



SCHOOL
FOR ADVANCED
STUDIES
LUCCA

Data Science & Statistical Learning | II Level Master

Policy Evaluation and Impact Analysis



UNIVERSITÀ
DEGLI STUDI
FIRENZE

Is there a relationship between refugees migration flows and residents support for far-right parties?

Prof. *Armando Rungi*

Gianmarco Santoro

05/02/2024

- Between 2010 and 2020, some elections took place in Europe while experiencing a refugee crisis. Right-parties performed well in many cases.
- This work presents an analysis to **assess causal effect between** the fact that certain municipalities are more affected by **migration flows** and the inclination to vote for **far-right parties**.
- **Data** shows information about political **elections results** of the far-right Golden Dawn party in **96 greek municipalities** close to the Turkish border, for a total amount of **384 records**.
- Elections have taken place in **4 different years, the last one after a big migration flow** (treatment) from Turkey to Greece in summer 2015.
- **Variables** in the dataset:
 - *municipality*: sequential number assigned to the municipality (1, 2, ..., 96)
 - *ever_treated*: TRUE if at least one *treatment* is 1, FALSE if all *treatment* records for that municipality are 0
 - *treatment*: whether there has been a large migration flow or not in the municipality (1 or 0)
 - *year*: election year (2012, 2013, 2015, 2016)
 - *gdvote*: outcome, percentage of votes for the main far-right party, Golden Dawn (%)
 - *trarrprop*: number of refugees per person (continuous number)
 - *logdist*: distance of the municipality from the Turkish border (logarithmic)

- When data have a time dimension with a group of units receiving a treatment at a given point in time and the other group does not, **causal effects** can be estimated by implementing **diff-in-diff estimators**, using both linear regression and fixed-effects regressions, given certain assumptions.
- These methods can serve as useful alternatives when conducting a randomized control trial is not practical, as in this case, and causal inference needs to be performed from observational data.
- General basic assumptions:
 - Linear model parameters
 - Random sampling (Rank condition; full rank)
 - Exogenous generated data (which implies parallel trends assumption in diff-in-diff)
 - Conditional homoskedasticity (spherical disturbance; normal distribution of the errors)
- SUTVA assumptions for Causal Inference:
 - No interference
 - No heterogeneous treatment

- First analysis considering only values in 2016, after treatment: **linear regression model** to analyze the **relationship between** the **treatment** (dummy variable measuring 1 if treated observation and 0 if not) and the **outcome** (gdvote, the share of votes of the far-right party):

Model:

```
lm(formula = gdvote ~ treatment, data = post_treat_data)
```

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	5.6591	0.2513	22.517	< 2e-16 ***
treatment	2.7448	0.7109	3.861	0.000207 ***

Multiple R-squared: 0.1369, Adjusted R-squared: 0.1277

- The regression estimated is the **difference in means** (GD support) between:
 - Treatment group:** municipalities with large inflows
 - Control group:** municipalities without big inflows
- The difference in means is:
 - Significant:** p-value for the treatment variable 0.000207 is very low, suggesting a significant relationship.
 - Positive:** treated municipalities were on average **2.74 % more supportive** for GD than non-treated municipalities in the post-treatment period.

Considering possible **selection bias** in assignment of *treatment* among municipalities, there is insufficient justification to conclude that the observed difference in means accurately reflects the causal effect of interest.

- **Difference-in-differences** between 2015 and 2016, among treatment and control groups:

PRE-treatment period, diff in means treat vs control:

$$\text{pre_diff} = 0.6212737$$

POST-treatment period, diff in means treat vs control:

$$\text{post_diff} = 2.744838$$

Diff in Diff (post_diff - pre_diff):

$$\text{diff_in_diff} = 2.123564$$

- There is a **positive average treatment effect on the treated** municipalities:
 - Difference in means in post-treatment period 2.74 is greater than the pre-treatment period 0.62.
- Diff-in-diff implies that refugee flow **increased support** for far-right party on treated municipalities of about **2.1 % on average**.
- This **result** suggests a substantial increase in the difference between the treatment and control groups after the treatment, **indicating a potential causal effect associated with the treatment**.

- Linear regression **including interaction term** to estimate the DiD effect, considering 2015 and 2016:

Model:

```
lm(formula = gdvote ~ ever_treated * post_treat, data = muni_from_2015)
```

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	4.3854	0.2415	18.162	< 2e-16 ***
ever_treatedTRUE	0.6213	0.6829	0.910	0.364147
post_treatTRUE	1.2737	0.3415	3.730	0.000253 ***
ever_treatedTRUE:post_treatTRUE	2.1236	0.9658	2.199	0.029120 *

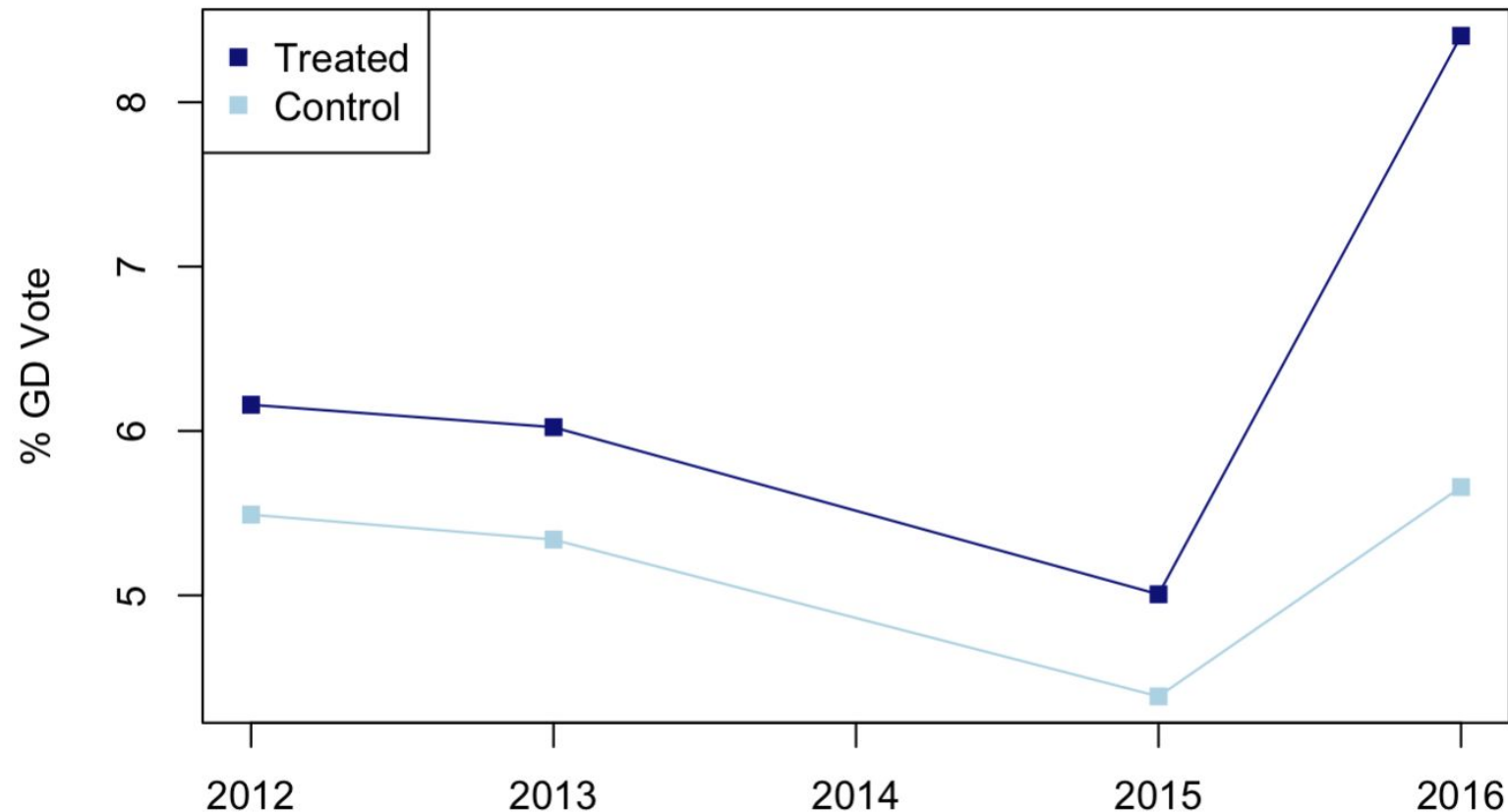
Multiple R-squared: 0.1762, Adjusted R-squared: 0.163

- Interaction term helps estimating whether the treatment effect varies between the pre vs post-treatment periods.
- post_treat variable consider only 2016.**
- post_treat has p-values 0.000253, suggesting its effects on gdvote is statistically significant.
- ever_treatedTRUE:post_treatTRUE, si significant but with low confidence.
- ever_treatedTRUE:post_treatTRUE** coefficient is 2.1236, representing the diff-in-diff estimate, indicating the **average additional change** in gdvote for treated municipalities during in post-treatment period.

This analysis shows a potential causal impact of the treatment on the outcome.

- **Verify parallel trends assumption:** without treatment, treated and untreated units would evolve similarly in the outcome variable, so selection bias between the two unit groups is stable over time:

Trends



- Assess the assumption by plotting the mean value of the dependent variable by year, among the two groups.
- The plot shows **parallel** trends during the 3 pre-treatment periods, **confirming and verifying the assumption**.
- The far-right vote share then **diverges** in the post-treatment period, **supporting the causal effect** of migrant flows in far-right vote shares.

- **Fixed-effects** (within-group) **regression** to estimate the diff-in-diff, considering *municipality*, *year* and *treatment*:

Model:

```
lm(formula = gdvote ~ as.factor(municipality) + as.factor(year) + treatment,
    data = muni)
```

```
summary(fixed_eff_model)$coefficients['treatment']:
```

Estimate	Std. Error	t value	Pr(> t)
2.087002e+00	3.932824e-01	5.306624e+00	2.252636e-07

- Fixed-effects regression is used in diff-in-diff analysis when there might be individual or group-specific characteristics that are constant over time (time invariant heterogeneity) and may influence both the treatment assignment and the outcome.
- The *treatment* coefficient is the most relevant in this model, which represents the diff-in-diff, given the fixed-effects.
- In this model, the **average treatment effects on the treated is 2.087**, confirming the same causal effect.
- This value is a little different from the ones calculated before, probably because it considers all the periods before treatment and not only 2015.

- Fixed-effects regression to estimate the diff-in-diff, considering *municipality*, *year* and *trarrprop*:

Model:

```
lm(formula = gdvote ~ as.factor(municipality) + as.factor(year) + trarrprop,  
    data = muni)
```

```
summary(fixed_eff_model_Prop)$coefficients['trarrprop',]:
```

Estimate	Std. Error	t value	Pr(> t)
6.061088e-01	1.324367e-01	4.576591e+00	7.082297e-06

- Similarly to the analysis before, the *trarrprop* coefficient shows that for every refugee per capita increase, the far-right party vote grows of 0.6 % on average.

- **Conclusions:**

- It is expected that nothing related with treatment is left outside of the model. In this case, considering the possible effect of other external variables (**possible endogeneity**) as economical and financial ones, the **2.1 % increase in far-right party vote shares** can be considered as a mild effect.
- In this context, it can be underlined that the **only exposure**, without any other implication for residents, **led to vote more for far-right** parties.

- **Considerations:**

- **External validity:** it is difficult to assess the generalizability of the results in other countries and contexts. Given this assumption the result can suggest some validity in context with a similar cultural, demographic, economic situation, like Italy or Spain. In any case, this consideration should be verified.
- **Further studies:**
 - It can be further analyzed if the proximity to the Turkish border has some influence on the outcome.
 - It could be also possible to use other tools, such as instrumental variables, for example using the distance of the municipalities to the Turkish border. The idea is that municipalities closer to the border may experience higher migration flows.



SCHOOL
FOR ADVANCED
STUDIES
LUCCA



UNIVERSITÀ
DEGLI STUDI
FIRENZE



<https://ethicaljournalismnetwork.org/media-and-migration>

Thanks for Attention