Moderation analysis is not as easy as you might have thought

Katrin Auspurg, LMU

Josef Brüderl, LMU

Lena Jost, LMU

Seminar “Analytical Sociology: Theory and Empirical Applications”

San Servolo (Venice), November 18th 2024

Standard moderation analysis

- How does a moderator M affect the effect of a treatment T on an outcome Y?
- Standard (linear) moderation model:
  - all constitutive terms
  - plus (multiplicative) interaction term

$$Y = \alpha + \beta\, T + \gamma\, M + \delta\, T \times M$$

  - (Conditional) treatment effects (**TE**)

$$TE[T \mid M = 0] = \beta$$
$$TE[M \mid T = 0] = \gamma$$

  - Moderation effect (**ME**)

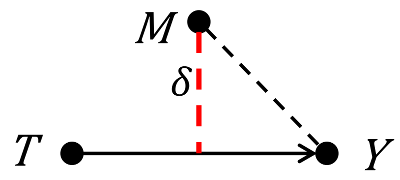
$$ME[M] = \delta$$

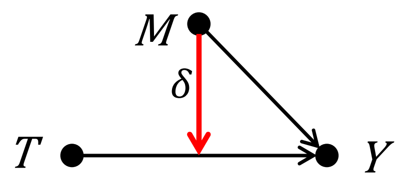
- In most standard moderation analyses
  - the theoretical estimand is not defined
  - identification/estimation assumptions are not discussed
- We will see
  - that there are many informative estimands
  - that the standard moderation model rests on (too) strong identification assumptions
- Content
  - Part I: observational data
  - Part II: experimental data

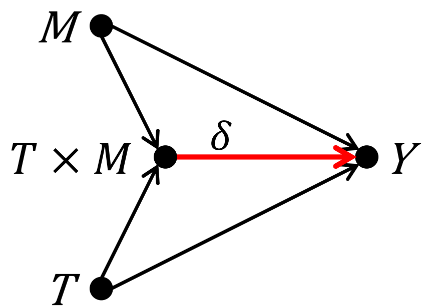
Many moderation estimands

- Descriptive estimand
  - Effect heterogeneity: sub-group specific difference in causal TE[T]
- Causal estimands
  - Causal moderation: (total) causal ME[M] on the (total) causal TE[T]
  - Causal interaction: the causal effect of two simultaneously applied treatments (joint treatment effect, JTE)

$$\boxed{\text{JTE} = \text{ME[M]} = \text{ME[T]}}$$







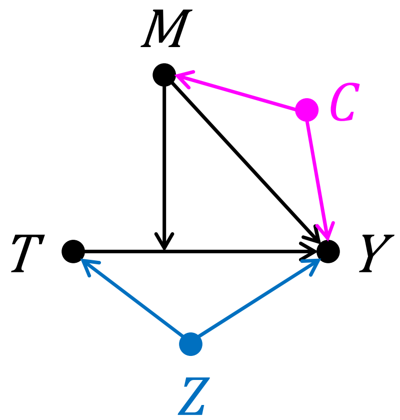
Identification of a causal moderation effect

<u>Identification</u> (Hernán/Robins 2023, pp. 58 ff.)

- TE of TM (combined treatment) must be identified
  - TE of T must be identified
  - TE of M must be identified
  - JTE must be identified

- Key identification assumption: conditional independence
  - After conditioning on confounders the potential outcomes and $\{M, T\}$ must be independent
  - No unobserved confounding

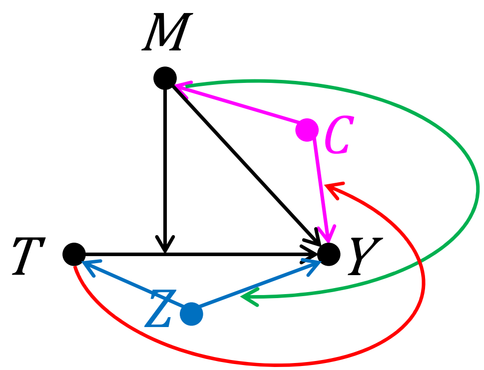
Situation: only TE confounding



- Constant confounding
- Estimation (linear model)

$$Y = \alpha + \beta\, T + \gamma\, M + \delta\, T \times M$$
$$+\, a\, C$$
$$+\, c\, Z$$

In addition JTE confounding

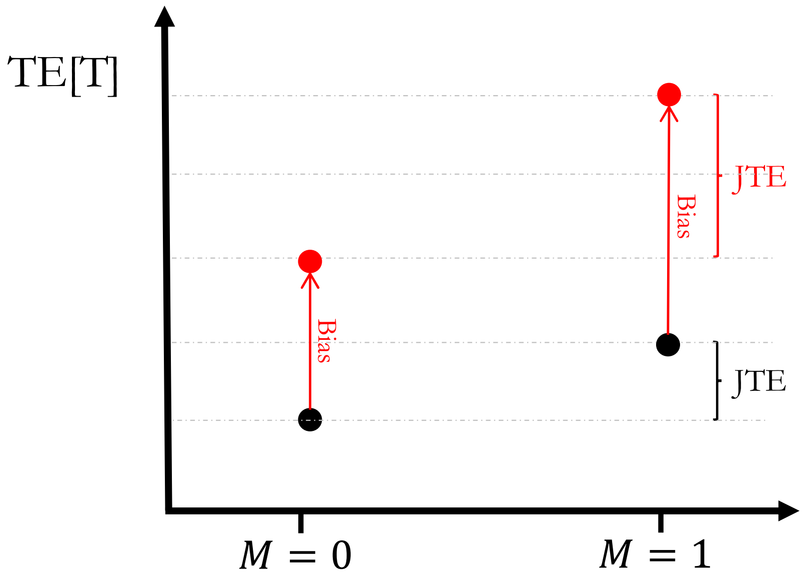


- Differential confounding
- Estimation (linear model)

$$Y = \alpha + \beta\, T + \gamma\, M + \delta\, T \times M$$
$$+\, a\, C + b\, C \times T$$
$$+\, c\, Z + d\, Z \times M$$

Differential confounding

- Why would the interaction $Z \times M$ bias the JTE?



- In this example, TE[T] is differentially confounded by Z for different levels of M
  - JTE is biased upwards

- Analogue arguments apply for differential confounding by $C \times T$

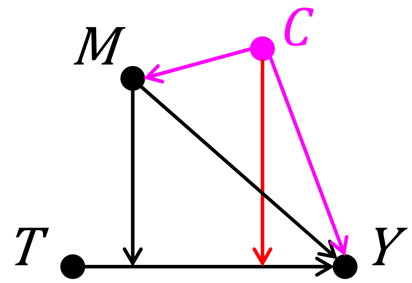
● No confounding by Z
● Differential confounding by $Z \times M$

Intuition: omitted interaction bias

- Due to differential confounding C is also a moderator of TE[T]
  - If we do not account for this, ME[M] will „pick up“ the moderation by C

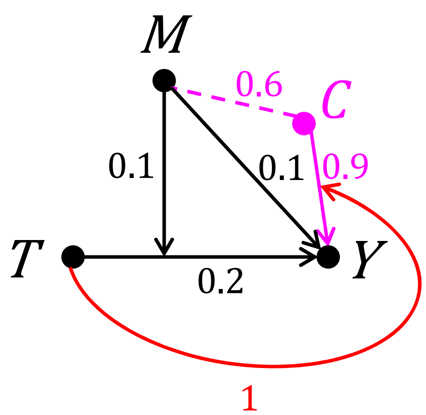
➢ “**Omitted interaction bias**” (Beiser-McGrath/Beiser-McGrath 2020, Blackwell/Olson 2022; Breen et al. 2015; Nilsson et al. 2021)

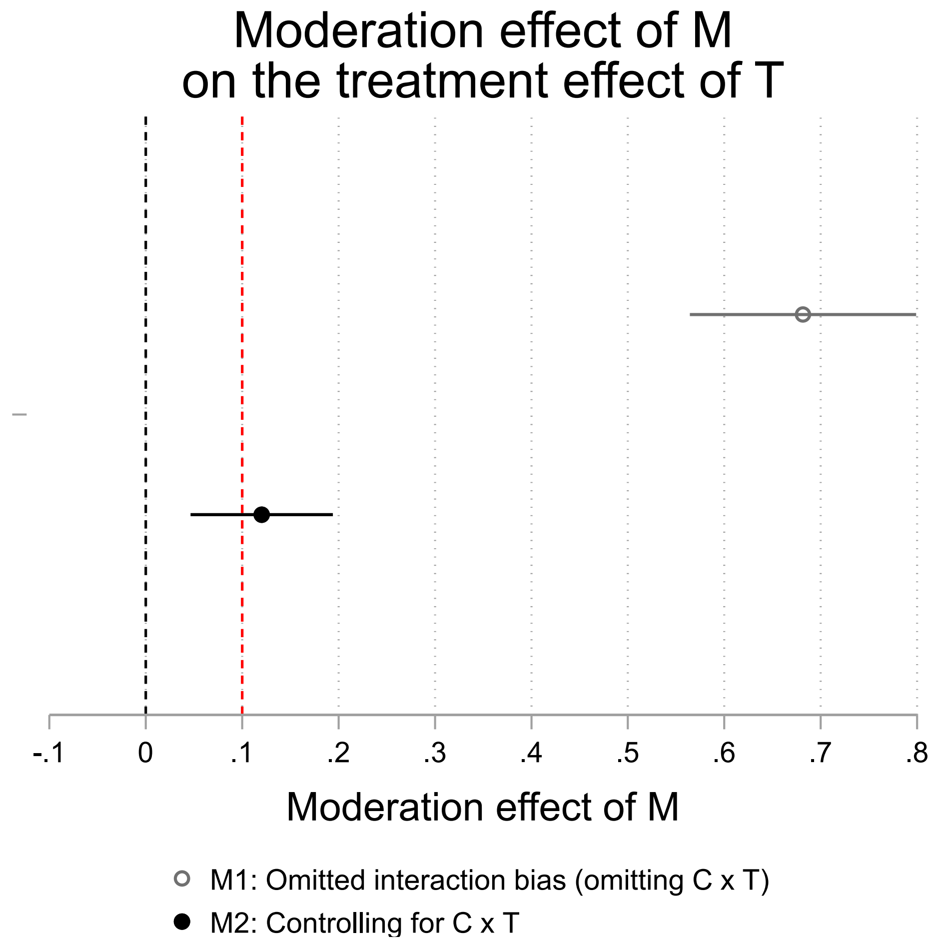
  - We can avoid an omitted interaction bias by controlling for C and $C \times T$
    - And due to symmetry analogous arguments also apply for Z and $Z \times M$



Simulation: omitted interaction bias

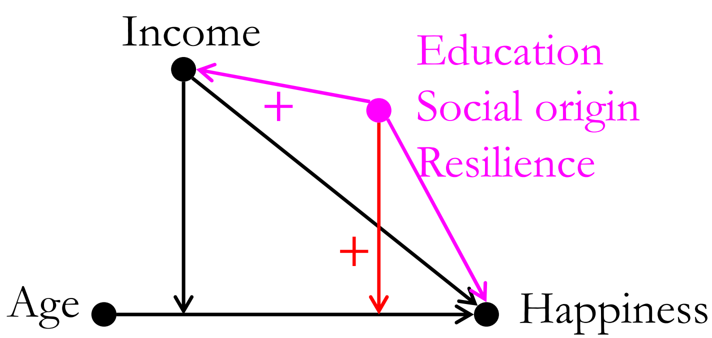
- DGP
  - $T, M, C \sim \boldsymbol{N}(0,1)$, $(N = 1{,}000)$
  - Moderation effect of M: $+\mathbf{0.1}$
  - Differential confounding by C





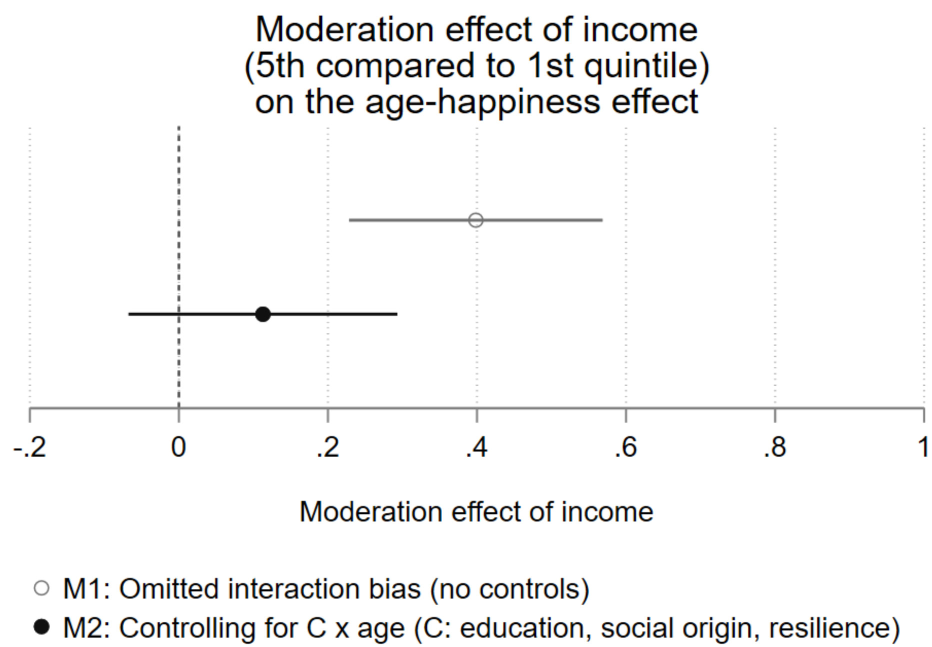
Omitted interaction bias in real world data

- Happiness declines with age
  - (Almost) linearly between ages 25 and 55 (Kratz/Brüderl 2021)
- Does income causally moderate the happiness decline?
- Potential omitted interaction bias

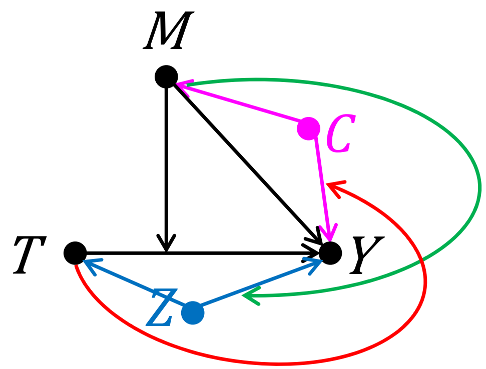


All moderators are time-constant, and therefore not affected by age

- SOEP v37, only ages 25 – 55, N=30,073
- Linear FE model
- Age / 30
- Personal net income in quintiles



Summary: different estimands, different estimation



- Estimation for causal moderation (resp. causal interaction)
- Est. for effect heterogeneity
  - TE[T] must be identified

$$Y = \alpha + \beta\, T + \gamma\, M + \delta_1\, T \times M + a\, C + b\, C \times T + c\, Z + d\, Z \times M$$

Causal moderation
$ME = \delta_1$

$$Y = \alpha + \beta\, T + \gamma\, M + \delta_2\, T \times M + c\, Z + d\, Z \times M$$

Effect heterogeneity
$EH = \delta_2$

A way out?
Do experimental designs make it easier to identify causal moderation effects?

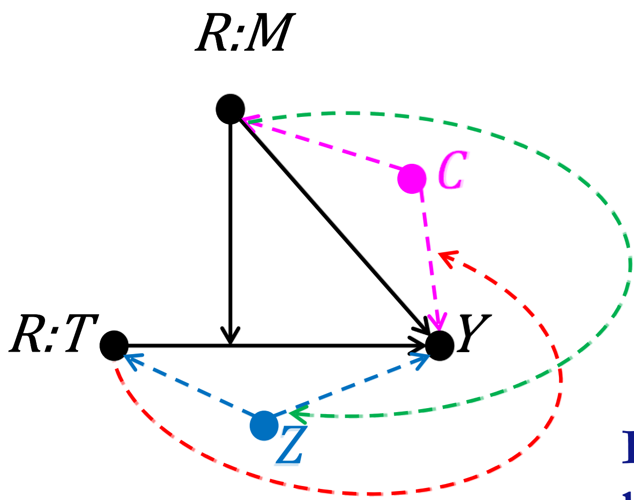
Design A: Randomized moderator

- Multifactorial experiment
  - Both T and M are assigned randomly
- Example: Test for statistical discrimination: Smaller effect of ethnicity when there is more information about employment?
- 2 x 2 experimental design
  - T = Ethnicity (e.g., Turk vs. German)
  - M = Info on employment (yes vs. no)
- JTE (T x M) is estimand for statistical discrimination

> Dear Ms./Mr.,
> I am very interested in the apartment you advertised. My name is Cem Güleryüz. I am permanently employed as an electrician. I would be very grateful if you could offer me a viewing.

Design A: Randomized moderator

- Identification?
- Due to the randomization (denoted by 'R:'): No confounders for T and M

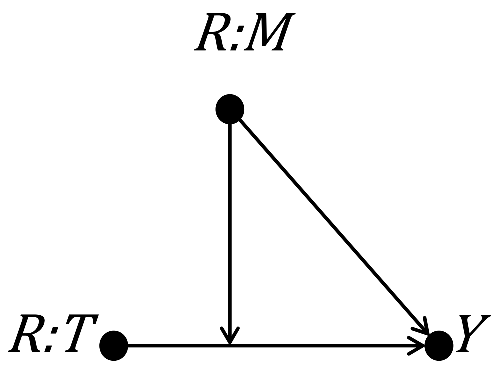


Eliminated by design!

- $Y = \alpha + \beta\, T + \gamma\, M + \delta\, T \times M$
  $+ a\, C + b\, C \times T$
  $+ c\, Z + d\, Z \times M$

Design A: Randomized moderator

- Identification?
- Due to the randomization (denoted by ‘R:’): No confounders for T and M



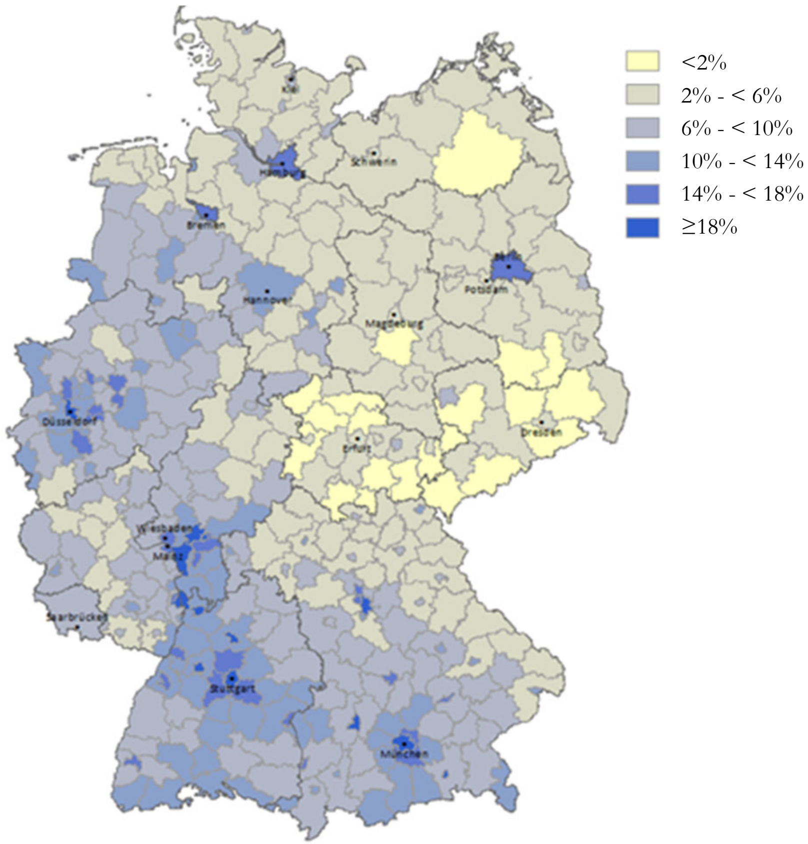
➢Both TEs and JTE $T \times M$ correctly identified by standard approach

- $Y = \alpha + \beta\, T + \gamma\, M + \delta\, T \times M$

Design B: Non-randomized moderator

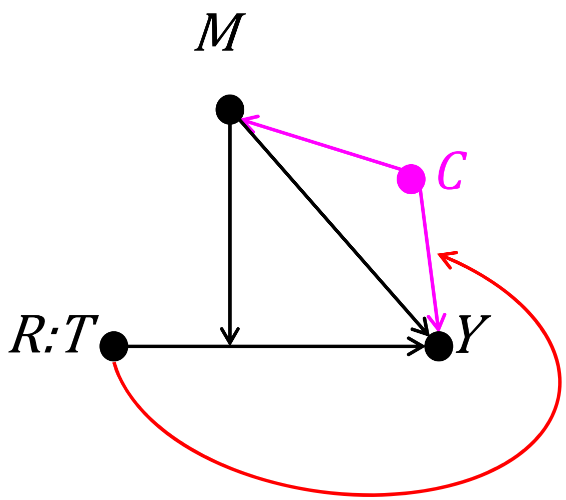
- For many research questions, a random assignment of M is hard to achieve, e.g.
  - Characteristics of participants
  - Different (regional) contexts
- Example: More/less discrimination in local contexts with many foreigners?
  - Effect heterogeneity: does discrimination vary by %foreigners? (correlation with segregation)
  - Causal moderation: does %foreigners *per se* make a difference? (evidence for the contact hypothesis)

% Foreigners



Design B: Non-randomized moderator

- Identification?
- Estimation of causal moderation effect: C and $C\ x\ T$ have to be included for successful identification!

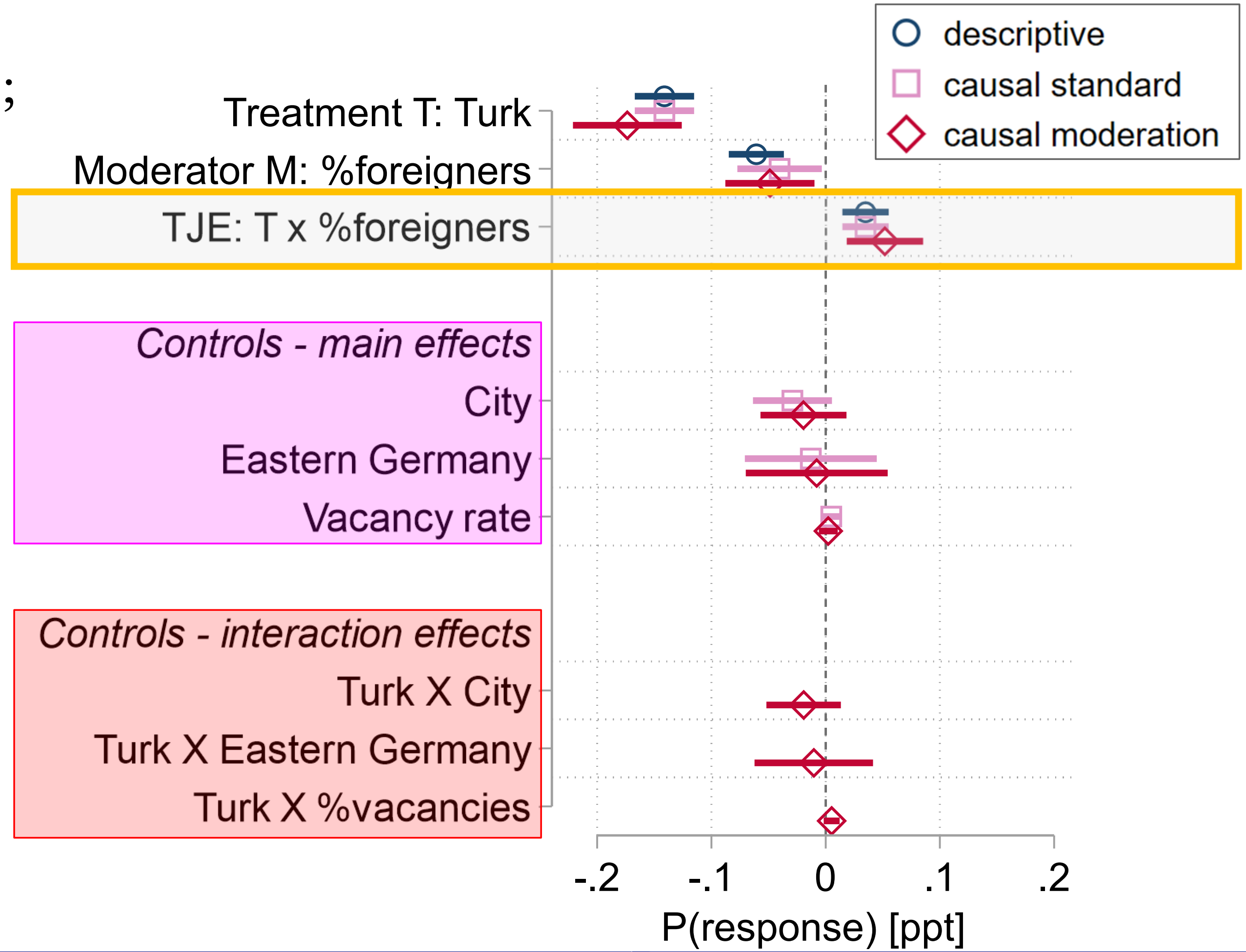


- $Y = \alpha + \beta\, T + \gamma\, M + \delta\, T \times M$
  $+\, a\, C + b\, C \times T$

Empirical example for Design B

- Field experiment on ethnic discrim.; discrimination depending on size of foreign population in local area?
  - *R: T* = Ethnicity (Turk vs. Ger)
  - *M* = %foreigners in county

- Data on the German housing market 2015
  (Auspurg/Lorenz/Schneck 2023, *N* = 9.450)



Unobserved moderation confounders

- But be aware:
  - There might be unobserved confounders
  - Especially in cross-sectional designs
- Think of randomizing M
  - Natural experiments
  - Randomized controlled trials
- Panel data on T *and* M can help
  - E.g., survey experiments implemented in a panel study

Example: Two-wave field experiment; in between: natural experiment of ↑% foreigners due to the "refugee crisis"

Figure 2: Monthly Number of Asylum Seekers and Timing of the Two Waves of our Field Experiment.
Note: This figure shows the number of newly registered asylum seekers in Germany from January 2014 to January 2016 and the timing of the two waves of our repeated field experiment on ethnic discrimination (each lasting one week). On average, about 1 refugee per 100 inhabitants arrived in the 401 counties in Germany between the dates of the two waves. Data: German Federal Ministry of the Interior and Community, own illustration.

Source: Auspurg/Lorenz/Schneck 2023, p. 647

Take-home messages for Venice-seminar (and beyond)

- If you are going to present a moderation analysis, please provide the following information:
  - What is your estimand?
    - Descriptive effect heterogeneity
    - Causal moderation effect
  - Discuss your identification and estimation assumptions
    - Are there potential moderation effect confounders?

Thank you for your attention!

References

Auspurg, K.; Lorenz, R., & Schneck, A. (2023) Does Unprecedented Mass Immigration Fuel Ethnic Discrimination? A Two-Wave Field Experiment in the German Housing Market. *Sociological Science* 10(1), 640-666.

Bansak, K. (2021) Estimating causal moderation effects with randomized treatments and non-randomized moderators. *J R Stat Soc Series A*, 184, 65–86.

Beiser-McGrath J, Beiser-McGrath LF (2020) Problems with products? Control strategies for models with interaction and quadratic effects. Political Science Research and Methods 8, 707–730. https://doi.org/10.1017/psrm.2020.17

Blackwell, M., & Olson, M. P. (2022) Reducing Model Misspecification and Bias in the Estimation of Interactions. *Political Analysis*, 30(4), 495–514.

Breen, R., Choi, S., & Holm, A. (2015) Heterogeneous Causal Effects and Sample Selection Bias. *Sociological Science*, 2(17), 351–369.

Hernán, M. A., Robins, J. M. (2020) *Causal Inference: What If*. Boca Raton: Chapman & Hall/CRC.

Kratz, F., & Brüderl, J. (2021) The Age Trajectory of Happiness.PsyArXiv. April 18. doi:10.31234/osf.io/d8f2z.

Nilsson, A., Bonander, C., Strömberg, U., & Björk, J. (2021) A directed acyclic graph for interactions. *International Journal of Epidemiology*, 50(2), 613–619.