Multiverse analysis of natural experiments

Systematic execution, presentation and interpretation of robustness analyses

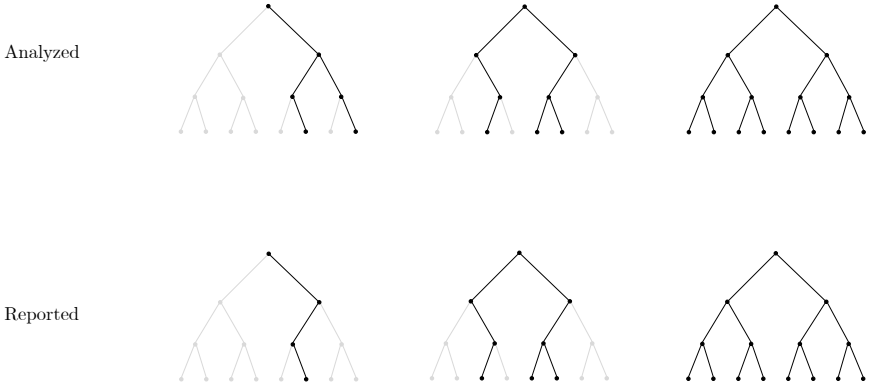
Dave Balzer and Nico Sonntag

November 17, 2020

BERGISCHE
UNIVERSITÄT
WUPPERTAL

JOHANNES GUTENBERG
UNIVERSITÄT MAINZ

Forking paths and researcher degrees of freedom



Multiverse analysis

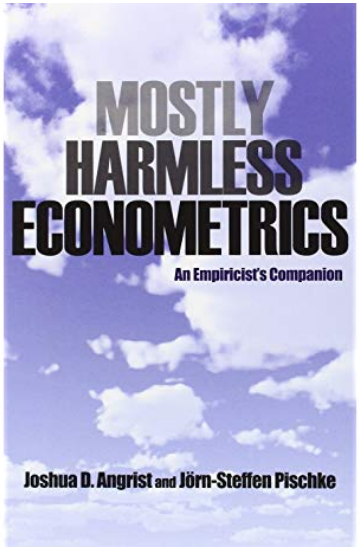
It is possible to calculate all plausible model combinations automatically. We build on approaches that have been around for some time. For example:

- "I Just Ran Four Million Regressions" (Sala-i-Martin 1997)
- "Multimodel analysis" und Stata module `mrobust` (Young & Holsteen 2017)
- "Multiverse analysis" (Steegen et al. 2016)
- "Specification curves (Simonsohn et al. 2015)
- "Coefficient stability plots" (Rao 2020)

Our own approach emphasizes

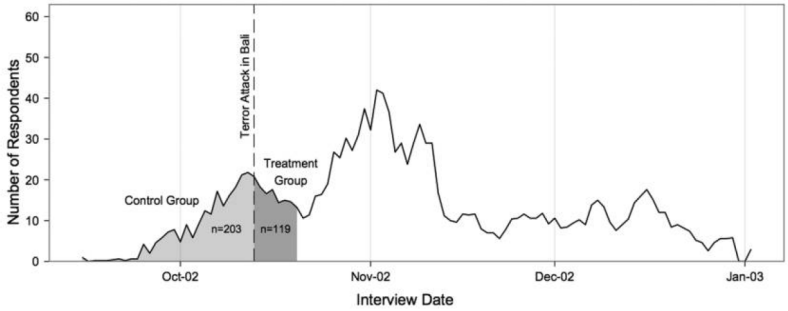
- Relevance for sociology and natural experiments
- Includes further degrees of freedom
- The aim is to assess which decisions are particularly critical for results

Harmless?


MOSTLY
HARMLESS
ECONOMETRICS
An Empiricist's Companion
Joshua D. Angrist and Jörn-Steffen Pischke

Unexpected Event during Surveys Design

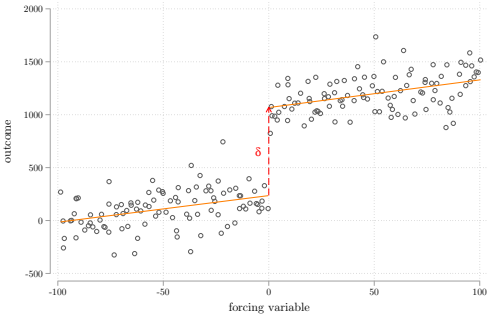
Terrorist attack during ESS field work as a natural experiment (Legewie 2013):



- Randomisation of whole periods (before/after attack)
- Estimate: ATE (?)

Unexpected Event during Surveys Design

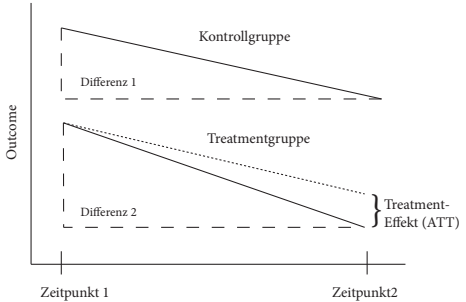
"In many ways, this identification strategy resembles a regression discontinuity design."



- Randomisation only around the threshold
- Correctly specified functional form
- Estimate: LATE (*local average treatment effect*)

Unexpected Event during Surveys Design

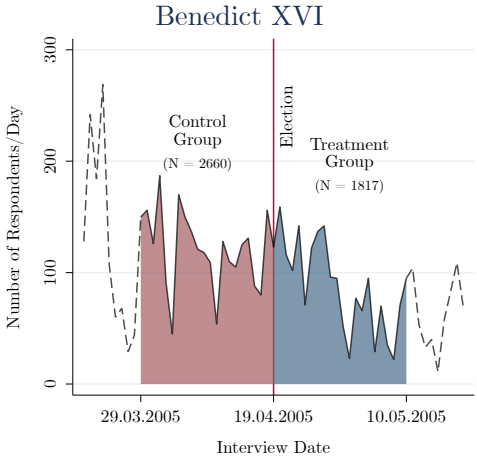
If longitudinal data are available, a difference-in-differences model could also be estimated:

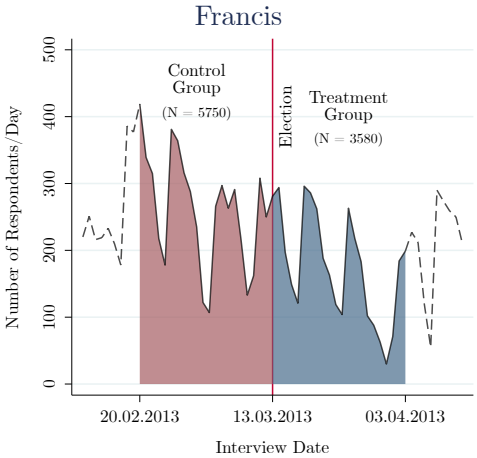


- (As-if) randomisation with regard to time trends (common trends assumption)
- Estimate: ATT

Our research question

Our research question





Our research question

To what extent did the 2005 and 2013 papal elections influence reported religious activity?

Features of our research project

- GSOEP allows the analysis of longitudinal data
- We are able to compare two similar events with the same data set
- Flexible study design: illustrative application of multiverse analysis

Religious and national identity

Mechanisms:

- Interplay between religious and national identity
- Do publicly visible religious leaders increase the salience of religion in their country of origin?

Hypotheses:

H1: Only the 2005 papal election, but not the 2013 papal election, should have increased reported religious activity.

H2: The 2005 papal election primarily influences respondents with low religious activity.

H3: The 2005 papal election primarily affects respondents without established religious identity.

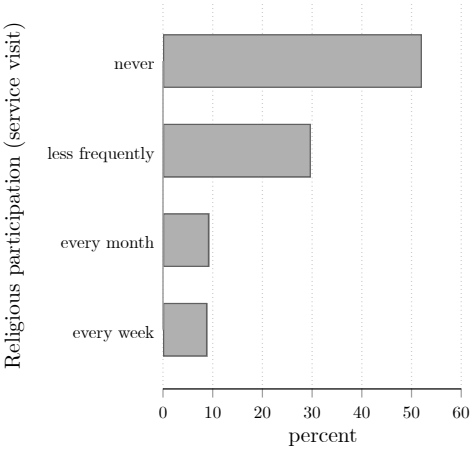
Data summary

Election Benedict XVI (2005):

- GSOEP v33: years 2001, 2005, 2007
- Total N: 64 342
- N $\pm$ 3 weeks: 20 296

Election Francis (2013):

- SGOEP v33: years 2011, 2013, 2015
- Total N: 86 872
- N $\pm$ 3 weeks: 44 078
- Outcome 2013 scaled differently!



Years 2001, 2005, 2007. N = 63,940.

Multiverse analysis: Election Benedict XVI (2005)

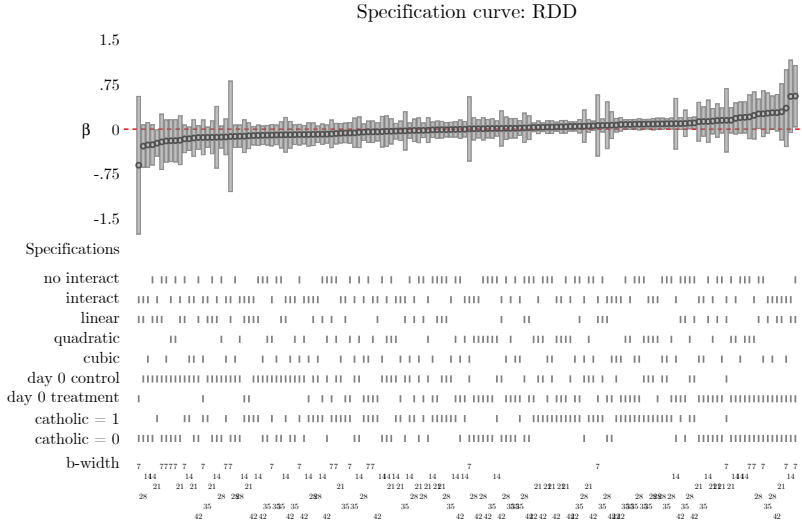
An exemplary multiverse analysis includes several variants of a regression discontinuity design:

- Time trend: linear, quadratic or cubic
- Constant or changing slope after treatment (interaction time trend $\times$ treatment)
- Day 0: treatment or control group
- Sub-sample: Catholics vs. non-Catholics
- Bandwidth: $[7; 42]$ days before and after the election

$$3 \times 2 \times 2 \times 2 \times 36 = 864 \text{ models}$$

We focus on 144 models with bandwidths at weekly intervals (7, 14, 21 ... 42 days).

Multiverse analysis: Election Benedict XVI (2005)



Multiverse analysis: Election Benedict XVI (2005)

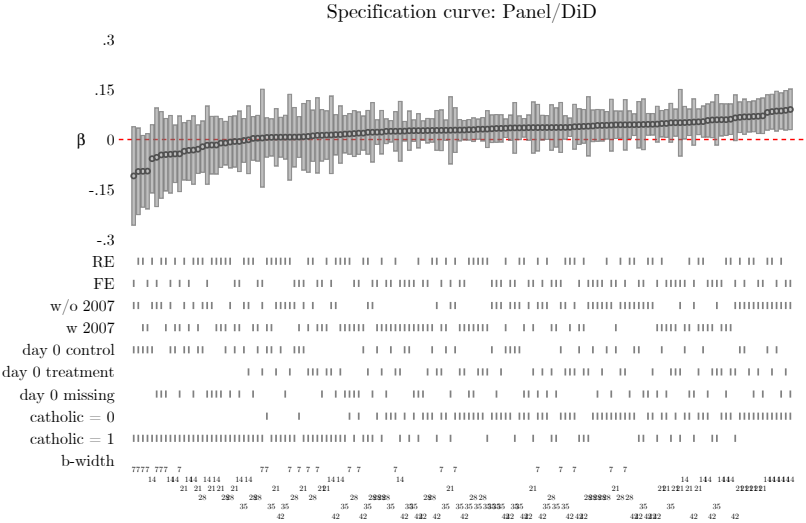
Another exemplary multiverse analysis covers variants of the panel models:

- Fixed- vs. random effects
- Year 2007 in analysis sample yes/no
- Day 0: treatment or control group or exclusion
- Sub-sample: Catholics vs. non-Catholics
- Bandwidth: $[7; 42]$ days before and after the election

$$2 \times 2 \times 3 \times 2 \times 36 = 864 \text{ models}$$

Again, we focus on 144 models with bandwidths at weekly intervals (7, 14, 21 ... 42 days).

Multiverse analysis: Election Benedict XVI (2005)



Multiverse analysis: Crucial decisions

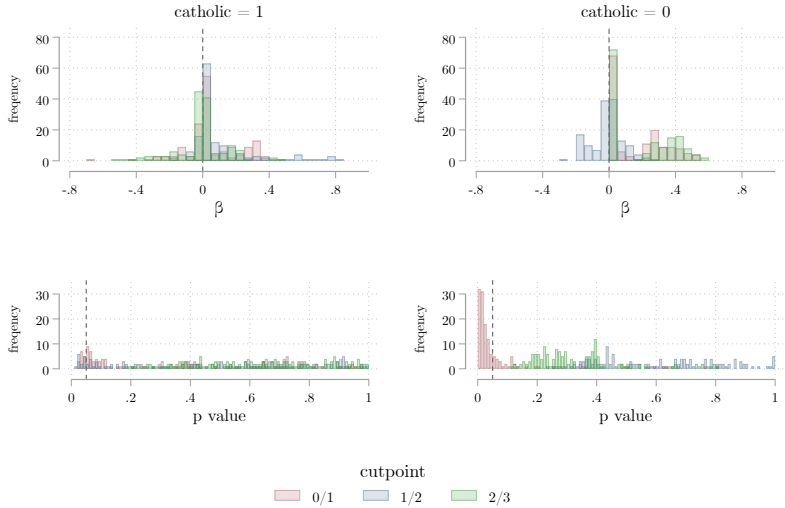
RD specifications:

- Overall, the results do not confirm the hypothesis
- Little systematic relationship between decisions and effect size
  - Non-Catholic respondents
  - Day of election = treatment

Panel specifications:

- Some evidence of a positive treatment effect
- Three patterns
  - Non-Catholic respondents
  - Bandwidth of 14 or 21 days
  - Inclusion of the year 2007

Multiverse analysis: Election Benedict XVI (2005)



Conclusion and discussion

Multiverse analyses help to increase transparency

- Very few natural experiments are "harmless" in the sense that they unambiguously call for a particular research design
- Thus, there are many researcher degress of freedom in the analysis of natural experiments

Open questions

- Shifting the problem to another level?
- Overburdening the readers?
- Best possible (graphic) presentation of results?
- Difficulty of model comparisons
- Too pessimistic about theory-driven model selection?

Thank you for your attention! Pax et bonum!

References

Angrist, J. D. and J.-S. Pischke (2008). *Mostly harmless econometrics: An empiricist's companion*. Princeton university press.

Legewie, J. (2013). Terrorist events and attitudes toward immigrants: A natural experiment. *American journal of sociology 118*(5), 1199–1245.

Rao, A. (2020). *Coefficient Stability Plots*. `https://github.com/AakaashRao/starbilit`.

Sala-i-Martin, X. X. (1997). I just ran four million regressions. *American Economic Review 87*.

Simonsohn, U., J. P. Simmons, and L. D. Nelson (2020). Specification curve: Descriptive and inferential statistics on all reasonable specifications. Available at SSRN: `https://ssrn.com/abstract=2694998`.

Steegen, S., F. Tuerlinckx, A. Gelman, and W. Vanpaemel (2016). Increasing transparency through a multiverse analysis. *Perspectives on Psychological Science 11*(5), 702–712.

Young, C. and K. Holsteen (2017). Model uncertainty and robustness: A computational framework for multimodel analysis. *Sociological Methods & Research 46*(1), 3–40.