Netherlands Organisation for Scientific Research
Social Sciences

Young Researcher Causal Inference Application Form 2023/2024

Registration form

1a. Details of applicant
-Name, title(s): F. Grozovskiy, intern-researcher at the Institute for Statistical Studies and Economics of Knowledge of HSE Moscow & M. Grados, PhD (Data Scientist)
-University, Department (or Institute): HSE Moscow, Faculty of Social Sciences & AI Factory at BBVA Mexico
-E-mail(s): fedor.grozovski@icloud.com & daygoro.grados@gmail.com

1b. Title of research proposal
Measuring Causal Impact of Compulsory Voting on Electoral Turnout

1c. Abstract
The global political landscape continuously evolves and faces new challenges. A significant concern in democratic societies is the decline in voter turnout, which undermines political representation and exacerbates political inequality. One potential solution to this issue is the introduction of compulsory voting (CV). In our study, we use data from the International IDEA website, comprising 2,463 records, and employ staggered difference-in-differences estimators (as an alternative approach to a TWFE model). Our hypothesis posits that the adoption of CV increases turnout only in the first post-treatment period, followed by a modest decline thereafter. We find evidence against this hypothesis and we also provide a measurement of the ATT to determine the magnitude of this effect.

Word Count: 115

Research proposal

2a. Description of the proposed research and societal results

Recent opinion polls indicate that "around 20-30 percent of Americans support compulsory voting" (Singh, 2024). There is a belief that compulsory voting (CV) would incentivize people to vote, shape their voting habits in the long-run, and contribute to the quality (maturity) of democracy in general: "for democracy to work, citizens must vote, yet turnout rates are low or declining in many democratic countries" (Singh, 2024). Indeed, one of the main arguments for CV is that it stimulates more representative (or, in other words, 'fair') outcomes of the elections and, consequently, "reduces socioeconomic inequalities" (Singh, 2024).

However, besides an elevated turnout rate, CV adoption also brings some side effects, such as an increased number of invalid ballots or votes in favor of the 'protest parties' (Singh, 2024) and greater support for left-wing parties, candidates, and policies, since more people with lower income are mobilized to vote (Bechtel et al, 2016; Singh, 2024).

The empirical research on the issue of CV gained momentum in the late 2000's – early 2010's. A number of influential political science papers pointed out the effects of CV in citizen's participation. In particular, Ferwerda (2014) notes that since CV raises turnout "through an artificial process", countries with and without CV (but with high levels of political participation) differ quite significantly. For example, despite its magnitude, the effect of CV disappears swiftly after it is abolished (Bechtel et al., 2018). In addition, Dunaiski (2021) did not find any evidence that CV forms voting habits and that citizens are eager to go to the polls after voting is not compulsory anymore. Moreover, in spite of widely spread claims, it is still unsubstantiated to argue that CV brings higher political and civic engagement (Singh, 2024).

An essential for our paper research by Kostelka et al. (2024) focused on democratic states in 1945-2017 and utilized the FE-regression models to discover the importance of sanctions for abstention on the CV's effect on turnout. Namely, they found that in countries where CV is *de-jure* present but *de-facto* is not enforced, the effect on turnout is considerably lower compared to the states with any legal enforcement of this rule.

In this study, we test a hypothesis and answer a research question, presented below:

- **Hypothesis**: the introduction of CV increases citizen participation in elections only in the first post-treatment period, followed by a modest decline of turnout in the subsequent periods. This is a one-tailed test.
- **Research question**: by how much, on average, does the CV introduction increase the turnout? This question statement implies that we need to estimate the ATT (Average Treatment Effect on the Treated).

The proposed research would benefit both society and academia. The societal results of the study would be useful for policy implications as the findings contribute to the debates on electoral reforms and designs of policies aimed at enhancing democratic participation. We demonstrate that the consideration of legal CV enforcement mechanisms is necessary to maintain high levels of political engagement over time.

It is also quite important to acknowledge that we find a positive effect of CV on turnout which (in contrast with the proposed hypothesis) persists even after the first post-treatment period. In other words, this study also provides robust empirical evidence that the cumulative elevating effect that CV has on turnout persists over time if all countries that have held elections at some period since 1945 are considered. We discuss the reasons why we achieved such findings in the section 2c and ibid argue for additional measures or reforms.

Our research furthermore includes an ATT measurement, quantifying the magnitude of CV's impact on voter turnout globally. This allows for a clear understanding of the causal influence of CV on turnout (on average). The staggered difference-in-differences methodological approach can be applied to other studies examining the effects of similar interventions in different (for example country-specific, region-specific, regime-specific, etc.) contexts.

2b. Approach: what is your methodological or experimental approach?

We use panel data from the International IDEA website. The cleaned dataset includes both presidential and legislative elections ranging from 1945 and 2023 *(N = 2,463 records)*. It should be noted that our panel is not balanced: elections are held at different times with varying time intervals in different countries.

As per design specification, the subjects in our research are countries, the outcome variable is *Turnout* (in a percentage scale), and the treatment variable is the *adoption of CV* (binary). We employ the TWFE model to estimate the causal effect and verify it using a staggered difference-in-differences approach (see de Chaisemartin & D'Haultfoeuille, 2020; de Chaisemartin & D'Haultfoeuille, 2024).

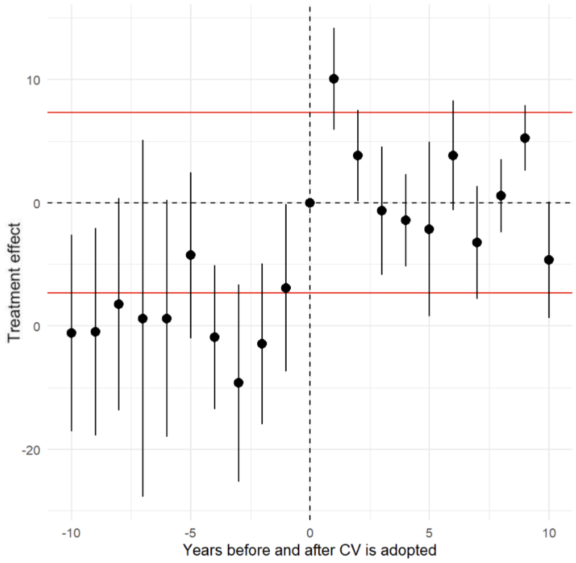
Let us now provide more details on the statistical models employed in the study. Wooldridge (2021) specifies that the TWFE (Two Way Fixed Effects) estimator includes both unit fixed effects "to remove unit-specific time averages" and time fixed effects "to remove secular changes in the economic environment that have the same effect on all units".

The TWFE assumptions are the following:

- Parallel Trends
- No anticipation
- Treatment responsiveness is homogeneous across treated subjects
- Conditional on treatment group status, treatment exposure is homogeneous across subjects
- ATTs are homogeneous over time
- All treated subjects are treated at the same time.

In our set up some TWFE assumptions are violated, while others are not. Thus, there is *no anticipation* (treatment takes effect on the first elections only after CV is introduced). Besides, in our set up *conditional on treatment group status, treatment exposure is indeed homogeneous across subjects* (CV is either adopted or not). As for the parallel trends assumption, it was tested with plotting placebo and post-treatment effects (see Figure 1). On the one hand, the assumption is not fulfilled because point estimates and their confidence intervals are outside the lower bound of the ATT. On the other hand, the negative placebos are aligned almost on the same straight line. Therefore, our estimation of the positive ATT will only be strengthened by this fact.

Figure 1
PTA plot



Yet some assumptions do not hold in our set up (making TWFE results not robust). Such as, *the treatment responsiveness is not homogeneous across treated subjects* (there may be some countries whose turnout was high enough before CV adoption, making the change not substantial). Moreover, A*TTs are also not homogeneous over time* (the turnout in the first elections after the adoption of CV (should, on average) reach a peak and decline afterwards). Finally, *not all treated subjects are treated at the same time* (due to the time imbalance in our panel data).

Moving on, the staggered difference-in-differences estimator from (de Chaisemartin & D'Haultfoeuille, 2020) is defined with the following formula:

$$\mathrm{DID}_M = \sum_{t=2}^{T} \left(\frac{N_{1,0,t}}{N_S} \mathrm{DID}_{+,t} + \frac{N_{0,1,t}}{N_S} \mathrm{DID}_{-,t} \right),$$

where $N_S = \Sigma(g,t): t \geq 2, D_{g,t} \neq D_{g,t} - 1 N_{g,t}$, $\mathrm{DID}_{+,t}$ accounts for the treated, and $\mathrm{DID}_{-,t}$ accounts for the quitters (which are also present in our data).

The staggered DID (*did_multiplegt_old* in R) assumptions (by de Chaisemartin & D'Haultfoeuille, 2020) are:

- Balanced panel of groups
- Sharp design
- Common trends for $Y(0)$ and $Y(1)$
- Strong exogeneity for $Y(0)$ and $Y(1)$
- Existence of "stable" groups
- Mean independence between a group's outcome and other groups' treatments.

Once again, some staggered DID assumptions are violated, while others are not. On the one hand, the *sharp design* (all units belonging to the same group in the same period are exposed to the same treatment) and the *mean independence between a group's outcome and other groups' treatments* (mean turnout in one group does not depend on the treatments of the other group) assumptions hold in the set up of this study. On the other hand, *our panel is not balanced* (in time) and *there are no "stable" groups* (not always there exists a country that remains treated at both time periods, if there is a country that switches from being treated to untreated).

Regarding the *common trends* and *strong exogeneity for Y(0) and Y(1)*, these assumptions are *probably* not violated. Common trends for *Y(0)* and *Y(1)* assumption poses that "between each pair of consecutive periods, the expectation of the outcome without treatment follow the same evolution over time in every group untreated at t−1, and that the expectation of the outcome with treatment follow the same evolution in every group treated at t−1" (de Chaisemartin & D'Haultfoeuille, 2020). This is quite problematic to test in our data, but we still suppose that these expectations evolve similarly in each of the groups. The strong exogeneity assumption "requires that the shocks affecting a group's *Y(0)* and *Y(1)* be mean independent of that group's treatment sequence" (de Chaisemartin & D'Haultfoeuille, 2020). This is another assumption which is not really feasible to test, yet we suppose that it is not violated in our set up.

Overall, the limitations of our approach are: 1) the absence of covariates and 2) time imbalance. While it is difficult to overcome the first issue due to the limits of the data itself, we propose a solution to the second one. We bin the years of elections for each country in intervals of 10 years (e.g. the year 1945 is transformed into the range [1945-1954], while the year 2023 – into the interval of [2015-2024]). This way, we also fulfil two more assumptions for staggered DID: *balanced panel of groups* and *"stable" groups*, leaving only two assumptions in question (*common trends* and *strong exogeneity for Y(0) and Y(1)*).

As a result, this data transformation allows us to use the *did_multiplegt_dyn* R package (de Chaisemartin & d'Haultfoeuille, 2024) and to compare the results produced by it with the output of *did_multiplegt_old* (de Chaisemartin & d'Haultfoeuille, 2020).

2c. Provisional results: What are your findings?
The TWFE model (with effects fixed on both year and country) estimated the ATT at 14.366 (significant at $\alpha = 0.001$). Thus, if we were eager to rely on the TWFE estimations, we would argue that the turnout in the treated group of countries was approximately 14% higher (on average) than in the untreated group. The overall output of the model is presented in Table 1.

Table 1
TWFE estimation[1]

	Estimate	**Cluster s.e.**	**P-value**
CV_idea	14.366	3.601	0.000973

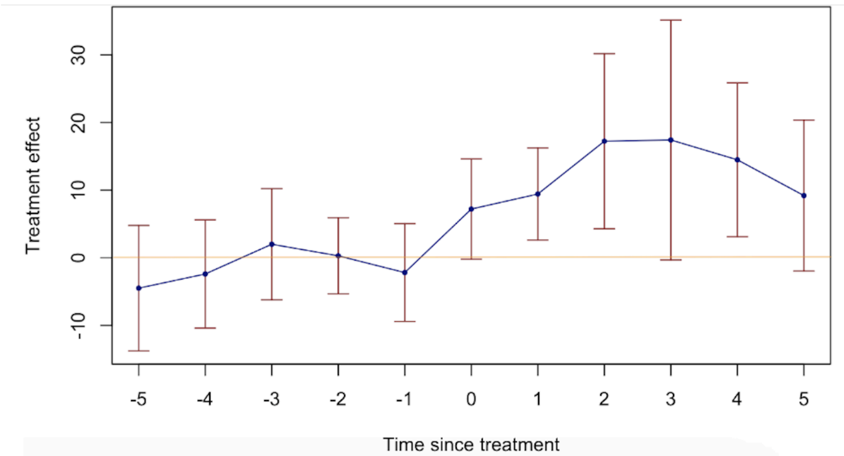
However, due to the numerous violations of the model's assumption, we cannot rely on these unrobust results. Therefore, we turn to more advanced estimators by de Chaisemartin & d'Haultfoeuille. Running a *did_multiplegt_old* R package produced an ATT of 7.030113, whereas *did_multiplegt_dyn* estimated the ATT of 10.64258. Still, both these ATT estimates are lower than the TWFE's one.

Figure 2 demonstrates the pre-treatment lags and post-treatment leads plot (by *did_multiplegt_old)* with time since treatment on the x-axis and treatment effect on the y-axis. All the placebo effects (lags) are not significant at $\alpha = 0.05$ (their 95% confidence intervals include zero), while three of the leads turned out to be significant at $\alpha = 0.05$ (lead 1, 2, and 4)[2].

[1] Estimator *felm* in R; SE's are clustered on a Country id ('cid') and 'year' variables; weighted on population.
[2] There is an orange line at y = 0 for the sake of clarity.

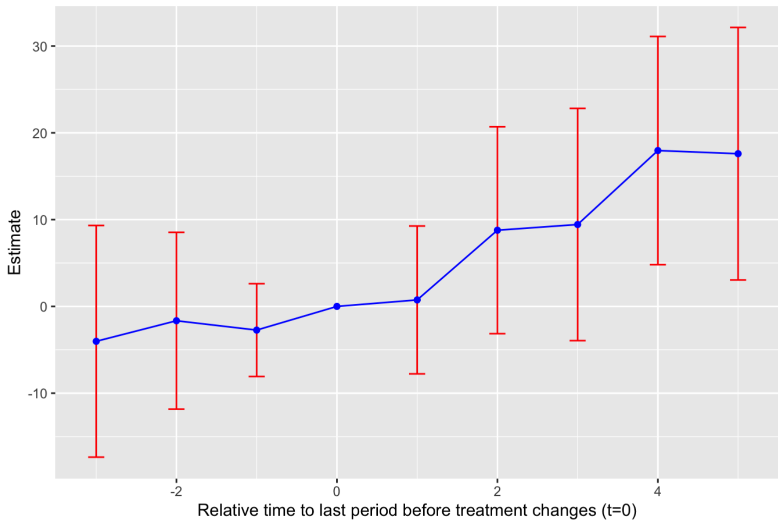
Young Researcher Causal Inference Application Form 2023/2024

Figure 2
Lags and Leads plot[3]



The same plot of the pre-treatment lags and post-treatment leads by *did_multiplegt_dyn* is presented in Figure 3. All the placebo effects (lags) are still not significant at α = 0.05 (their 95% confidence intervals include zero), while only the last two leads are significant at α = 0.05 (lead 4 and 5).

Figure 3
Lags and Leads plot[4]



Given the statistical insignificance of the placebos (pre-treatments lags) and the growth of the treatment effect in the post-treatment periods we can argue for the existence of a positive causal effect of the CV introduction (treatment) on the turnout (outcome). To answer the research question *(by how much, on*

[3] Estimator *did_multiplegt_old* in R by (de Chaisemartin & d'Haultfoeuille, 2020); time variable – 'year'; 5 lags; 5 leads; 30 bootstraps (to compare with the output of *did_multiplegt_dyn* – see Figure 3); SE's are not clustered (*did_multiplegt_old* does not allow clustering SE with our data).
[4] Estimator *did_multiplegt_dyn* in R by (de Chaisemartin & d'Haultfoeuille, 2024); time variable – 'interval' (decades, such as [1945-1954], [1955-1964], etc.); 3 lags (maximum in *did_multiplegt_dyn*); 5 leads; 30 bootstraps (maximum that *did_multiplegt_dyn* allows to use on our data); SE's are clustered on a Country id ('cid') variable.

average, does the CV introduction increase the turnout?), we will turn to the *did_multiplegt_dyn* R package (de Chaisemartin & d'Haultfoeuille, 2024) ATT estimation. Its output is superior and is the most credible compared to the one by basic TWFE model, as well as to the one by *did_multiplegt_old* R package (de Chaisemartin & d'Haultfoeuille, 2020). Since the TWFE model requires a number of assumptions which are rarely satisfied in the real world data set ups (just as in our case) and the *did_multiplegt_old* estimator (de Chaisemartin & d'Haultfoeuille, 2020) is claimed to perform worse than the *did_multiplegt_dyn* (de Chaisemartin & d'Haultfoeuille, 2024) by de Chaisemartin and d'Haultfoeuille themselves, we believe that our preference of *did_multiplegt_dyn* is legitimate.

Thus, the overall output of the *did_multiplegt_dyn* is presented in Table 2, providing the estimations of pre-treatment (Placebo_1, etc.) and post-treatment effects (Effect_1, etc.), as well as the ATT measurement. **Consequently, on average, CV introduction increases the turnout by 10.64258%**. Congruently, all the placebos are negative and insignificant at $\alpha = 0.05$, while all the effects are greater than 0 (though not all of them are statistically significant).

Table 2
Staggered difference-in-differences estimation[5]

	Estimate	**SE**	**LB CI**	**UB CI**	**N**	**N Switchers**
ATT	10.64258	5.91776	-0.95602	22.24118	591	39
Placebo_1	-2.72918	2.72802	-8.07601	2.61765	278	12
Placebo_2	-1.65203	5.19688	-11.83772	8.53366	79	6
Placebo_3	-4.02214	6.80972	-17.36895	9.32467	34	3
Effect_1	0.74568	4.34641	-7.77312	9.26448	376	13
Effect_2	8.77782	6.08155	-3.14180	20.69744	222	9
Effect_3	9.43592	6.82629	-3.94336	22.81520	217	9
Effect_4	17.95765	6.70801	4.81018	31.10512	92	5
Effect_5	17.59296	7.42234	3.04544	32.14048	89	3

As one may notice, **contrary to our hypothesis, the cumulative effect of the CV adoption persists over time** (in the 4 post-treatment periods – see Figure 3). This is likely due to the fact that our dataset includes all the possible country-specific scenarios, with both enforced and unenforced CV legislations. Perhaps due to the fact that in certain countries CV was initially introduced as a mere formality and only afterwards the legal mechanisms to compel compulsory attendance at the polling stations appeared, the average treatment effect over time continues to grow in subsequent periods after treatment. Such an assumption is consistent with the empirical study by Kostelka et al. (2024). Using FE-regressions, this paper shows that unenforced (unsanctioned) CV has a much smaller effect on turnout than enforced (sanctioned) one. Thus, in the model without covariates (with the dependent variable being log-transformed turnout), the coefficient for the predictor of unsanctioned CV was 9.63 versus 22.72 for the sanctioned CV. Adding the covariates into the model makes the differences smaller, but still

[5] See the previous footnote.

substantial (10.13 against 18.62). All the estimates of these coefficients were statistically significant at $p < 0.001$.

The fact that the cumulative effect of the CV introduction remains and growth over time could also be conditioned by the presence of countries with different political regimes (democratic and undemocratic) in our data, since authoritarian states may be interested in manipulating turnout. Although, as empirical research suggests, this is definitely not the key reason: the differences in turnout in *de-facto* competitive authoritarian elections and uncompetitive ones are quite small (see i Coma & Morgenbesser, 2020). The existence of this cumulative effect may also occur due to the variations in time, in which the CV was adopted in different countries.

Hence, the introduction of CV should be supported by additional legal measures (such as sanctions for abstaining from expressing one's will in elections). If only countries with enforced CV were represented in our data, we would be able to observe a much larger ATT (measuring the effect of CV on turnout) than the 10.64258 we obtained. Besides, as Kostelka's paper suggests, it is generally not important which sanctions for abstention are utilized to support CV (severe or not severe, monetary or non-monetary).

Word Count (Sections 2a – 2c): 2468 (using Google Docs word count; excluding figures and tables; including all footnotes)

References

Bechtel, M. M., Hangartner, D., & Schmid, L. (2018). Compulsory voting, habit formation, and political participation. *Review of Economics and Statistics*, *100*(3), 467-476. https://doi.org/10.1162/rest_a_00701

Bechtel, M. M., Hangartner, D., & Schmid, L. (2016). Does compulsory voting increase support for leftist policy?. *American Journal of Political Science*, *60*(3), 752-767. http://dx.doi.org/10.7910/DVN/29591

De Chaisemartin, C., & d'Haultfoeuille, X. (2024). Difference-in-differences estimators of intertemporal treatment effects. *Review of Economics and Statistics*, 1-45. https://doi.org/10.1162/rest_a_01414

De Chaisemartin, C., & d'Haultfoeuille, X. (2020). Two-way fixed effects estimators with heterogeneous treatment effects. *American economic review*, *110*(9), 2964-2996. https://doi.org/10.1257/aer.20181169

Dunaiski, M. (2021). Is compulsory voting habit-forming? Regression discontinuity evidence from Brazil. *Electoral Studies*, *71*, 102334. https://doi.org/10.1016/j.electstud.2021.102334

Ferwerda, J. (2014). Electoral consequences of declining participation: A natural experiment in Austria. *Electoral studies*, *35*, 242-252. https://doi.org/10.1016/j.electstud.2014.01.008

Gaebler, S., Potrafke, N., & Roesel, F. (2020). Compulsory voting and political participation: Empirical evidence from Austria. *Regional Science and Urban Economics*, *81*, 103499. https://doi.org/10.1016/j.regsciurbeco.2019.103499

i Coma, F. M., & Morgenbesser, L. (2020). Election turnout in authoritarian regimes. Electoral Studies, 68, 102222. https://doi.org/10.1016/j.electstud.2020.102222

Kostelka, F., Singh, S. P., & Blais, A. (2024). Is compulsory voting a solution to low and declining turnout? Cross-national evidence since 1945. *Political Science Research and Methods*, *12*(1), 76-93. https://doi.org/10.1017/psrm.2022.57

Young Researcher Causal Inference
Application Form 2023/2024

Singh, S. P. (2024). Compulsory voting. *The University of Chicago Center for Effective Government's Democracy Reform Primers.* https://blockyapp.s3.eu-west-2.amazonaws.com/store/d0a6ab8dafdb4ca4518ee5decb4222c3.pdf

Wooldridge, J. M. (2021). Two-way fixed effects, the two-way mundlak regression, and difference-in-differences estimators. Available at SSRN 3906345. http://dx.doi.org/10.2139/ssrn.3906345