

Dr. Christian Czymara

FORSCHUNGSPRAKTIKUM I UND II: LÄNGSSCHNITTDATENANALYSE IN R

Natural experiments
session x

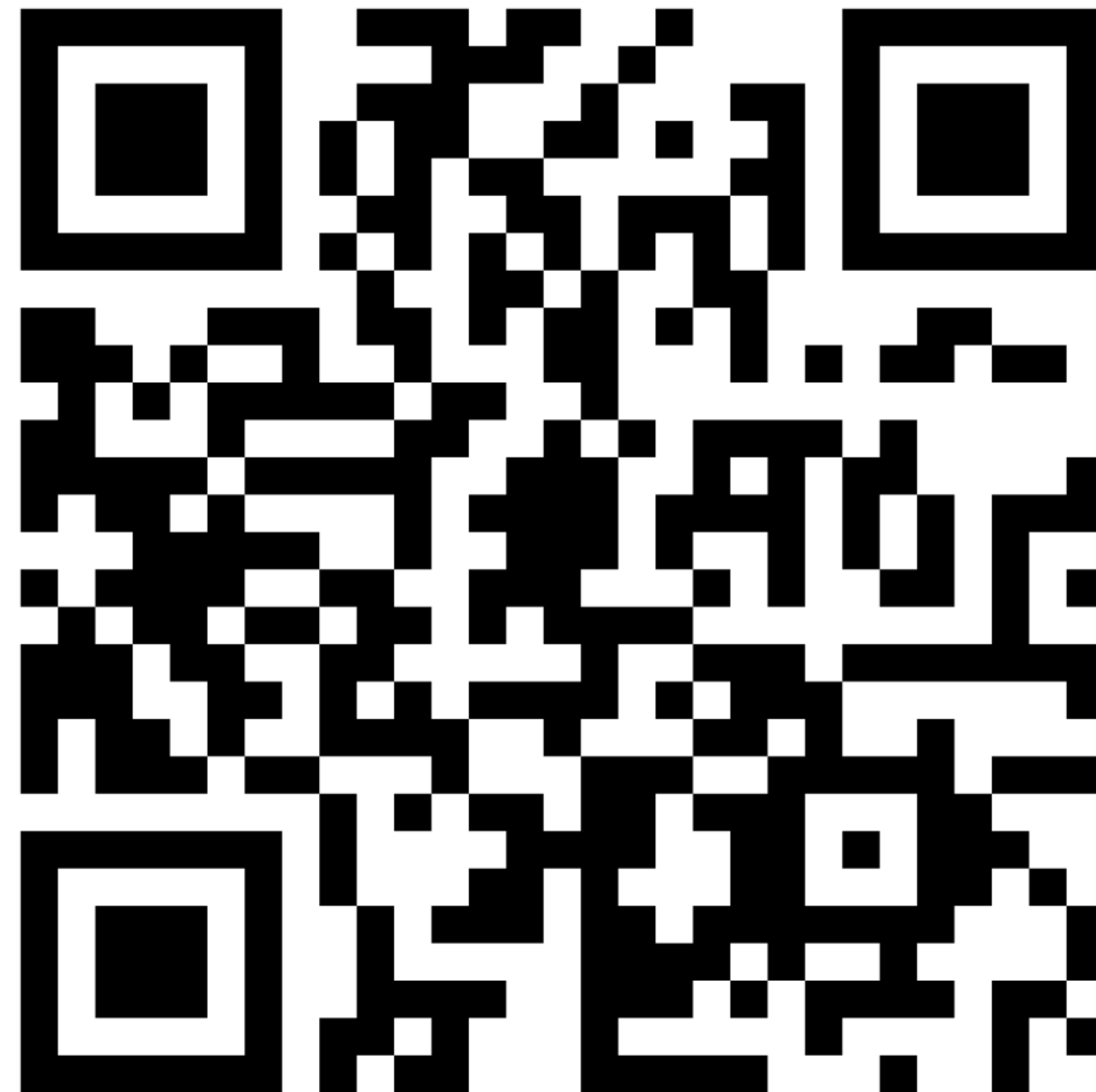
AGENDA

- Evaluation
- So far, we attributed causal inference mainly to intra-individual comparisons
- Today: Causality based on comparisons between respondents

EVALUATION

LINK

- <http://r.sd.uni-frankfurt.de/7c144c9b>



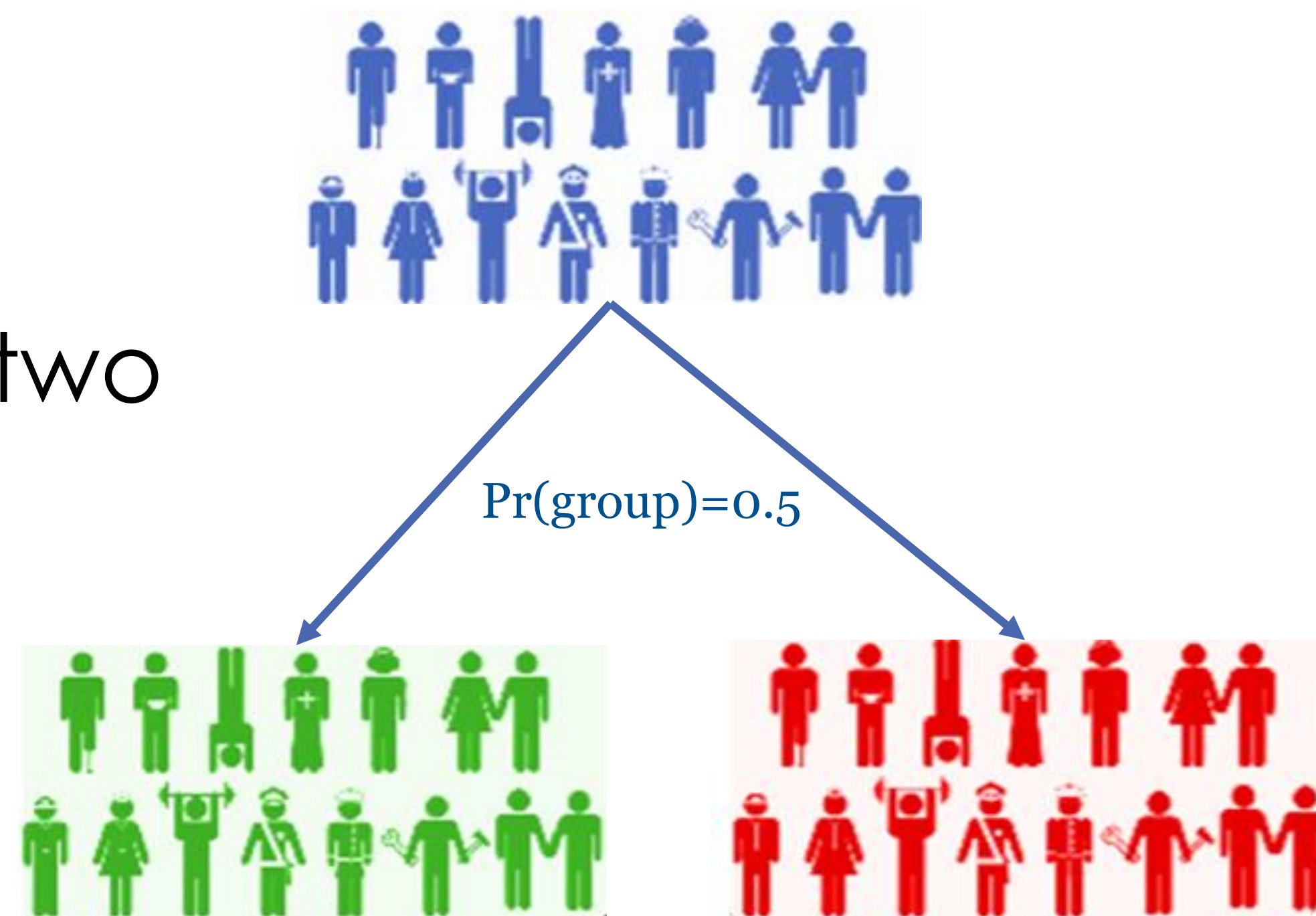
THE GOLD STANDARD: RANDOMIZED EXPERIMENTS

EXPERIMENTS

- We want to know if x causes y in a population

- Split the population in two

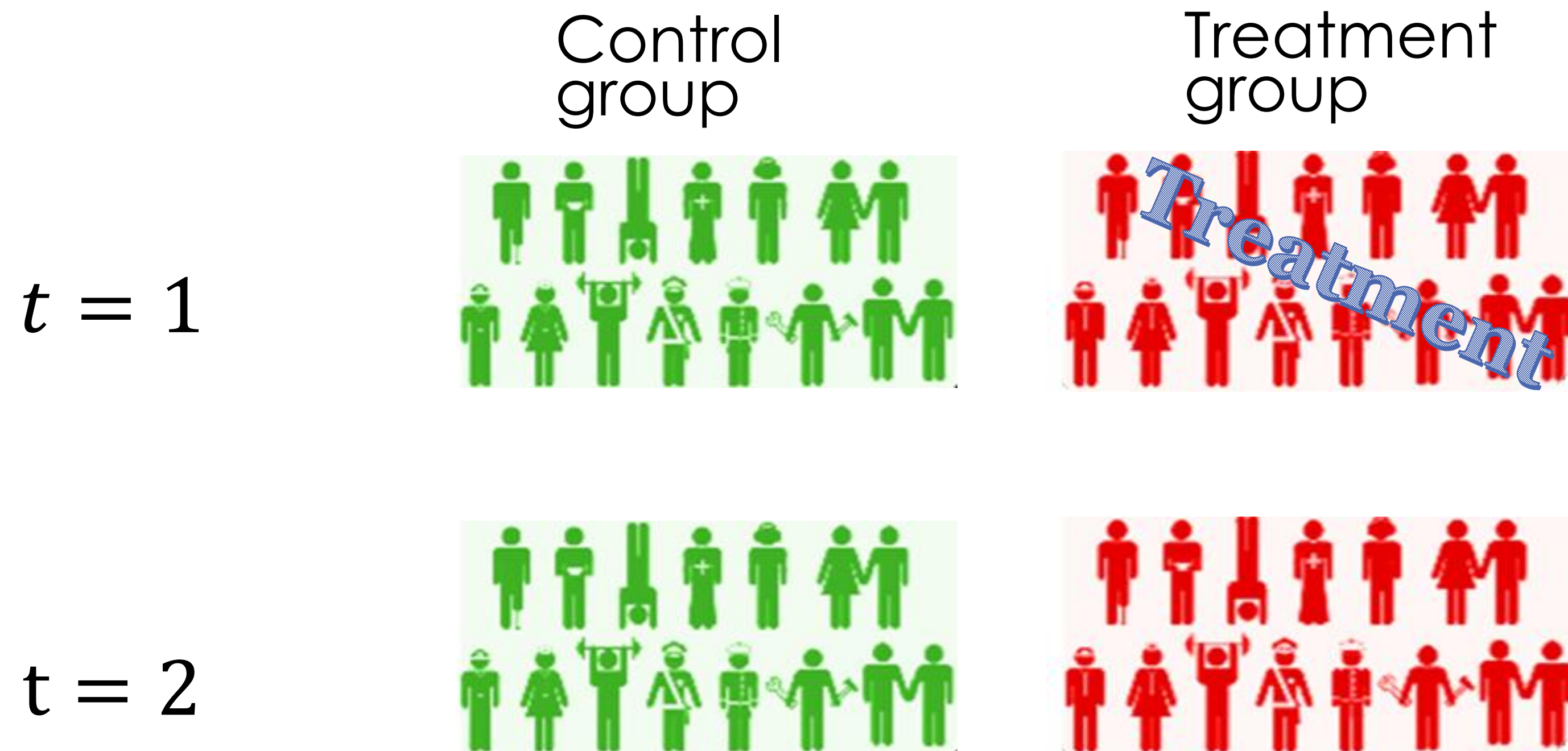
- ... By *randomization*!



RANDOMIZATION

- Assignment into group is random
- $\Pr(\text{Group}_1) = \Pr(\text{Group}_2) = 0.5$
- If assignment into group is perfectly random, all covariates are approximately balanced between both groups
- Differences between group should be small and just by chance

TREATMENT ASSIGNMENT



→ $y?$

TREATMENT EFFECT

- Calculate difference for treatment group: $\bar{y}_{2,T} - \bar{y}_{1,T}$
- But...
 - Might this difference be a placebo effect?
 - If large gap between $t = 1$ and $t = 2$: general time trend?
- We need a second comparison
- Ideally: *what if*-situation → What would be the value of $\bar{y}_{2,T}$ if people in treatment group hadn't actually received the treatment? (*counterfactual scenario*)
- Calculate difference for control group: $\bar{y}_{2,c} - \bar{y}_{1,c}$
- Effect is given by so-called difference-in-differences (DiD) equation:
 $(\bar{y}_{2,T} - \bar{y}_{1,T}) - (\bar{y}_{2,c} - \bar{y}_{1,c})$
- Average treatment effect (ATE)

TREATMENT EFFECT

- Such an effect is causal because all disturbing third variables are equally distributed across treatment and control group
- Put differently, both groups differ in nothing but treatment status
- Hence, differences in y must be due to treatment

RANDOMIZATION IN SOCIAL SCIENCE

- Randomized experiments are a nice ideal and feasible in, e. g., some clinical studies
- But rarely in social sciences
- In very little cases, researchers can randomly assign people into two groups regarding a social characteristic
- E.g.: Who gets into higher education / receives higher income / lives in homogeneous neighborhoods / has an authoritarian personality / ... is far from random
- For most research questions, social scientists must *observe* individuals as they are
- How can we draw valid conclusion regarding causal processes from that?

CAUSALITY WITH OBSERVATIONAL DATA

- Most popular idea is statistical control: If we have measured relevant aspects, we can hold them constant a posteriori using regression models
- However, rarely the case that everything is measured
- Our second approach: Don't compare different individuals, but changes within individuals
- Each person acts “as their own control” (at a different time point)
- Automatically accounts for everything that is constant over time
- However, many (relevant) things change over time
- In this case, within estimates still biased

NATURAL EXPERIMENTS

- So-called “natural” experiments mimic randomized trial
- Treatment assignment is not manipulated by researcher, yet (ideally) random
- Exposure to treatment determined by “nature”
- Benefit: Allows to examine aspects that one could not directly manipulate
- Idea: Find something that is random and cuts population into two parts
- Compare observations above and below the threshold

EXAMPLES

DINAS ET AL. (2019). WAKING UP THE GOLDEN DAWN: DOES EXPOSURE TO THE REFUGEE CRISIS INCREASE SUPPORT FOR EXTREME-RIGHT PARTIES? *POLITICAL ANALYSIS*.

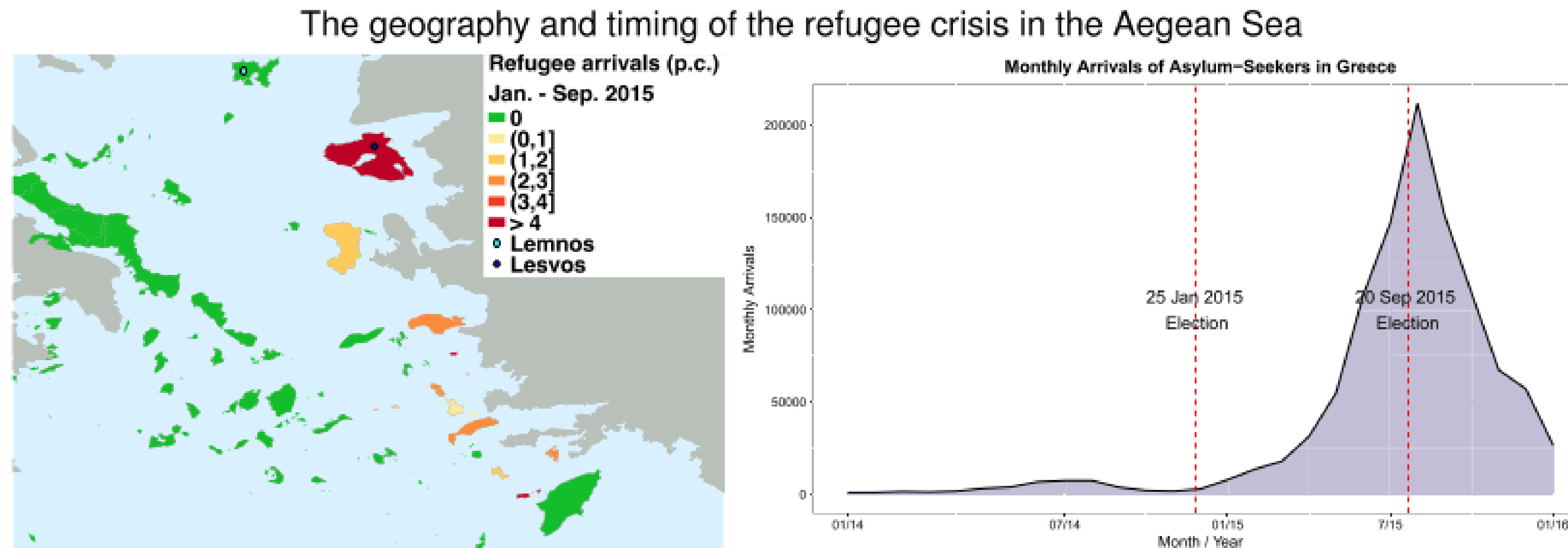


Figure 1. Panel A shows that islands close to the Turkish border received the most refugee arrivals per capita. Panel B shows the monthly number of asylum seekers arriving at Greek islands over the period from January 2014 to March 2016. During the study period, the first election took place just before the onset of the refugee crisis on January 25, 2015. A second election took place at the height of the refugee crisis on September 20, 2015.

DINAS ET AL. (2019). WAKING UP THE GOLDEN DAWN: DOES EXPOSURE TO THE REFUGEE CRISIS INCREASE SUPPORT FOR EXTREME-RIGHT PARTIES? *POLITICAL ANALYSIS*.

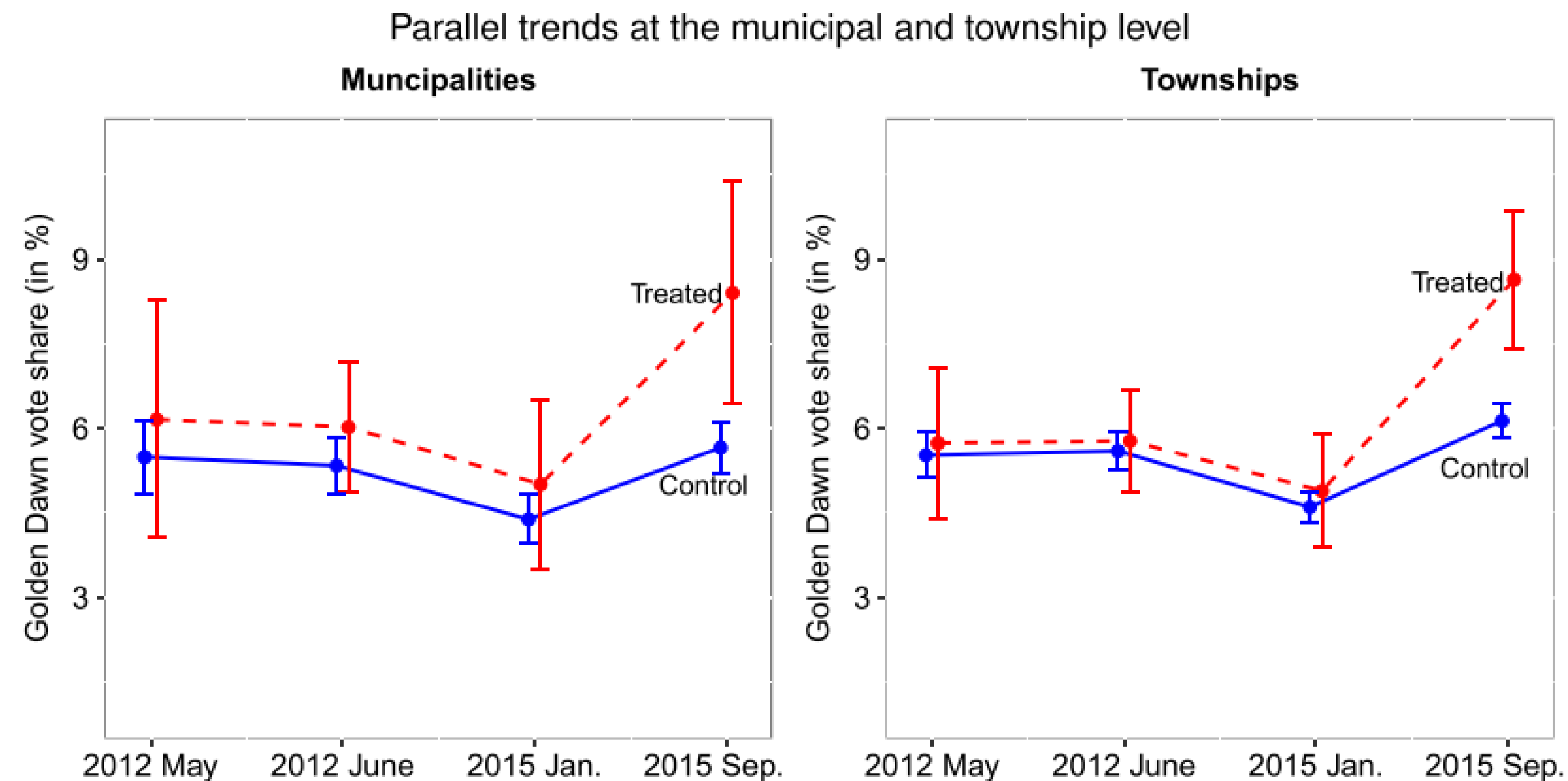
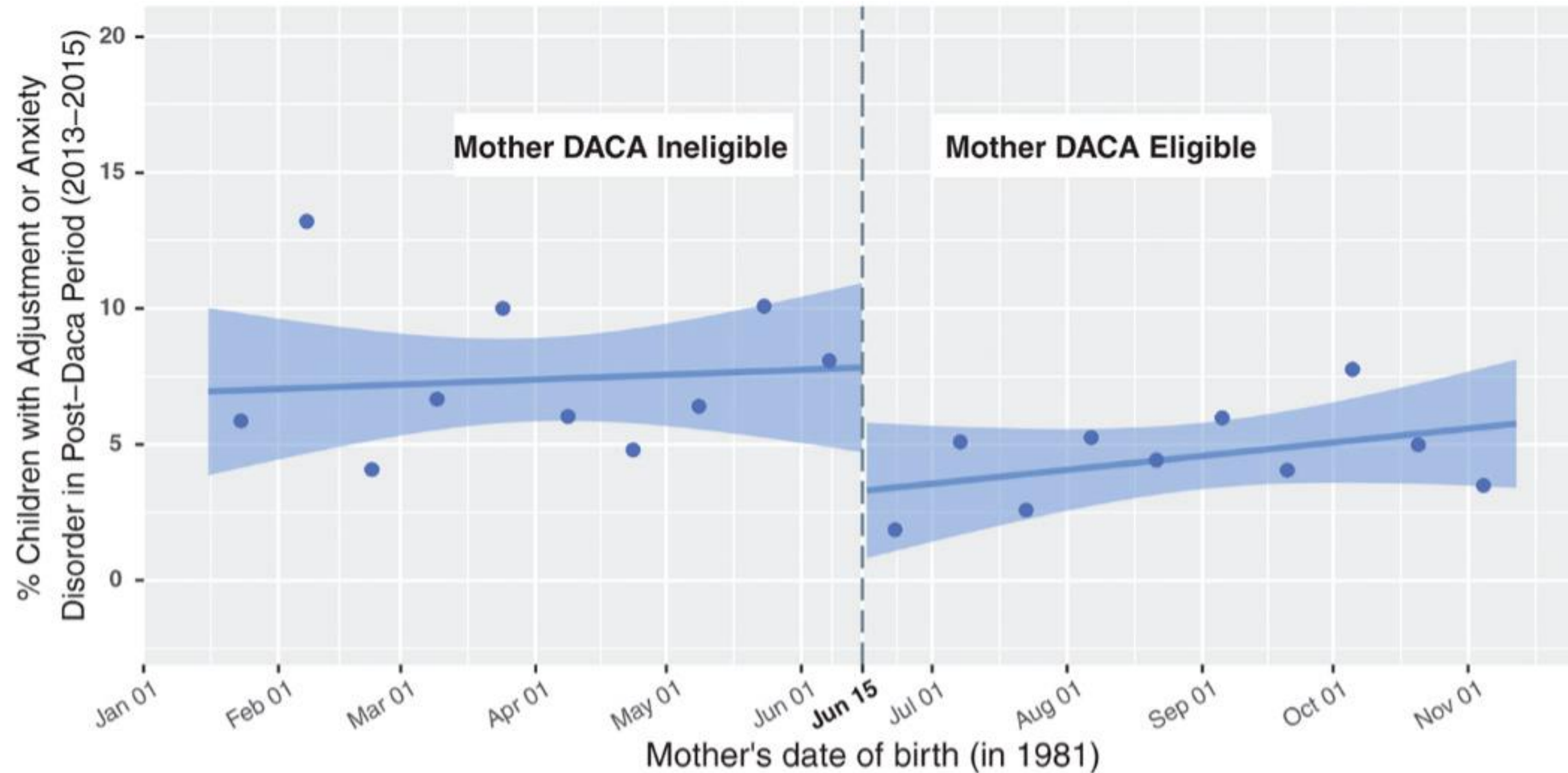


Figure 2. The analyses at the municipality (left panel) and township level (right panel) show that treated and control islands experience highly similar changes in support for GD prior to the refugee crisis, thereby strengthening our confidence in the parallel trend assumption. The blue connected line indicates the average vote share for GD in the municipalities (left panel) and townships (right panel) that received refugees. The red dashed line denotes the average GD vote share in municipalities and townships without refugee exposure.

HAINMUELLER ET AL. (2017). PROTECTING UNAUTHORIZED IMMIGRANT MOTHERS IMPROVES THEIR CHILDREN'S MENTAL HEALTH. *SCIENCE*.



HAINMUELLER, HANGARTNER, & WARD (2019). THE EFFECT OF CITIZENSHIP ON THE LONG-TERM EARNINGS OF MARGINALIZED IMMIGRANTS: QUASI-EXPERIMENTAL EVIDENCE FROM SWITZERLAND. *SCIENCE ADVANCES*.

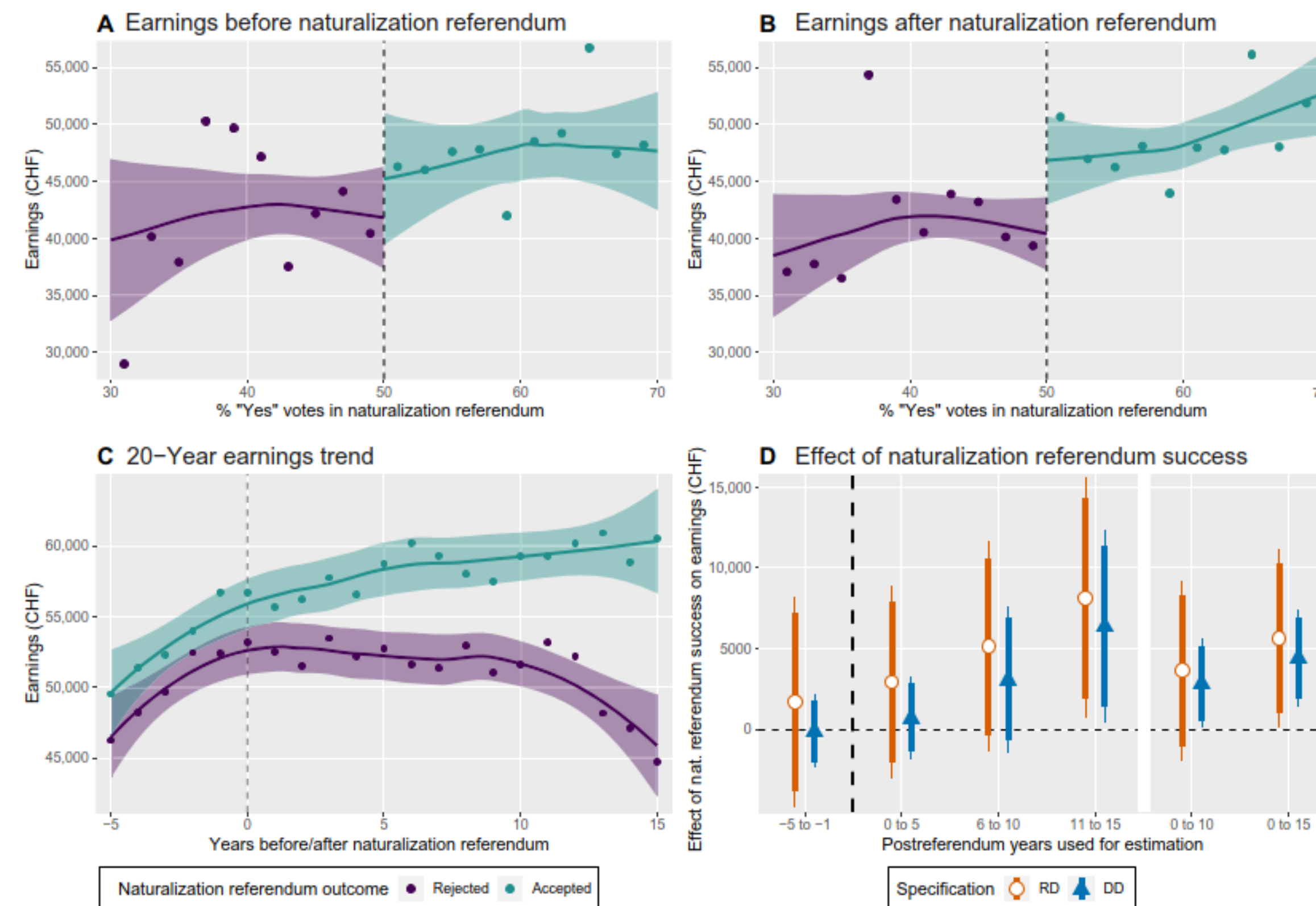
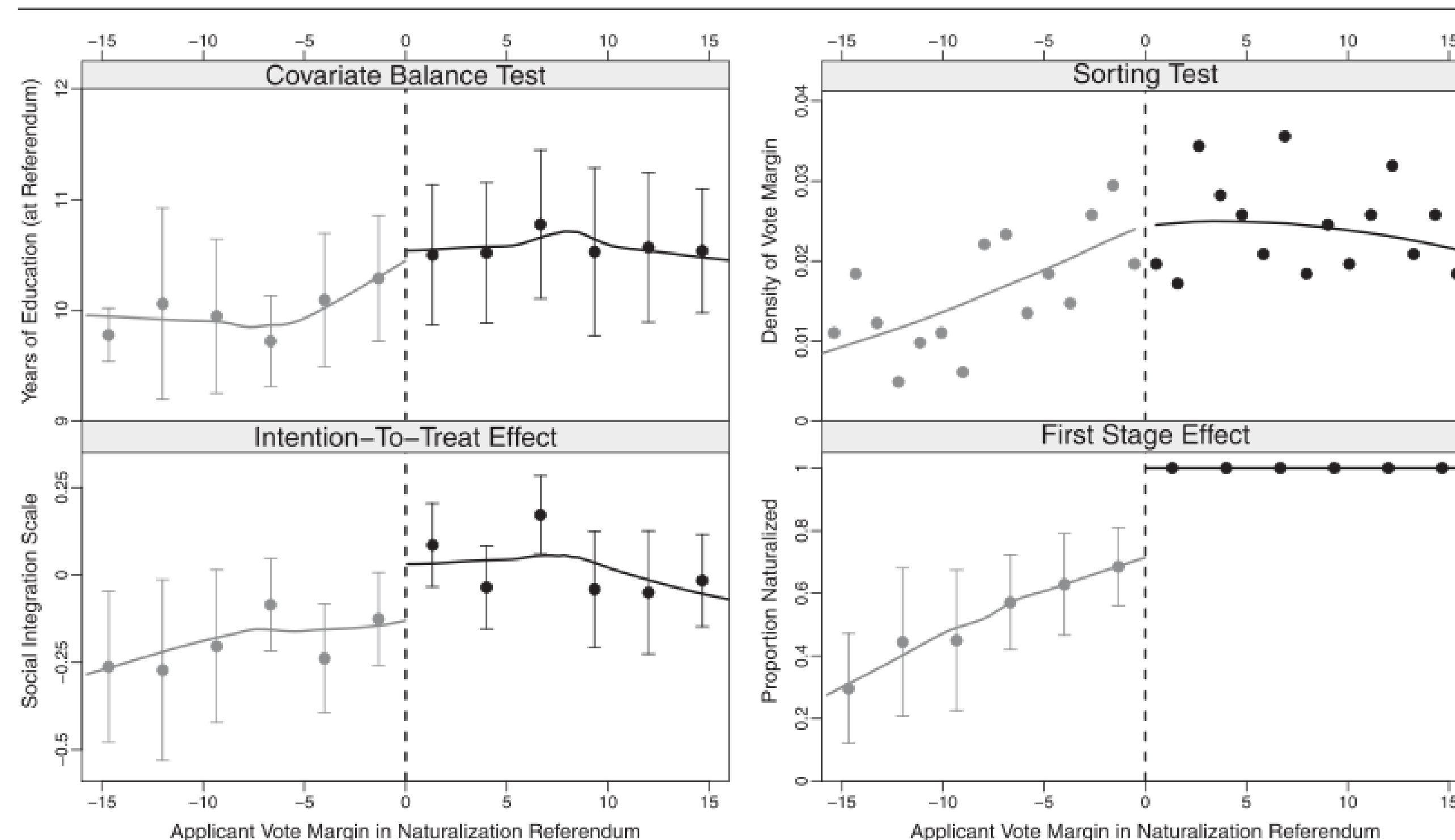


Fig. 1. Winning citizenship in the referendum increases immigrant earnings in the long term. (A) Placebo test shows no significant differences in earnings between immigrants who barely won or lost their citizenship referendum in the last 5 years before the referendum ($n = 1337$). (B) RD estimate shows sizable and significant differences in earnings between immigrants who barely won or lost their citizenship referendum in the years after the referendum ($n = 2262$). (C) The 20-year earnings trends spanning the time period before and after the naturalization referendum shows an increasing earnings gap between immigrants who won or lost their referendum (applicants in 40 to 60% yes-vote range; $n = 10,731$). (A) to (C) show loess smoother and 95% confidence intervals. (D) Point estimates for the RD and DD regressions and 90% (thick line) and 95% (thin line) confidence intervals.

HAINMUELLER, HANGARTNER, & PIETRANTUONO (2017). CATALYST OR CROWN: DOES NATURALIZATION PROMOTE THE LONG-TERM SOCIAL INTEGRATION OF IMMIGRANTS?. *AMERICAN POLITICAL SCIENCE REVIEW*.

FIGURE 3. Fuzzy Regression Discontinuity Design: Identification Checks and the Effect of Naturalization on Long-Term Social Integration



Notes: Upper left panel indicates that years of education (a pretreatment covariate) is balanced at the victory threshold in the naturalization referendums. Upper right panel shows that the density is smooth across the victory threshold, suggesting that there is no evidence of sorting. Lower left and right panels show that long-term social integration and the probability of naturalization sharply increase when comparing applicants who barely won and barely lost their first naturalization referendum. (Loess lines; 95% confidence intervals for binned averages.)

HAINMUELLER, HANGARTNER & PIETRANTUONO (2015).
NATURALIZATION FOSTERS THE LONG-TERM POLITICAL INTEGRATION
OF IMMIGRANTS. *PROCEEDINGS OF THE NATIONAL ACADEMY OF
SCIENCES*.

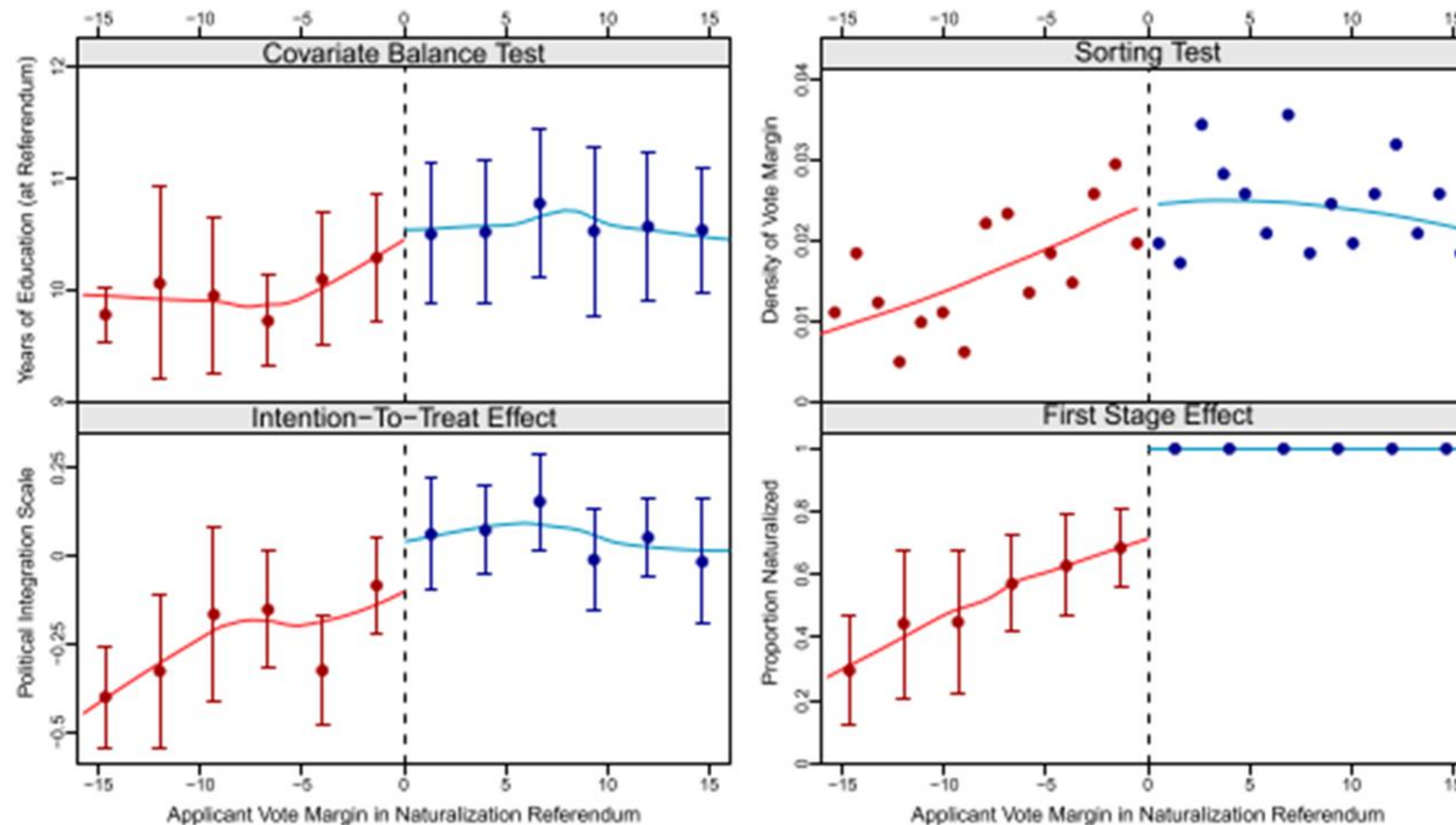


Fig. 1. Fuzzy RD design: identification checks and the effect of naturalization on political integration. *Upper Left* shows that the applicants' (pretreatment) years of education are well balanced at the vote threshold for winning the naturalization referendum. *Upper Right* shows that there is no discontinuity in the density of the vote margin variable, indicating that applicants are not sorting around the threshold of winning. *Lower Left* and *Lower Right* show that barely winning versus barely losing the referendum increased levels of political integration and the probability of naturalization, respectively. Loess lines; 95% confidence intervals for binned averages (dots).

LIPPMANN & SENIK (2018). MATH, GIRLS AND SOCIALISM. *JOURNAL OF COMPARATIVE ECONOMICS*.

A B S T R A C T

This paper argues that the socialist episode in East Germany, which constituted a radical experiment in gender equality in the labor market and other instances, has left persistent tracks on gender norms. We focus on one of the most resilient and pervasive gender gaps in modern societies: mathematics. Using the German division as a natural experiment, we show that the underperformance of girls in math is sharply reduced in the regions of the former GDR, in contrast with those of the former FRG. We show that this East–West difference is due to girls’ attitudes, confidence and competitiveness in math, and not to other confounding factors, such as the difference in economic conditions or teaching styles across the former political border. We also provide illustrative evidence that the gender gap in math is smaller in European countries that used to be part of the Soviet bloc, as opposed to the rest of Europe. The lesson is twofold: (1) a large part of the pervasive gender gap in math is due to social stereotypes; (2) institutions can durably modify these stereotypes.

UNEXPECTED EVENT DURING SURVEY DESIGN

UNEXPECTED EVENT AS TREATMENTS

- Sometimes, unpredicted events coincide with the fieldwork period of surveys
- This resembles a natural experiment, where cutoff point is the event
 - Treatment group: Interviews after event
 - Control group: Interviews before event
- Assuming the interview timing does not depend on the event, assignment to treatment and control group is random
- Comparing outcome between groups gives causal effect of event
- Terror attacks are such events, since they are naturally unpredictable for the public

ASSUMPTIONS: EXCLUDABILITY

- *Differences between respondents interviewed before and after only due to event*
 1. No collateral events: Reactions only triggered by the event itself (e. g. policy changes due to terror attack → might be considered an indirect effect of the event)
 2. No simultaneous events: No other (relevant) events happening at the same time
 3. No unrelated time trend: E.g. humans tend to become more negative in fall (Harmatz et al., 2000)
 4. Exogenous event timing: Events such as terror attacks involve human planning and a human decision of timing

ASSUMPTIONS: IGNORABILITY

- *Interview date as good as random*
- 1. Fieldwork organization: Sampling designs might be clustered or stratified or include quota sampling → If design is known, it could be controlled (leading to *conditional ignorability*)
- 2. Reachability: E. g. urban population might be easier to reach than rural one (or vice versa)
- 3. Attrition: Nonresponse might be related to attack (e.g. willingness to respond might increase after attack)

NONCOMPLIANCE

- *Assignment to treatment group corresponds to actually receiving the treatment*
 1. Heterogeneous effects: Effect might only be relevant for certain subgroup of population
 2. Posttreatment bias: Control variables for treatment group always measured after event, for control group always before

ARE ASSUMPTIONS PLAUSIBLE?

- Researchers must be informed about
 - The event: Nature, timing, salience, development etc.
 - Process of data collection: Sampling procedure, how were respondents contacted? How did the fieldwork develop?
- In addition: robustness checks

ROBUSTNESS CHECKS FOR IGNORABILITY ASSUMPTION

- Compare socio-demographics in treatment and control group (balance test)
- Try different before- / after-time range:
 - In shorter time frames, fewer “other” things happen
 - But also fewer observations (less statistical power)
 - Moreover, some events might take time to unfold
- Statistical control of variables: Be careful, variables that might be influenced by event can introduce *post-treatment bias*
- Attrition: Differences in the share of “don’t know” answers between treatment and control group?

ROBUSTNESS CHECKS FOR EXCLUDABILITY ASSUMPTION

- Preexisting time trends: Test control group (those interviewed before event) for effect at its median (hoping there is none)
- Placebo test 1: Test for differences at same date in prior survey wave (or other country, if applicable)
- Placebo test 2: Test for effect on different outcomes (ideally: close to y but unrelated to event)

EXAMPLES OF UNEXPECTED EVENTS

FINSERAAS & LISTHAUG (2013). IT CAN HAPPEN HERE: THE IMPACT OF THE MUMBAI TERROR ATTACKS ON PUBLIC OPINION IN WESTERN EUROPE. *PUBLIC CHOICE*.

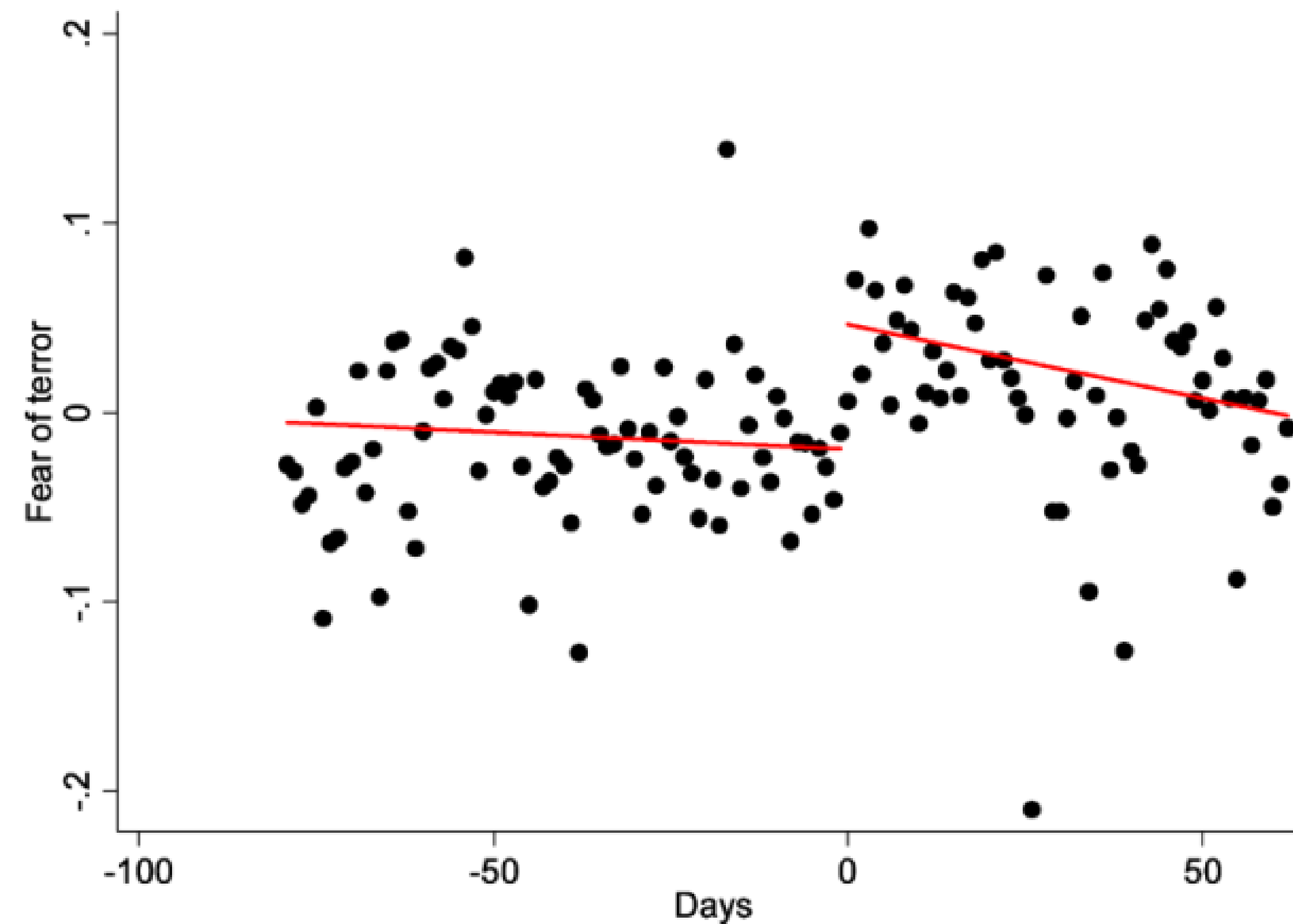
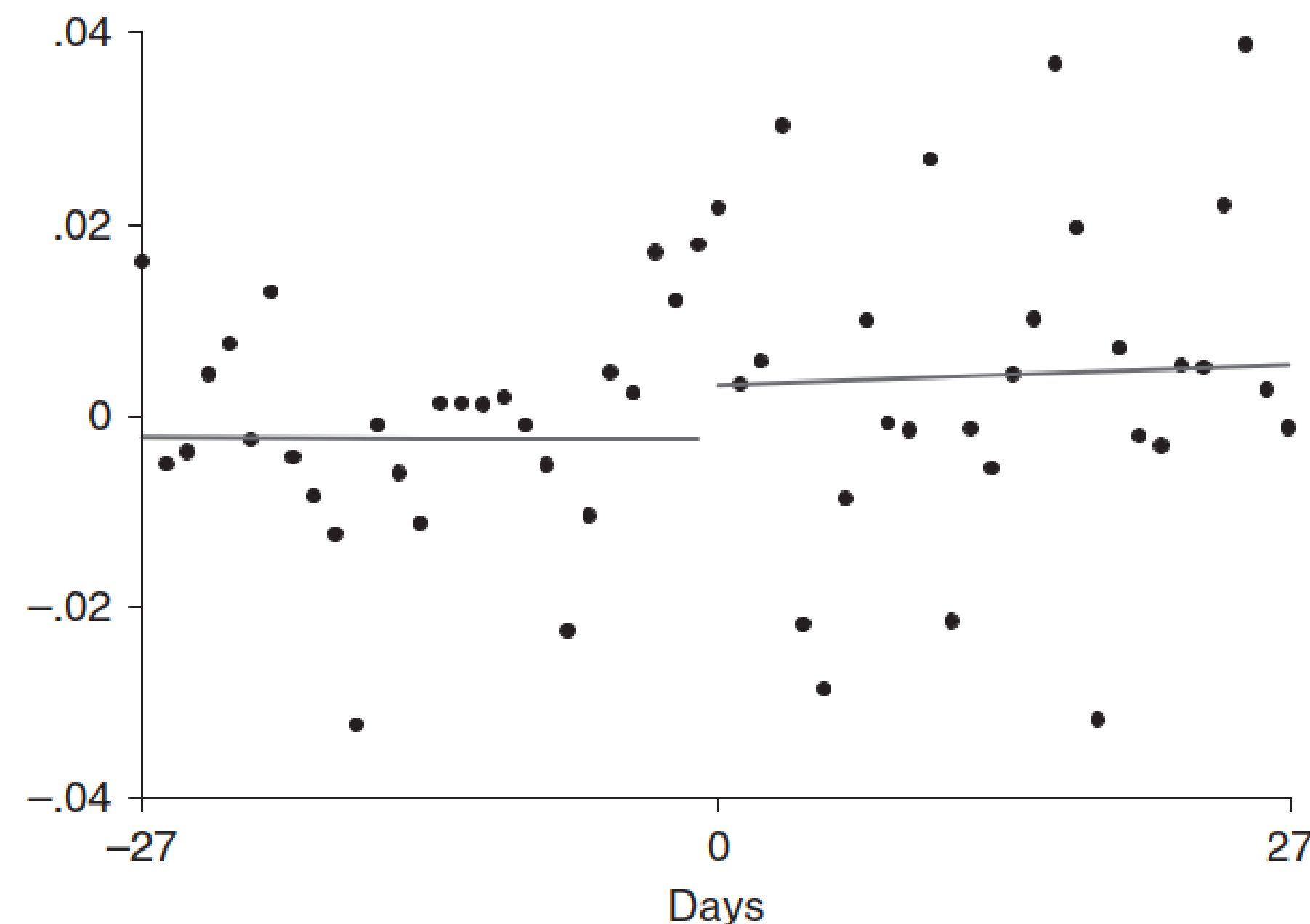


Fig. 1 Mean score on fear of terrorism by days

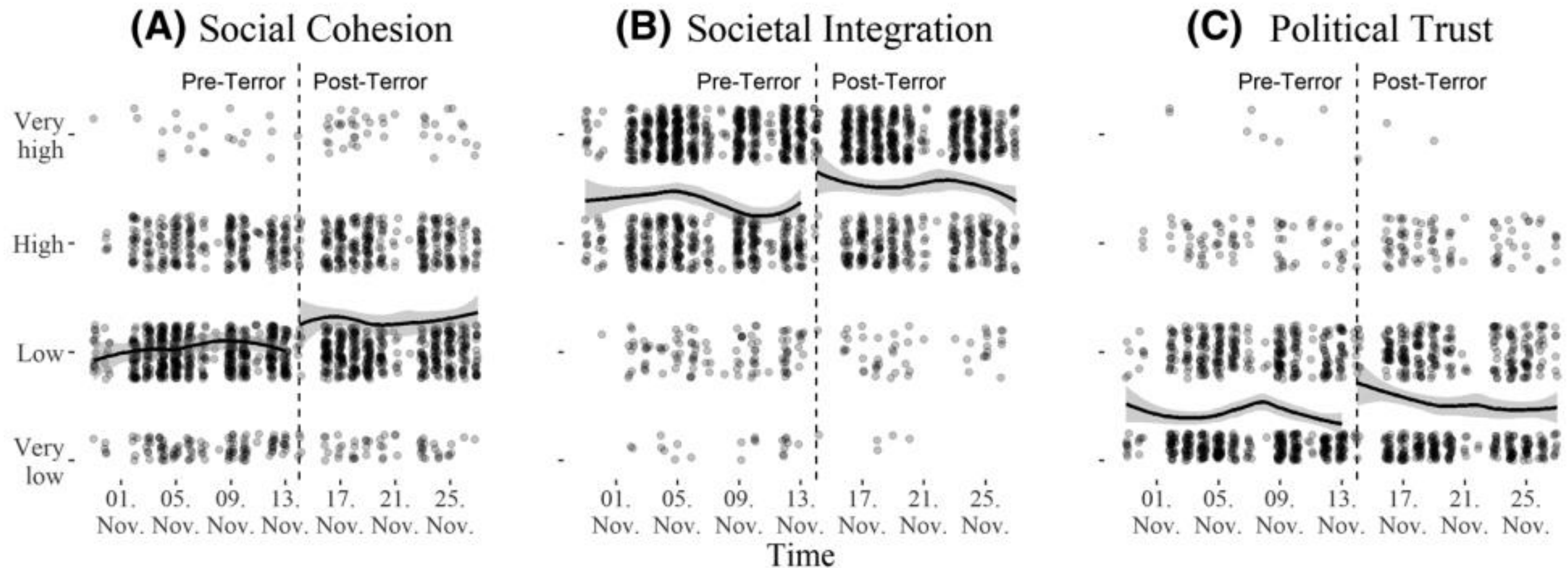
FINSERAAS, JAKOBSSON & KOTSADAM (2011). DID THE MURDER OF THEO VAN GOGH CHANGE EUROPEANS' IMMIGRATION POLICY PREFERENCES? *KYKLOS*.

Figure 1

Mean scores on support for restrictive immigration policy by days. Note: The X-axis refers to the date of the interview where 0 refers to the day after the murder, 27 refers to 28 days after the murder, and -27 refers to 27 days before the murder. The Y-axis is the average of the country-specific mean scores on support for restrictive immigration policy. The mean scores are adjusted for country-specific differences by subtracting the country mean. The dots in the figure are the average of the country-specific mean scores by days, i.e. the figure displays how policy preference changes over the period we study. The lines are OLS regression lines where the line to the left is estimated on the data before the murder, while the line to the right is estimated on the data after the murder.



VAN HAUWAERT & HUBER (2020). IN-GROUP SOLIDARITY OR
OUT-GROUP HOSTILITY IN RESPONSE TO TERRORISM IN FRANCE?
EVIDENCE FROM A REGRESSION DISCONTINUITY DESIGN. *EUROPEAN
JOURNAL OF POLITICAL RESEARCH*.



HANSEN ET AL. (2016). INCREASED INCIDENCE RATE OF TRAUMA-AND STRESSOR-RELATED DISORDERS IN DENMARK AFTER THE SEPTEMBER 11, 2001, TERRORIST ATTACKS IN THE UNITED STATES. *AMERICAN JOURNAL OF EPIDEMIOLOGY*.

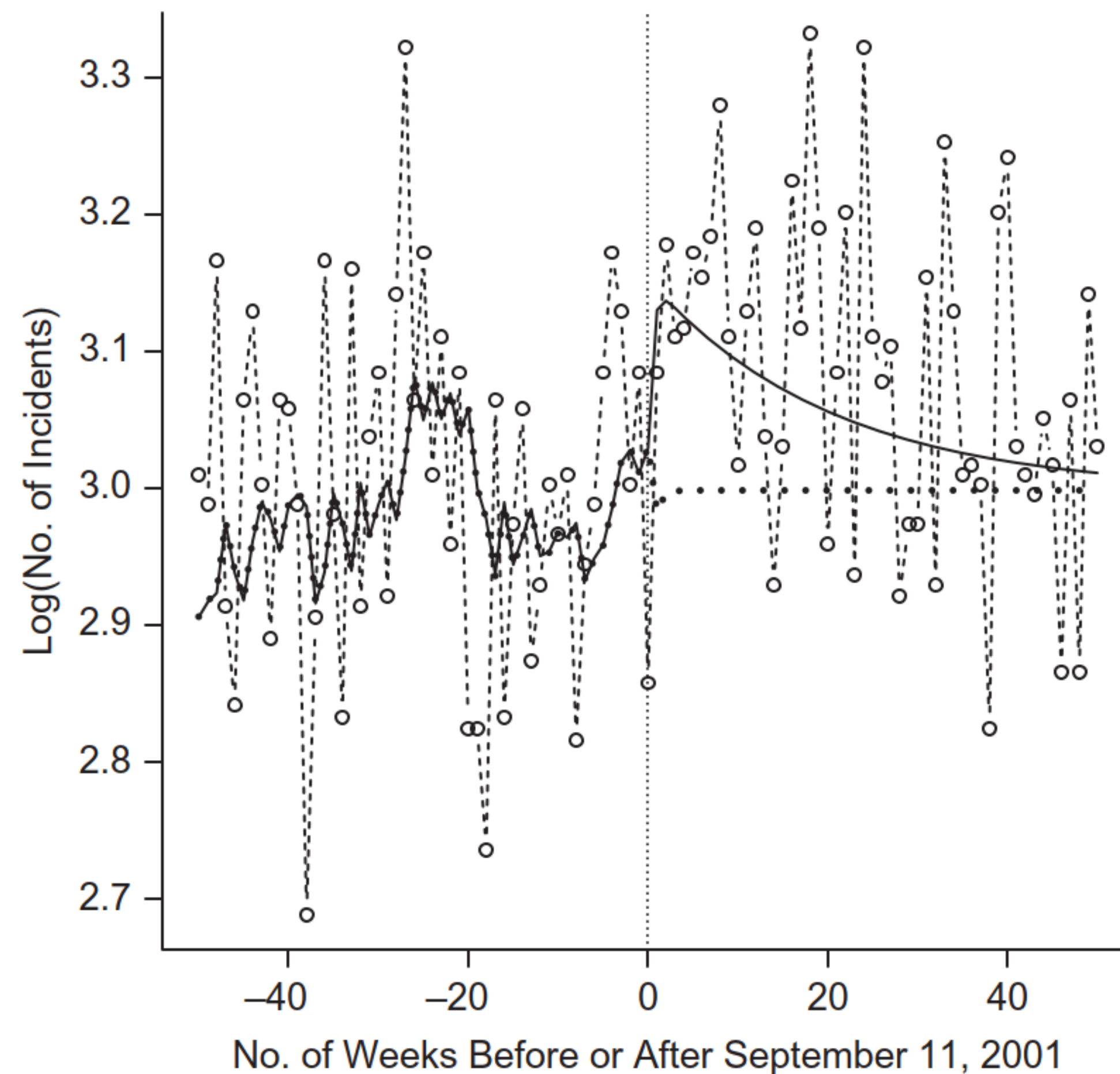
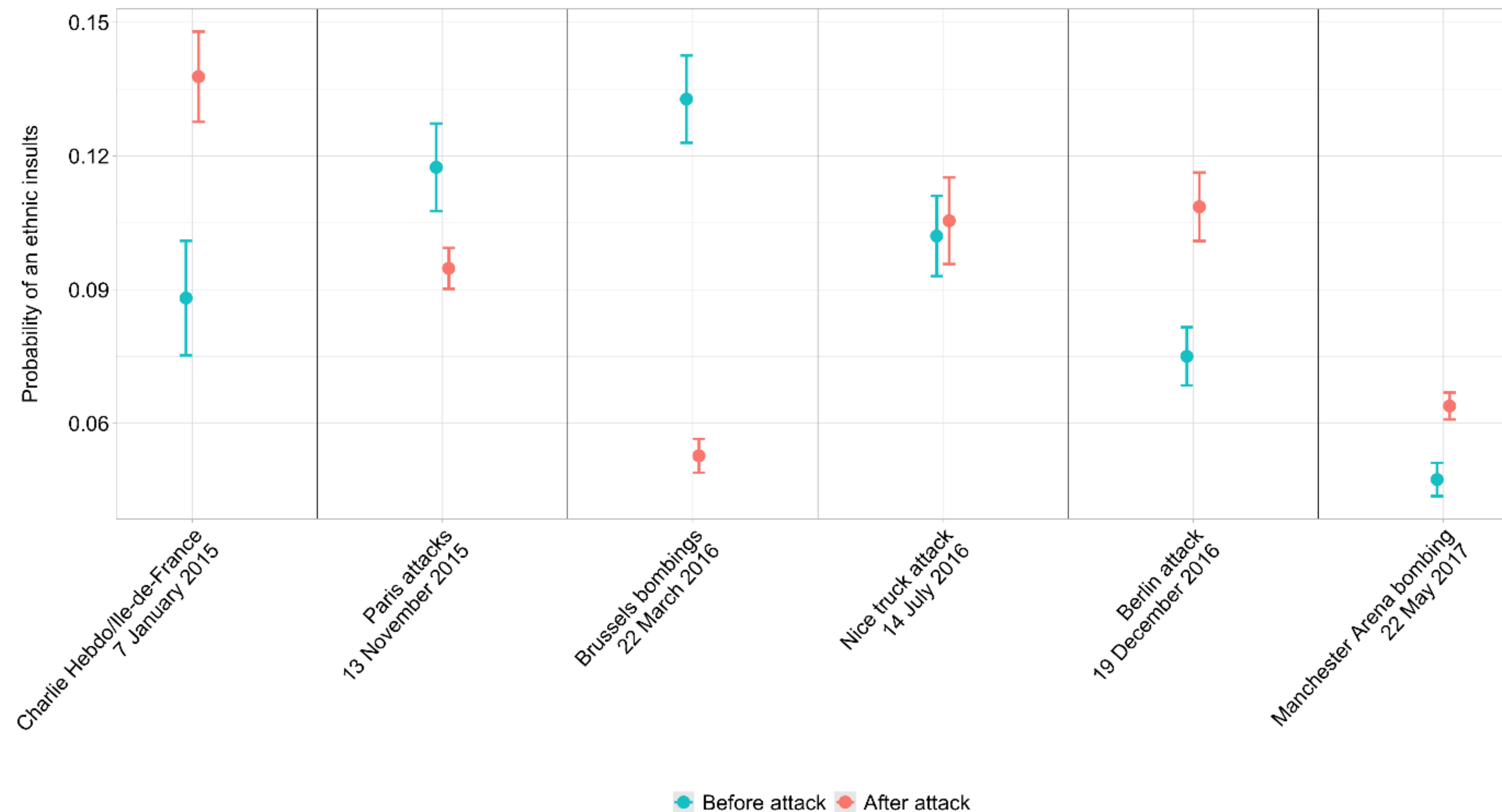


Figure 1. One-step and dynamic predictions of incidence rate for trauma- and stressor-related disorders, Denmark, 1995–2012. Using a time window of approximately 1 year before and after the September 11, 2001 (hereafter referred to as 9/11) attacks, we created a plot showing the observed number of trauma- and stressor-related disorders (dashed line) along with predictions from both the transfer function model (solid line) and a pure autoregressive integrated moving average model (large dotted line). Up until the first week after 9/11 (week = 0), the predictions are 1-step, meaning that the models used data from all relevant previous periods when predicting the value of the series in a given period. After 9/11, the predictions become dynamic and are thus based solely on data from before 9/11 and the parameters of the transfer function.

CZYMARA ET AL. (FORTHCOMING): CATALYST OF HATE? ETHNIC INSULTING ON YOUTUBE IN THE AFTERMATH OF TERROR ATTACKS IN FRANCE, GERMANY, AND THE UNITED KINGDOM 2014-17. *JOURNAL OF ETHNIC AND MIGRATION STUDIES*.



FREY (2020). 'COLOGNE CHANGED EVERYTHING'—THE EFFECT OF THREATENING EVENTS ON THE FREQUENCY AND DISTRIBUTION OF INTERGROUP CONFLICT IN GERMANY. *EUROPEAN SOCIOLOGICAL REVIEW*.

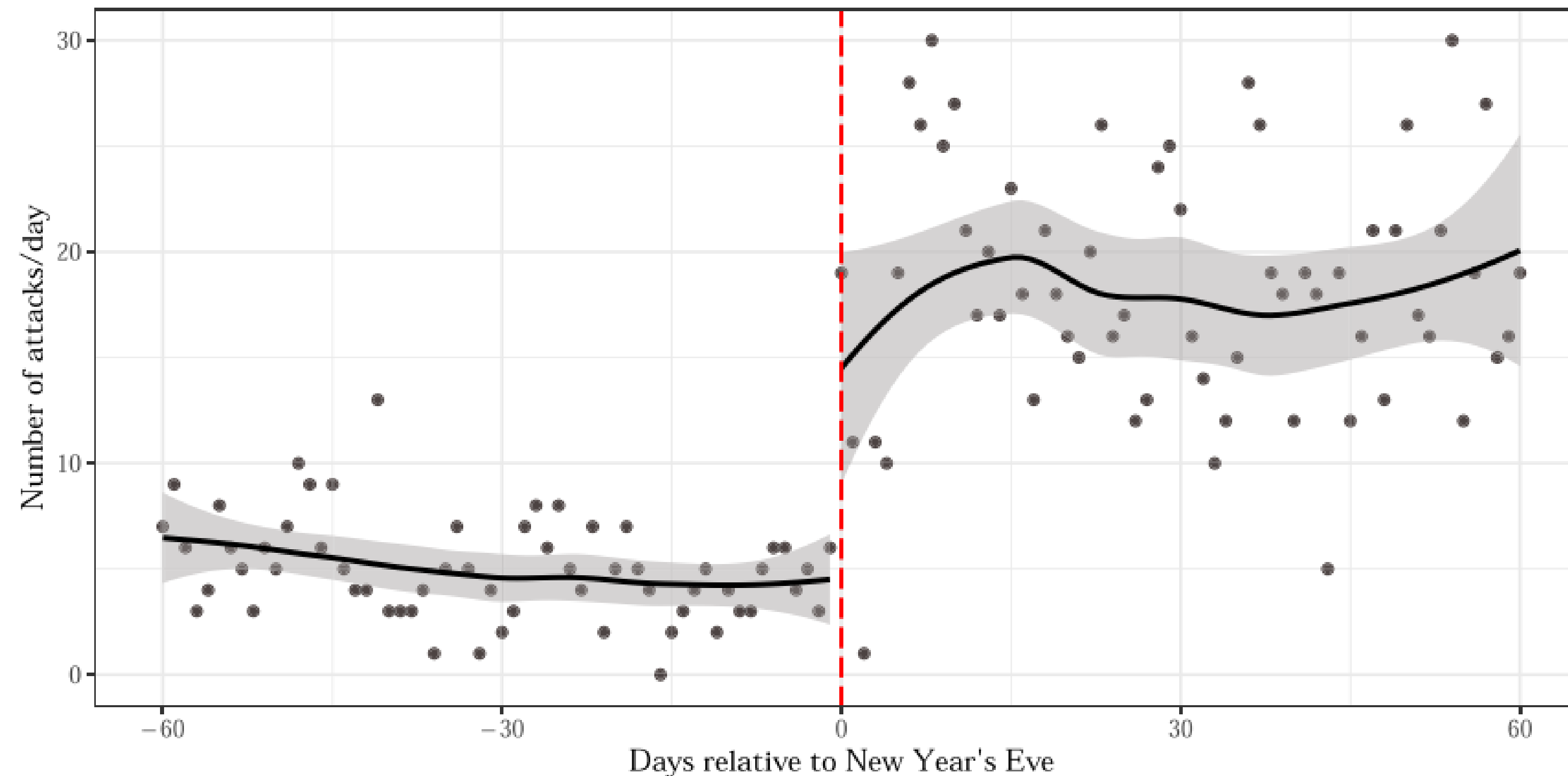


Figure 1. Daily anti-refugee attacks in Germany in the 2 months before and after 1 January 2016

Note: Local regression estimates with 95 per cent confidence intervals.

FLORES (2017). DO ANTI-IMMIGRANT LAWS SHAPE PUBLIC SENTIMENT? A STUDY OF ARIZONA'S SB 1070 USING TWITTER DATA. *AMERICAN JOURNAL OF SOCIOLOGY*.

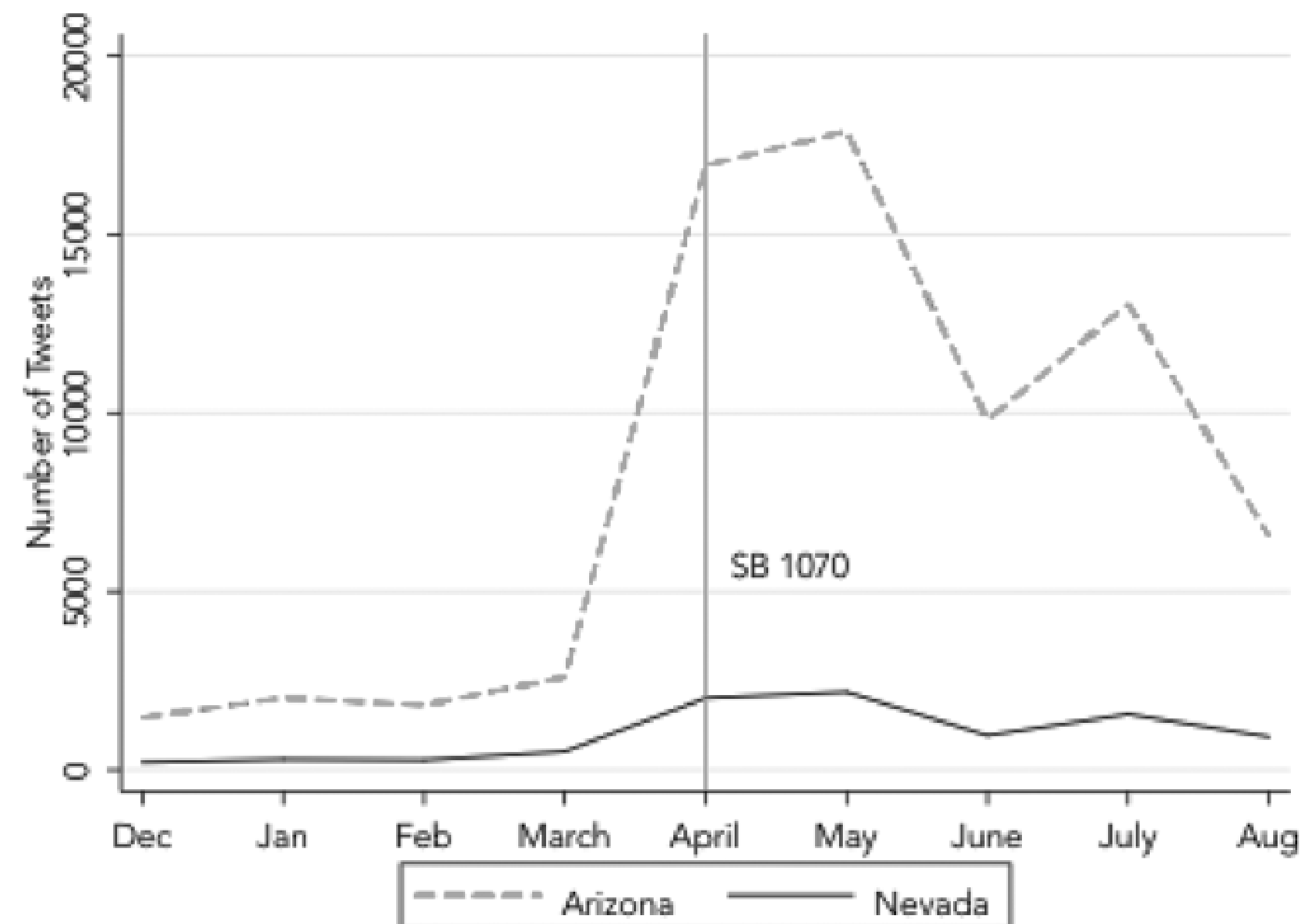


FIG. 2.—Number of Twitter messages related to immigrants per month in Arizona and Nevada (December 2010–August 2011). The vertical line on April 2010 indicates when the Arizona governor approved SB 1070.

MINKUS, DEUTSCHMANN, & DELHEY (2019). A TRUMP EFFECT ON THE EU'S POPULARITY? THE US PRESIDENTIAL ELECTION AS A NATURAL EXPERIMENT. *PERSPECTIVES ON POLITICS*.

Figure 1
The election of Donald Trump as an “external shock”

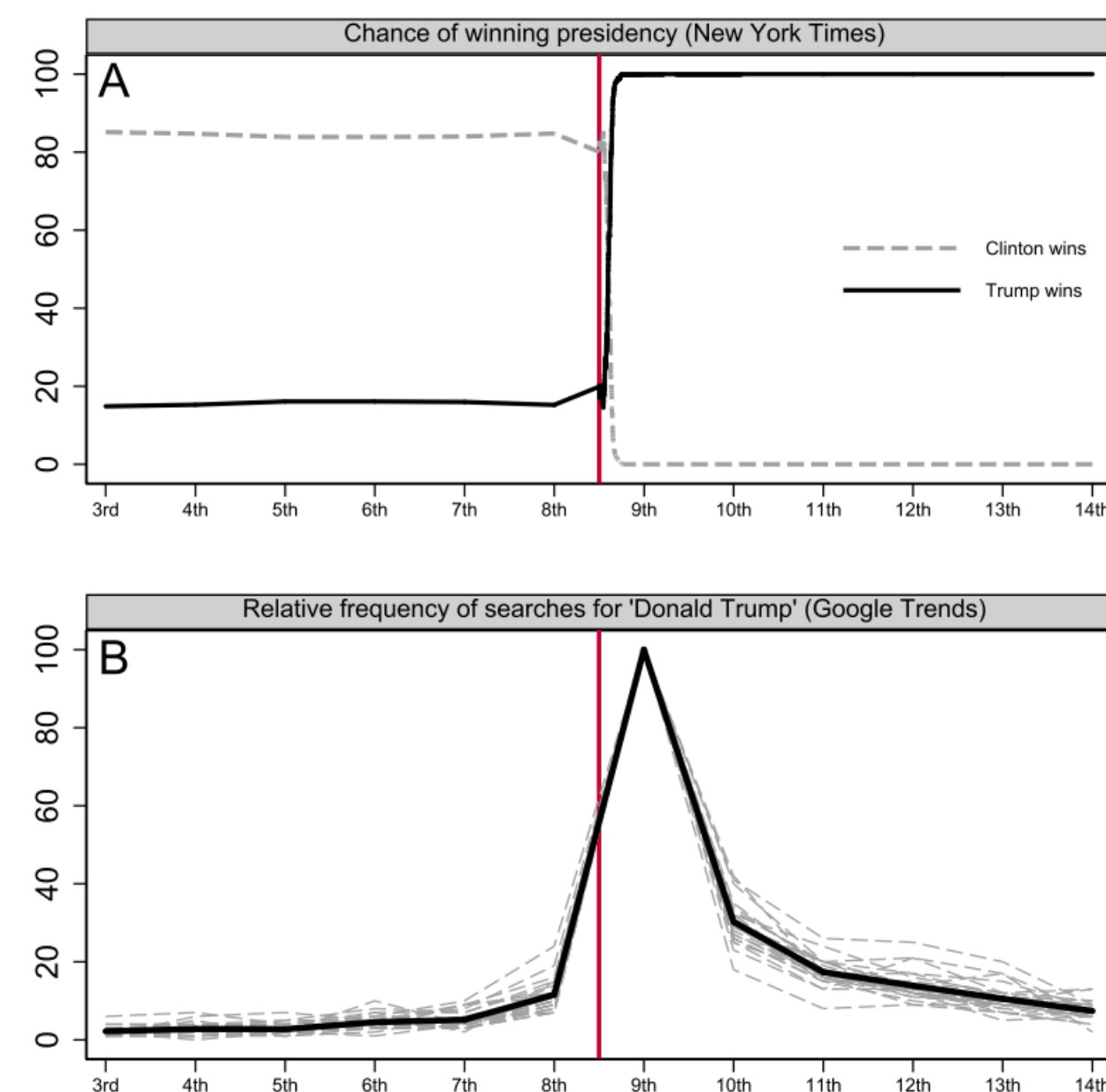
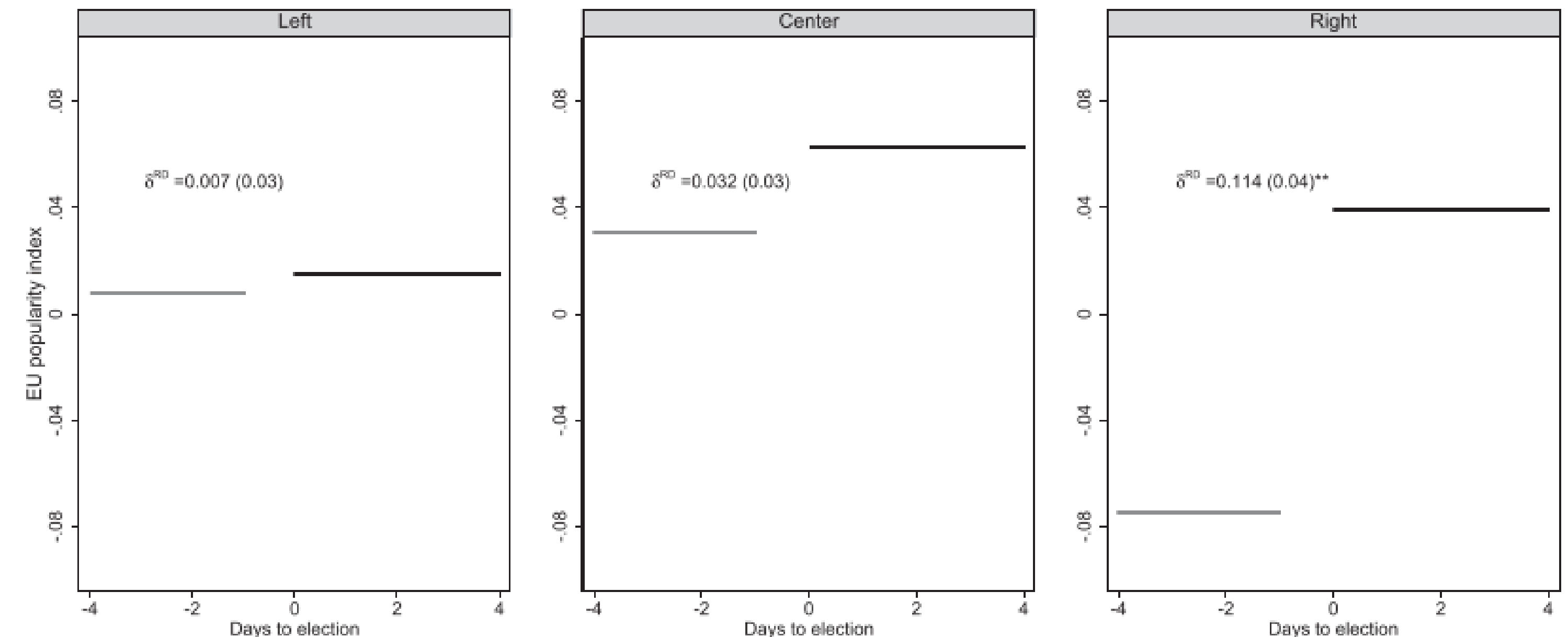


Figure 1A combines scraped data from two *New York Times* websites (nyti.ms/2miH2qZ and nyti.ms/2gOIYTY, last accessed 8/8/2017);

Figure 1B is based on Google Trends (trends.google.com, last accessed 10/10/2017).

Dashed lines in Figures 1B denote trends for the 27 EU countries in our sample, while the solid line denotes the average across these countries.

Figure 5
Treatment effect of the U.S. presidential election; subgroup analysis by *partisanship*



Note: Based on the regressions presented in table A2.

Control variables, including the country dummies, are set at their mean value.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

LITERATURE

- Muñoz, Falcó-Gimeno & Hernández (2020). Unexpected event during survey design: Promise and pitfalls for causal inference. Political Analysis 28 (2): 186-206.