

**Final Paper: The Impact of E-Voting on Citizen Perception**

Minerva University

CS130: Statistical Modeling - Prediction and Causal Inference

Prof. A. Diamond

April 22, 2022

## Final Paper: The Impact of E-Voting on Citizen Perception

**To:** R. Michael Alvarez, Ines Levin, Julia Pomares and Marcelo Leiras

**From:** William Nguyen and Mishaal Lakhani

**Date:** April 22, 2022

### Executive Summary

After trying a different method from the one used in the paper by Alvarez, Levin, Pomares and Leiras, we believe that greater and more reliable balance between treatment and control groups can be achieved thanks to:

1. Replacing *Propensity Score Matching* with *Genetic Matching* to improve covariate balance and verify results reported in *Table 2. Balance Statistics* in the paper
2. Extending our study validity check by conducting a sensitivity analysis to better understand strengths and weaknesses of *Propensity Score Matching* and *Genetic Matching*

### Background Information

This paper evaluates the adoption of e-voting in Salta, Argentina. Specifically, it looks at the evaluation, and overall perception voters have of e-voting compared to regular voting. This exploration is because of the intuition that the different mediums of voting create a different level of trust/transparency regarding the election, which is formed by the degree of satisfaction. From this inspiration, the answer is framed as, **“Does a change from traditional voting to e-voting improve voter perceptions?”**. The study examines different outcomes, ranging from people’s satisfaction with the alternative voting medium to their feelings of transparency and

confidentiality of the voting system. However, in our response, we will only focus on one outcome only (*agree.evoting*), specifically **whether or not the experience with e-voting makes citizens advocate for substituting traditional voting by e-voting**.

The ideal experiment would be a randomized assignment to voting technology, but the realistic scenario is that the voting technologies have been distributed by socio-demographics. Therefore, the study is subject to the influence of confounding variables. There are 7 covariates assessed in this study, namely gender, age, education, full-time employment status, technology familiarity, political information and type of employment. Therefore, matching methods are appropriate to correct both groups' imbalance.

## Data Collection

In terms of data collection, a long questionnaire was administered to 1,502 voters, 887 of whom used the e-voting method. The raw and processed data can be both accessed in Harvard Dataverse. Most questions were qualitative and quantified by converting ordinal categorical variables into numerical values. For example, the variable about ease of voting procedure was coded as 1 if respondents' answers were "*quick*" or "*very quick*" and 0 otherwise ("*slow*", "*very slow*").

The analysis breaks down into exploring two areas that help understand voter sentiments. One is understanding the ease of use of the voting procedures. The other is the confidence each method gives voters. These are interrelated because if there is increased complexity in usage, it can become unfamiliar and untrustworthy, thus negatively impacting perception and confidence. Note that we chose to work with the outcome variable about citizen's consent with the e-voting transformation which turns out to be binary.

## Replication: Propensity Score Matching

We tried to replicate *Table 2. Balance Statistics*, both before and after Matching. The original paper uses propensity score matching, a method using logistic regression to predict the probability that a unit is assigned to treatment given observed covariates. Then, when each unit has a score of subjectivity to treatment, we just match treated units with control ones with the closest scores. We tried to use the same model as the original paper, which was pretty complicated and included so many interaction terms. We also maintained the same pre-processing strategy (e.g. set NA values to massive numbers so that this observation does not get matched on). We also chose to set the same caliper as in the original paper which is 0.05. However, while the pre-matching balance statistics are the same, there is a difference between post-matching balance statistics. The original paper reports that the Minimum After Matching KS p-value is 0.55 while our result shows a value of 0.04 which is much smaller. We were aware that we used the Matching package which was different from the statistical tool in the original paper (MatchIt). This difference can cause our setting arguments to the same value to actually input different signals for the machine.

	ev	tv	diff	pvalue
age.group	2.4757506	2.4433498	0.032400824	0.52600000
educ	4.7713626	4.1428571	0.628505444	0.00000000
white.collar	0.3025404	0.2758621	0.026678347	0.29287524
not.full.time	0.2771363	0.3349754	-0.057839111	0.01998267
male	0.4965358	0.4909688	0.005566995	0.87472467
tech	4.1836028	3.9096880	0.273914758	0.00000000
pol.info	1.4745958	1.3103448	0.164251015	0.00000000

Table 1. Pre-Matching Balance Statistics between two groups (E-Voting vs Traditional Voting). We confirm the reported table in our original paper.

Covariates	<u>Original Paper</u>			<u>Our results</u>		
	Mean Treatment	Mean Control	KS P-value	Mean Treatment	Mean Control	KS P-value
<i>age.group</i>	2.5	2.5	1	2.5	2.4	0.36
<i>educ</i>	4.2	4.2	0.72	4.75	4.75	0.34
<i>white.collar</i>	29.2	28.4	0.80	30.5	29.7	0.68
<i>not.full.time</i>	30.8	32.0	0.80	27.4	25.5	0.26
<i>male</i>	49.0	49.0	1	49.2	47.8	0.48
<i>tech</i>	4.0	3.9	0.59	4.2	4.2	0.13
<i>pol.info</i>	1.4	1.3	0.55	1.5	1.4	0.04

Table 2. Post-Matching Balance Statistics between two groups (treatment vs control) in the original paper and in our results. The Minimum After Matching p-value is different, yet the mean and the variable share consistency

The bad result of Matching can be attributed to setting missing values to massive values, which affects matched units. We experimented with some variations to make Propensity Score Matching achieve better results. We first changed the model to a much simpler one. The logistic regression model in the original paper includes so many interaction terms and in total more than 20 predictors, which can involve a lot of bias and model dependence. Therefore, we only tried with the simplest form of model. We also set the caliper to a stricter value (0.01) so that matching will be more effective. We also cleaned the data differently, specifically, only extracting relevant columns (covariates, treatment and 1 outcome) and deleting rows with

missing values instead of changing to big values. We also set BiasAdjust to TRUE to adjust bias in our model. With the least matched covariate still being political information, Minimum P-value improves from 2.22e-16 to 0.135, much better than 0.04 before.

## Genetic Matching

We now changed to the method of Genetic Matching to improve the matching procedure. For Genetic Matching, we used the function GenMatch to find a set of optimal scalars assigned to different predictors so that the multi-dimensional space is adjusted most appropriately and matching is most efficient between treated and control units. After we achieve a set of optimal scalars for all predictors, we will use the Match function to conduct matching procedures and estimate the average treatment effect because now we have two quite similar groups for valid comparison. Finally, we use MatchBalance to recheck once again the degree of covariate balance between two groups.

Here, we used data where we only included relevant columns and removed missing values. We aimed for better matching by increasing the number of bootstraps in the Matching procedure (setting *nboots* to 500). The larger *nboots*, the better but we need to balance with time to run the code. We also tried with exact matching on binary values, namely gender, employment type and technology familiarity. However, the results with exact matching seem very not promising.

Covariates	<b><u>Before Matching</u></b>			<b><u>After Matching</u></b>		
	<b>Mean Treatment</b>	<b>Mean Control</b>	<b>P-value</b>	<b>Mean Treatment</b>	<b>Mean Control</b>	<b>P-value</b>
<i>age.group</i>	2.5	2.4	0.76	2.5	2.4	0.41

<i>educ</i>	4.76	4.1	5.4e-09	4.76	4.75	0.58
<i>white.collar</i>	30.5	28.1	0.33	30.5	30.3	0.31731
<i>not.full.time</i>	27.6	33.4	0.019	27.6	26.7	0.37
<i>male</i>	0.49	0.49	0.9	0.49	0.49	0.53
<i>tech</i>	4.2	3.9	0.0015	4.2	4.2	0.87
<i>pol.info</i>	1.46	1.3	3.3e-05	1.46	1.46	1

Table 3. Before and After Matching Balance Statistics using Genetic Matching. The results can be seen to have a sharp improvement from 5.4e-09 to 0.37 as the Minimum P-value

## Discussion

We can clearly see that with the same data pre-processing strategy, Genetic Matching performs much better than Propensity Score Matching. However, the estimate for treatment effect in both methods did not differ much, still around 0.18 and the standard error is around 0.02. Therefore, we should be really careful about matching results in our work.

Genetic Matching appears to be a more suitable and accurate technique because it seeks a set of optimal scalars to adjust for the multi-dimensional space, thus making the process of finding the nearest neighbor more comprehensive. Instead, Propensity Score Matching maps all units which already involve so many covariates to only 1 dimension which is somehow not fulfilling enough. Furthermore, using logistic regression introduces more extrapolation, bias and model dependence for Propensity Score Matching, which possibly prints out wrong results of matching.

## Sensitivity Analysis

We extend our study by conducting sensitivity analysis to check the robustness of our results. We only need the estimate of average treatment effect, the standard error and degrees of freedom to get statistics about sensitivity.

	Treatment Effect	Standard Error	Degrees of Freedom	Robustness Value of t-value	Robustness Value of point estimate
Propensity Score Matching	0.18273	0.023639	1402	14.26 %	18.62 %
Genetic Matching	0.18634	0.029299	1402	11.08 %	15.6 %

Table 4. The Table reports statistics about robustness of our results to hidden bias.

Although Genetic Matching performs better at matching treated units with control ones, this method is subject to more hidden bias. The Robustness Value of t-value and point estimate shows the minimum strength of unobserved confounders needed to bring the treatment effect estimate down to not a statistically significant range and to 0, respectively. Here, we see that values for Propensity Score Matching are higher, which means this method is less subject to influence of unmeasured confounders.

## Conclusion

Overall, even though Genetic Matching did better at achieving the optimal covariate balance, the method is less robust to hidden bias compared to Propensity Score Matching. This trade-off prevents us from suggesting the superiority of Genetic Matching with full confidence. Furthermore, we are also aware of constraints with regards to technologies (different packages used) and preprocessing data, which makes our replication not successful.



## Contribution Statement

William and Mishaal were the coauthors of this assignment. We both researched on and finalized on a paper to choose together. We then collectively read through and understood the paper, and outlined the key things that would be worked on for each part. The replication and extension code was adeptly written by William. The summary (executive summary, background) was done by Mishaal. The writeup for replication and extension was outlined by William, drafted by Mishaal. The figures were co-ideated and created. The editing was done collaboratively.

## Appendix

The code for the entire assignment can be found [here](#).

Sensitivity Analysis Test from the Online Calculator:

Regression Sensitivity Analysis: the Robustness Value and the partial  $R^2$ 
ENGLISH

Please use the following citation: Cinelli, C. and Hazlett, C. 'Making Sense of Sensitivity: Extending Omitted Variable Bias.' (2018)

**Input your data here**

**Coefficient estimate:**

**Standard error:**

**Degrees of freedom:**

☐ Compute bounds on confounding

☐ Change null hypothesis

REFRESH

INSTRUCTIONS   THE RV AND THE PARTIAL  $R^2$    CONTOURS (POINT ESTIMATE)   CONTOURS (T-VALUE)   EXTREME SCENARIOS

**Sensitivity statistics for routine reporting: the Robustness Value and the partial  $R^2$**

Use the sidebar on the left to enter the values for your study. If you need help, check the instructions tab first.

Partial $R^2$ of the treatment with the outcome	Robustness Value for the point estimate	Robustness Value for the t-value
2.8%	15.6%	11.08%

Sensitivity analysis statistics for routine reporting

Click on *Show interpretation* for an explanation of each value. Below you can also find LaTeX code for generating a minimal sensitivity reporting table.

☐ Show interpretation

☐ Latex code for sensitivity analysis table

Regression Sensitivity Analysis: the Robustness Value and the partial  $R^2$ 

ENGLISH

Please use the following citation: Cinelli, C. and Hazlett, C. 'Making Sense of Sensitivity: Extending Omitted Variable Bias.' (2018)

Input your data here

Coefficient estimate:

0.18273

Standard error:

0.023639

Degrees of freedom:

1402

☐ Compute bounds on confounding

☐ Change null hypothesis

REFRESH

INSTRUCTIONS

THE RV AND THE PARTIAL  $R^2$

CONTOURS (POINT ESTIMATE)

CONTOURS (T-VALUE)

EXTREME SCENARIOS

Sensitivity statistics for routine reporting: the Robustness Value and the partial  $R^2$

Use the sidebar on the left to enter the values for your study. If you need help, check the instructions tab first.

Partial  $R^2$  of the treatment with the outcome

Robustness Value for the point estimate

Robustness Value for the t-value

4.09%

18.62%

14.26%

Sensitivity analysis statistics for routine reporting

Click on *Show interpretation* for an explanation of each value. Below you can also find LaTeX code for generating a minimal sensitivity reporting table.

☐ Show interpretation

☐ Latex code for sensitivity analysis table

## References

Cinelli, C. and Hazlett, C. 'Making Sense of Sensitivity: Extending Omitted Variable Bias.' (2018) [https://carloscinelli.shinyapps.io/robustness\\_value/](https://carloscinelli.shinyapps.io/robustness_value/)

R. Michael Alvarez, Ines Levin, Julia Pomares and Marcelo Leiras (2013). *Voting Made Safe and Easy: The Impact of e-voting on Citizen Perceptions*. Political Science Research and Methods, 1, pp 117-137 doi:10.1017/psrm.2013.2