University of Oxford: MPhil in Politics

Causal Inference: Take-Home Exam

1090063

Contents

1 Problem 1: Natural Disasters and Voting Behaviour

Natural disasters are often used as exogenous shocks to examine the extent to which incumbents are rewarded for providing services to their constituents (Wolfinger and Rosenstone 1980; Mettler and Stonecash 2008; Bechtel and Hainmueller 2011). In this exercise, we aim to evaluate two possible explanations for this phenomenon. On the one hand, voters may reward incumbents after a natural disaster because they are grateful for any help they may have received. On the other hand, voters may perceive politicians as more competent after a natural disaster, especially if they demonstrate skills that helped mitigate the consequences of the disaster.

1.1 Data Collection

Explain what kind of data you would collect. Clearly define your units of analysis as well as your treatment and control group.

To investigate the causal effect of the *Prestige* oil spill on the incumbent party's vote share, whilst also disentangling the mechanisms of of the government's response to the disaster and the subsequent transfer of aid, a panel data set measured at the municipality unit of analysis is required.

To help disentangle the mechanisms, there will be two treatments: (i) the oil spill treatment and (ii) the payment treatment. The oil spill treatment is a binary variable indicating whether a municipality was affected by the oil spill disaster, while the payment treatment is a binary variable indicating whether a municipality received government aid in response to the disaster. Therefore, the control groups are those municipalities that were not affected by the oil spill and did not receive government aid.

The primary dependent variable is the vote share of the incumbent party in each municipality, measured at three time points: before the oil spill treatment, after the oil spill treatment, and after the payment treatment, with the pre-election time periods being measured as vote intention.

Below, `Table 1` is a mock-up of the panel data set that would be collected:

Table 1: Example Panel Dataset of Data Collected

Municipality ID	**Time Period**	**Controls**	**Oil Spill Treated**	**Payment Treated**	**Vote Share**
1	0	$\mathbf{X}_{1t}$	0	0	0.72
2	2	$\mathbf{X}_{2t}$	0	0	0.35
3	2	$\mathbf{X}_{3t}$	1	1	0.15
.	.	.	.	.	.
.	.	.	.	.	.
.	.	.	.	.	.
i	$t = [0\|1\|2]$	$\mathbf{X}_{it}$	$spill = [0\|1]$	$pay = [0\|1]$	$\mathbf{Vote}_{it}$

- `Municipality ID`: A unique identifier for each municipality (unit of analysis), indexed by i.
- `Time Period`: Indicates the time point t at which the observation is recorded. The dataset is structured as a panel, with multiple observations for each municipality over time. Before the oil spill treatment $t = 0$, after the oil spill treatment $t = 1$, and after the payment treatment $t = 2$. The time periods are indexed by t.
- `Controls`: Denotes a vector of time-variant observed covariates for each municipality at time t, represented as $\mathbf{X}_{it}$. This vector of covariates would include socio-economic, geographic, or political variables used to adjust for confounding.
- `Oil Spill Treated`: A binary indicator for whether a municipality was affected by the oil spill natural disaster. Equal to 1 if treated, 0 otherwise. This indicates whether a municipality is in the initial treatment or control group.
- `Payment Treated`: A binary variable indicating whether a municipality received a government payment as aid in response to the disaster. Equal to 1 if aid was received, 0 otherwise. This indicates whether a municipality is in the second treatment or control group.
- `Vote Share`: Vote share is the proportion of voting intentions measured for the incumbent party in each municipality at time t. This variable is used to measure electoral outcomes and public sentiment towards the incumbent after the disaster. It is denoted as $\mathbf{Vote}_{it}$.

1.2 Identification Strategy

Explain and justify your identification strategy for estimating the causal effect of interest. If you are estimating a model (e.g. OLS, DiD, etc.), provide the equation and explain all its terms. What are the identification assumptions of your research design and what do they mean in the context you are studying here? Can you test them? If so, how?

To estimate the causal effect of the *Prestige* oil spill disaster on the vote share of the incumbent party, a generalised Difference-in-Differences (DiD) with two-way fixed effects will be used. In particular, along with the panel dataset, this DiD model is used to allow for the estimation of the causal effect of the oil spill treatment and the payment treatment on the vote share of the incumbent party simultaneously in one regression specification, while controlling for time-variant covariates.

The generalised DiD approach works by calculating how much the treated group's outcomes changed compared to the control group's outcomes before and after the treatment. The two-way fixed effects approach accounts for unobserved time-invariant characteristics of the municipalities and common shocks that affect all municipalities at the same time.

The equation for the generalised DiD model is as follows:

$$\mathbf{Vote}_{it} = \alpha_i + \lambda_t + \beta_1(\text{spill}_i \times \text{post\_spill}_t) + \beta_2(\text{pay}_i \times \text{post\_pay}_t) + \gamma\mathbf{X}_{it} + \varepsilon_{it} \tag{1}$$

Where:

Table 2: Explanation of Terms in theGeneralised DiD Regression Model

Term	Description
$\mathbf{Vote}_{it}$	Share of vote intentions for the incumbent party in municipality i at time t.
α_i	Municipality-level fixed effect, capturing all time-invariant characteristics of municipality i.
λ_t	Time fixed effect, capturing all time-specific shocks that affect all municipalities at time t.
β_1	Coefficient for the interaction between the oil spill treatment and the post-treatment period.
spill_i	Binary indicator for whether municipality i was affected by the oil spill.
post_spill_t	Binary indicator for post-treatment period for the oil spill.
β_2	Coefficient for the interaction between the payment treatment and the post-treatment period.
pay_i	Binary indicator for whether municipality i received government aid.
post_pay_t	Binary indicator for post-treatment period for the payment.
γ	Vector of coefficients for the covariates.
$\mathbf{X}_{it}$	Vector of time-variant covariates for municipality i at time t.
ε_{it}	Error term of unobserved factors affecting vote share in municipality i at t.

For this DiD model to be valid, the core assumption of parallel trends must hold. This assumption states that if there was no treatment, the average outcomes for the treated and control groups would have followed the same trend over time. In this context, it means that the vote share of the incumbent party in municipalities affected by the oil spill would have followed the same trend as those not affected by the oil spill if the disaster had not occurred.

This assumption cannot be directly tested due to the fact that we cannot observe the counterfactual of whether the oil spill did not happen. However, we can conduct a pre-treatment analysis to check whether the trends in vote share for the treated and control groups were similar before the oil spill occurred. This can be done by plotting the average vote share over time for both groups and checking to see if they follow the same trend. This is potentially a problematic assumption due to spillover effects from the oil spill treatment. For example, if the oil spill affected the views and thus vote share of the incumbent party in neighbouring municipalities that were not directly affected by the spill, this could violate the parallel trends assumption.

We also are assuming that we are able to control for all time-variant confounders that may affect the vote share of the incumbent party. This is done by including a vector of time-variant covariates $\mathbf{X}_{it}$ in the regression model. If these are unobserved confounders that are not included and are correlated with both the treatment and the outcome, this could bias our estimates. For example, there can be other political events other than the oil spill that could affect the vote share of the incumbent party, especially in the run up to an election. If these events are not controlled for, they could confound the relationship between the oil spill and the vote share. However, psephological research suggests that voter intentions are generally stable over time and it is hard to persuade a change in behaviour in the run up to an election. Therefore, it is likely that the time-variant covariates will be sufficient to control for confounding.

1.3 Disentangling the Mechanisms

Can your research design disentangle the effects of the two proposed mechanisms, namely (i) the initial central government response to the emergency and (ii) the subsequent transfer of aid? If yes, explain your reasoning in detail. If not, explain how you could modify your originally proposed research design or data collection to separate the two mechanisms.

As previously mentioned, the design of the research design is to use a generalised DiD approach with two-way fixed effects. Using the panel data spanning multiple time periods, this allows for a generalised model to simultaneously estimate the causal effect of the oil spill treatment and the payment treatment on the vote share of the incumbent party. This is done by including both treatments in the regression model as separate interaction terms with their respective post-treatment indicators.

To estimate the initial central government response, we want to see how the share of vote intentions for the incumbent party changes after the oil spill treatment between those municipalities that were affected by the oil spill and those that were not. This is calculated by the interaction term of $(spill_i \times post\_spill_t)$. By running an OLS regression, the coefficient β_1 estimates the average difference in the change in voting intentions from the pre-spill period ($t = 0$) to the post-spill periods ($t = 1$, $t = 2$) for the spill-affected municipalities, relative to the change experienced by the control municipalities over the same periods.

Then, to estimate the subsequent transfer of aid, we want to see how the share of vote intentions for the incumbent party changes after the payment treatment between those municipalities that received government aid and those that did not. This is calculated by the interaction term of $(pay_i \times post\_pay_t)$. Here, β_2 estimates the average additional difference in the change in voting intentions from the pre-compensation period ($t = 1$) to the post-compensation period ($t = 2$) for the compensated towns, relative to the change experienced by the uncompensated spill-affected towns over that same specific period ($t = 1$ to $t = 2$). The control group here are the municipalities that were affected by the oil spill but did not receive government aid, whilst also controlling for time-based trends through the use of time fixed effects.

By using two-way fixed effects, we can control for the unobserved time-invariant characteristics of the municipalities and common shocks that affect all municipalities at the same time. This allows us to isolate the effects of the two treatments on the vote share of the incumbent party through the specific interaction terms and their coefficients to be comparing only the relevant control and treatment groups for each mechanism.

1.4 Limitations and Threats

Elaborate on your research design's limitations and threats to causal identification.

With this research design comes a few possible limitations and threats to having non-biased and consistent OLS estimates, as well as accurately estimating the causal nature of the effects. As previously mentioned, the parallel trends assumption is a key assumption of the DiD model. If this assumption does not hold, the estimates of the treatment effects may be biased. This could happen if there are other confounding factors that affect the vote share of the incumbent party in the treated and control groups differently over time. For example, if there were other political events or changes in public opinion that affected the vote share of the incumbent party in one group but not the other, this could violate the parallel trends assumption.

Another threat is the extent and scale of the treatments of the oil spill and the compensation received. A core assumption thus far is that the effects of the oil spill and the compensation are homogenous across all municipalities. However, for example, the severity of the oil spill may have varied across municipalities, leading to different levels of impact on businesses and livelihoods, which could affect voting intentions differently across municipalities. Variation in compensation levels should be less of an issue as all municipalities within Galicia received compensation, and as the unit of analysis is the municipality, the average compensation received should be similar across municipalities.

One of the biggest limitations to the research design comes from the quality and accuracy of the panel data set created to analyses these effects. The nature of the panel data requires the same municipalities to be observed over time, which can be difficult to achieve in practice. If there are missing data points, or differences in how the data is collected over time, this can lead to bias in the estimates. In particular, this will be problematic when estimating the share of the voting intentions for the incumbent party. This will require accurate sampling from an identically and independently distributed sample of the municipality each time period is measured. If the sample is not representative of the population, this can lead to bias in the estimates, with only the final election results showing the true vote shares. This means that the estimates of **Vote**_{it} may be biased.

2 Problem 2: Voter Turnout and the Number of Candidates

Following the 2016 US presidential election, Jill Stein, the defeated Green Party candidate, said in an interview that voter turnout would have been lower if she had not run. When asked about the possibility that her candidacy contributed to Hillary Clinton's defeat, Stein claimed that her supporters would have abstained if they had not been able to vote for Stein. This intuition is consistent with some political science research: in countries where political competition revolves around a few candidates (such as the United States, the United Kingdom or Canada), people are more likely to abstain from voting if they have preferences that do not align with any of the major candidates, e.g. because they feel alienated from the system (for a review of the literature, see Blais, 2006). As a result, some scholars argue that elections with (i) more than two competitive candidates and (ii) flexible state regulations that facilitate candidacy can increase voter turnout and reduce inequalities in political representation (Gallego, 2014). Several existing comparative and cross-sectional studies have examined this issue. However, their findings are mixed: some find that the number of candidates reduces turnout (e.g. Jackman, 1987), others that it increases turnout (e.g. Taagepera et al., 2013), while others find null effects (e.g. Fornos et al., 2004).

2.1 Data Summary

Familiarise yourself with the data by answering the following questions: i) What are the units of observation (rows) in the dataset? ii) What types of elections does the dataset cover? iii) How many candidates stood in the first and second rounds in each election? iv) What is the average turnout and null & blank votes in the second round of elections when two and three candidates stand in the second round? v) How many candidates who get the third highest share of votes in the first round qualify for the second round by exceeding the qualifying threshold? How many of them actually run in the second round?

The following provides a brief summary of the dataset so that we can clearly understand the data and elections. Firstly, each row within the dataset represents a single canton-level (district) election in France at a given time period, thus the unit of observation is the canton-level election. The dataset includes election types of **`cantonal`** and **`national`** elections from **`1978`** to **`2012`**. The elections are held in two rounds, with the first round being a preliminary election to determine which candidates will advance to the second round. Across the **`13974`** district elections in the dataset, the number of candidates in each election in each of the first and second rounds are summarised in **`Table 3`** below:

Table 3: Summary Statistics for Number of Candidates in Each Round

Statistic	First Round	Second Round
Min	3.0	1.0
Median	6.0	2.0
Mean	6.8	2.1
Max	26.0	4.0
Sum	95721.0	29018.0

When looking at the second round of elections, in particular when two or three candidates stand, the average turnout and null & blank votes are as follows in **`Table 4`**. The key takeaway here is that the increased number of candidates gives an initial indication that there may be a positive relationship between the number of candidates and the average turnout which will be later tested.

Table 4: Average Turnout and Null & Blank Votes in the Second Round (in %)

Number of Candidates	Average Turnout	Average Null & Blank Votes
2	61.4%	3.3%
3	64.9%	2.5%
2 or 3 Candidates	63.1%	2.9%

To get a final sense of the dataset, we can look at the number of candidates who get the third highest share of votes in the first round and qualify for the second round by exceeding the qualifying threshold. We can also look at how many of these qualified candidates go onto run in the second round. The results are summarised in `Table 5` below, where we see the key finding that only `47.5%` of third-place candidates who qualified went on to run in the second round. This means that the threshold for qualifying for the second round is clearly non-deterministic of whether a candidate will run in the second round, showing a large amuont of one-sided non-compliance at the threshold.

Table 5: Number of Third-Place Candidates Qualifying and Running for Round Two

Election Type	No. Qualified	No. Running	% of Qualified Running	Qualifying Threshold
National	698	272	39.0	12.5%
Cantonal (2011)	48	24	50.0	12.5%
Cantonal	3350	1651	49.3	10%
Total	4096	1947	47.5	NA

Note: The qualifying threshold for the national election and cantonal election in 2011 is 12.5%, while for other cantonal elections it is 10%. No elections were won outright in the first round to cause a candidate to be automatically elected.

2.2 Fuzzy Regression Discontinuity Design

Think about a causal inference method covered in the course that allows you to investigate, using the data provided, whether increasing the number of candidates standing in an election reduces the abstention rate. Explain why this design is a good choice in this setting. Discuss its identification assumption(s) in the context you are studying here.

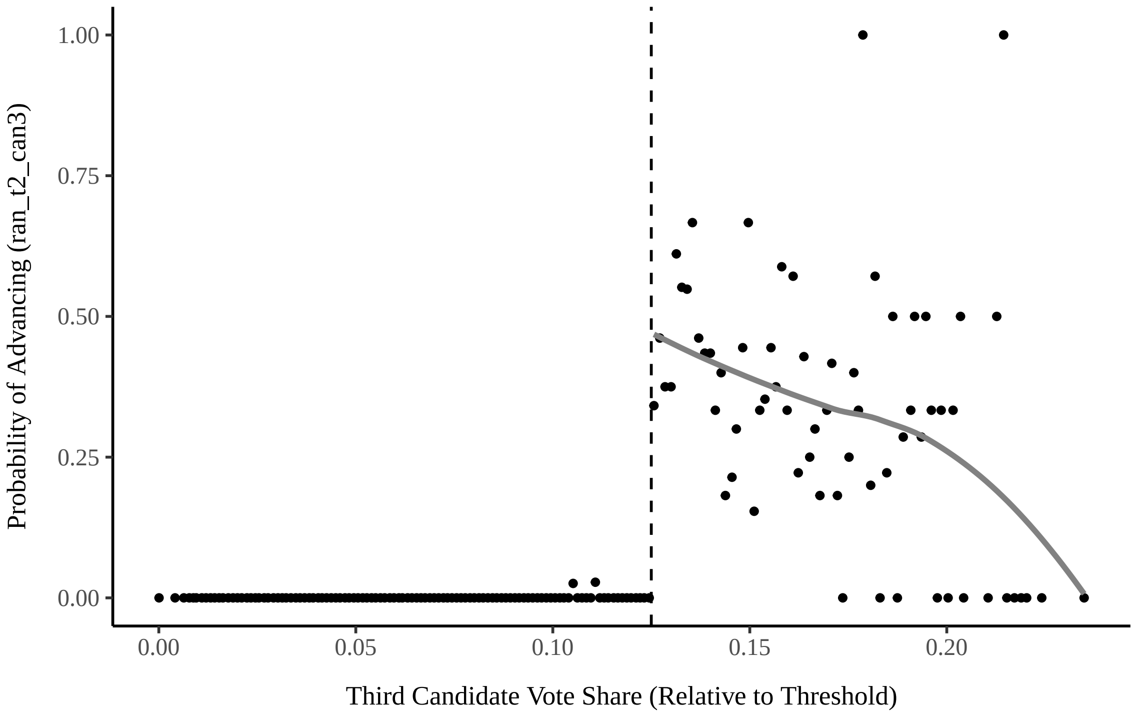
We are investigating whether increasing the number of candidates standing in an election reduces the abstention rate (increases turnout). As shown by the analyses of the data so far, we have seen that there is a possible positive relationship between candidate numbers and turnout. In the elections analysed, there is also a clear threshold which determines whether additional candidates are allowed to run in the second round of elections. Based on this, a regression discontinuity design (RDD) is likely best placed to estimate the causal effect of increasing the number of candidates on the abstention rate as the threshold acts as a random assignment mechanism to generate exogeneity in the independent variable. However, in this case, the threshold is not deterministic of whether a candidate will run in the second round, as shown by the large amount of one-sided non-compliance at the threshold meaning a fuzzy regression discontinuity design is most appropriate. Here, we are interested in the intent-to-treat effects, focussing on treatment assignment rather than treatment exposure.

With fuzzy RDD designs, the threshold acts as an instrument to generate exogenous variation in the independent variable and therefore must satisfy the first-stage relevance assumption where:

$$\text{cov}(Z_i, X_i) \neq 0 \tag{2}$$

This assumption clearly holds as the threshold is a determinintistic requirement for whether additional candidates can choose to run in the second round. This is shown below for national elections where the cut-off is `12.5%`, with a clear discontinuity at the threshold but also a number of `0` values where there was non-compliance, demonstrating the need for a fuzzy RDD.

Figure 1: First Stage: Discontinuity in Treatment Take-Up



Notes: This example has been done on the set of national elections. Other election types are analysed below.

For a fuzzy RDD design, we also need to ensure that there is local independence of the elections just above and below the threshold. This is necessary so that we can confidently argue that turnout changes with candidate numbers as a causal effect of additional candidates rather than some observed or unobserved heterogeneity between parties standing. In this case, this would mean that 3rd-place candidates just above (e.g., 12.6% vote share) are similar to those below the threshold (e.g. 12.4% vote share) in every way other than the fact that one can stand in the second round and one cannot. This is tested for in the next section.

Next, the exclusion restriction assumption should hold whereby the threshold only affects the outcome through the treatment of an additional 3rd-place candidate standing, rather than the threshold affecting turnout through another mechanism such as tactical voting and protest votes. This cannot be directly tested, but effects can be shown to be minimal through a pre-treatment placebo. However, in this context voting bahviour in different rounds (where the first acts as a 'qualification round' more than an outright winner-takes-all) may vary as voters and interest groups mabilise to employ tactical and protest voting behaviours.

The monotonicity assumption cannot be tested either but is likely to hold due to electoral rules. Monotonicity assumes that no candidate is less likely to run after qualifying for the second round after receiving the threshold treatment. This holds as no unqualified candidates can run (hence the one-sided non-compliance), and that there is an implict assumption that candidates who run in the first round do so with the objective of winning the election; therefore, with this objective, candidates would be more likely to run than to abstain upon recieving the treatment in order to maximise their chances of running.

Finally, we assume that there is no sorting and the running variable remains continuous across the threshold. In this case, this means that candidates cannot manipulate their vote share to artificially cross the threshold. If this were to happen, we would see a clustering of candidates with vote shares near to the threshold. This will next be tested to ensure the density of candidate vote shares is smoothly distributed across the threshold.

2.3 Falsification Checks

Conduct at least two empirical tests to show that your choice of research design is internally valid (falsification checks).

Note: after assessing the following falsification checks with Cantonal (2011) data, the sample size is too small for many meaningful results. Therefore, the following analyses focus on subsets of **`france_data`** *for national and cantonal elections (exc. 2011) only.*

Falsification tests can be used to probe the validity of the identifying assumptions of sorting, the exclusion restriction, and local independence stated above.

Firstly, a density test can help show a smoothly distributed running variable which does not cluster at the threshold such that the sorting assumption holds. This test uses a null hypothesis that the running variable **`threshold_party_can3`** is continuous at the threshold.

Table 6: Density Test Results by Election Type

Election Type	Test Statistic	P-Value	Significance
National Elections	-2.374	0.0176	*
Cantonal Elections	-0.960	0.3369	-

Note:

Significance stars: *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$.

`Table 6` shows the results of these density tests across each election type and its threshold. Cantonal elections show no sign of sorting with a p-value of `0.337` meaning we cannot reject the null. However, the national elections show a slightly significant p-value of `0.018`. This indicates that there is some evidence of sorting at the threshold in national elections but only with weak significance.

This continuity in the running variable can also be seen in the visual plots below. Note that given the low number of observations around the threshold for Cantonal elections in 2011, this plot is not displayed. Given the low signficance and visually continuous nature of the running variables, it is concluded that sorting at the threshold is not a major concern for using a fuzzy RDD research desgin.

Figure 2: Density plot of national vote share relative to the 12.5% threshold

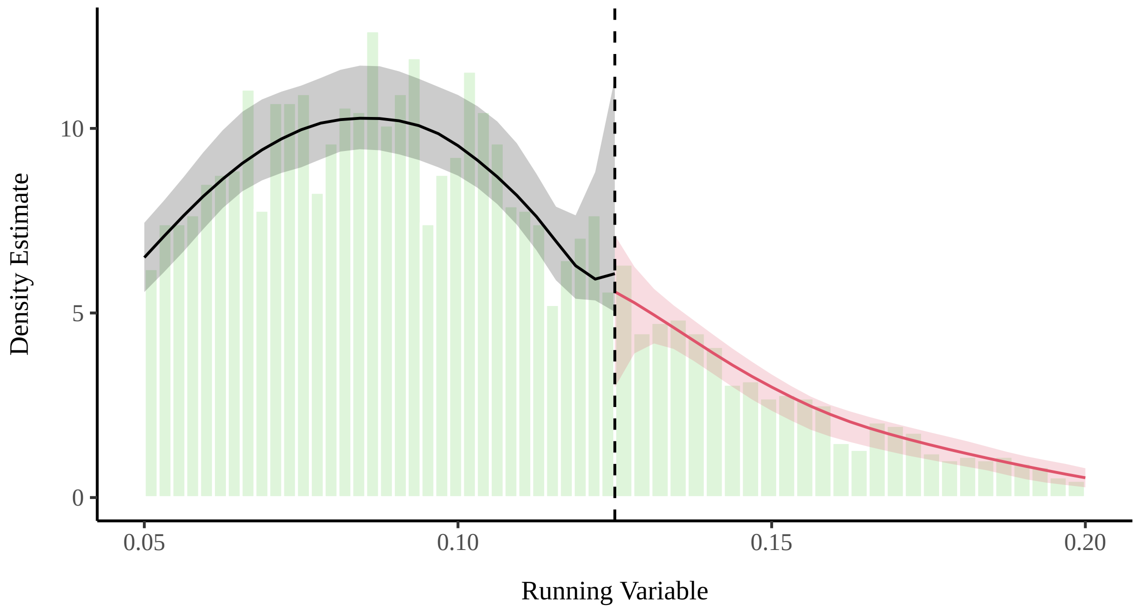
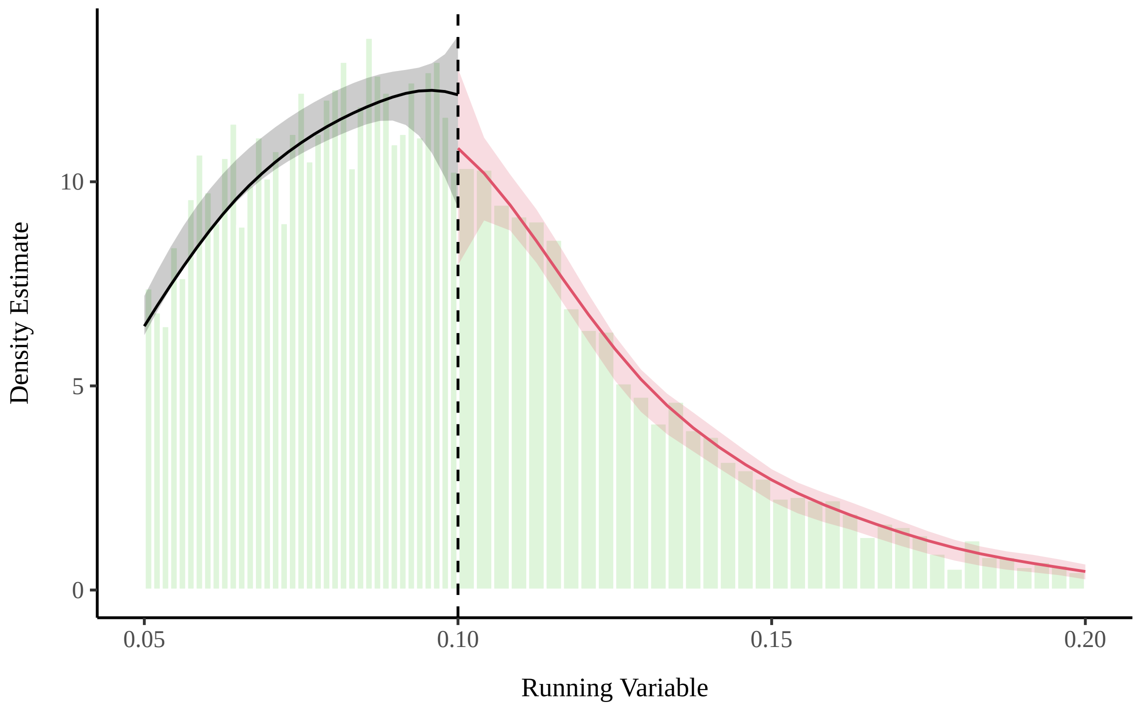


Figure 3: Density plot of Cantonal vote share relative to the 10% threshold



Next, balance and placebo tests can be used to assess local independence. **`Table 7`** below tests the local balance near the threshold across the set of pre-treatment covariates to see whether the treated and control groups near the threshold are comparable.

Table 7: Balance Test Results: Covariate Discontinuities by Election Type

	National			Cantonal		
Covariate	Estimate	P-Value	Significance	Estimate	P-Value	Significance
ideology_can1	-0.06	0.4477	-	-0.044	0.3962	-
ideology_can2	-0.038	0.7257	-	0.071	0.2121	-
ideology_can3	-0.477	0.0063	**	-0.115	0.1654	-
turnout_t1	-0.004	0.6578	-	0.009	0.1056	-
margin_t1	0	0.9909	-	0.001	0.8075	-

Note: Significance stars: *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$. Estimate values are the average difference in covariate values at the threshold.

The covariates included are variables which could theoretically affect voting behaviour and the turnout/null votes, and are pre-determined independently of the threshold. The `Estimate` coefficients in the table show the average difference between the covariate values before and after the threshold to see whether there is a discontinuity, and therefore a possibility that the covariate could be a factor - other than the threshold - which affects the outcome, thus violating the local independence assumption. The only covariate where we see a statistically signficant difference at the threshold is `ideology_can3` in national elections with a coefficient of `-0.477` and p-value of `0.006`. The negative coefficient implies that third-place party candidates above the threshold are more likely to be more left wing (assuming lower on the ideology scale = left). As this covariate is statistically significantly different, the local independence assumption may not hold as stronly in national elections and is therefore included as a control in the fuzzy RDD calculations. However, across other covariates, there appears to be randomness at the threshold for both national and cantonal elections.

Finally, although we cannot test the exclusion restriction directly, placebo tests are used to help verify that the only pathway of affect on the outcome variable is that of an additional candidate given by the threshold treatment and not other factors correlated with the threshold of the running variable. The placebo tests look at the outcome variable at different threshold cut-offs, with the expectation that there would be no changes in the outcome at these different cut-offs - there should be no discontinuities.

The placebo tests are run at different cut-offs above and below the real cut-off of `12.5%` for national elections and `10%` for cantonal elections. Due to there only being one-sided non-compliance above the threshold, running a fuzzy RDD estimate for threshold values below the real cut-off will lead to zero values at the first-stage. Therefore, for thresholds below the actual cut-off, a sharp RDD is run, and a fuzzy RDD is run for above the cut-off. The placebo tests are run at intervals of `0.01` between each cut-off. The results are shown in `Table 8` and `Table 9` below.

Table 8: Placebo Test Results: National Elections

Direction	Cut-off	Estimate	P-value	Significance
Below	0.115	-0.009	0.4847	-
Below	0.105	0.008	0.5246	-
Below	0.095	0.002	0.8628	-
Below	0.085	0.002	0.8500	-
Above	0.135	-0.196	0.2268	-
Above	0.145	3.910	0.9270	-
Above	0.155	0.067	0.6560	-
Above	0.165	-0.193	0.7774	-

Note: Significance stars. *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$. Estimates above represent the LATE, while estimates below represent the ATE.

Table 9: Placebo Test Results: Cantonal Elections

Direction	Cut-off	Estimate	P-value	Significance
Below	0.09	-0.009	0.1859	-
Below	0.08	0.001	0.9394	-
Below	0.07	0.003	0.8009	-
Below	0.06	-0.014	0.2424	-
Above	0.11	0.004	0.9624	-
Above	0.12	0.431	0.5242	-
Above	0.13	0.420	0.8515	-
Above	0.14	-0.039	0.6886	-

Note: Significance stars. *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$. Estimates above represent the LATE, while estimates below represent the ATE.

As expected, the estimates from these tests are near zero and not statistically significant across each different cut-off values above and below the threshold, and for both national and cantonal elections. This suggests that the exclusion restriction holds as there is no jump in the outcome variable at the placebo cut-offs, and therefore the threshold does not affect the outcome through any other mechanism than the treatment of an additional candidate standing in the second round. The only exception to this is for national elections where there is a small jump in turnout at `0.145`; however, this is not a statistically significant jump, and we can concretely verify the use of a fuzzy RDD approach.

Based on the above three falsification checks, we have shown that the choice of a fuzzy RDD researcg design is generally internally valid. In particular, for cantonal elections we can be very confident of the interal validiy as there was no sign of sorting at the threshold, nor an imalance of covariates giving credence to the local independence assumption, and finally there are no statistically discontinuities at the threshold helping support the exlcusion restriction.

For the final fuzzy RDD analyses, some doubt may be cast on the validity of the results for national elections, but the inclusion of `ideology_can3` will be used to help mitigate any confounding issues.

2.4 Non-Parametric Fuzzy RDD Estimation

Finally, estimate the causal effect of a third candidate standing in the second round, using the research design of your choice. Consider two main outcome variables: (1) turnout and (2) null and blank votes. Select or construct appropriate variables to operationalise the main outcome variables and the running variable. Interpret the results and discuss the effect size of your estimates in the first/second stage and their statistical significance.

Finally, to estimate the effect of a third candidate standing in the second round on election turnout and null/blank votes, a fuzzy RDD test is calculated across both national and cantonal elections. As with the falsification tests above,the results are split across **`national`** and **`cantonal`** election types with each taking their cutoffs of **`12.5%`** and **`10%`** respectively. The fuzzy RDD estimates are calculated across two outcome variables of **`turnout_t2`** and **`blancsnull_t2`**, with the treatment status defined by **`ran_t2_can3`** and the running variable **`threshold_party_can3`**. For robust results, a **`triangular`** kernal and MSE-optimal bandwiths are used, along with clustered standard errors, clustered by **`id_canton`**. As noted above, the covariate of **`ideology_can3`** is controlled for in national election calculations.

Table 10: Fuzzy RDD Estimates by Stage and Outcome

		National			Cantonal		
Stage	Outcome	Estimate	P-Value	Significance	Estimate	P-Value	Significance
First	blancsnull_t2	-0.056	0.0005	***	-0.029	0.0000	***
Second	blancsnull_t2	-0.056	0.0005	***	-0.029	0.0000	***
First	turnout_t2	0.035	0.4772	-	0.026	0.0374	*
Second	turnout_t2	0.035	0.4772	-	0.026	0.0374	*

Note: Significance stars: *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$.

`Table 10` reports the first- and second-stage results of the Fuzzy RDD estimates. The coefficients provide a Local Average Treatment Effect (LATE) estimate for when third-place candidates near the threshold from the first-stage elections comply with the threshold treatment and stand in the second stage. The first-stage estimates show the effect of the treeshold treatment on uptake (standing in the second election), and the second-stage estimates show the effect of the treatment on the outcomes of turnout and null/blank votes.

The first-stage estimates are signficant for all models other than national turnout. This indicates that the threshold is a strong instrument for determining whether an additional candidate stands in the second round, thus supporting the relevance assumption discussed above. However, as shown in **`Figure 1`**, there is a clear discontinuity in treatment take-up at the threshold for national elections, indicating that the instrument should be strong for national turnout estimates too. Yet, the bandwidths for these estimates were specifically chosen based on the outcome varibles used in the calculations meaning different samples were used. This gives credence for the need to test the robustness of results at different bandwidths.

For `blancsnull_t2` estimates, in both national and cantonal elections, when a third candidate qualifies and runs, null and blank voting drops significantly. This is shown by effect sizes of `-5.6` and `-2.9` percentage points for national and cantonal elections respectively. These results are highly significant at the 99.9% confidence level, suggesting strong and robust support for the initial hypothesis that more candidate choice in elections reduces protest voting. Given this, we can argue that alienated voters are more responsive to a wider pool of candidates to choose from such that they can feel that their voice can be heard, rather than feeling the need to abstain in protest.

Now looking at `turnout_t2`, the presence of an additioanl candidate in the second round increases turnout in cantonal elections by `2.6` percentage points, which is statistically significant at the 95% confidence level. National election results also suggest a positive effect of increasing candidates on tunrout, with a `3.5` percentage point increase; however, this result is not statistically significant. These results provoide mixed evidence for the hypothesis that increasing candidate numbers increases turnout. The lack of significance in national elections can be due to substantive differences between national and cantonal elections, but also due to methodological issues. For example, the internal validity checks showed possible sorting at the threshold and imbalance of covariates which could confound the results. Sorting in national elections could come from tactical voting and campainging to push up the number of votes for a third-place candidates to just get them to the threshold, limiting the validity of the threshold instrument.

As a final robustness check, the fuzzy RDD models are estimated across a range of different bandwidths to see if changing the window of values around the threshold used in estimation alter results. The results shown below in `Figure 4` are consistent with those estimated in `Table 10`. For robustness and added credibility to our estimates, we want to see LATE estimates that are stable and flat across different bandwidths, that are statistically signficant with narrow confidence intervals. This is the case for all election types and outcome variables other than `turnout_t2` in national elections, the same as seen previously. This estimate remains not robustly signficant meaning possible substantive differences between national and cantonal elections. These checks therefore strengthern the credibility of the causal claim that an increase in the number of candidates on the ballot increases turnout and decreases the number of disaffected blank/null voters, an outcome which is particularly robust in cantonal elections.

Figure 4: Fuzzy RDD Bandwidth Sensitivity by Election Type and Outcome