

# M-FE Forschungsmethoden und Evaluation

## *Evaluation der Wirksamkeit psychologischer Interventionen anhand von Regressions-Diskontinuitäts-Analysen*

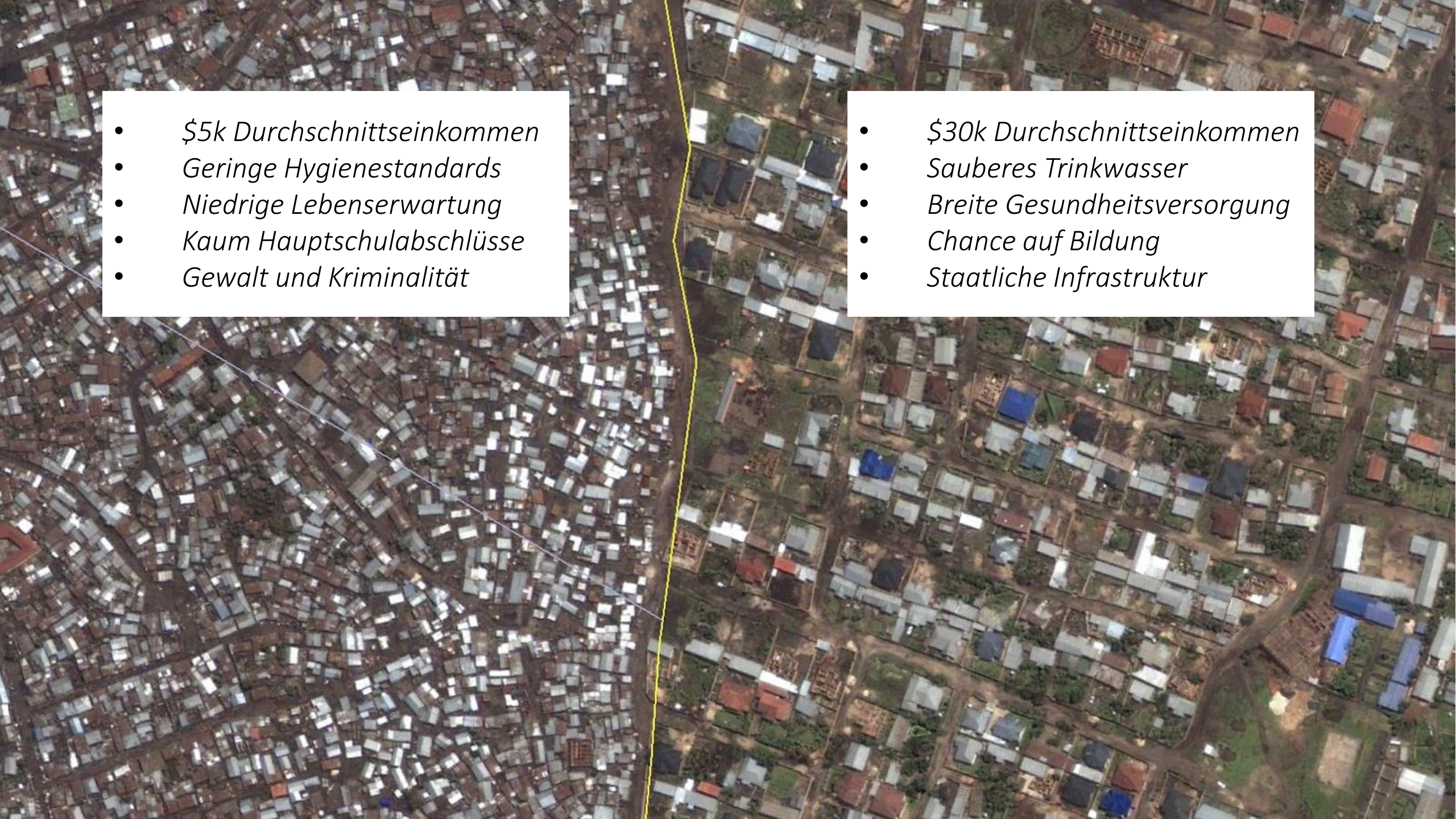
*Dominik Deffner*









- 
- *\$5k Durchschnittseinkommen*
  - *Geringe Hygienestandards*
  - *Niedrige Lebenserwartung*
  - *Kaum Hauptschulabschlüsse*
  - *Gewalt und Kriminalität*

- *\$30k Durchschnittseinkommen*
- *Sauberes Trinkwasser*
- *Breite Gesundheitsversorgung*
- *Chance auf Bildung*
- *Staatliche Infrastruktur*



An aerial photograph showing the border between Mexico and the United States in the city of Nogales. A yellow dashed line runs vertically through the center of the image, separating the two countries. To the left of the line is the Mexican side, and to the right is the US side. Both sides are densely populated with residential buildings, many of which have corrugated metal roofs. The Mexican side appears more densely packed with smaller structures, while the US side shows slightly more spacing between buildings.

*Nogales, Sonora,  
Mexiko*

*Nogales, Arizona,  
USA*





*Nogales, Sonora,  
Mexiko*

*Nogales, Arizona,  
USA*

*Beinahe identische...  
...Geographie  
...Klima  
...Rohstoffe  
...Sozialisation*



An aerial photograph of a city, likely Nogales, showing a dense residential area with many small houses. A yellow dashed line runs vertically through the center of the image, representing the border between Mexico and the USA. The left side of the image is labeled 'Nogales, Sonora, Mexiko' and the right side is labeled 'Nogales, Arizona, USA'.

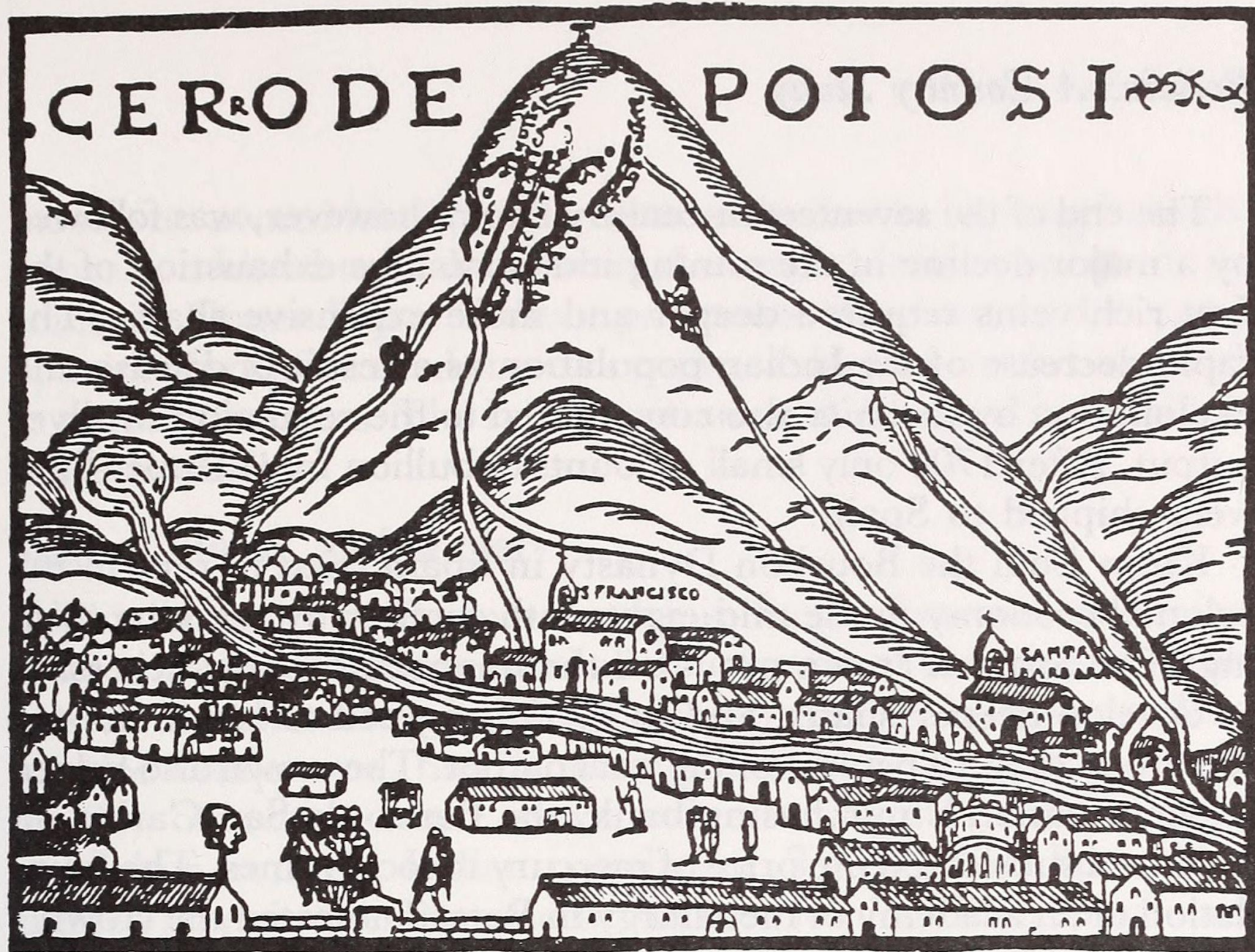
*Nogales, Sonora,  
Mexiko*

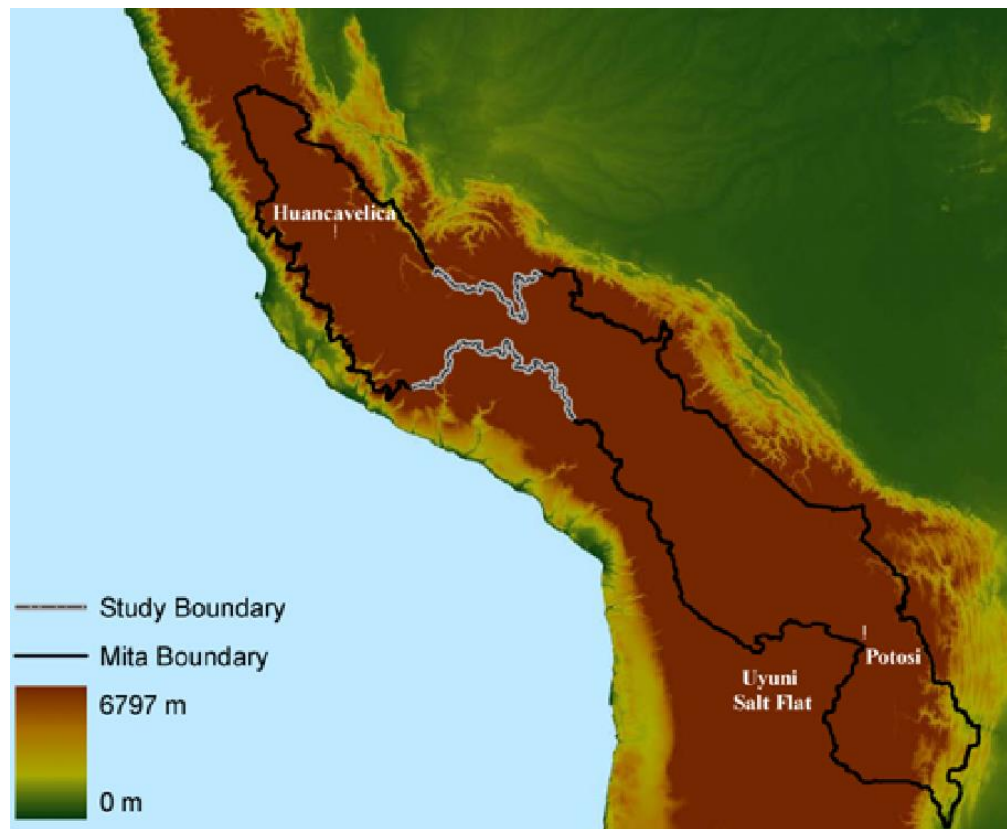
*Nogales, Arizona,  
USA*

*Beinahe identische...  
...Geographie  
...Klima  
...Rohstoffe  
...Sozialisation*

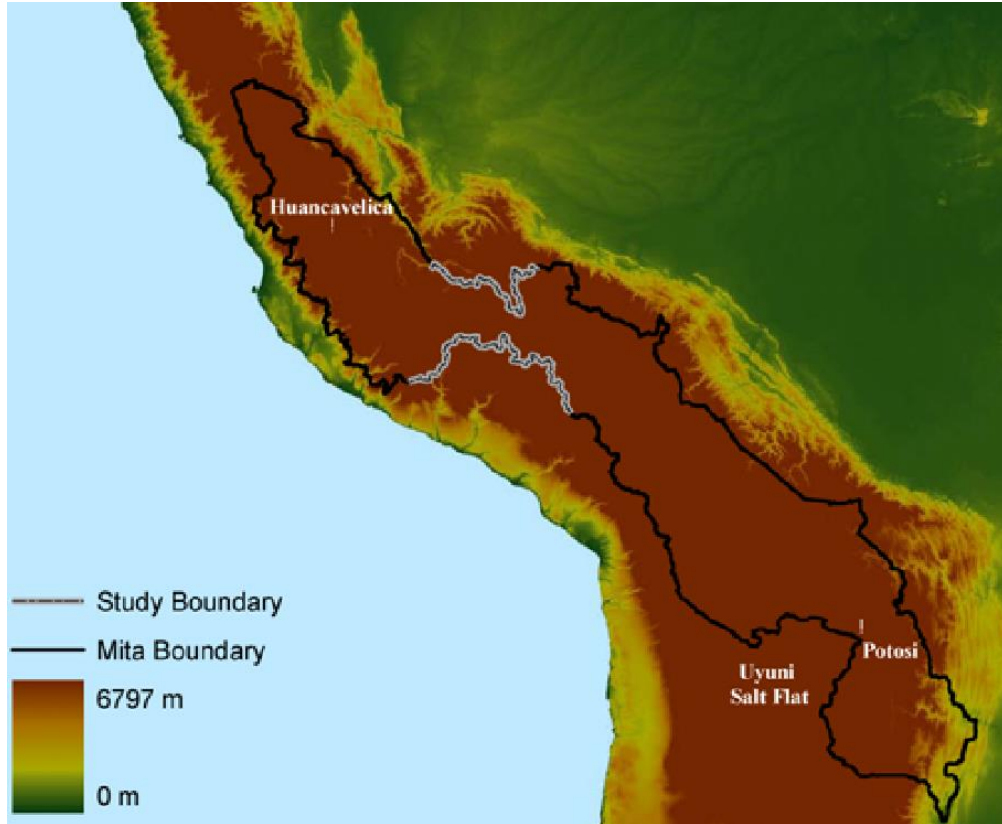
*Grenzen als natürliche  
Quasi-Experimente*





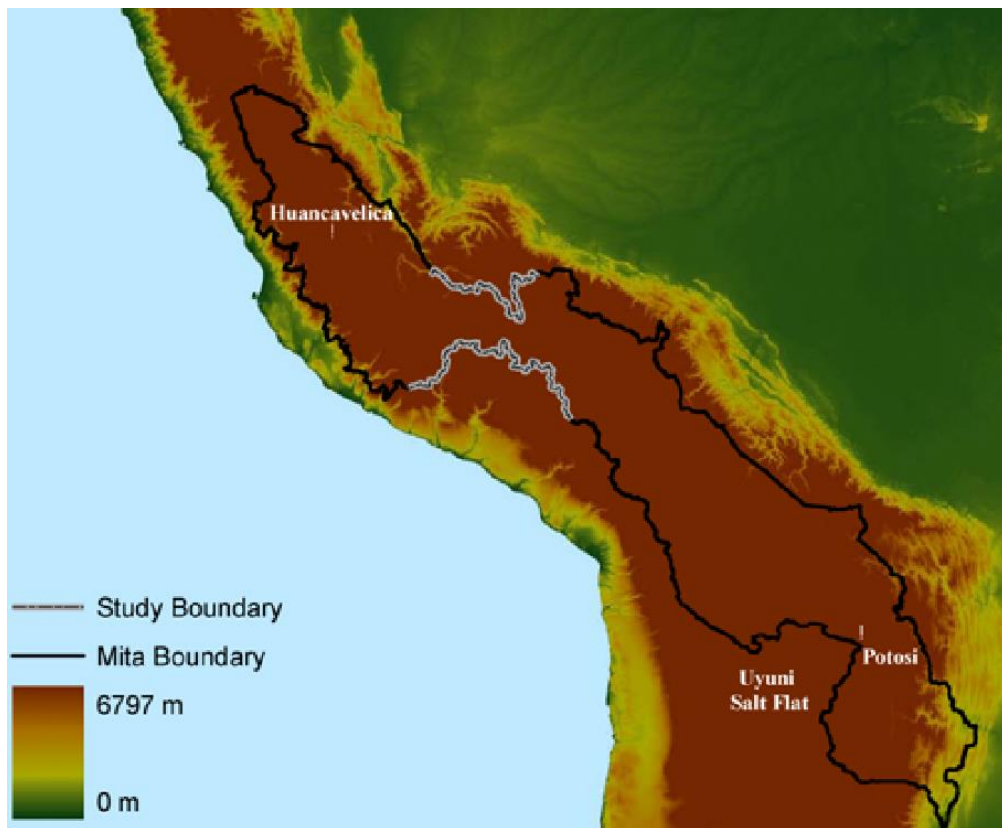






Welchen Einfluss hat das koloniale  
Mita-System auf heutige  
ökonomische, gesundheitliche und  
soziale Outcomes?

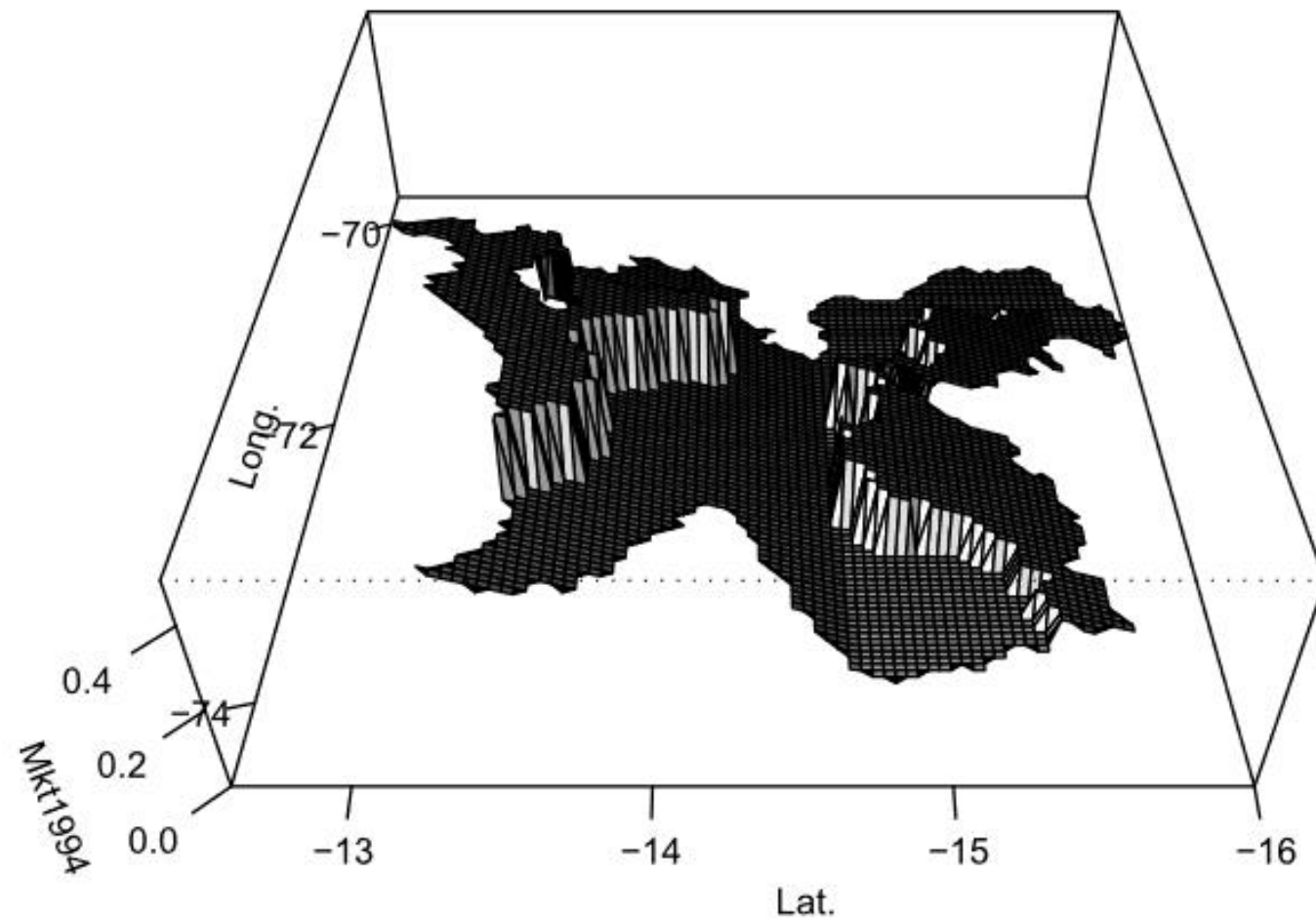
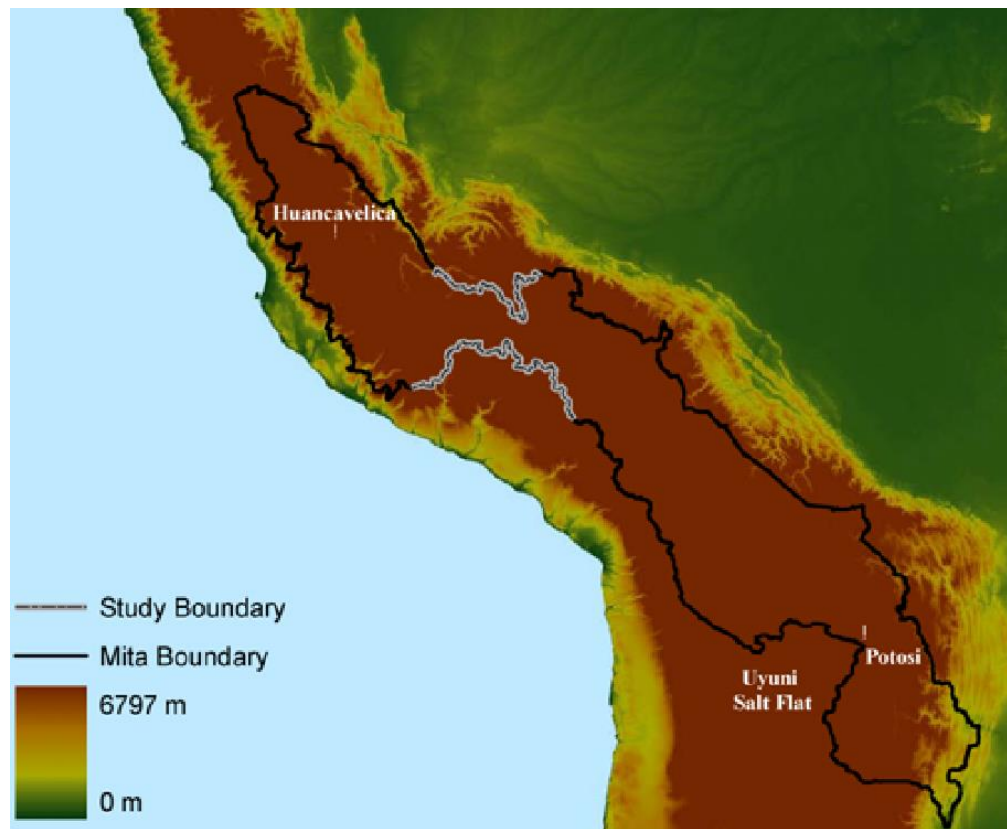




Welchen Einfluss hat das koloniale  
Mita-System auf heutige  
ökonomische, gesundheitliche und  
soziale Outcomes?

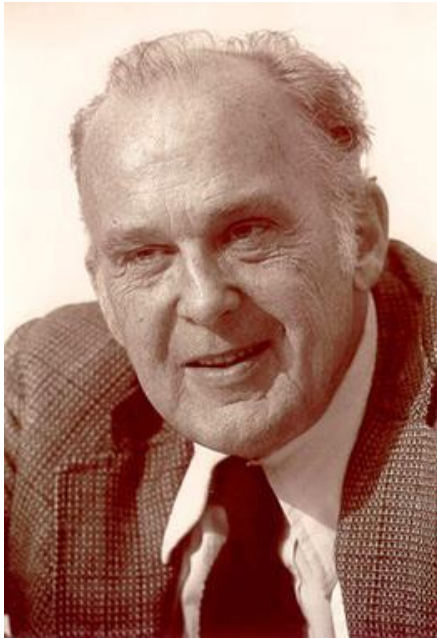
Grenzregion als natürliches Quasi-  
Experiment!



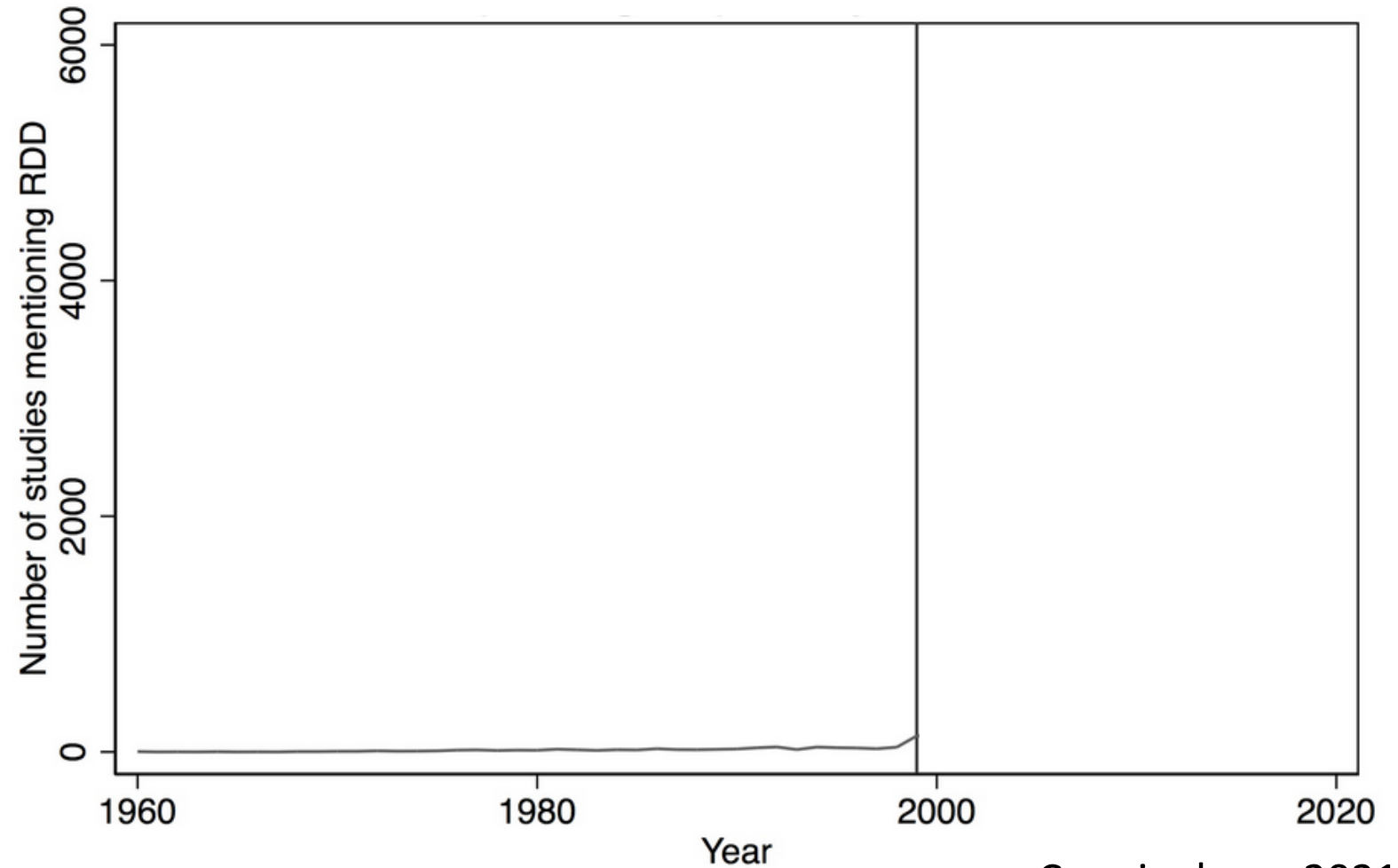




# Regressions-Diskontinuitäts-Analysen (RDA)



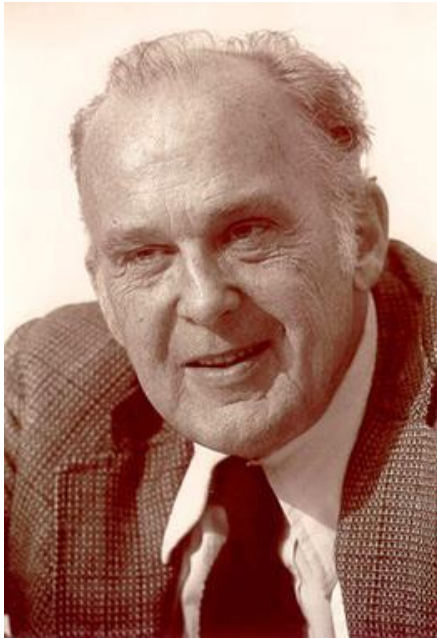
Donald T. Campbell



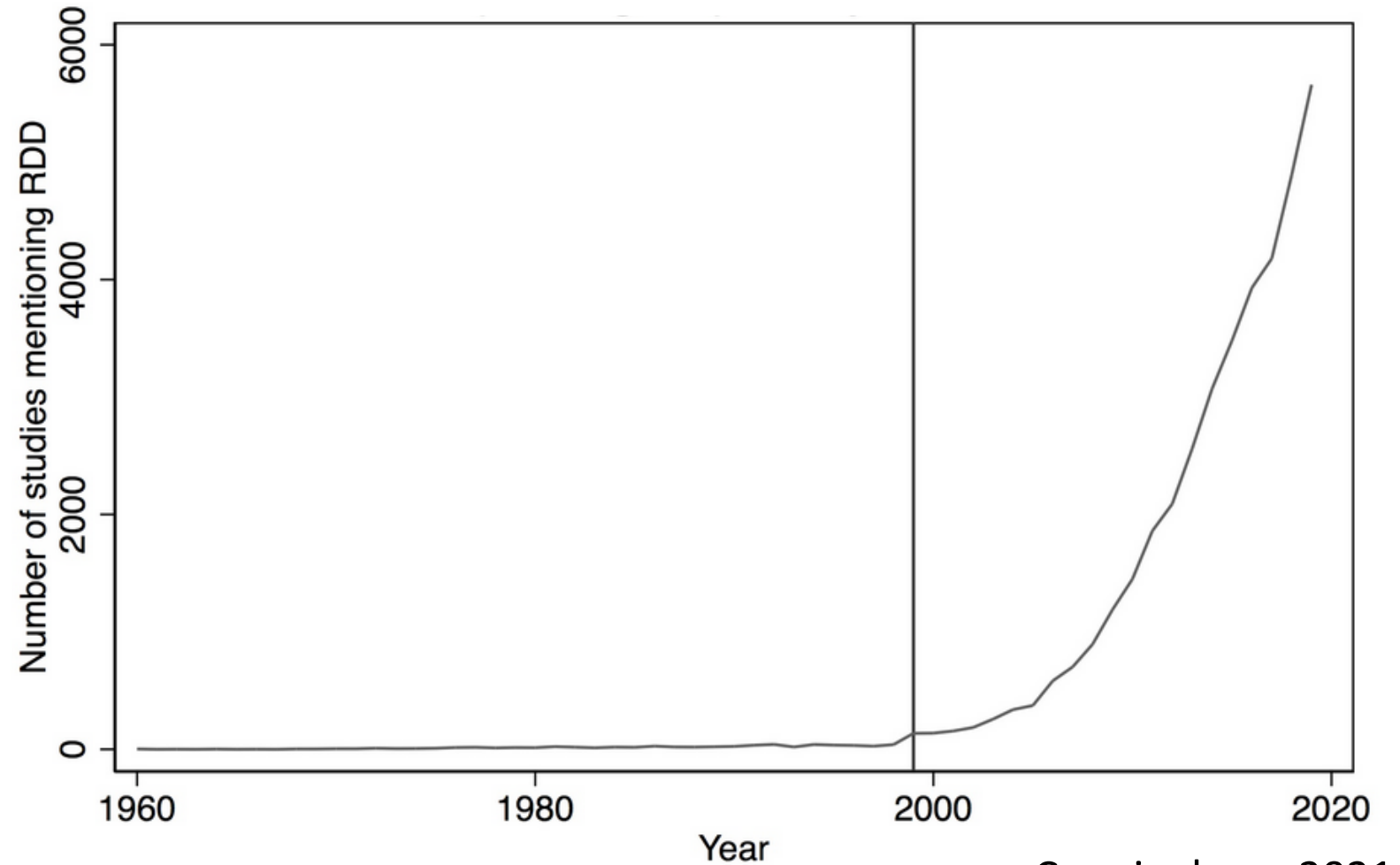
Cunningham, 2021



# Regressions-Diskontinuitäts-Analysen (RDA)



Donald T. Campbell



Cunningham, 2021



# Regressions-Diskontinuitäts-Analysen in der Evaluation psychologischer Interventionen





Am besten geeignet bei Regeln oder Vorgehen  
mit klaren Grenzwerten:

- Hochbegabtenförderung ab **IQ von 130**
- Medikamentöse Therapie i.d.R. bei schwerer Depression (**BDI  $\geq$  29**)
- Substitution bei Suchterkrankungen **ab bestimmten Marker-Scores**
- ...



HESSEN



**Hessisches  
Kultusministerium**

Die Hessische Landesregierung implementiert an Modellschulen verpflichtenden Nachhilfeunterricht für Kinder mit geringem Haushaltseinkommen.

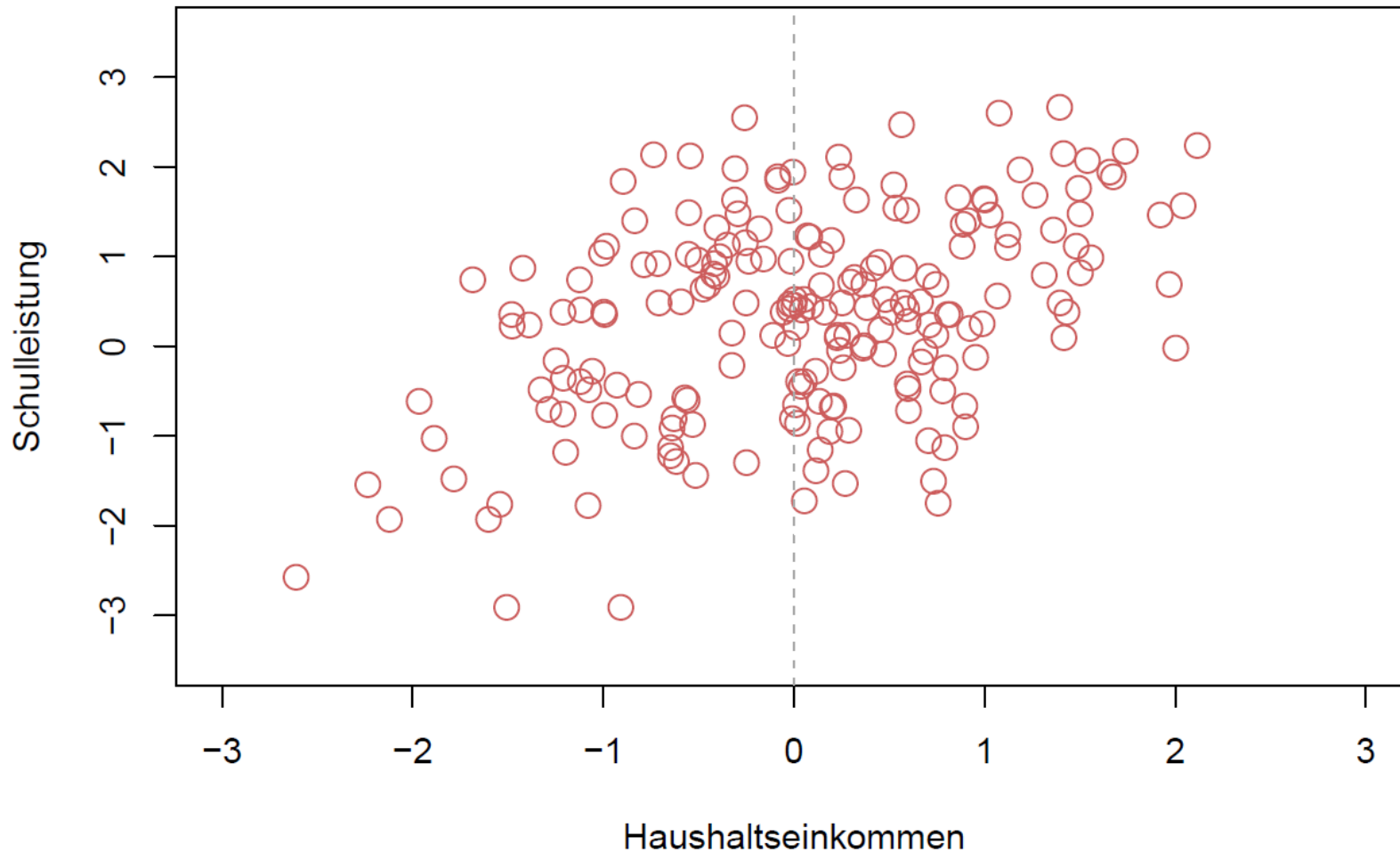
Führt diese Intervention zu verbesserten Schulleistungen?



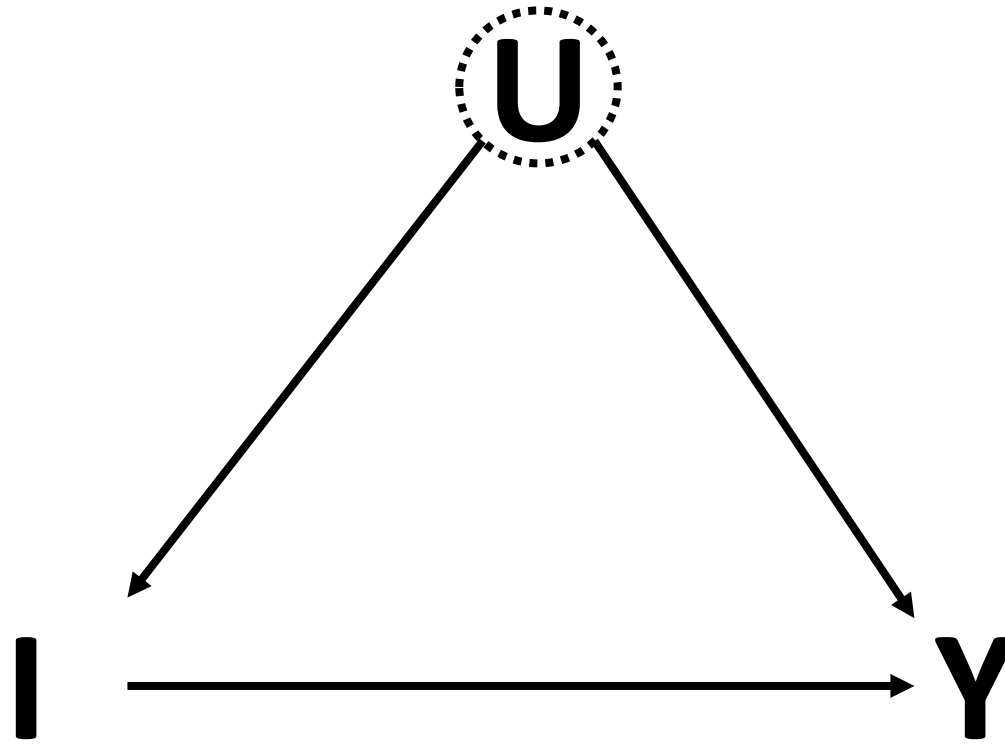
```
N <- 200 #Stichprobengröße
X <- rnorm(N) #Haushaltseinkommen ("running variable")
c0 <- 0 #Cutoff wert
I <- ifelse(X < c0, 1, 0) #Dummy Variable für Intervention
b_X <- 1 #Effekt von Haushaltseinkommen
b_I <- 1 #Effekt der Intervention
Y <- rnorm(N, b_X*X + b_I*I, 1) #Simuliere Schulleistung
```



```
N <- 200                                #Stichprobengröße
X <- rnorm(N)                            #Haushaltseinkommen ("running variable")
c0 <- 0                                  #Cutoff wert
I <- ifelse(X < c0, 1, 0)                 #Dummy Variable für Intervention
b_X <- 1                                 #Effekt von Haushaltseinkommen
b_I <- 1                                 #Effekt der Intervention
Y <- rnorm(N,b_X*X + b_I*I, 1)          #Simuliere Schulleistung
```



# Directed Acyclic Graphs (DAGs)

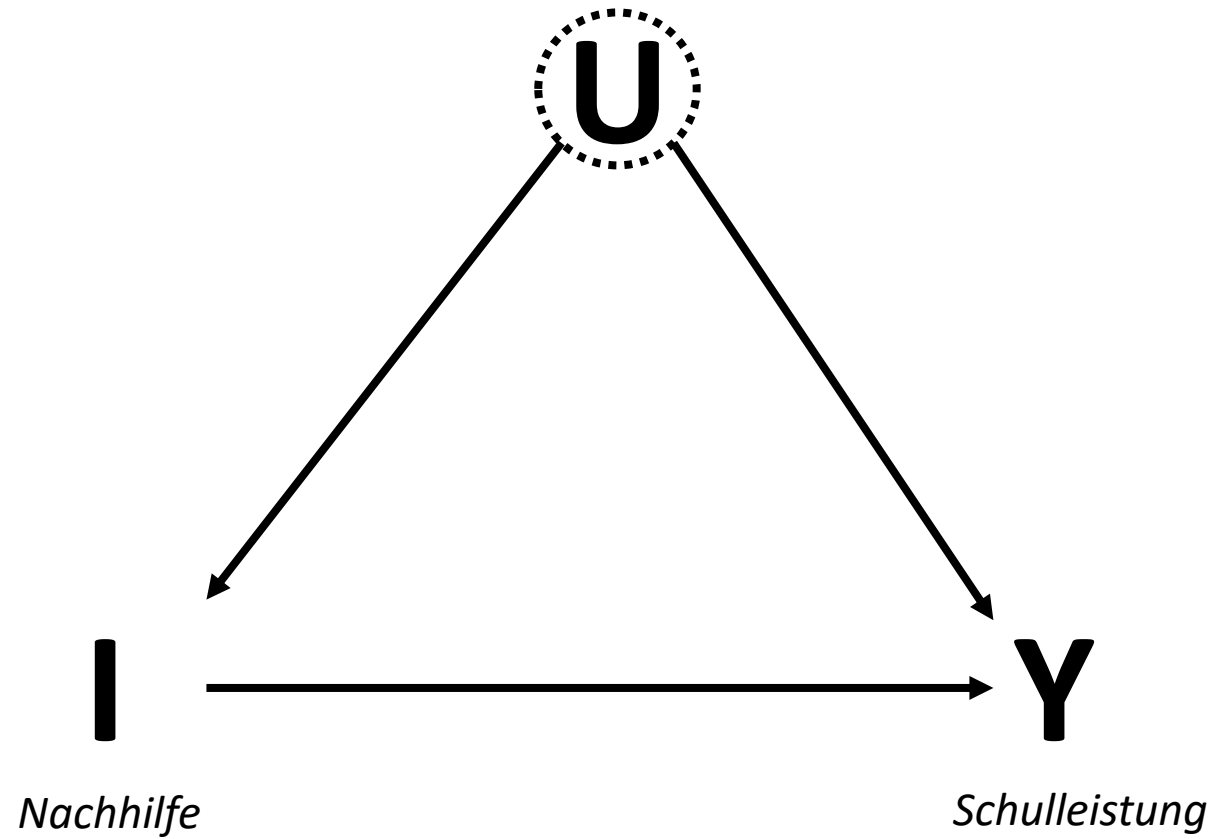


$$I = f(U)$$

$$Y = f(I, U)$$



*SES, Wohnort, soziales Netzwerk, Intelligenz, Kultur.....*



*EvaluatorInnen*

*SES, Wohnort, soziales Netzwerk, Intelligenz, Kultur.....*

**E**

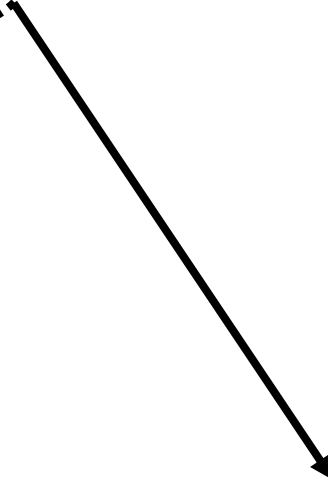
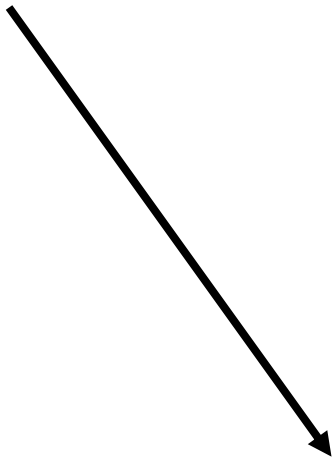
**U**

**I**

**Y**

*Nachhilfeintervention*

*Schulleistung*





*Evaluationsteam*

*SES, Wohnort, soziales Netzwerk, Intelligenz, Kultur.....*

**E** HESSEN



*Nachhilfeintervention*

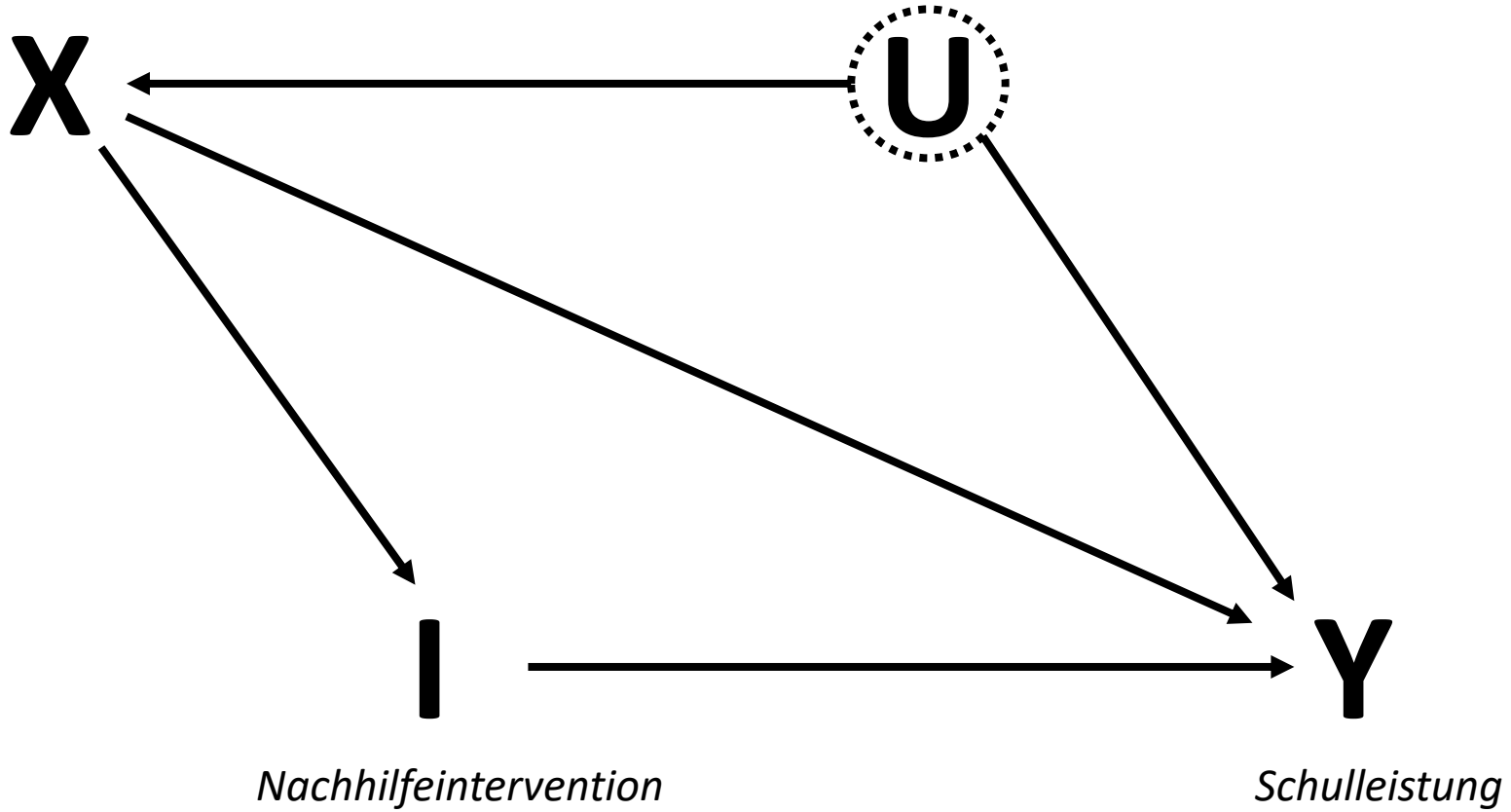


**Hessisches  
Kultusministerium**

*Schulleistung*

*Haushaltseinkommen*

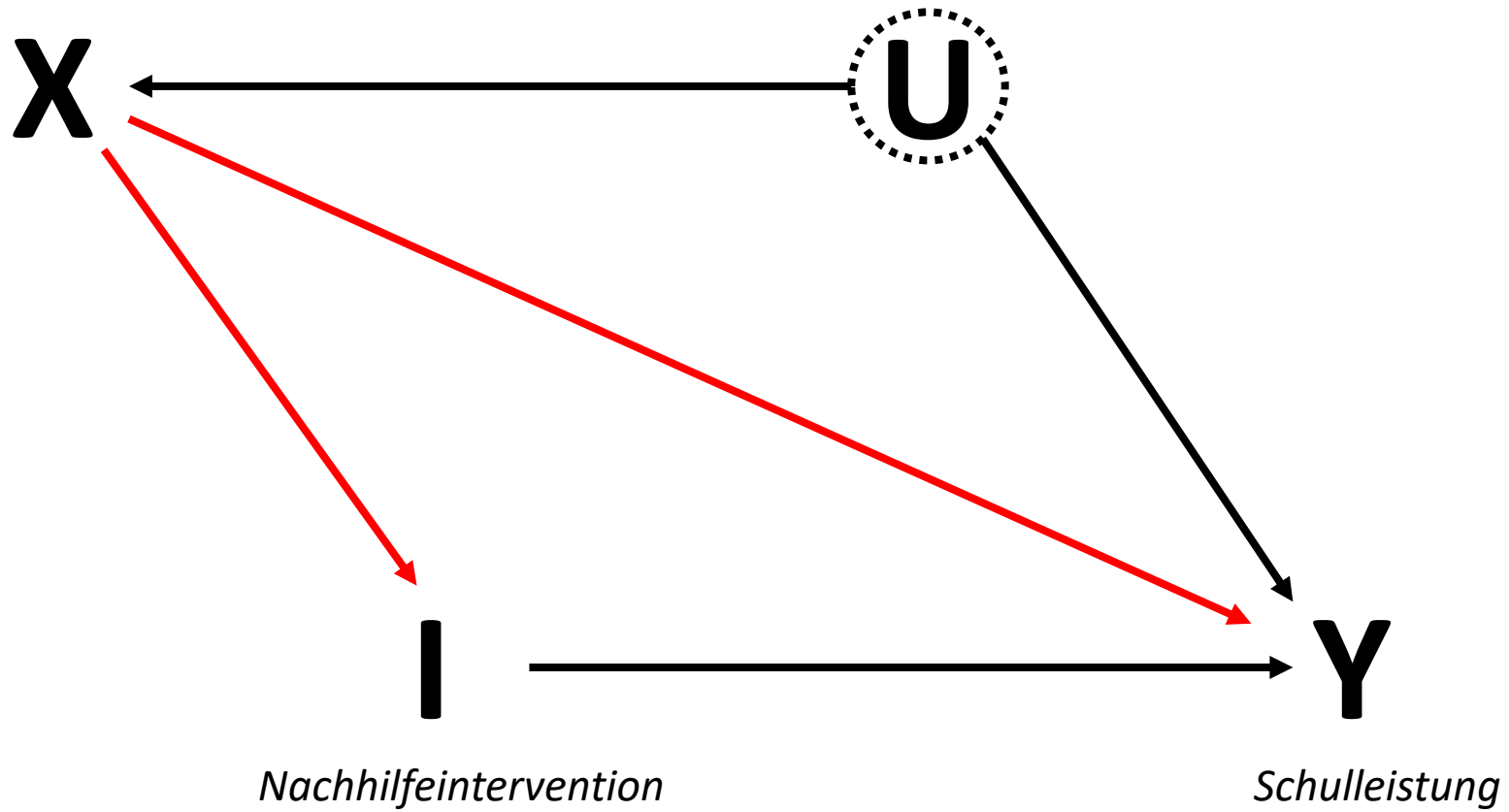
*SES, Wohnort, soziales Netzwerk, Intelligenz, Kultur.....*





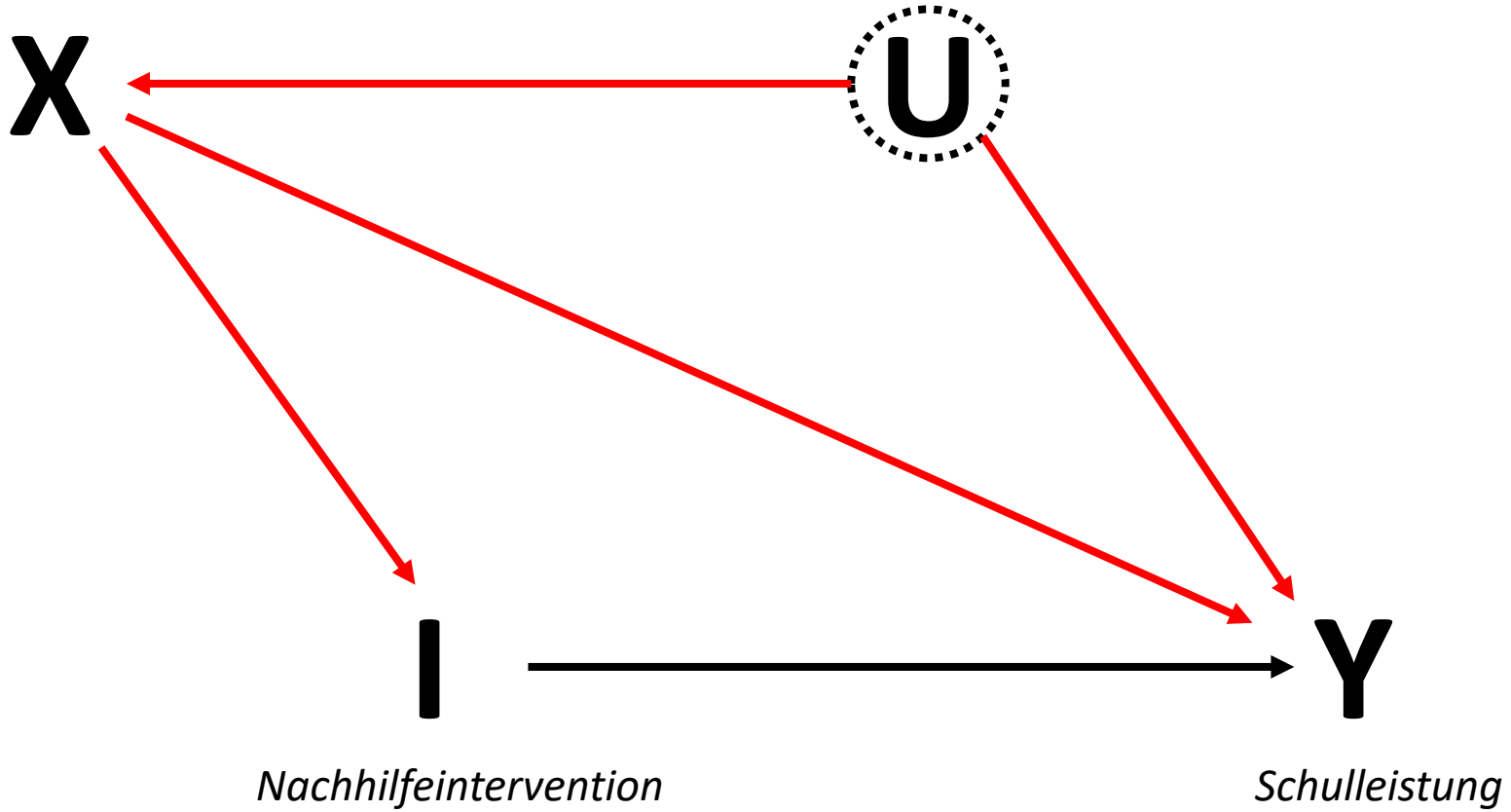
*Haushaltseinkommen*

*SES, Wohnort, soziales Netzwerk, Intelligenz, Kultur.....*

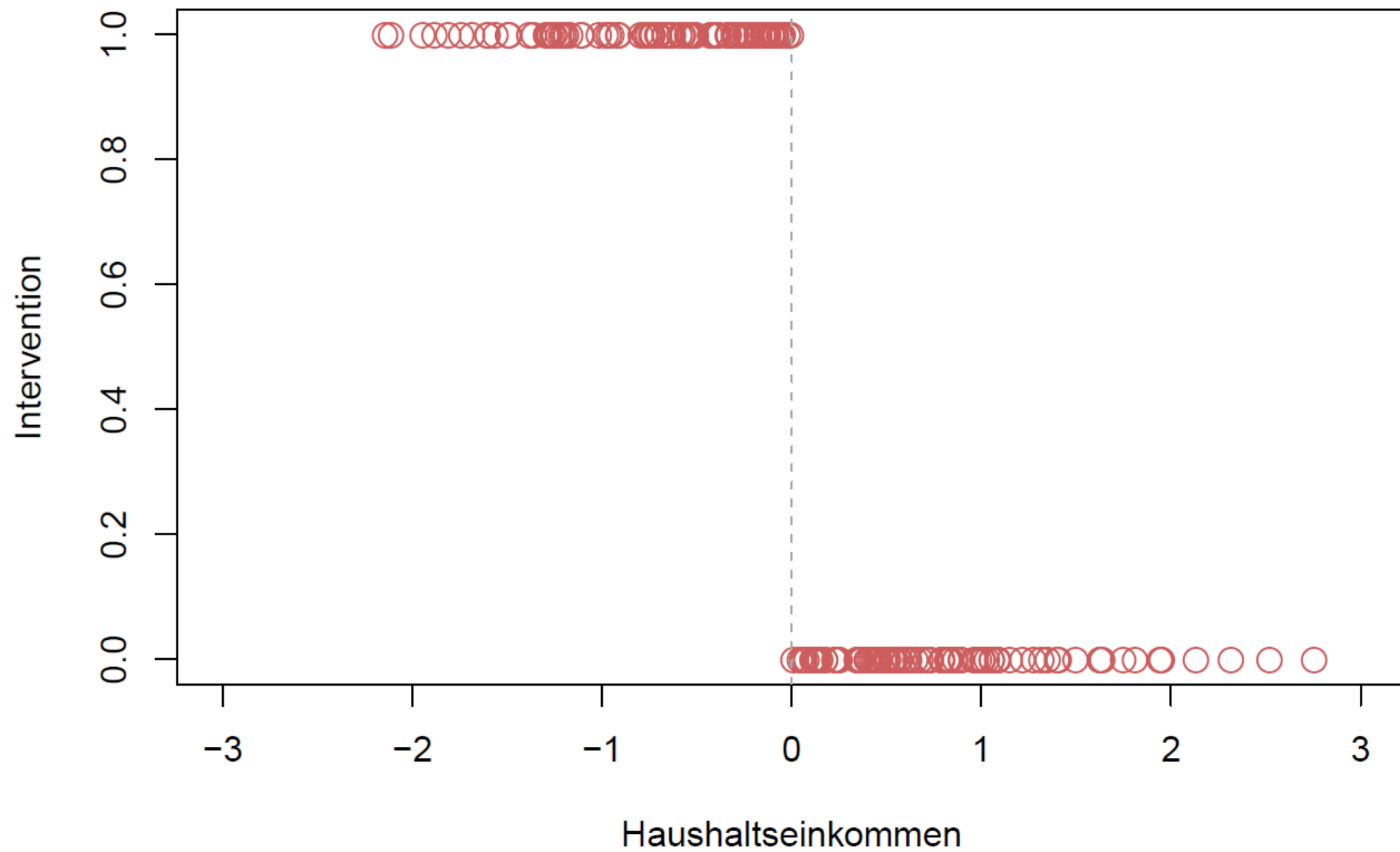


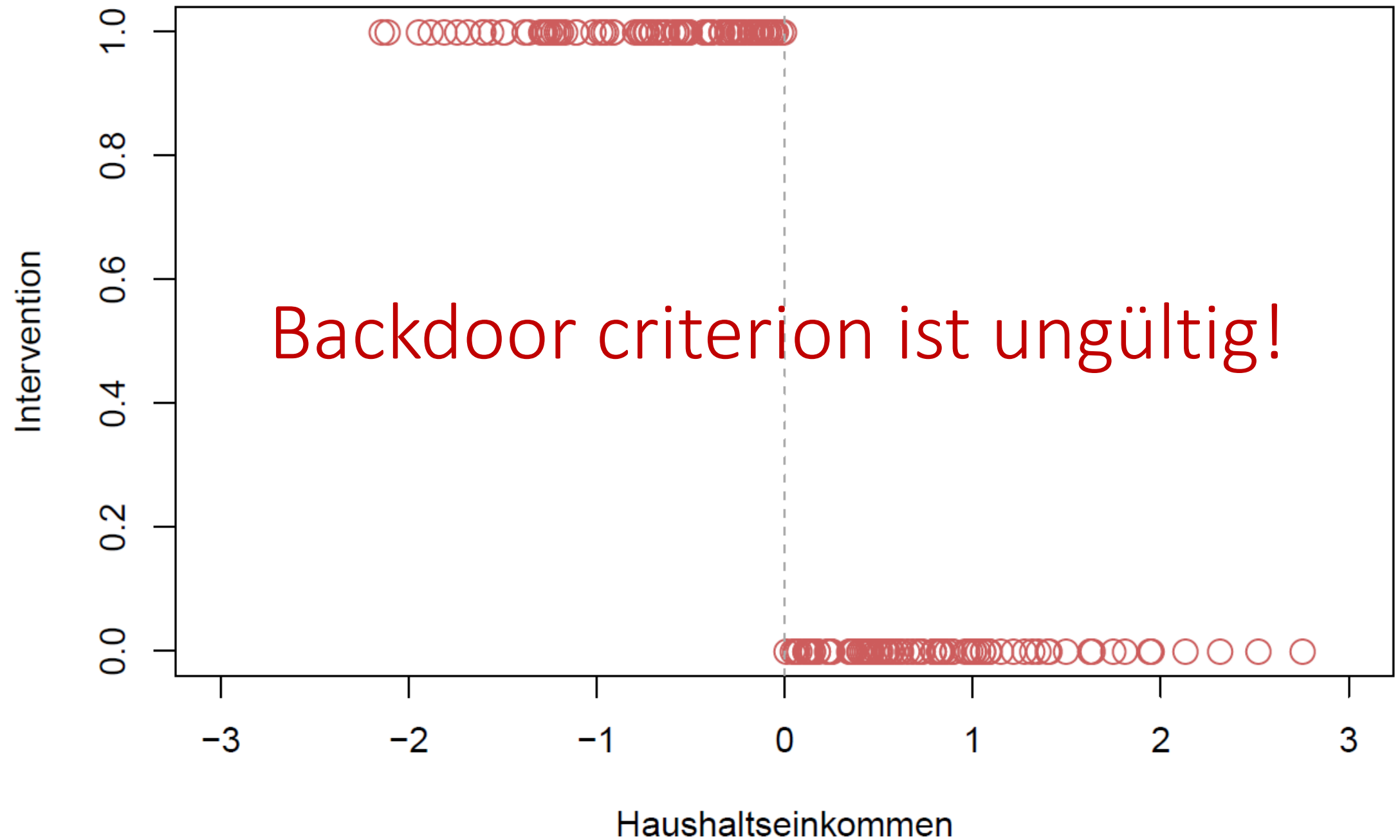
*Haushaltseinkommen*

*SES, Wohnort, soziales Netzwerk, Intelligenz, Kultur.....*











*EvaluatorInnen*

*SES, Wohnort, soziales Netzwerk, Intelligenz, Kultur.....*

**E**

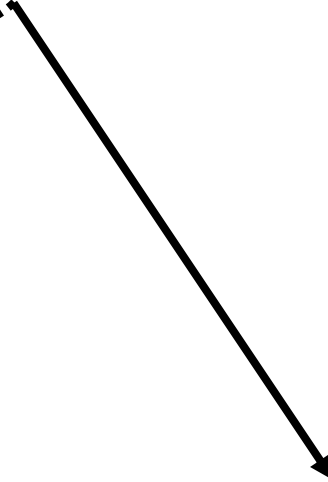
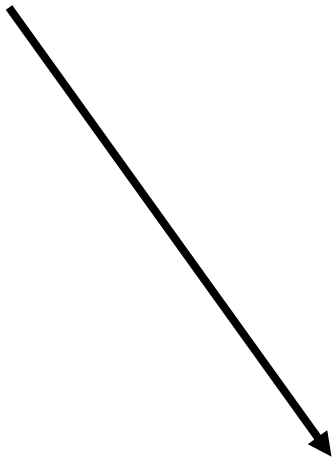
**U**

**I**

**Y**

*Nachhilfeintervention*

*Schulleistung*



*Haushaltseinkommen  
am cutoff*

*SES, Wohnort, soziales Netzwerk, Intelligenz, Kultur.....*

**X -> c0**

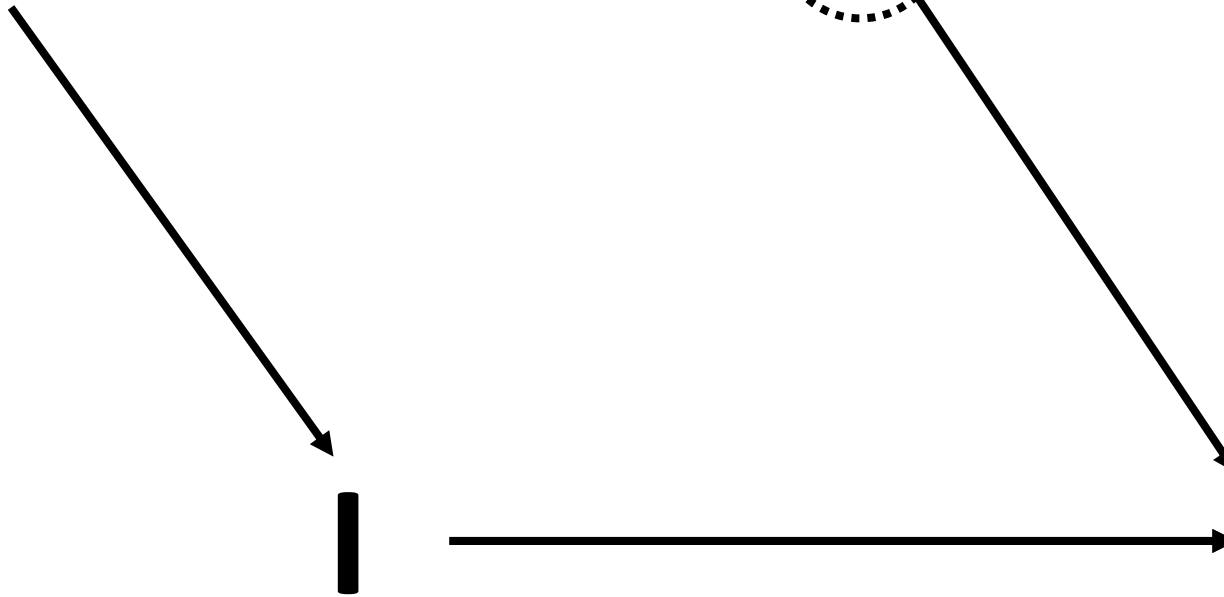
**U**

**I**

**Y**

*Nachhilfeintervention*

*Schulleistung*

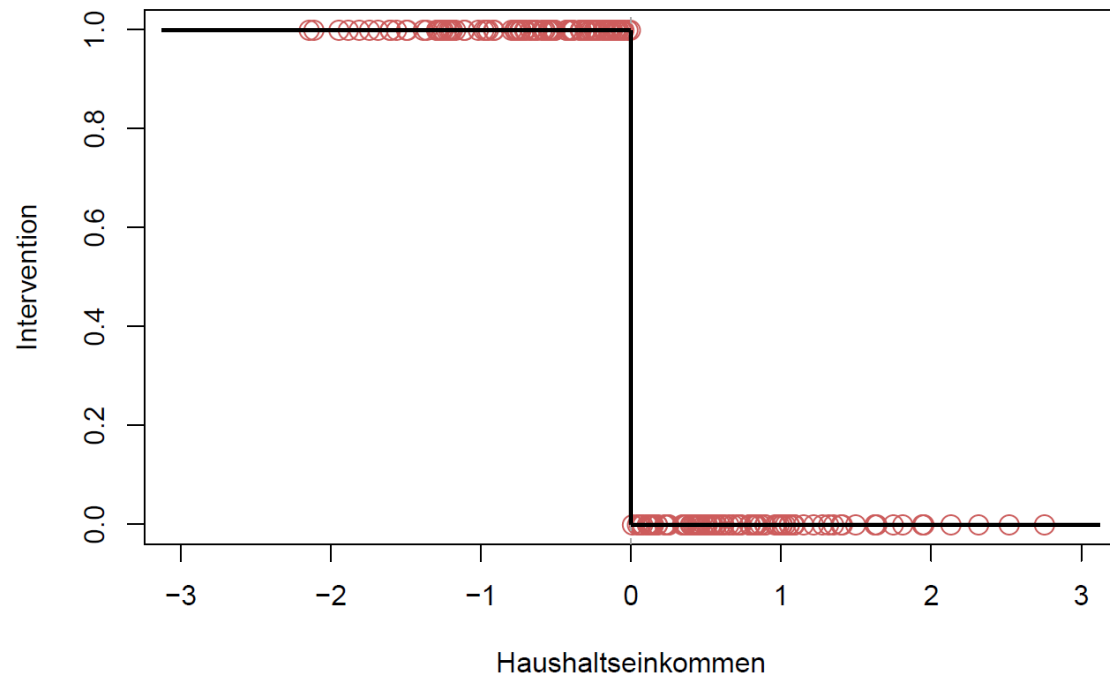




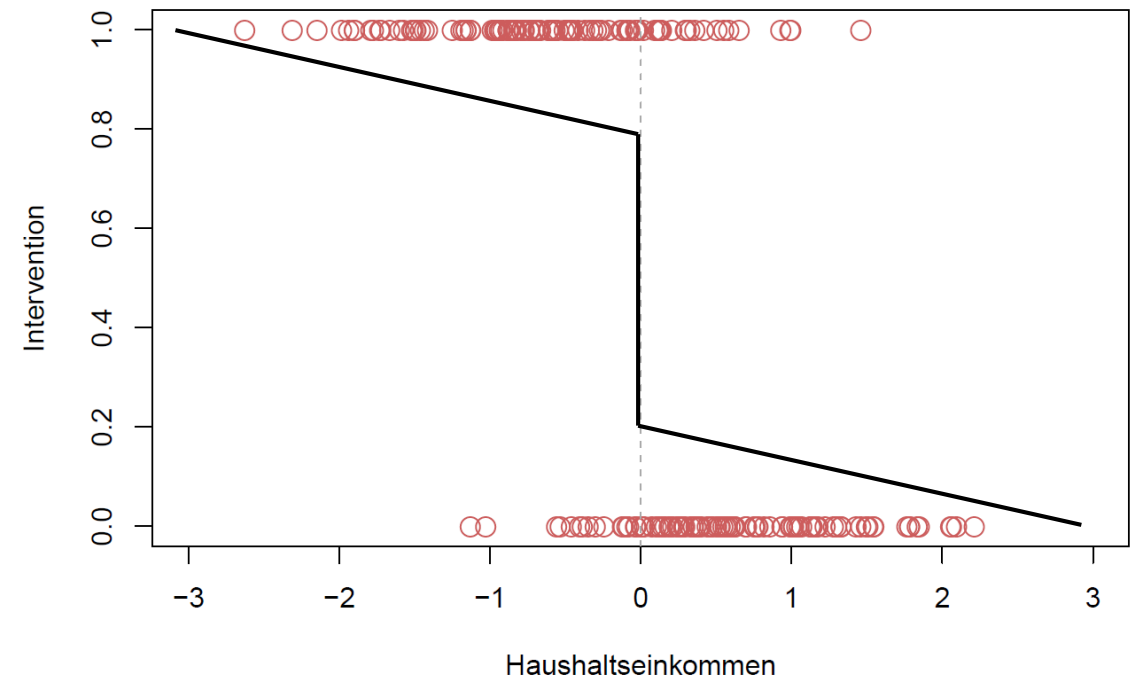




Scharfe (“sharp”) RD-Analyse

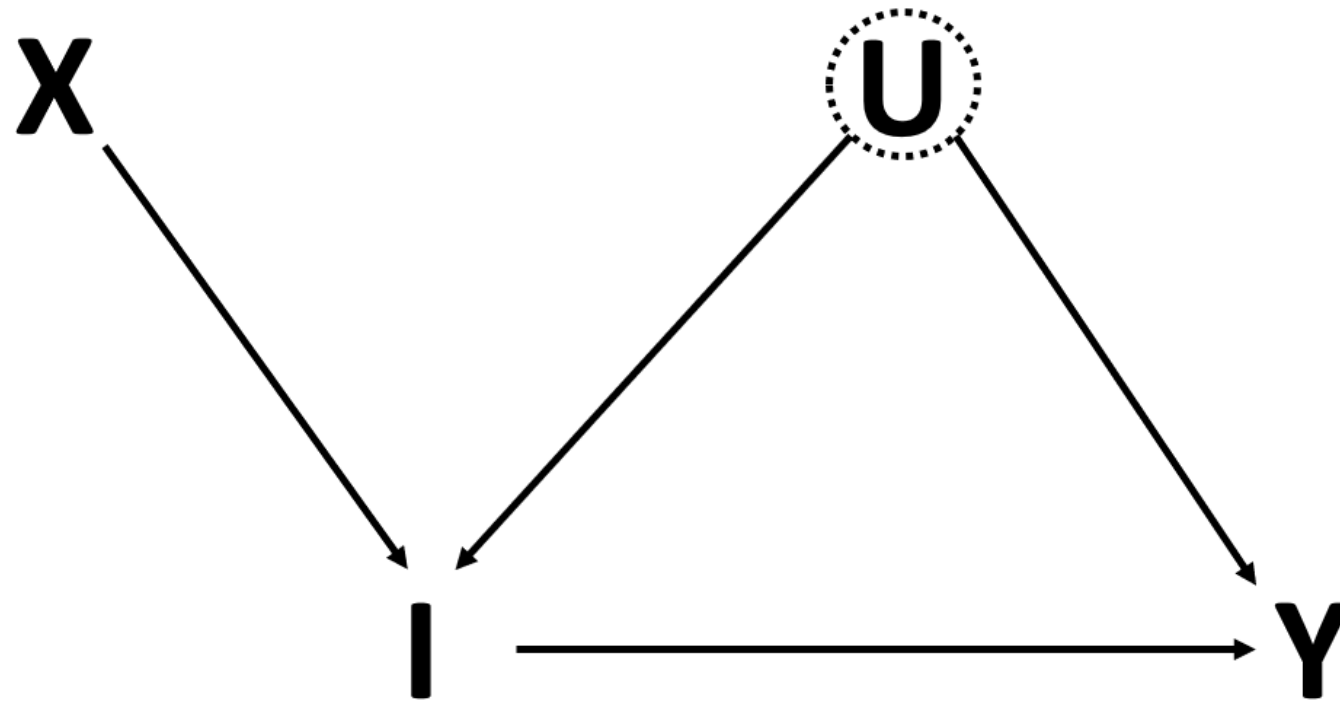


vs. Unscharfe (“fuzzy”) RD-Analyse

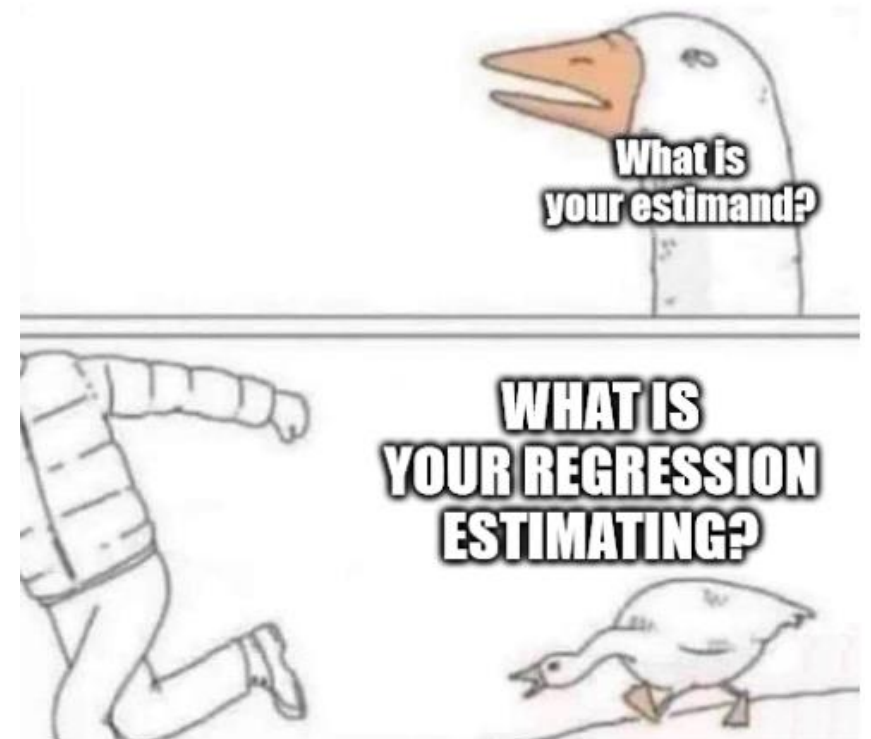




Nächste Woche: Instrumentvariablenschätzung

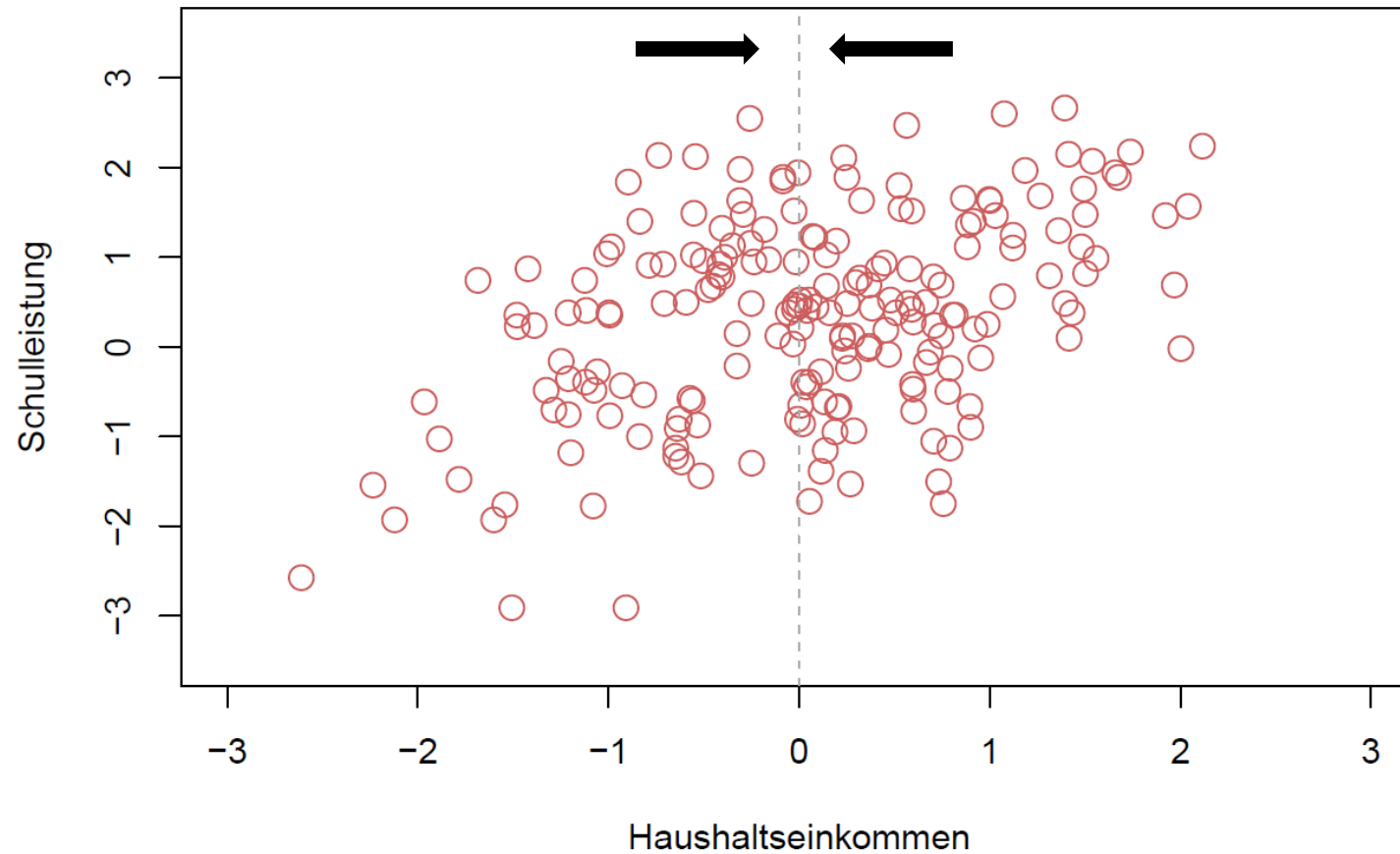


Was wollen wir schätzen?



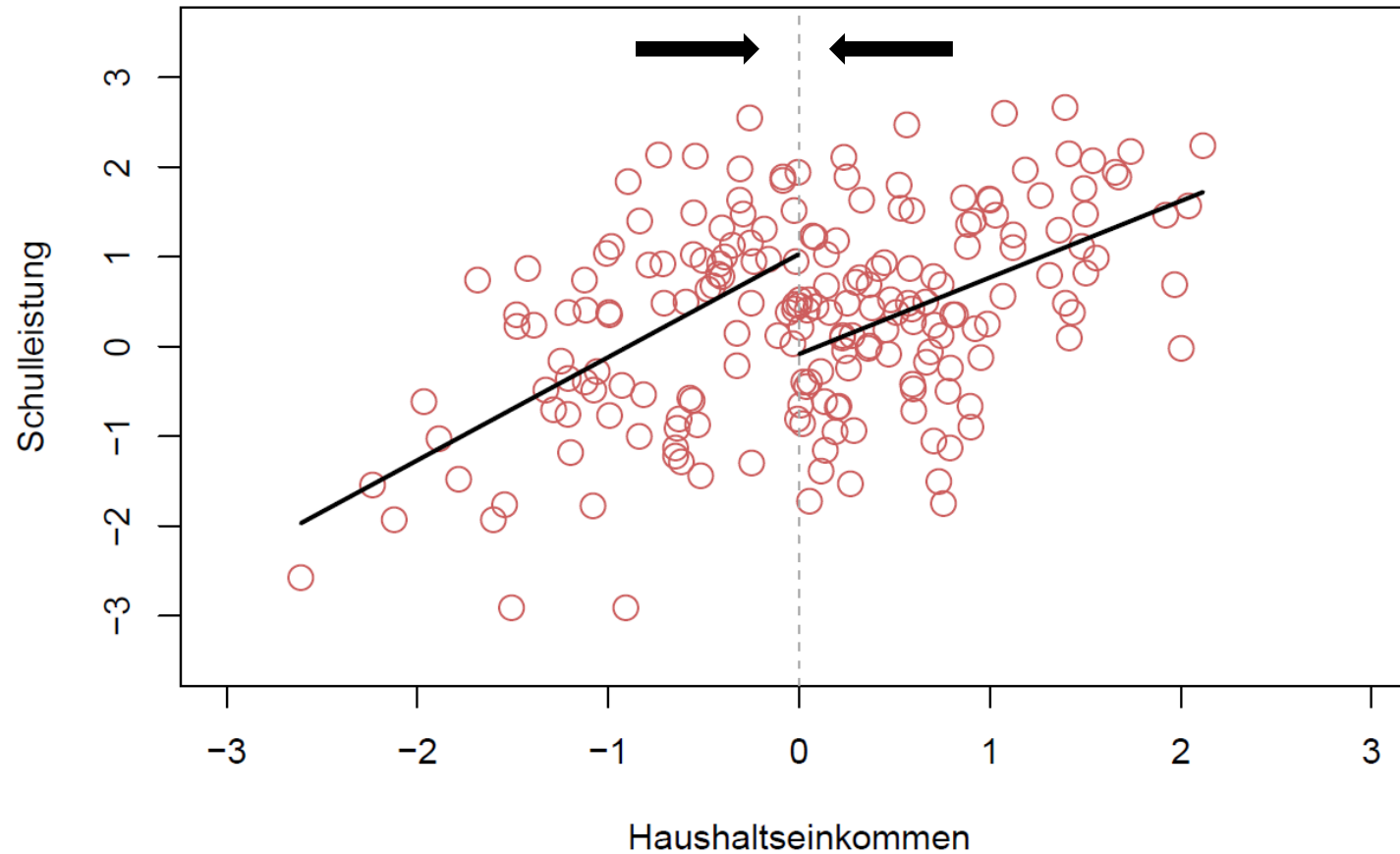
@krichard1212

# Local average treatment effect (LATE)



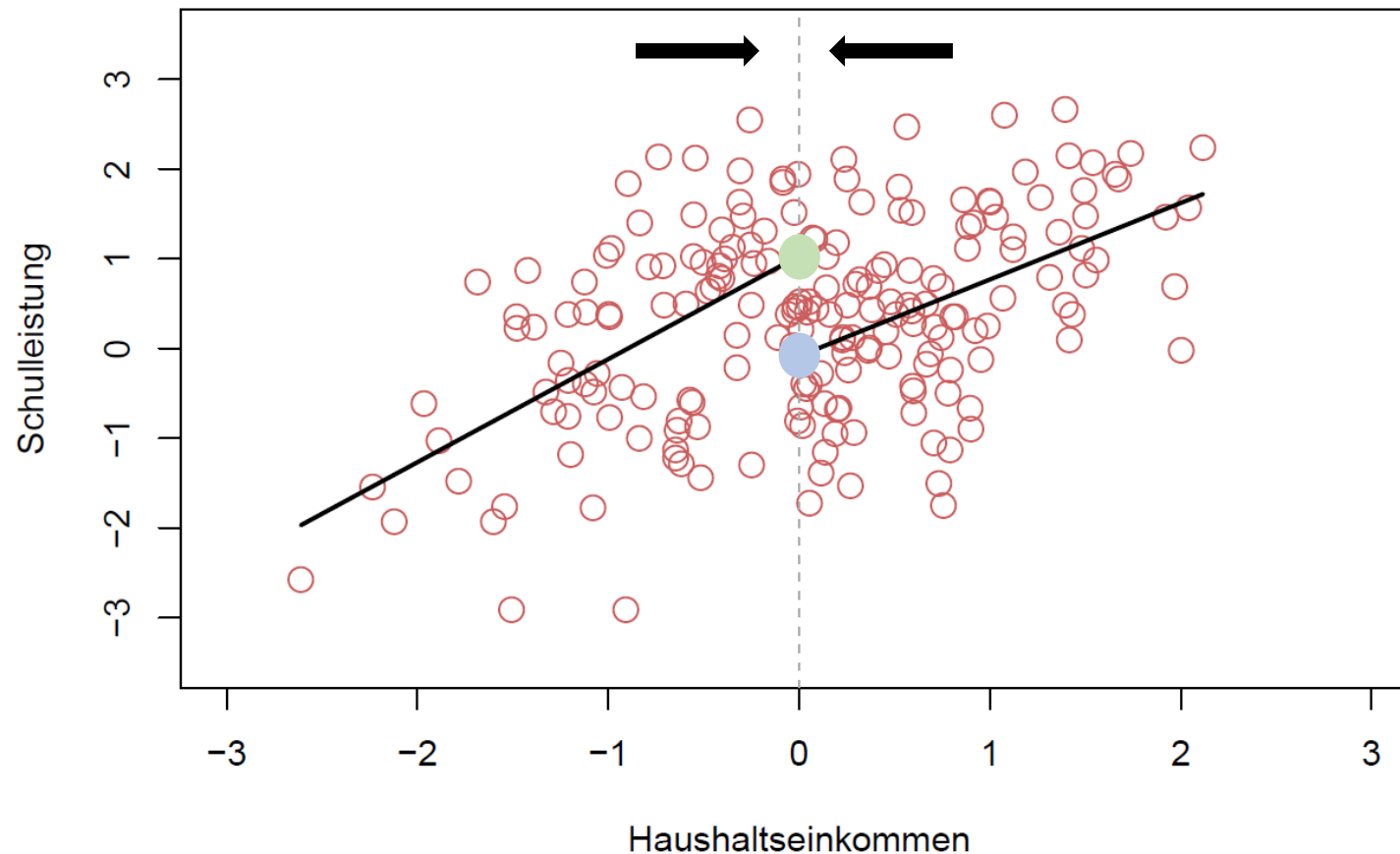


# Local average treatment effect (LATE)

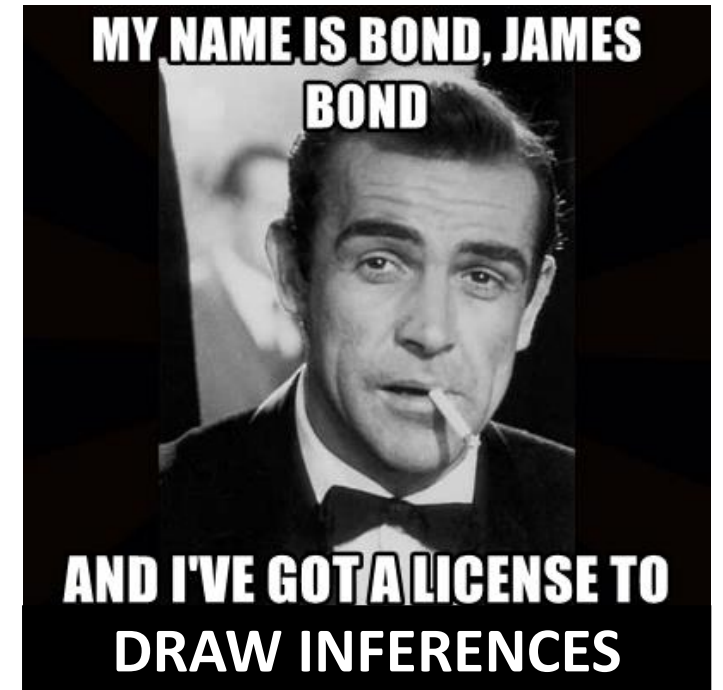
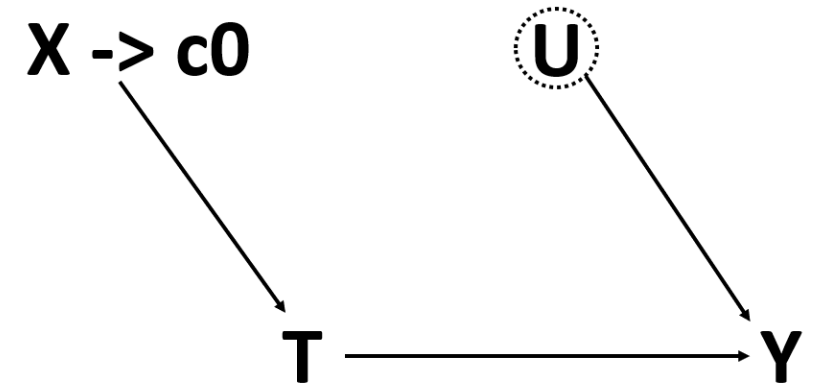


# Local average treatment effect (LATE)

$$\lim_{x \rightarrow 0} E[Y^1 | x] - \lim_{0 \leftarrow x} E[Y^0 | x]$$



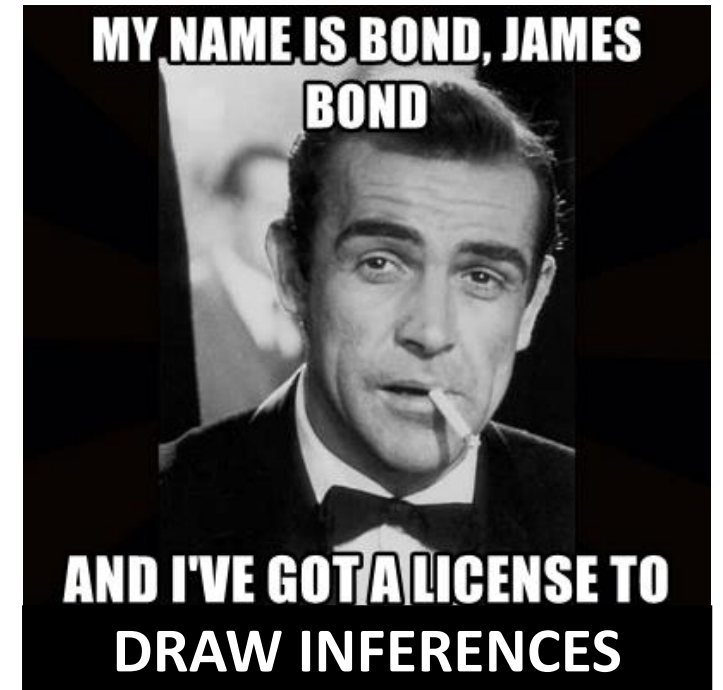
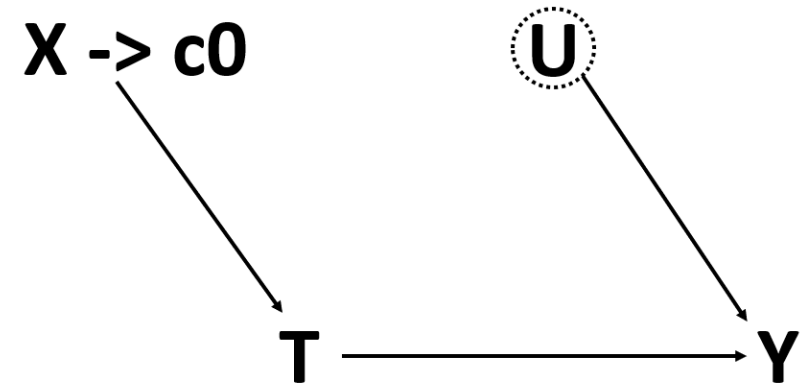
# Annahmen des LATE





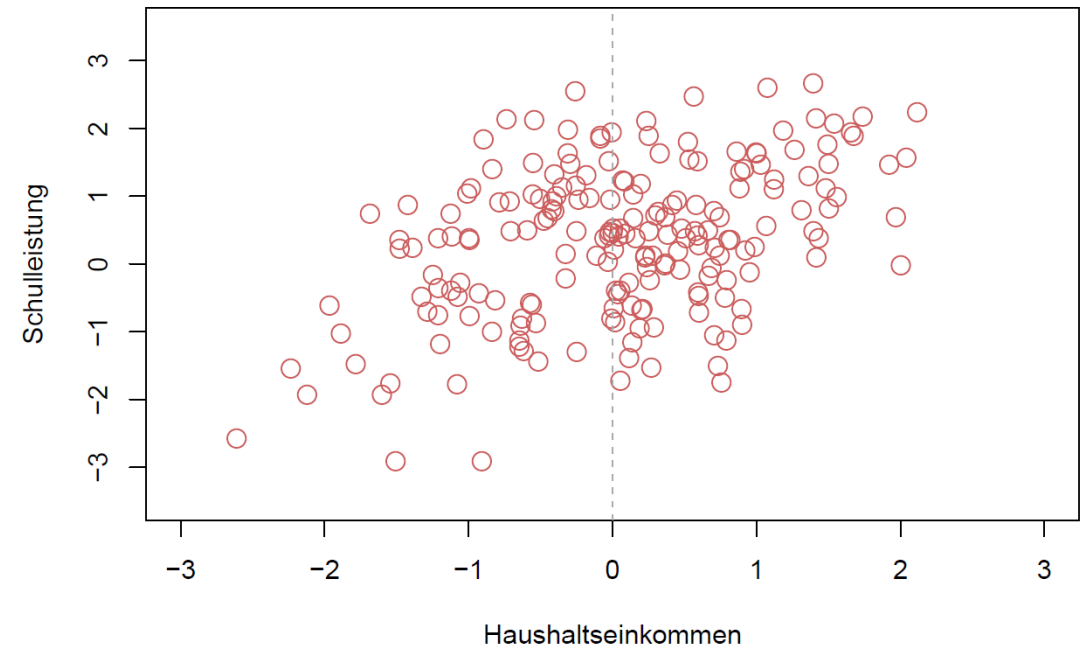
# Annahmen des LATE

- Kontinuität relevanter Kovariaten am Cutoff
- Kein “Sortieren” über den Cutoff hinweg



# Schätzung des LATE

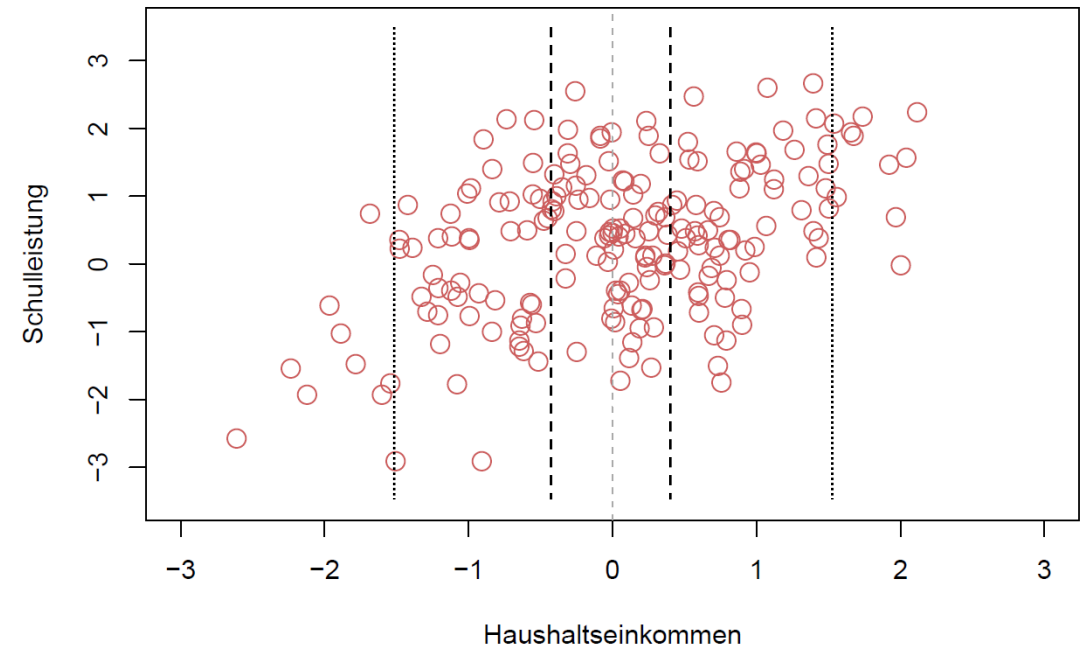
- Wie “lokal” soll die Schätzung sein?  
mehr Daten vs. präzisere Kontrolle



# Schätzung des LATE

- Wie “lokal” soll die Schätzung sein?  
mehr Daten vs. präzisere Kontrolle

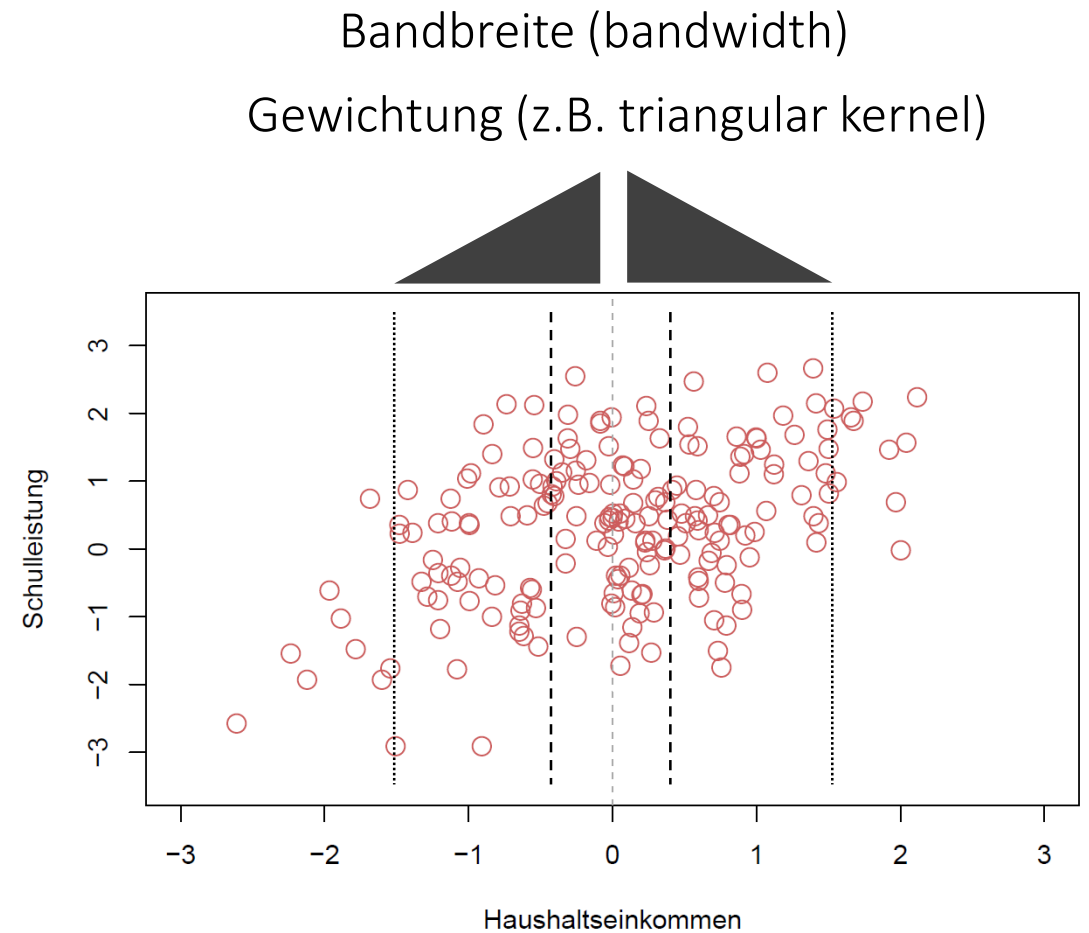
Bandbreite (bandwidth)





# Schätzung des LATE

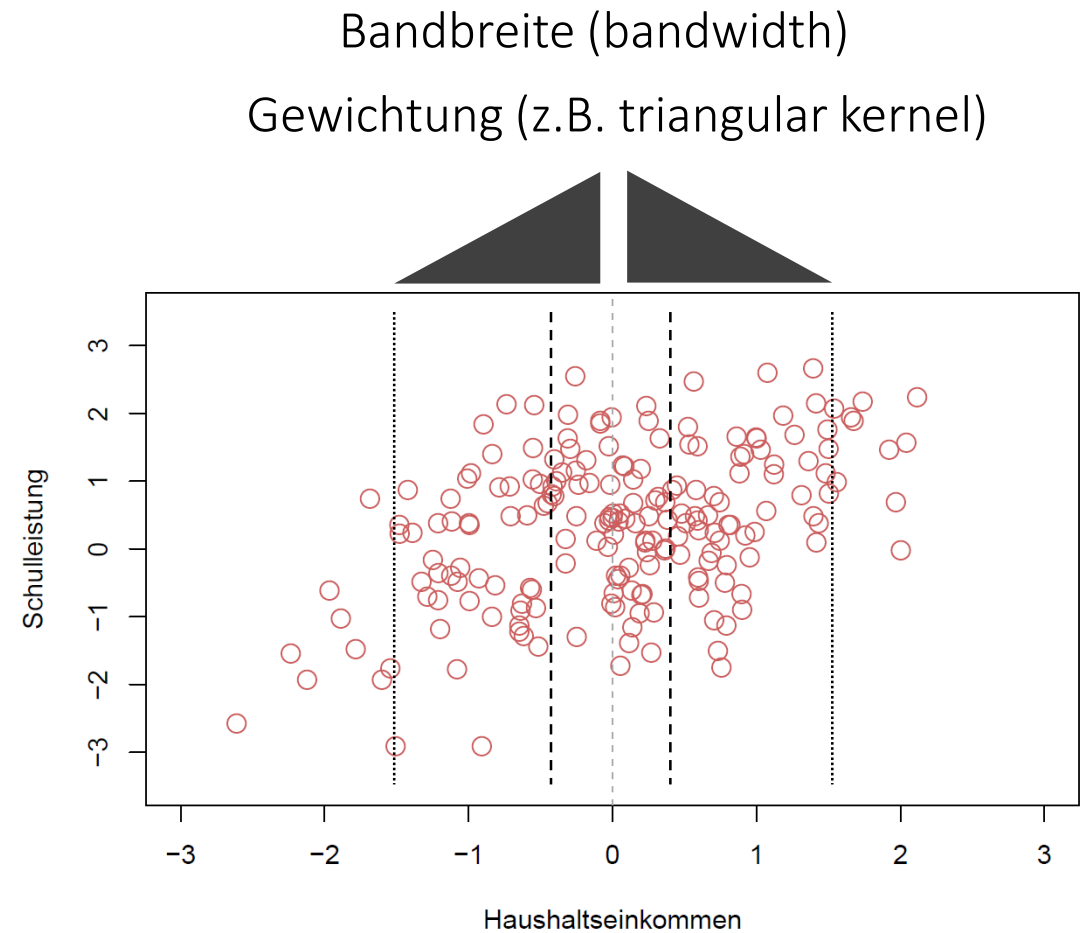
- Wie “lokal” soll die Schätzung sein?  
mehr Daten vs. präzisere Kontrolle



# Schätzung des LATE

- Wie “lokal” soll die Schätzung sein?  
mehr Daten vs. präzisere Kontrolle

In Praxis oft datengetriebene Methoden



# Schätzung des LATE

- Wie “lokal” soll die Schätzung sein?  
mehr Daten vs. präzisere Kontrolle
- Welche Form nehmen wir für den  
Zusammenhang an?  
overfitting vs. underfitting

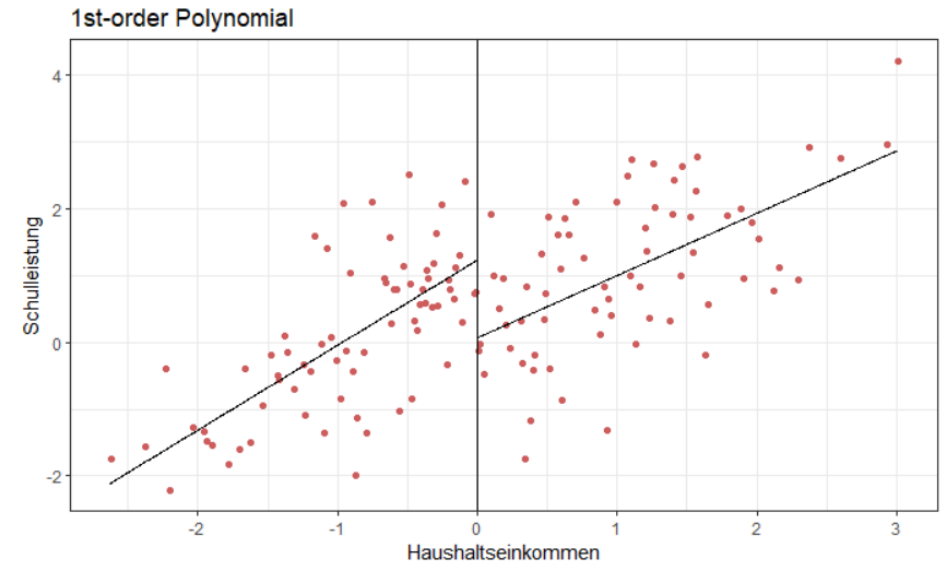


# Schätzung des LATE

- Wie “lokal” soll die Schätzung sein?  
mehr Daten vs. präzisere Kontrolle
- Welche Form nehmen wir für den  
Zusammenhang an?  
overfitting vs. underfitting

```
library(rdrobust)
llr <- rdrobust::rdrobust(Y,X,c = 0,p=1, kernel = "tri",bwselect = "mserd")
```

“triangular” “Mean Square Error”

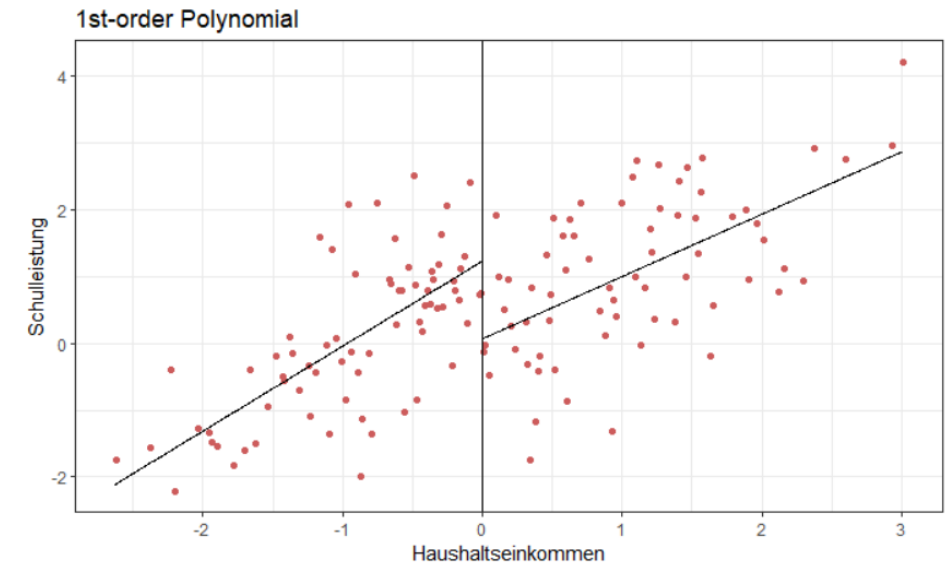


# Schätzung des LATE

- Wie “lokal” soll die Schätzung sein?  
mehr Daten vs. präzisere Kontrolle
- Welche Form nehmen wir für den  
Zusammenhang an?  
overfitting vs. underfitting

```
library(rdrobust)
llr <- rdrobust::rdrobust(Y,X,c = 0,p=1, kernel = "tri",bwselect = "mserd")
```

“triangular” “Mean Square Error”



Method	Coef.	Std. Err.	z	P> z	[ 95% C.I. ]
Conventional	-1.016	0.502	-2.023	0.043	[-2.000 , -0.032]
Robust	-	-	-1.818	0.069	[-2.262 , 0.085]

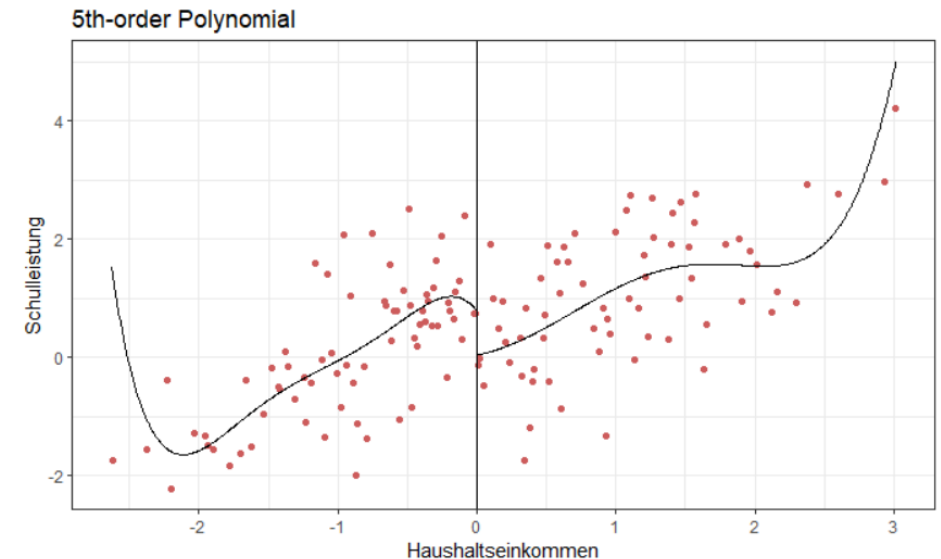
```
N <- 200 #Stichprobengröße
X <- rnorm(N) #Haushaltseinkommen ("running variable")
c0 <- 0 #Cutoff Wert
I <- ifelse(X < c0, 1, 0) #Dummy Variable für Intervention
b_X <- 1 #Effekt von Haushaltseinkommen
b_I <- 1 #Effekt der Intervention
Y <- rnorm(N,b_X*X + b_I*I, 1) #Simuliere Schulleistung
```

# Schätzung des LATE

- Wie “lokal” soll die Schätzung sein?  
mehr Daten vs. präzisere Kontrolle
- Welche Form nehmen wir für den  
Zusammenhang an?  
overfitting vs. underfitting

“triangular” “Mean Square Error”

```
library(rdrobust)
llr <- rdrobust::rdrobust(Y,X,c = 0,p=5, kernel = "tri",bwselect = "mserd")
```



Method	Coef.	Std. Err.	z	P> z	[ 95% C.I. ]
Conventional	-0.661	0.915	-0.723	0.470	[-2.454 , 1.131]
Robust		-	-0.706	0.480	[-2.612 , 1.228]

```
N <- 200 #Stichprobengröße
X <- rnorm(N) #Haushaltseinkommen ("running variable")
c0 <- 0 #Cutoff wert
I <- ifelse(X < c0, 1, 0) #Dummy Variable für Intervention
b_X <- 1 #Effekt von Haushaltseinkommen
b_I <- 1 #Effekt der Intervention
Y <- rnorm(N,b_X*X + b_I*I, 1) #Simuliere Schulleistung
```



# Praktische Übung in R



Lehrprobe\_RDA Private

main 1 Branch 0 Tags

Go to file Add file Code

DominikDeffner Include tasks and script 2e977fa · 17 minutes ago 2 Commits

- RDA\_simulation.R Include tasks and script 17 minutes ago
- README.md Include tasks and script 17 minutes ago

README

## Lehrprobe\_RDA

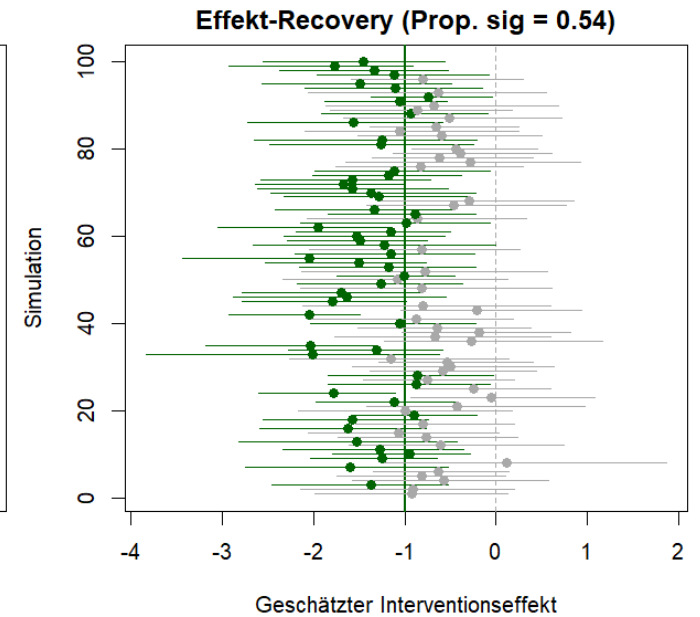
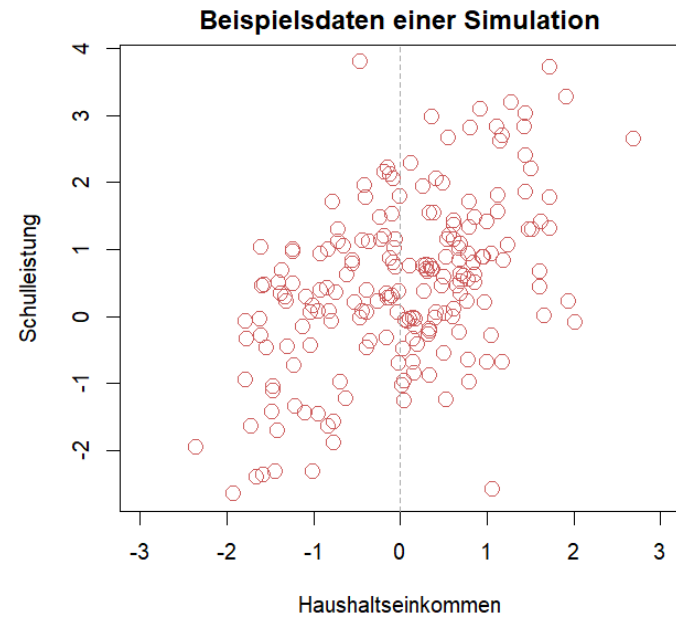
Liebe Studierende,

hier finden Sie die begleitenden Aufgaben zur Regressions-Diskontinuitäts-Analyse.

Öffnen Sie das Skript "RDA\_simulation" in R Studio und installieren Sie das Package "rdrbust" ("install.packages("rdrbust")" in der Konsole ausführen).

Bearbeiten Sie bitte selbstständig folgende Aufgaben. Dazu müssen Sie die Parameterwerte im oberen Teil des Skripts ändern, das gesamte Skript laufen lassen, und dann die Ergebnisgraphik interpretieren. Setzen Sie bitte nach jeder Aufgabe die Parameter wieder auf ihre Ursprungswerte zurück.

(1) Bedeutung der Stichprobengröße: Variieren Sie systematisch "N" und beschreiben Sie, was das für die Schätzung des Interventionseffekts bedeutet.



[https://github.com/DominikDeffner/Lehrprobe\\_RDA/](https://github.com/DominikDeffner/Lehrprobe_RDA/)