report

Manoel Galdino

2024-04-09

Introduction

Can Civil society monitoring on social media change Member of Parliament behavior? This report presents the result of the case of Rosie bot in Brazil.

We are interested in assessing the causal effect of monitoring on spending and reimbursement by MPs.

Data

To answer the question, we are analyzing data from Brazilian MPs allowance quota and the tweets by the Rosie Bot over time. We analyzed data from 2015 to 2022 for the lower chamber in Brazil.

To have a sense of the evolution of spending, it is useful to plot the data on our outcome variable.

Figure 1 below shows total spending per month over time.

```
## `geom_smooth()` using formula = 'y ~ x'
```

We see that there is no clear trend over time. However, there are periods in which the spending plummets. It's parliamentary recess.

Making comparisons

To understand the causal effect of Rosie, we need to specify what is the comparison of interest. Figure 2 below shows the evolution of treatment over time.

It is now known that in the presence of varying treatment periods for multiple units, the causal coefficient is a is the weighted average of all two-group/two-period DiD estimators in the data where some of the weights can be negative.

Standard estimators are only good if effects are homogeneous across units and across time periods, which in general is not credible. Thus, we need to use new DiD estimator. There are several estimators in the literature, but when units switch in and out of treatment, which is the case of our study, there is only one estimator (as far as I know): the one proposed by Imai and Kim (2021).

In the present analysis, we choose to compare MPs with four prior months of the same history of treatment and control status, and check if in up to 5 time periods from the treatment, there is a change in behavior.

Results

We ran a DiD estimator with matching, in which we are matching histories of treatment status, plus party and state as controls. Figure 3, 4 and 5 present the results for three different outcome variables: spending per Mp, net spending per MP and reimbursement per MP.

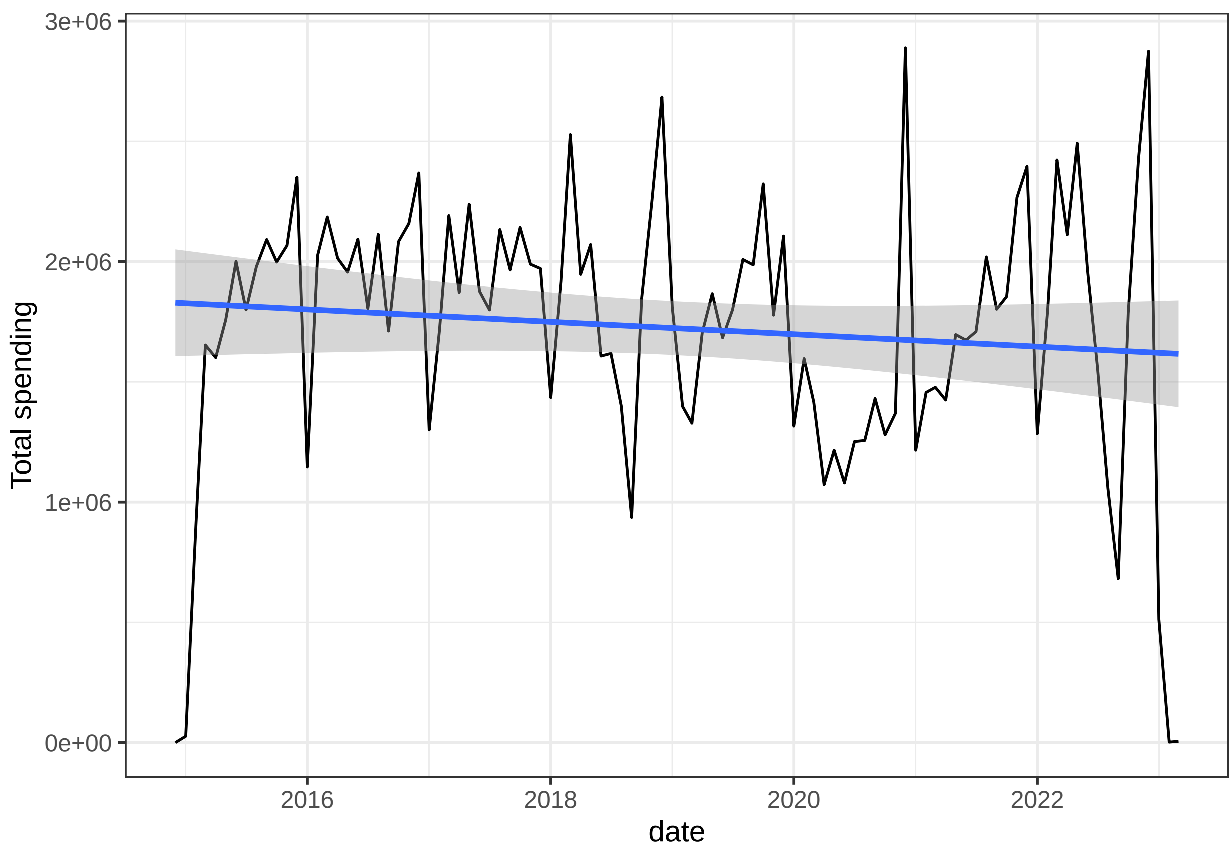

Figure 1: MPs spending per month

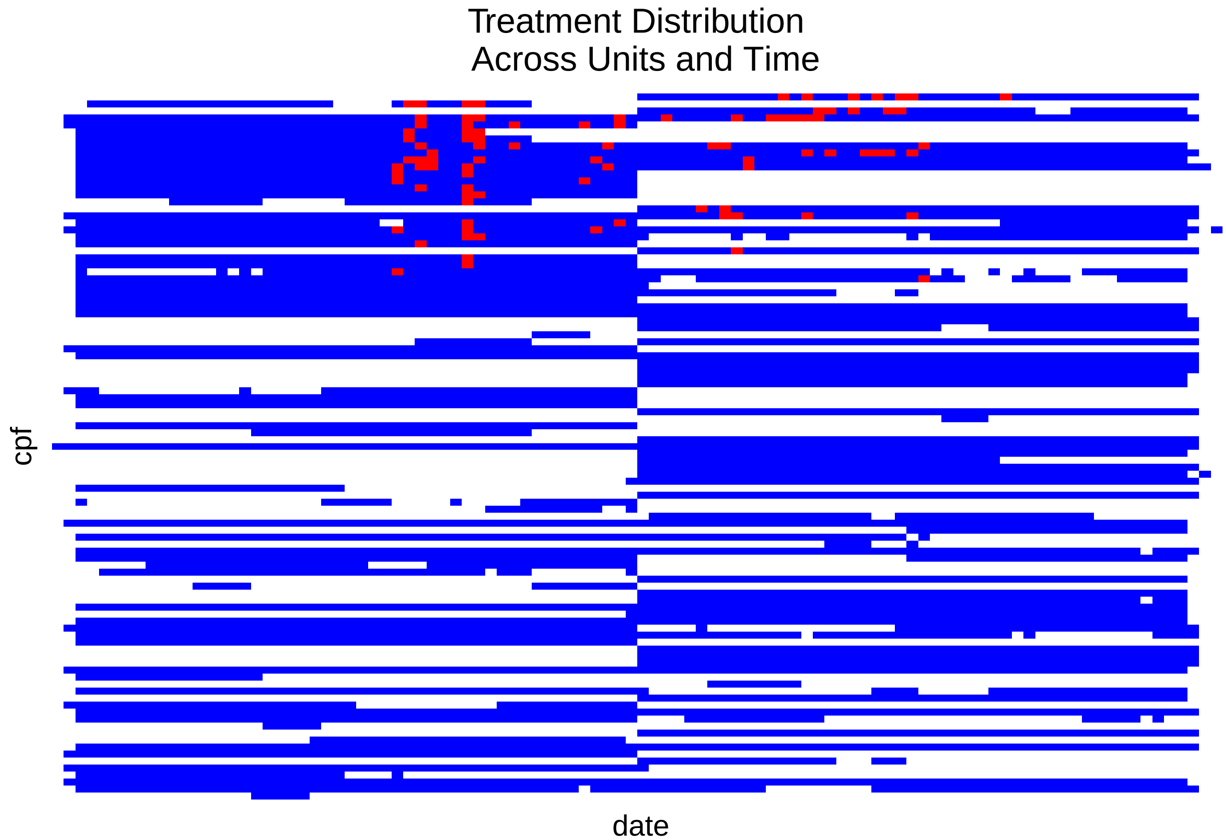

Figure 2: Treatment histories

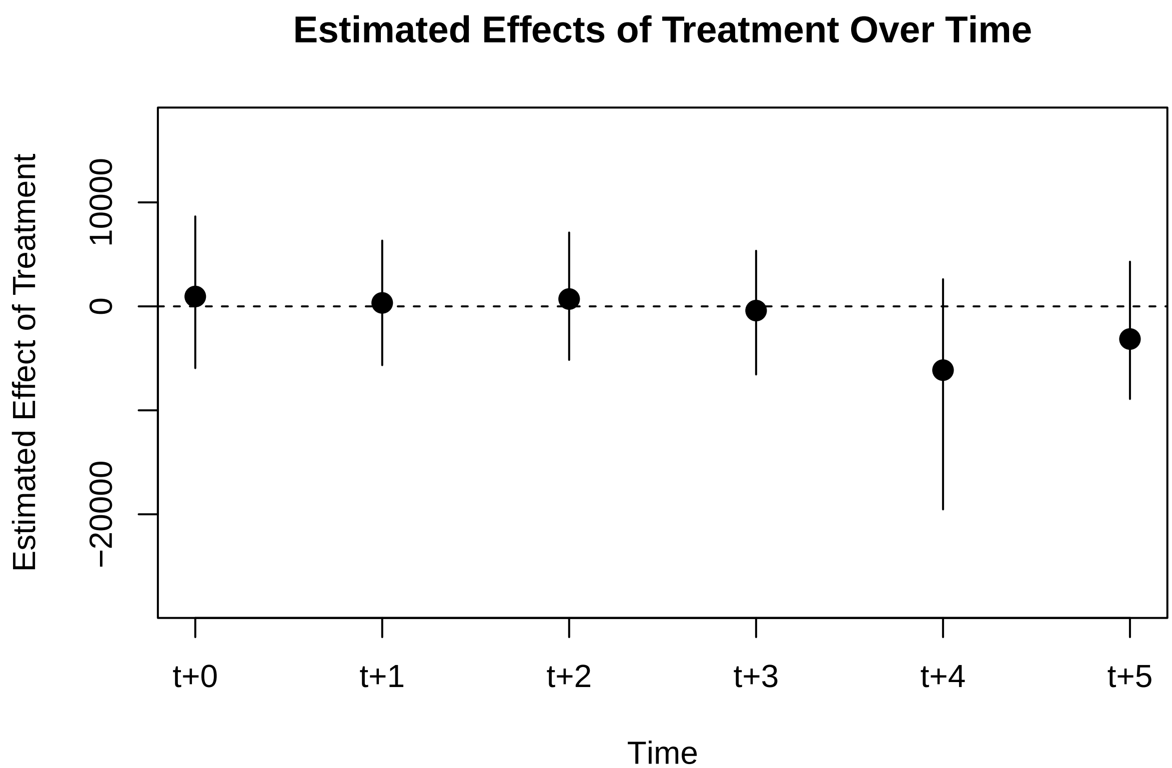

Figure 3: Effect of Rosie bot on spending by MPs

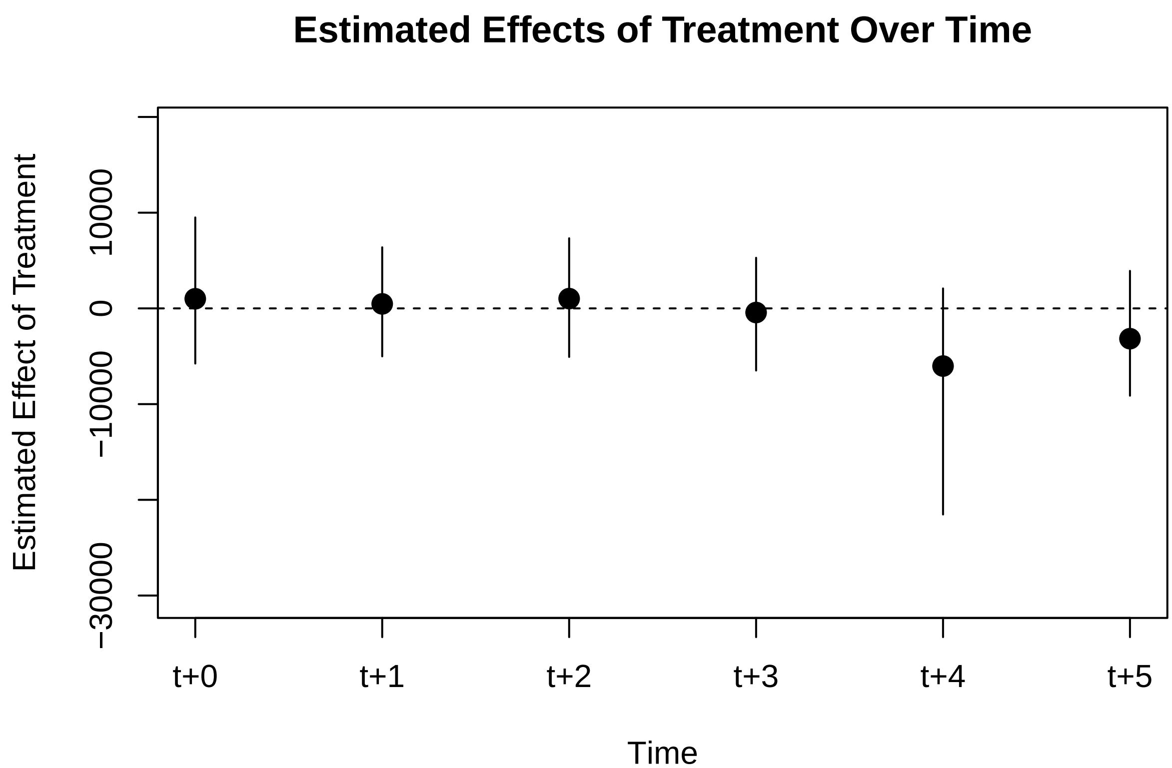

Figure 4: Effect of Rosie bot on net spending by MPs

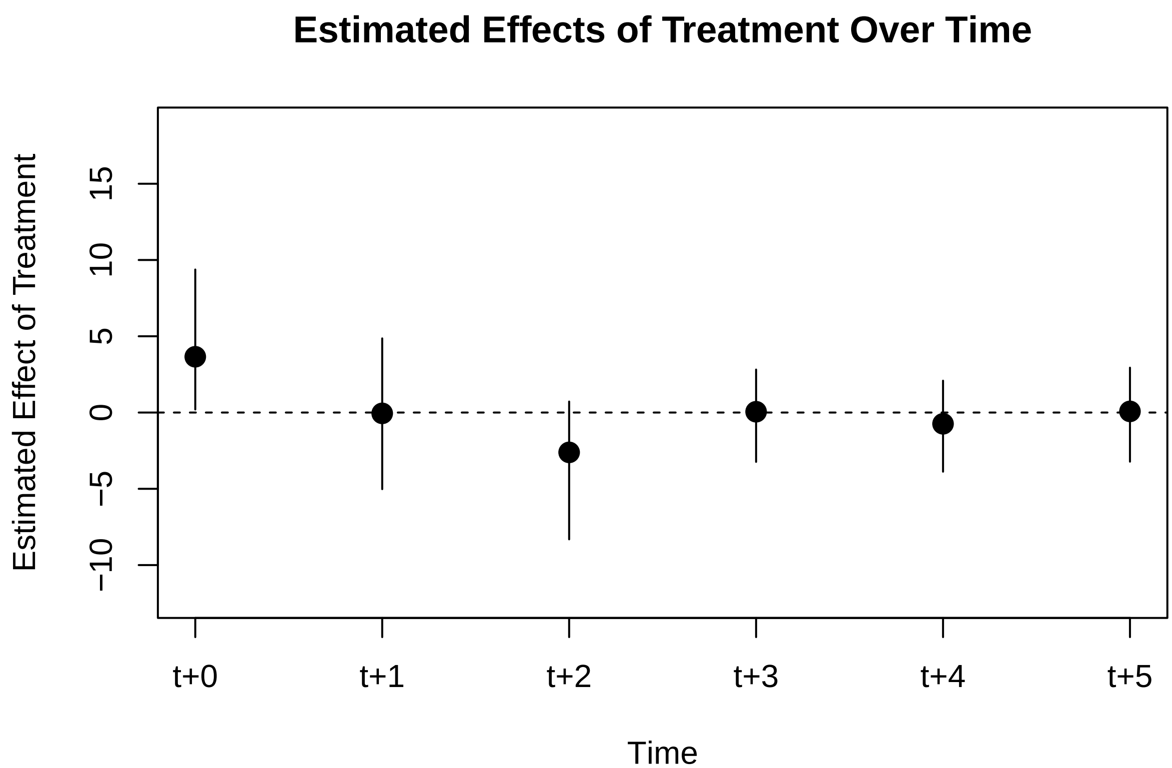

Figure 5: Effect of Rosie bot on reimbursement by MPs

What is missing?

Inclusion of all tweets

There are two data sets of tweets, the first one contains the personal identifier of MPs (cpf), the second don't. And I couldn't successfully match the tweet data set with the allowance data.

Definition of time period

The estimator is implicitly a comparison between treatment and control for one time period. And the time period is a month. But perhaps the time period should be longer (three months? 6 months?). For different time periods, we need to aggregate the data correspondingly and rerun the analysis.

We need to add covariates that are possible confoudings.

Mp's Parties. It is possible that, say, MPs from centrão spend more and are less influenced by twitter monitoring..

Government coalition. It is possible that, say, MPs from government coalition spending more and are more influenced by twitter monitoring.

Social media presence. It is possible that, say, MPs with more social media presence spending more (or less?) and are more influenced by twitter monitoring.

Next steps

Getting political data is easy and we can include in the regressions

Getting social media presence is harder, specially because ideally we should get the number of followers or number of posts or engagement from the the time window we are looking into. The simplest one is to get if the MP is on twitter at the time of monitoring. Still, it will take some time.

After getting all the data, rerun the analysis and see if it changes anything.

Include all tweets. See if it changes results.

References

Imai, Kosuke, and In Song Kim. 2021. "On the Use of Two-Way Fixed Effects Regression Models for Causal Inference with Panel Data." Political Analysis 29 (3): 405–15.