates in red are not statistically significant at .05 (two-tailed) while those in blue are. Figures S5.3 and S5.4 show, as one would expect, that there is some variation in the estimates of the effect of compulsory voting when using different bandwidths. But, the results are consistent in that the estimates reported in Table S3.3 are similar to those presented in Figures S5.3 and S5.4 with alternative bandwidths. More importantly, the substance of the findings remains unchanged: the inferences drawn in the paper are maintained when considering varying bandwidths. Only for very small bandwidths (generally less than 5 days) are the estimates not statistically significant, given the small number of observations used in the calculation of those effects.

Figure S5.3: Robustness to alternative bandwidth choices: Election Day, 2010 Brazilian Presidential Elections

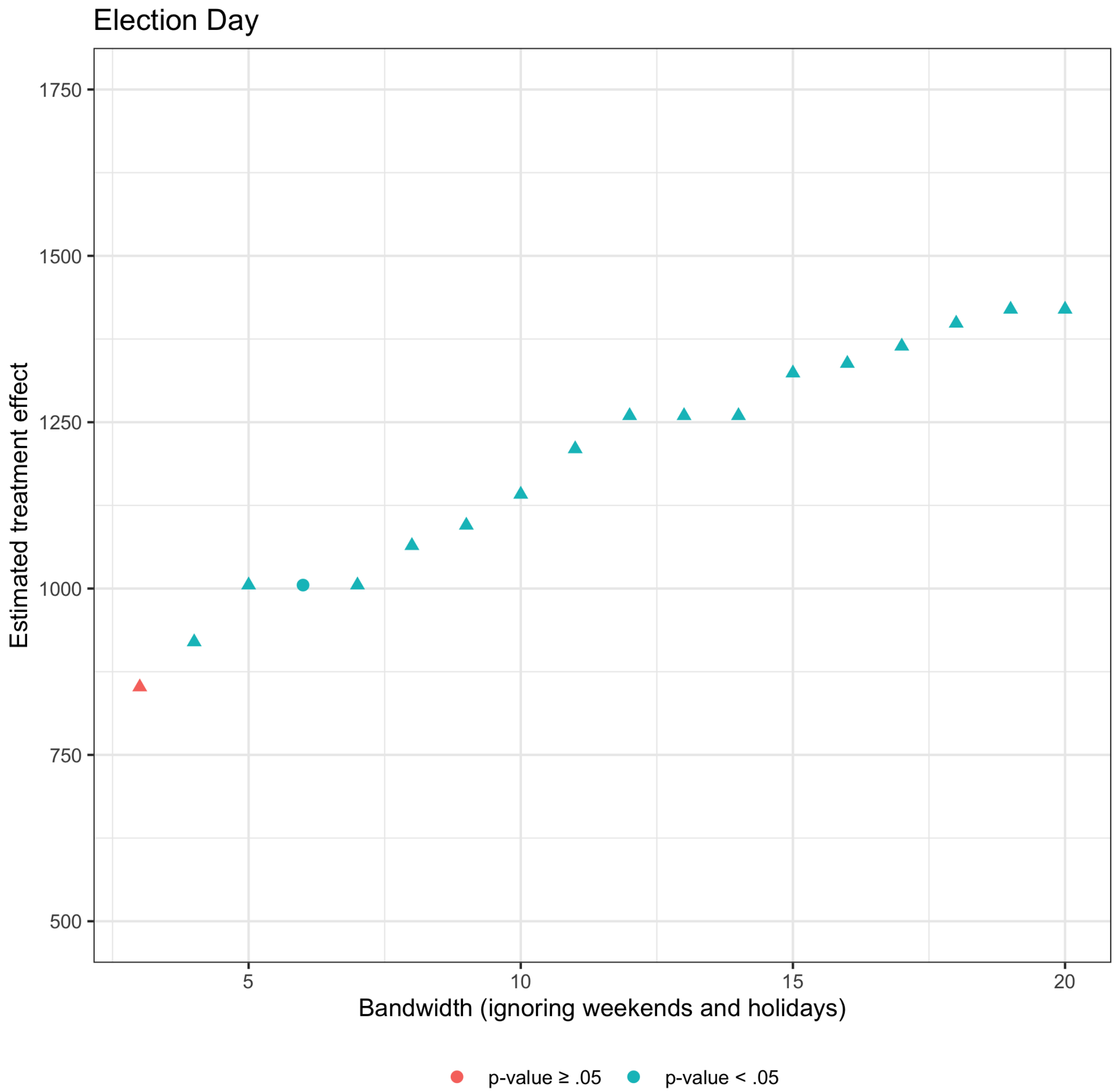
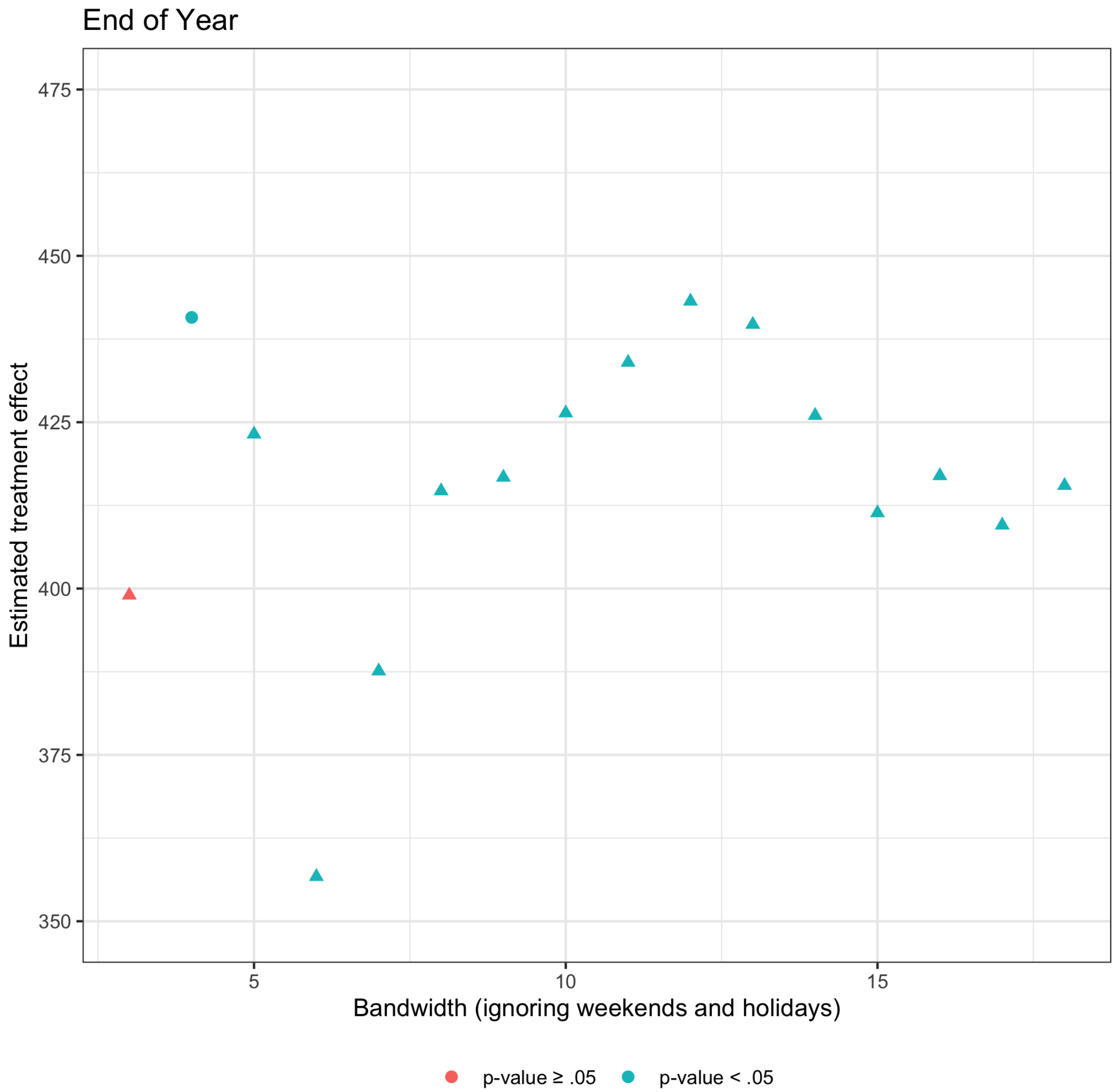


Figure S5.4: Robustness to alternative bandwidth choices: End-of-Year, 2010 Brazilian Presidential Elections



S5.4: Other possible cutoffs

Another question of interest concerns how confident we can be that the mistaken cutoff is actually at the end of the year. Or, in other words, are there reasons to believe that there could be other discontinuities on other days between Election Day and the end of the year and shortly after? The end of the year appears as a theoretically relevant cutoff—and the responses to the second question from our survey reported in Table 1 in the paper seem to support this point—which is why we considered it. But, this is an empirical question. To address it, we looked for other discontinuities between Election Day and the end of the year and a few weeks after. Specifically, we estimated discontinuities from October 18 to December 16 and from January 14 to February 1, adopting the local randomization approach using a bandwidth of 7 days. We chose these two periods as they exclude the dates used in Figure 2 in the paper. There is no need to look for dates prior to Election Day as nearly all voters know that voters who are 18 on Election Day are obliged to vote (first question from of our survey). The uncertainty concerns voters who turn 18 between Election Day and the end of the year. The results of this additional analysis are presented below in Figures S5.5 and S5.6. Figure S5.5 presents the estimated p-values for the various cutoffs, evaluating the null hypothesis that the estimated effect at these discontinuities is 0. The figure presents two different types of p-values. The first one considers each estimate separately (red dots) while the second one takes into account that the probability of type 1 error increases with multiple testing (blue dots), using the Holm correction. The results are quite telling: none of the adjusted p-values indicates rejection of the null hypothesis that any of the effect estimated at the numerous discontinuities between Election Day and the end of the year and a little after are statistically different from 0. Only a handful of p-values from the "raw" analysis, that is, the ones estimated in "isolation" are below the conventional .05 level, but these could happen by chance. Specifically, the probability of falsely rejecting at least one test among a set of m tests is given by $H = H_1, ..., H + m$ is $1 - (1 - \alpha)^m$ (Hothorn, Bretz and Westfall 2008). In our cases, that probability is .94.

But there is more. Figure S5.6 reports the estimated increase in percentage (like in Figure 2 in the paper) for those estimates with p-values below .05 (two-tailed, from the "raw" analysis) and compares them to the estimates for Election Day and the end of the year (20.1% and 12.3% increases, respectively). The average estimated increase in percentage is 5.7 with the highest estimate at 8.1% for the end of October. If anything, the higher estimates found at the end of October may indicate that some voters interpreted the election month and not the election year as the cutoff for the obligation to vote. The effect at the end of October, however, is much smaller than the one uncovered at the end of the year. More importantly, this result does not affect the substance of our findings as it also suggests a misinterpretation of the compulsory voting laws and, similarly, contributes to an underestimation of the effect of compulsory voting on turnout (when considering only the cutoff at Election Day).

Figure S5.5: P-values for estimated discontinuities (2010 elections), by date of birth

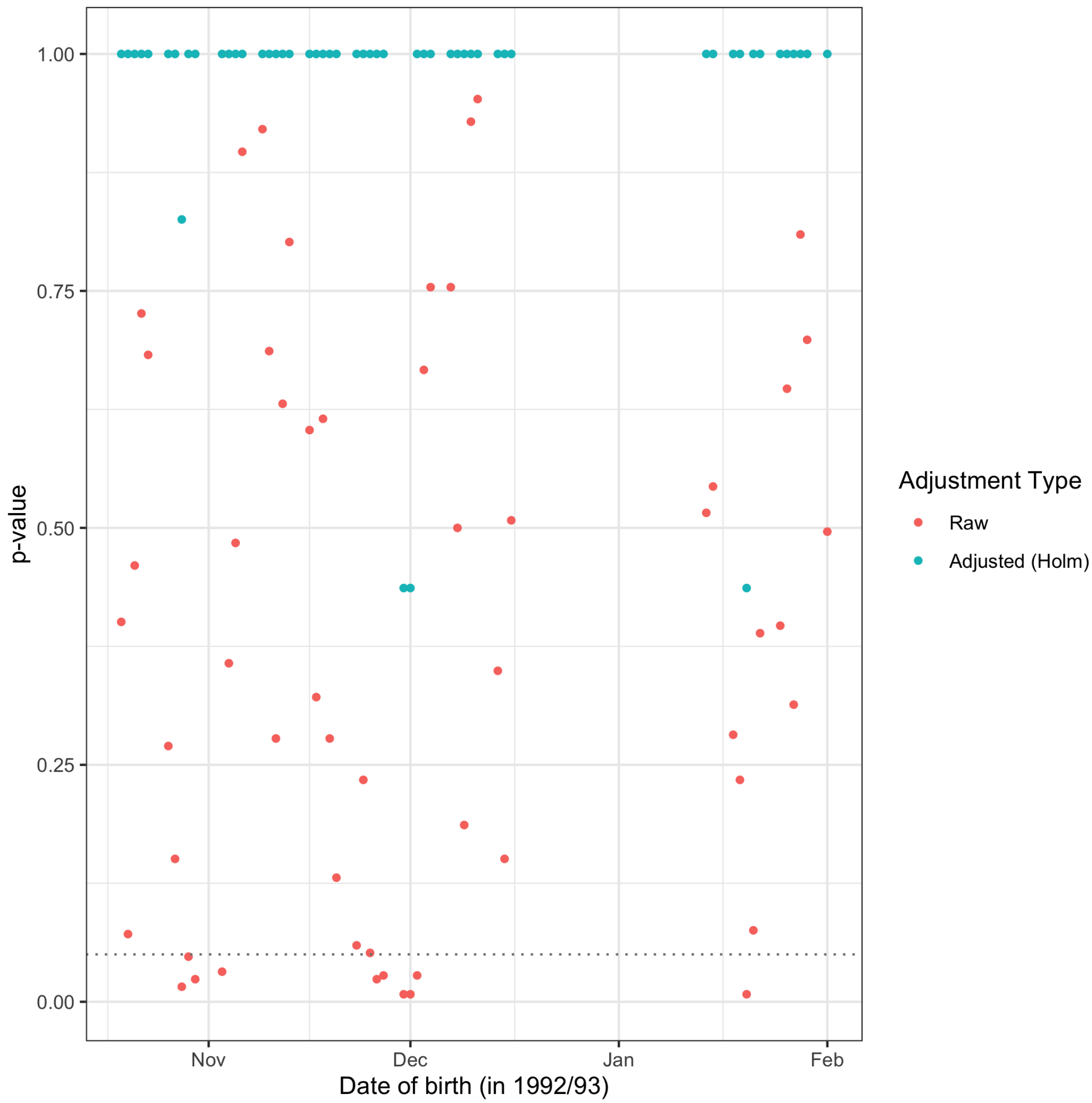
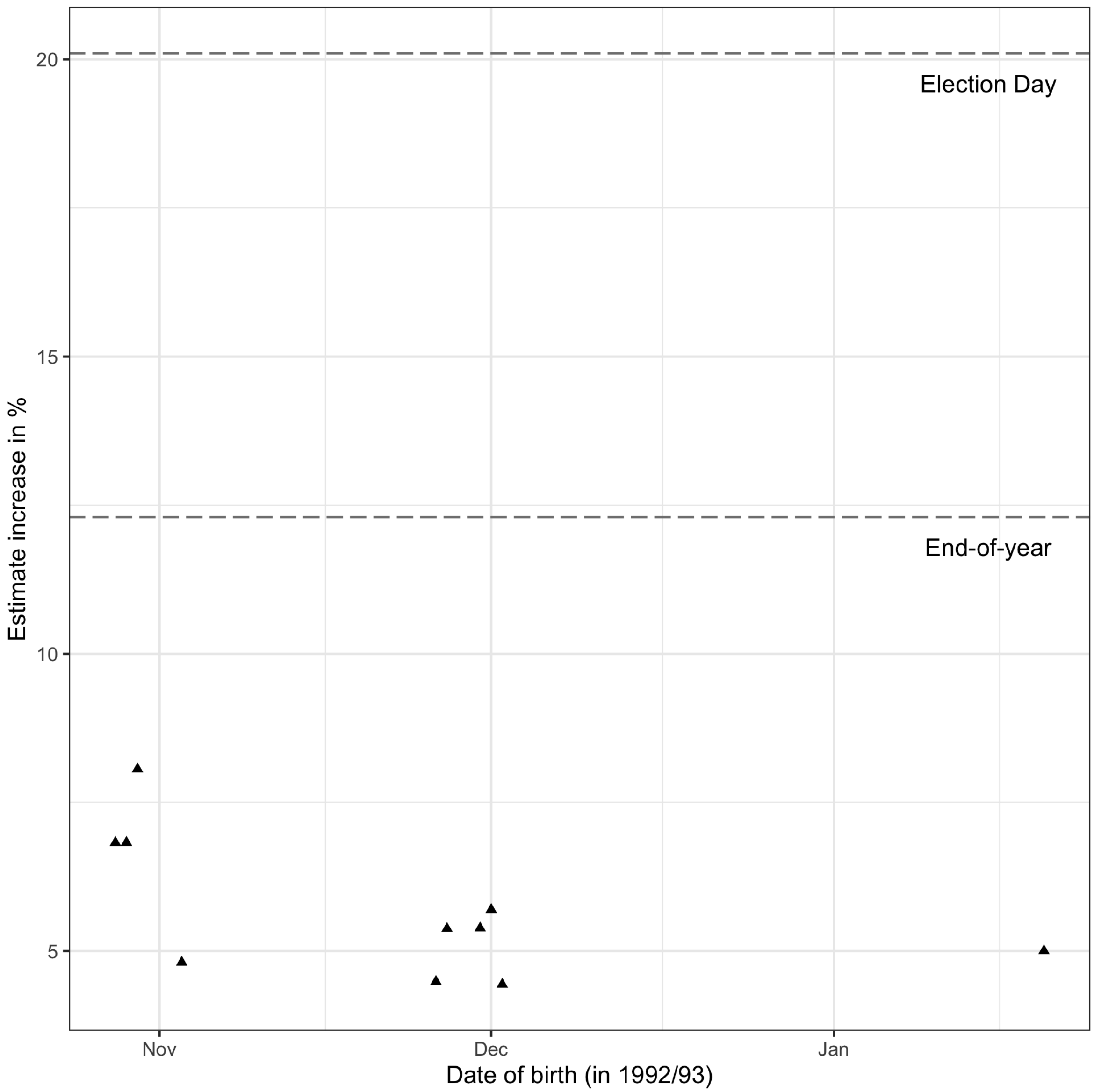


Figure S5.6: P-values for estimated discontinuities (2010 elections), by date of birth



References

Calonico, Sebastian, Matias D. Cattaneo, Max H. Farrell and Rocio Titiunik. 2018. *rdrobust: Robust Data-Driven Statistical Inference in Regression-Discontinuity Designs*. R package version 0.99.1.
URL: *https://CRAN.R-project.org/package=rdrobust*

Goodman, Michael J, Winnie W Nelson and Michael V Maciosek. 2005. "Births by day of week: A historical perspective." *Journal of midwifery & women's health* 50(1):39–43.
URL: *https://doi.org/10.1016/j.jmwh.2004.09.005*

Hothorn, Torsten, Frank Bretz and Peter Westfall. 2008. "Simultaneous inference in general parametric models." *Biometrical Journal: Journal of Mathematical Methods in Biosciences* 50(3):346–363.
URL: *https://doi.org/10.1002/bimj.200810425*

Lee, David S and Thomas Lemieux. 2014. "Regression discontinuity designs in social sciences." *Regression Analysis and Causal Inference, H. Best and C. Wolf (eds.), Sage* .

Mancuso, Peggy J, James M Alexander, Donald D McIntire, Emma Davis, Grace Burke and Kenneth J Leveno. 2004. "Timing of birth after spontaneous onset of labor." *Obstetrics & Gynecology* 103(4):653–656.