University of Bremen

A Brief Overview of

P-Values

Nate Breznau

Lecture prepared for

"Open Science in Social Sciences"

BA Blockseminar

University of Zurich, Fall 2021

socium

Research Center on
Inequality and Social Policy

The data and code to generate the figures in this lecture are available at

https://github.com/nbreznau/p_values

University of Bremen

What is a p-value?

- *Almost always*: A statistical test result used to evaluate the **null hypothesis**

socium
Research Center on Inequality and Social Policy

University of Bremen

What is a p-value?

- *Almost always*: A statistical test result used to evaluate the **null hypothesis**
  - Can be used to evaluate any hypothesis test (must not be null)
  - Short for **“probability”-value**
  - A test to determine if statistics calculated for a given dataset are likely *different* from some hypothesized value
  - “different” = based on sampling probability
- ‘what is the likelihood of observing this statistic if the hypothesized value is true’

University of Bremen

What is a null hypothesis?

- The claim that something measured in the world is equal to zero
  - For example:
  - That a vaccine has zero impact on infection rates (a controlled experiment)
  - That the average height difference of boys and girls at age 5 is zero (population observation)
  - That there is zero difference in intelligence between immigrants and natives (population observation)
  - That income has zero effect on happiness after adjusting for religious beliefs (population observation, w/ adjustment)

University of Bremen

What is a p-value?

- A statistical test result used to evaluate the **null hypothesis**
- It is based entirely on **probability**

socium
Research Center on
Inequality and Social Policy

University of Bremen

What is a p-value?

- A statistical test result used to evaluate the **null hypothesis**
- It is based entirely on **probability**
- It is the **probability of observing the data** assuming the null hypothesis is correct

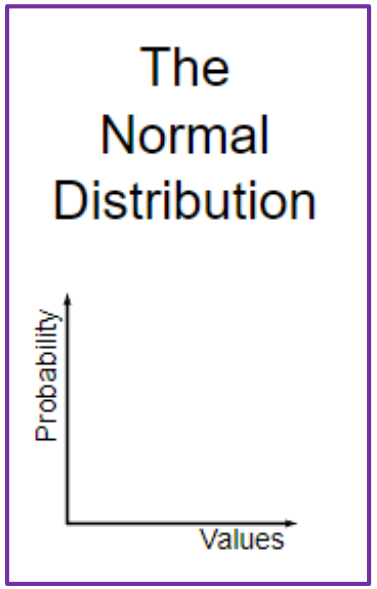
socium
Research Center on Inequality and Social Policy

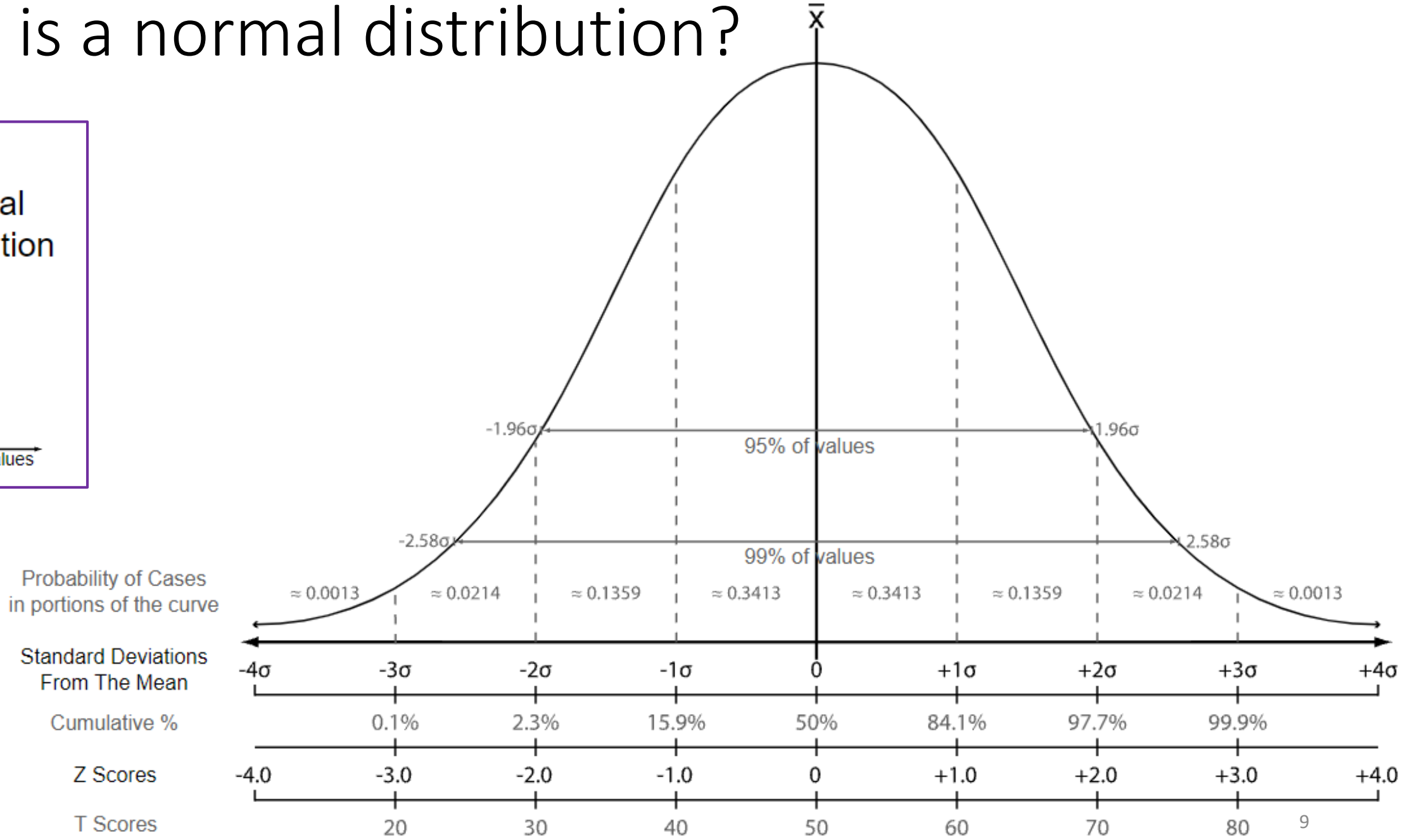
University of Bremen

What is a p-value?

- A statistical test result used to evaluate the **null hypothesis**
- It is based entirely on **probability**
- It is the **probability of observing the data** assuming the null hypothesis is correct
  - Assumes that the data were randomly sampled from a population
  - Assumes that the test is properly modeled (e.g., improper: testing if getting sick or not makes one more likely to get a vaccine)
  - Assumes a **normal distribution** of observing the null hypothesis if it were true

What is a normal distribution?





Credit:
Heds 1

What is a p-value?

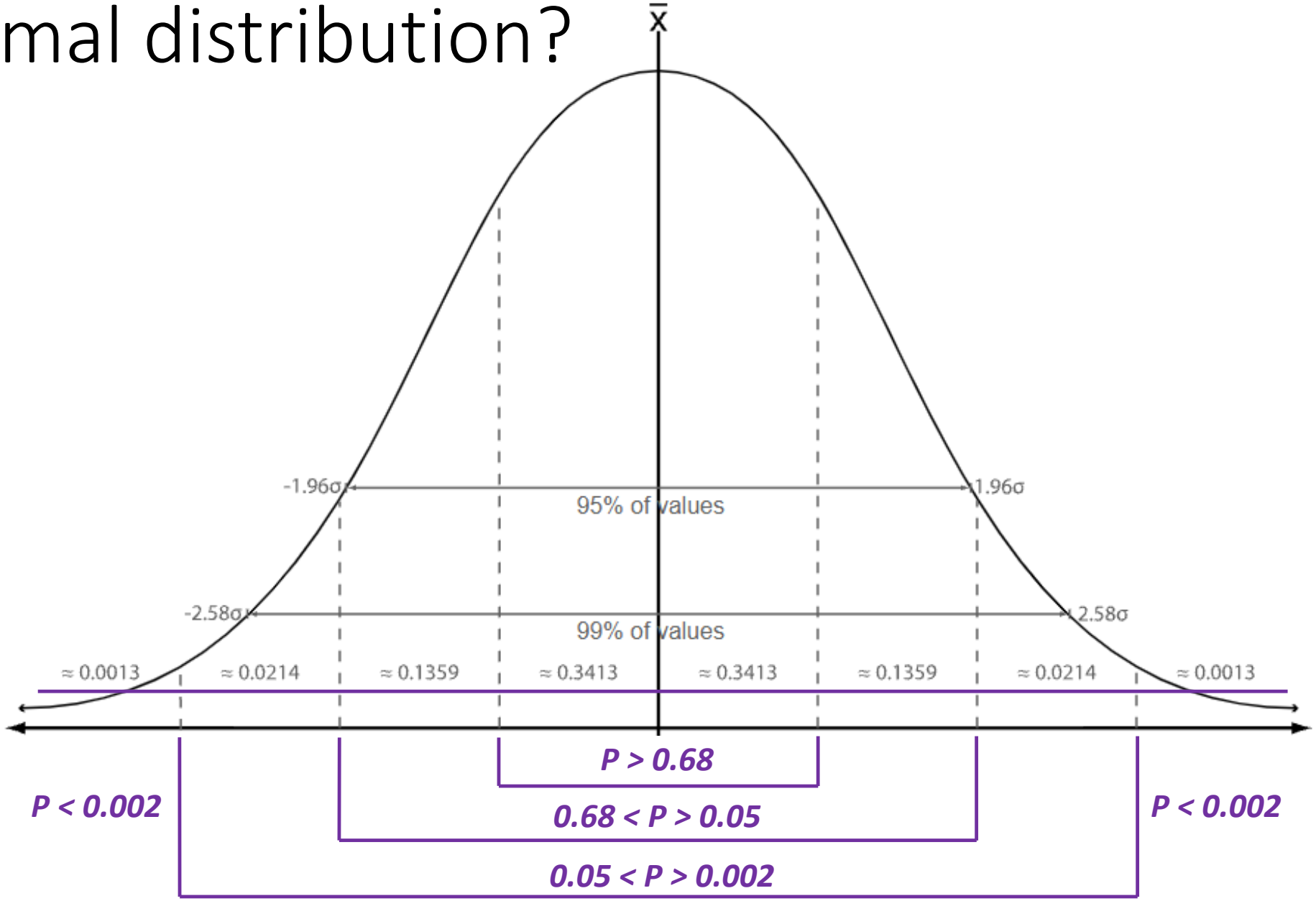
- A statistical test result used to evaluate the **null hypothesis**
- It is based entirely on **probability**
- It is the **probability of observing the data** assuming [illegible] hypothesis is correct
  - Assumes that the data were randomly sampled from a popu[illegible]n
  - Assumes that the test is properly modeled (e.g., improper: testing if getting sick or not makes one more likely to get a vaccine)
  - Assumes a **normal distribution** of observing the null hypothesis if it were true

Also known as "**NHST**" (Null Hypothesis Significance Testing)

What is a normal distribution?

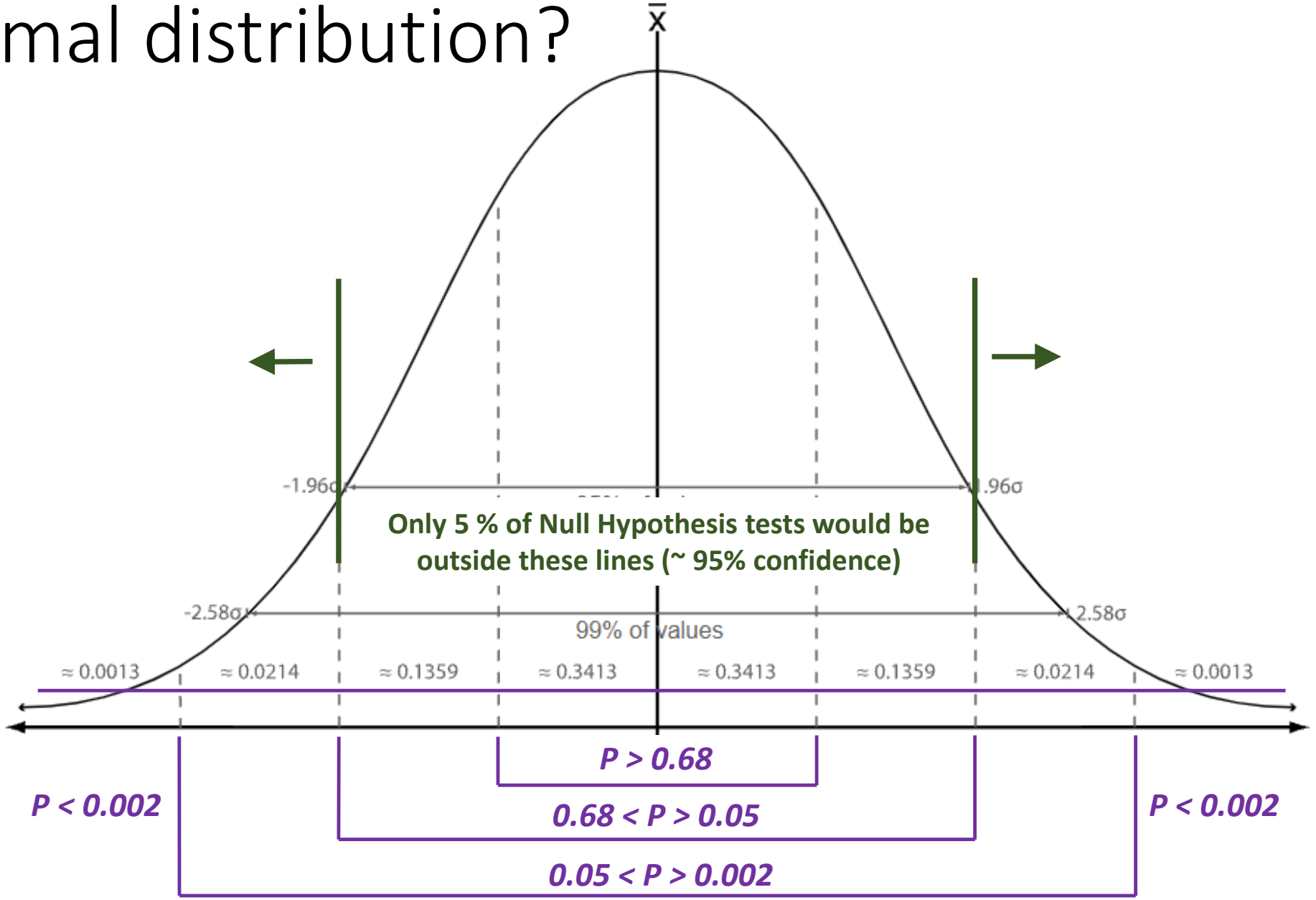
P-Values
(if all assumptions are true)

Probability of Cases in portions of the curve



What is a normal distribution?

P-Values
(if all assumptions are true)



Probability of Cases in portions of the curve

University of Bremen

Example 1. NHST – Comparing Means

Comparing means

- Experiment (randomized controlled trial)
  - Zero difference in infection rates between vaccine and placebo groups
  - Zero difference in employer call-back rates for job candidates with ethnic versus native surnames
- Experiment (intervention)
  - Zero difference between adolescent test scores between schools with the introduction of different teaching techniques

Example 1. NHST – Comparing Means

Comparing means

- Experiment (randomized controlled trial)
  - Zero difference in infection rates between vaccine and placebo groups
  - Zero difference in employer call-back rates for job candidates with ethnic versus native surnames
- Experiment (intervention)
  - Zero difference between adolescent test scores between schools with the introduction of different teaching techniques
- Demographic/population properties (non-experimental)
  - Zero difference between height of girls and boys at age five
  - **Zero difference between happiness between those with high incomes and those with low incomes**

Can Money Buy Happiness?

Example 1. NHST – Comparing Means

How would you test this?

Can Money Buy Happiness?

Example 1. NHST – Comparing Means

One idea:

- Sample from the populations of different countries
- Ask questions about happiness and income
- Compare the average happiness of those with high to those with low incomes

Can Money Buy Happiness?

Example 1. NHST – Comparing Means

World Values Survey

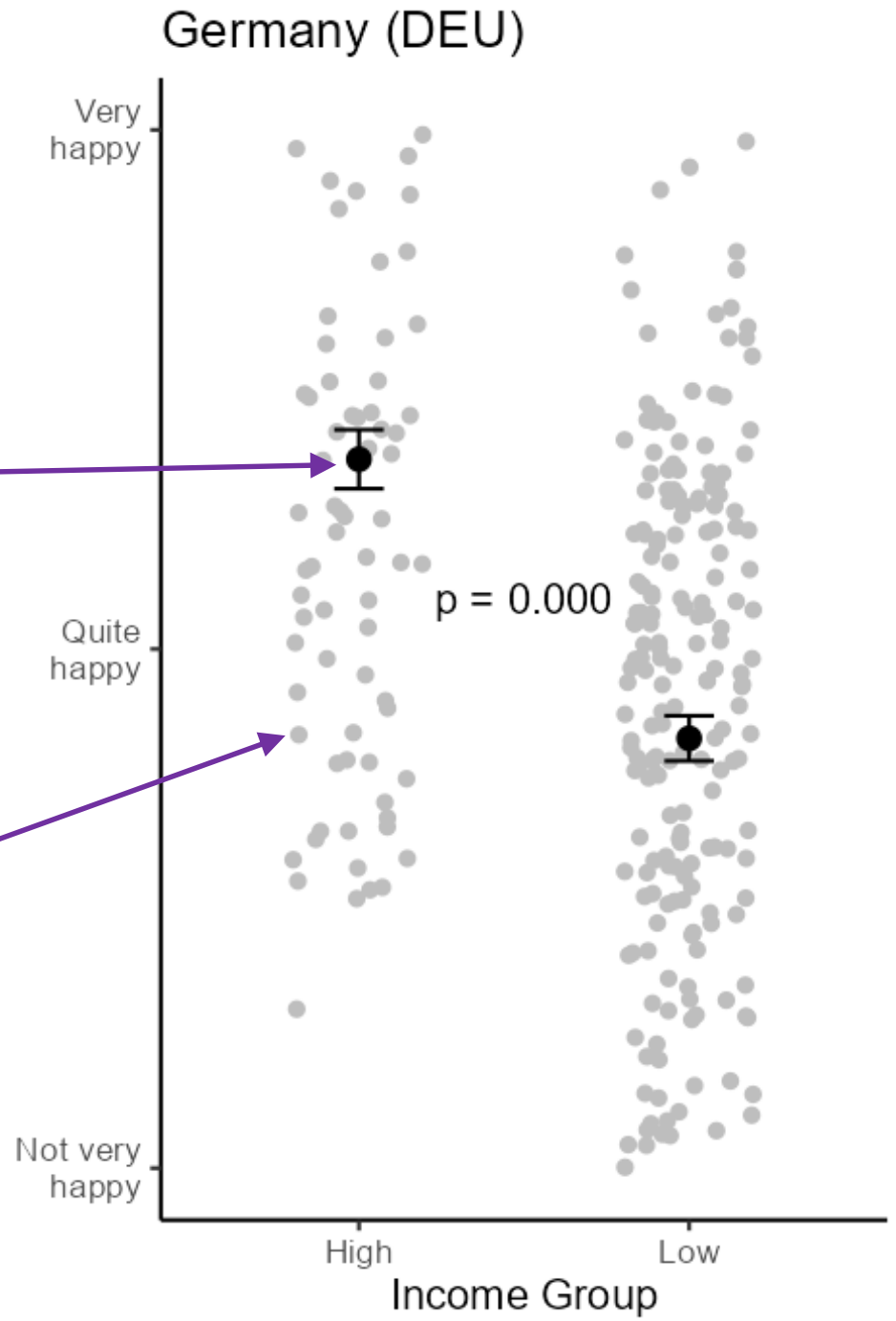
- Up to 80 countries
- 1985-2021 (7 waves)
- Basis for studies of societal values and change
- Roughly 1,000 or more respondents per country

Haerpfer, C., Inglehart, R., Moreno, A., Welzel, C., Kizilova, K., Diez-Medrano J., M. Lagos, P. Norris, E. Ponarin & B. Puranen et al. (eds.). 2020. World Values Survey: Round Seven - Country-Pooled Datafile. Madrid, Spain & Vienna, Austria: JD Systems Institute & WVSA Secretariat. doi.org/10.14281/18241.13

Can Money Buy Happiness?

Group average or mean

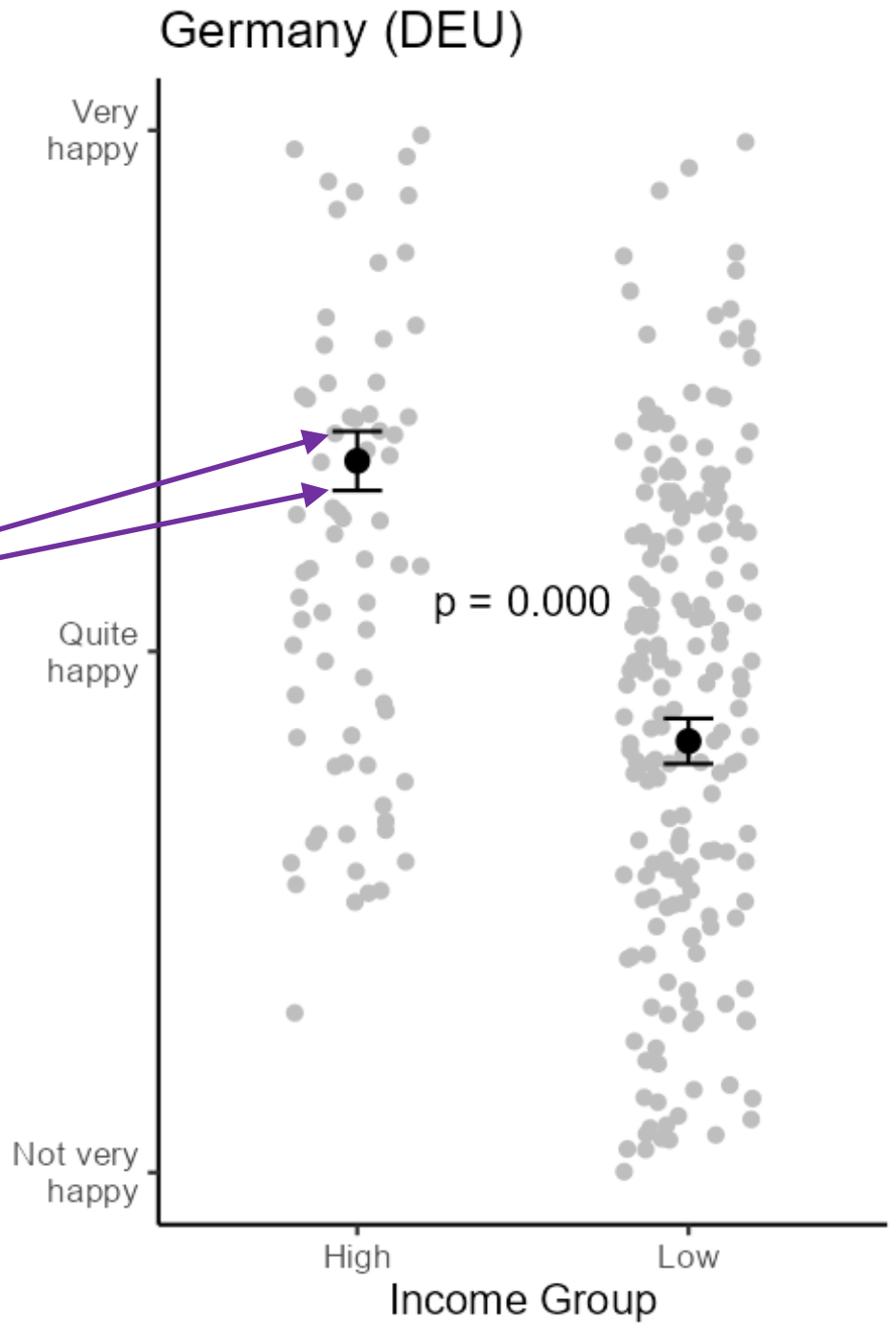
Individual survey respondent (data point)



socium
Research Center on Inequality and Social Policy

Can Money Buy Happiness?

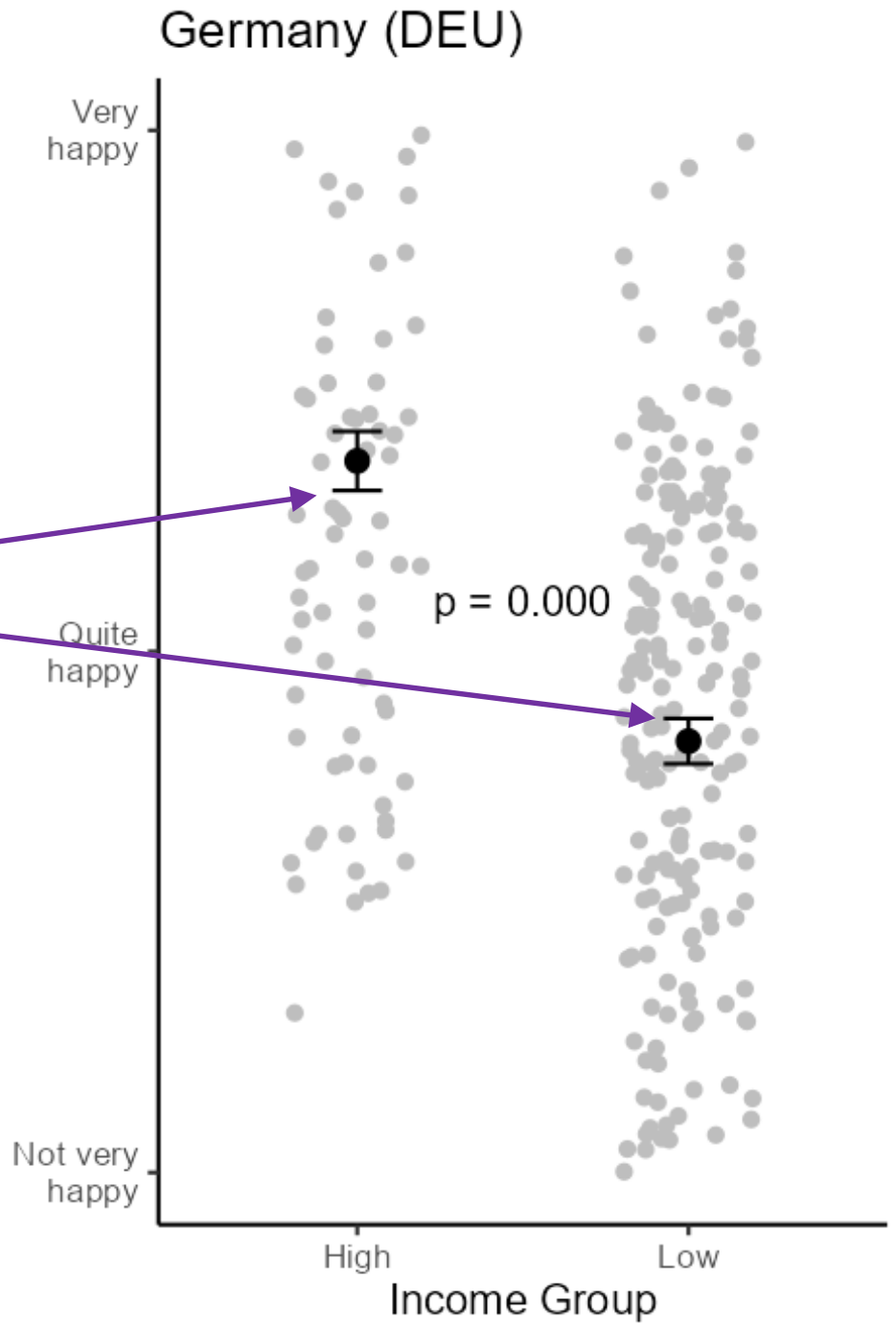
Standard error
=
Standard deviation / square root of sample size



socium
Research Center on Inequality and Social Policy

Can Money Buy Happiness?

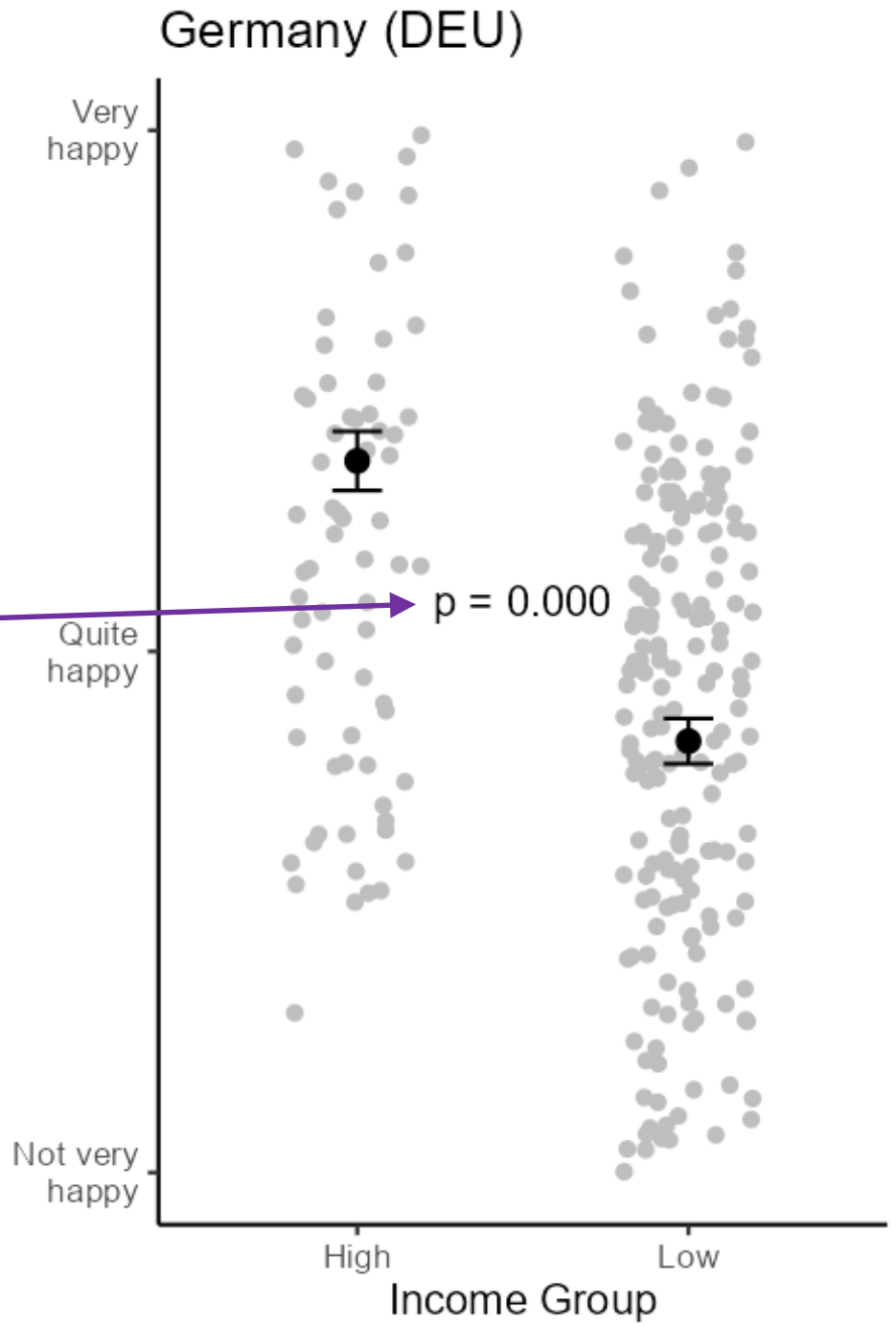
We can see this based on the standard error bars, they do not overlap



socium
Research Center on Inequality and Social Policy

Can Money Buy Happiness?

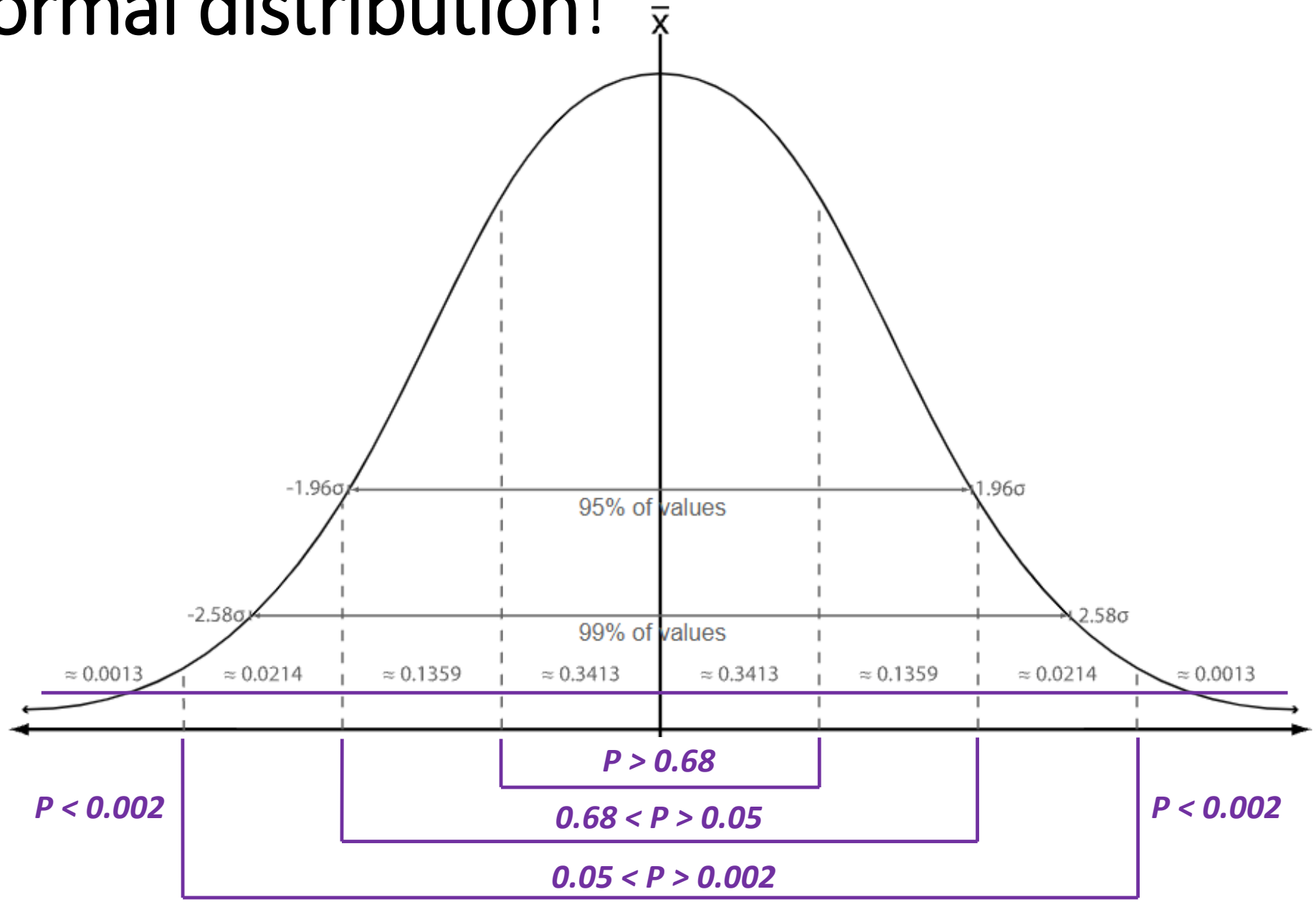
NHST that the difference in happiness between high and low income groups is zero



socium
Research Center on Inequality and Social Policy

Remember, **normal distribution**!

P-Values
(if all assumptions are true)



Probability of Cases in portions of the curve

Can Money Buy Happiness?

Example 1. NHST – Comparing Means

Group Task

1) Look at different country comparisons in the folder "…results/means_p" at https://github.com/nbreznau/p_values

2) Select two countries to compare, save the two figures on a slide or document

3) Be prepared to meaning of the p-values for each country

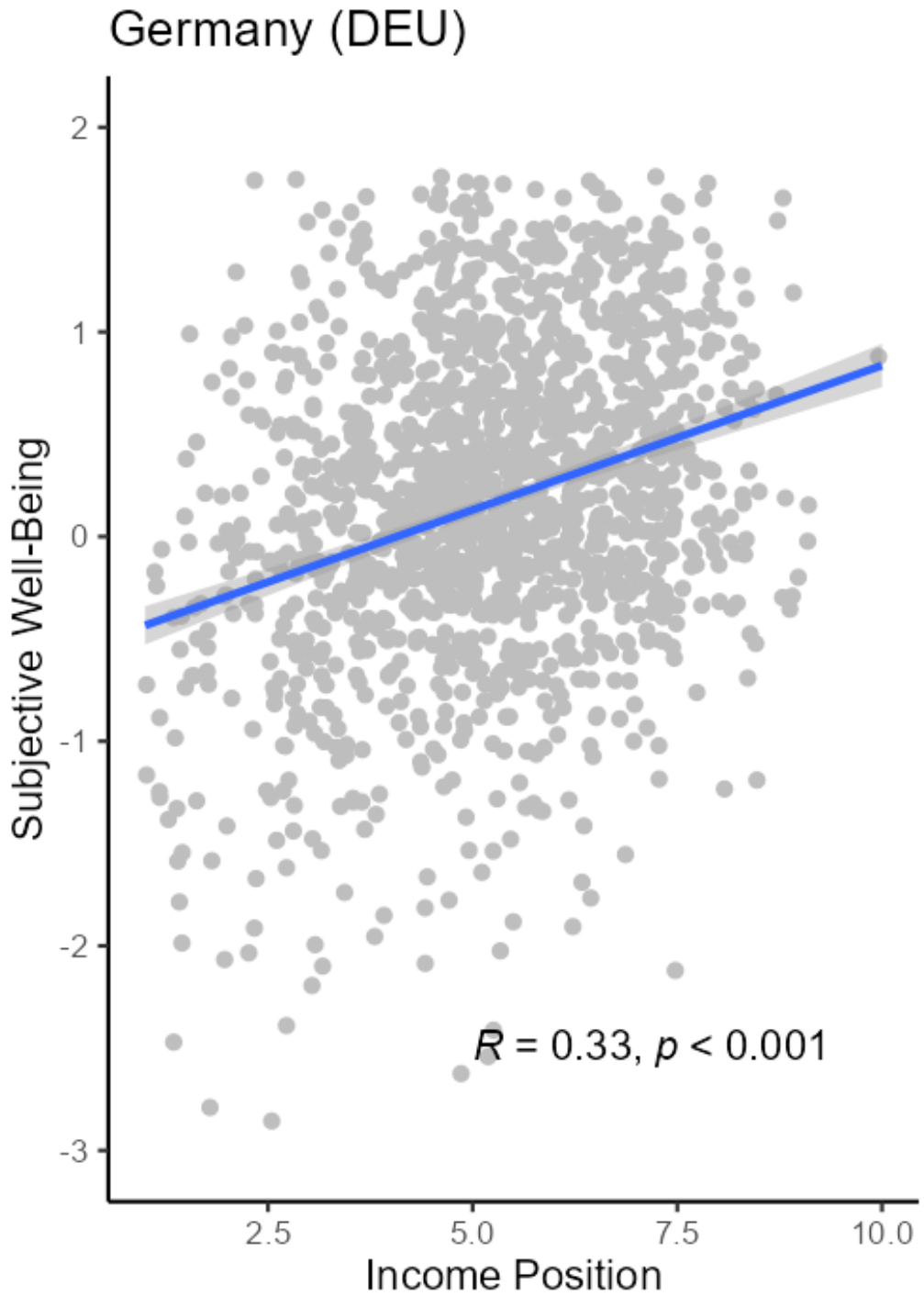
Can Money Buy Happiness?

Example 2. NHST – Linear Association

- Compare income and happiness as **continuous** measures.
- WVS has an 8-category income scale (quasi-continuous)
- WVS asks ‘how happy are you’ and ‘how satisfied are you with your life’. Using these two questions we can construct a **scale**.

Can Money Buy Happiness?

Example 2. NHST – Linear Association



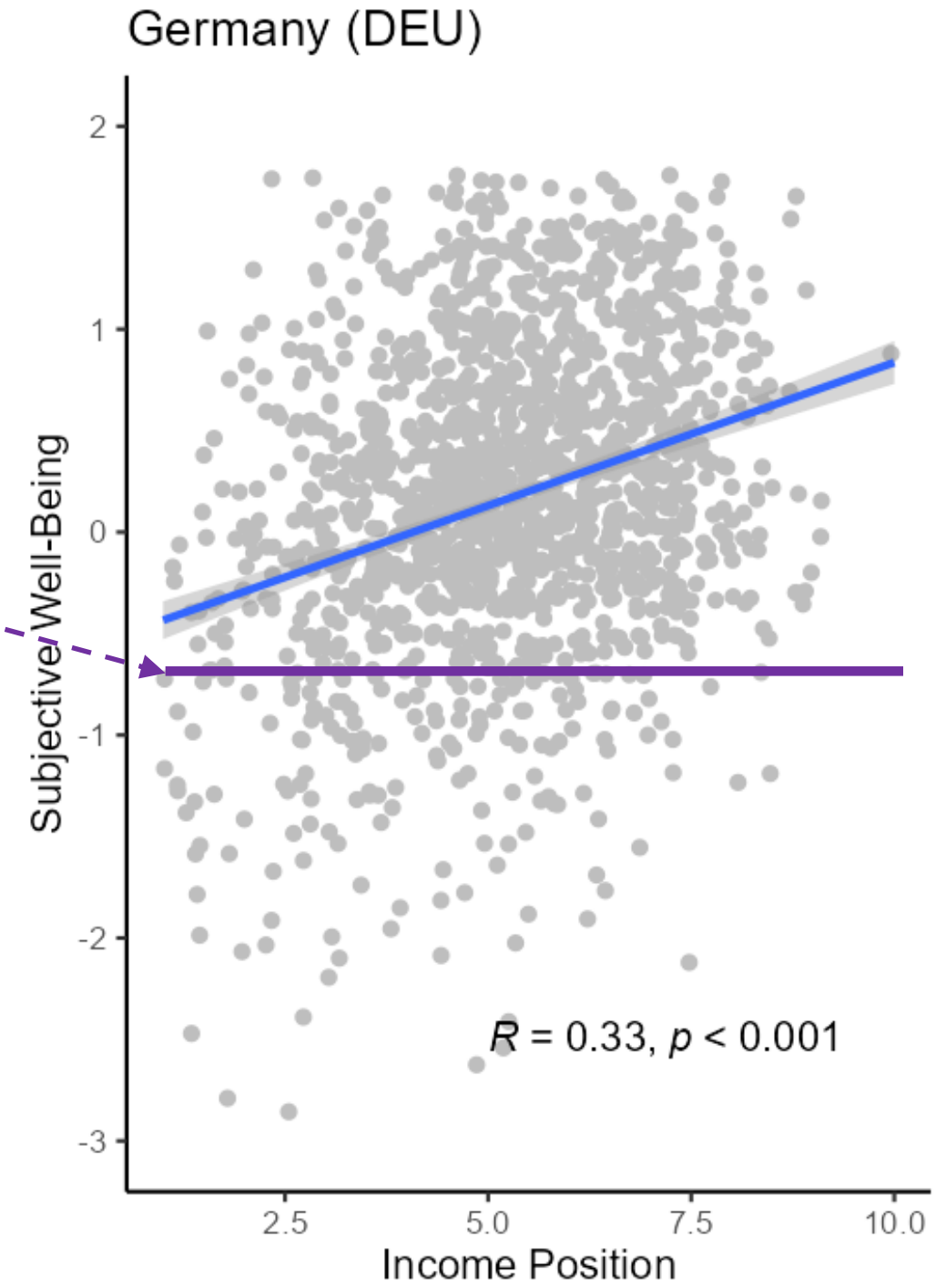
socium

Research Center on
Inequality and Social Policy

Can Money Buy Happiness?

Example 2. NHST – Linear Association

NHST is that the line of best fit has a slope of **zero**



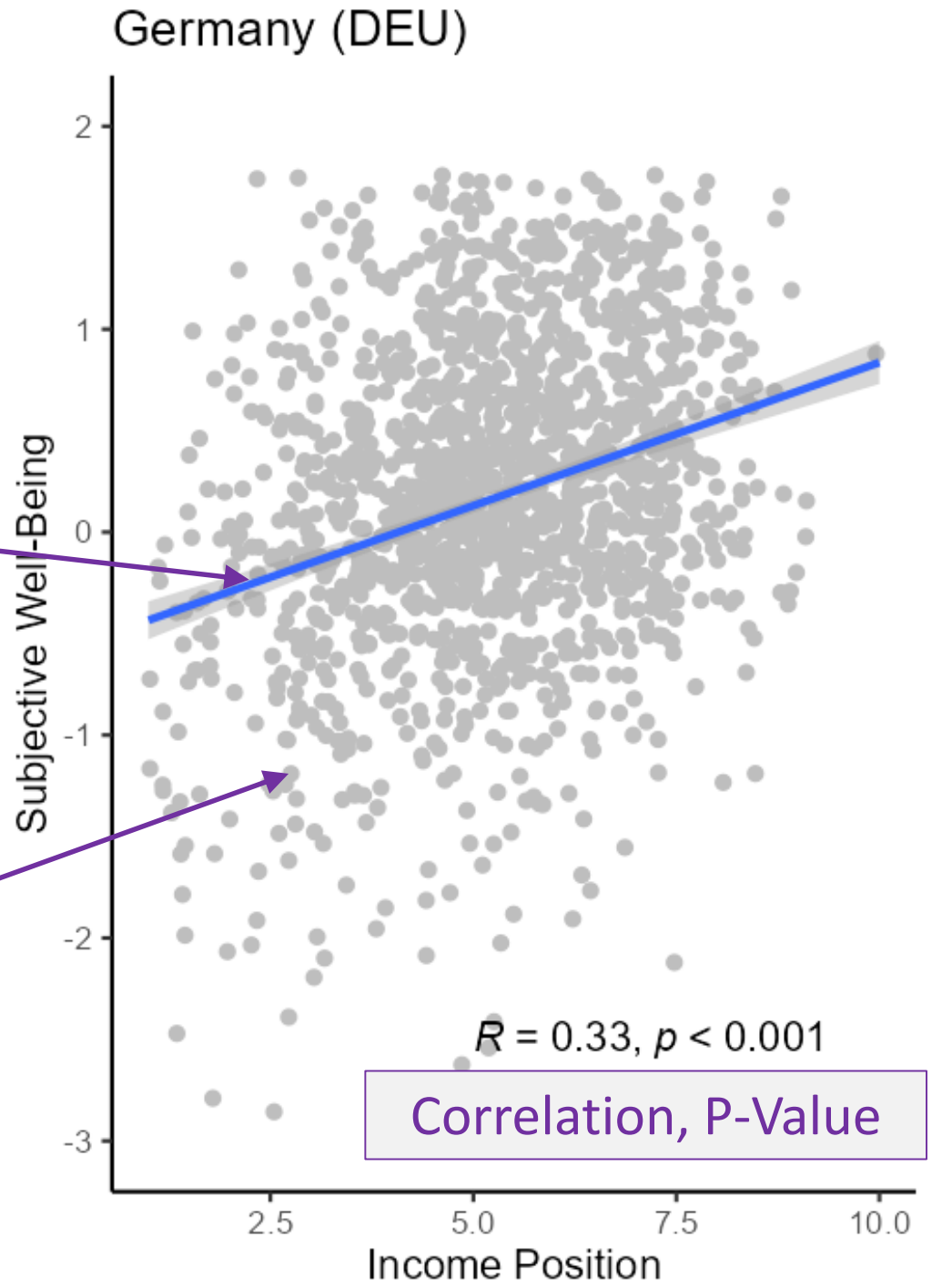
socium
Research Center on
Inequality and Social Policy

Can Money Buy Happiness?

Example 2. NHST – Linear Association

Line of best fit
(minimizes sum of squared error)

Individual survey respondent (data point)



Correlation, P-Value

socium
Research Center on Inequality and Social Policy

Can Money Buy Happiness?

Example 2. NHST – Linear Association, adjusted

University of Bremen

- Sometimes we want to adjust for other variables that might explain a linear association.
- With the *income-happiness association* for example, maybe people who are higher educated are happier, or people who have stronger religious beliefs.
- Other variables may explain a linear association better, or in part

Can Money Buy Happiness?

Example 2. NHST – Linear Association, adjusted

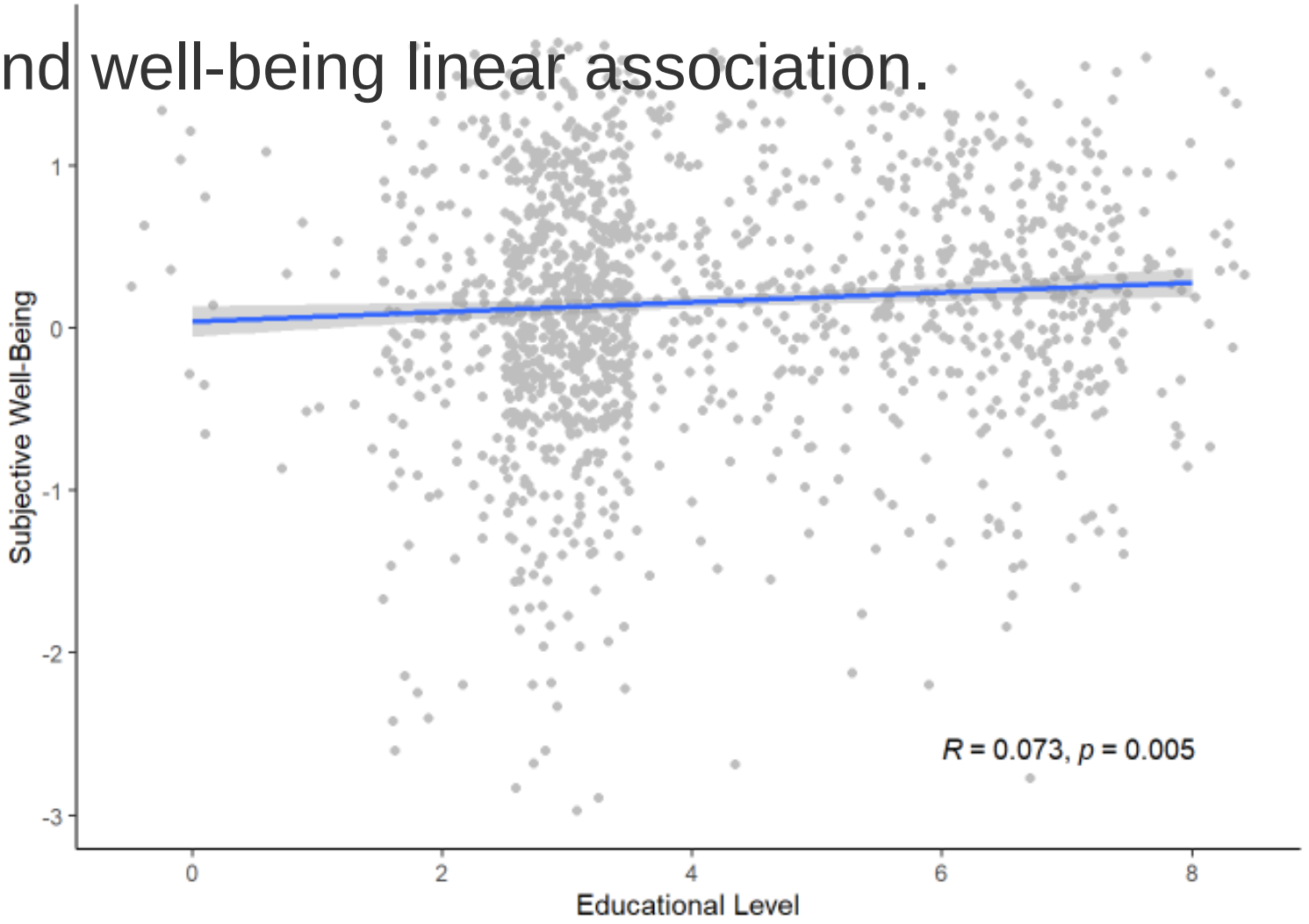
- Education* and well-being linear association.

*Note that ISCED is not intended as a linear measure, but we will use it that way for now just for the purposes of example (and we removed 9 for 'other').

Can Money Buy Happiness?

Example 2. NHST – Linear Association, adjusted

- Education* and well-being linear association.



*Note that ISCED is not intended as a linear measure, but we will use it that way for now just for the purposes of example (and we removed 9 for 'other').

Can Money Buy Happiness?

Example 3. NHST – Linear Association, adjusted

- One option is a **partial correlation**
- This procedure:
  - Compute residuals for x (dependent) ~ z (independent) model
  - Compute residuals for y (dependent) ~ z (independent) model
  - Compute correlation between the residuals

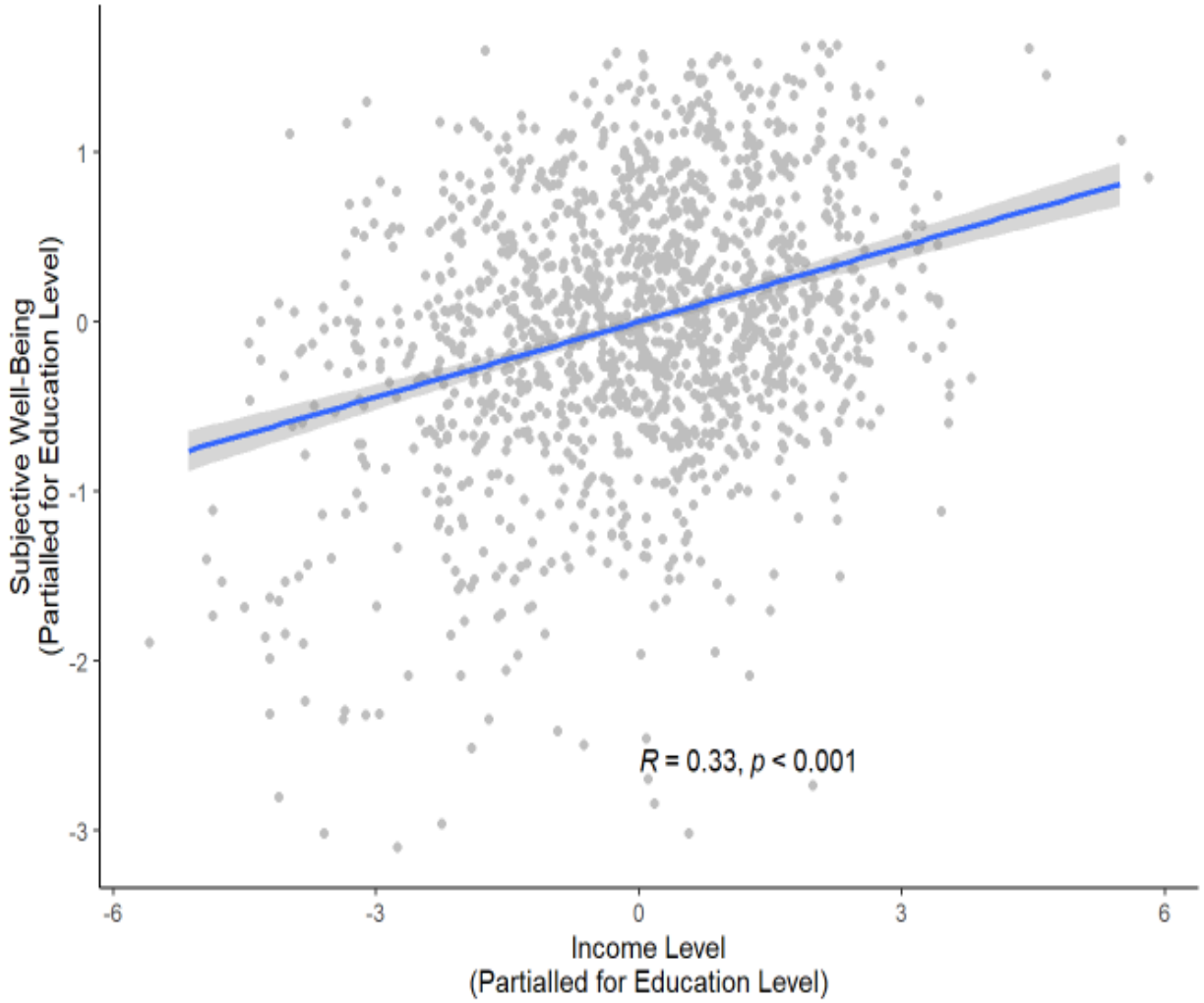
- In other words:

  A partial correlation is what remains of a linear association after removing the portion of that association that can be explained by a third variable

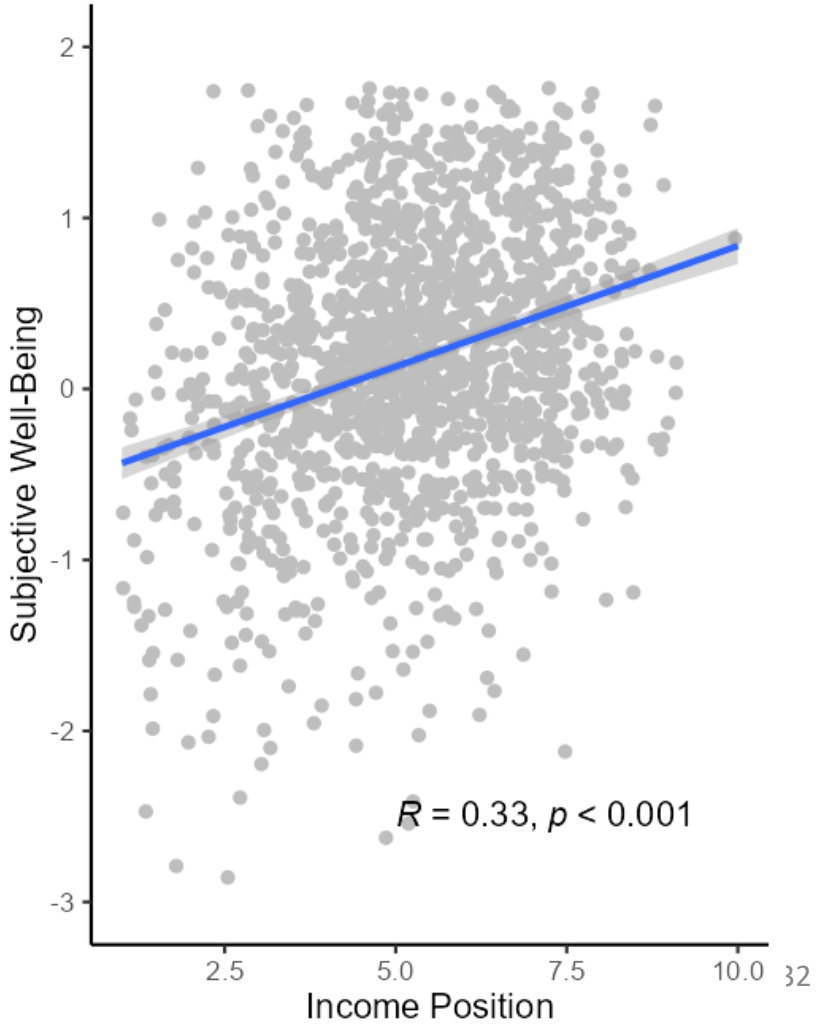
Does Education affect the *income-happiness* association in Germany?

University of Bremen

Adjusted Linear Association

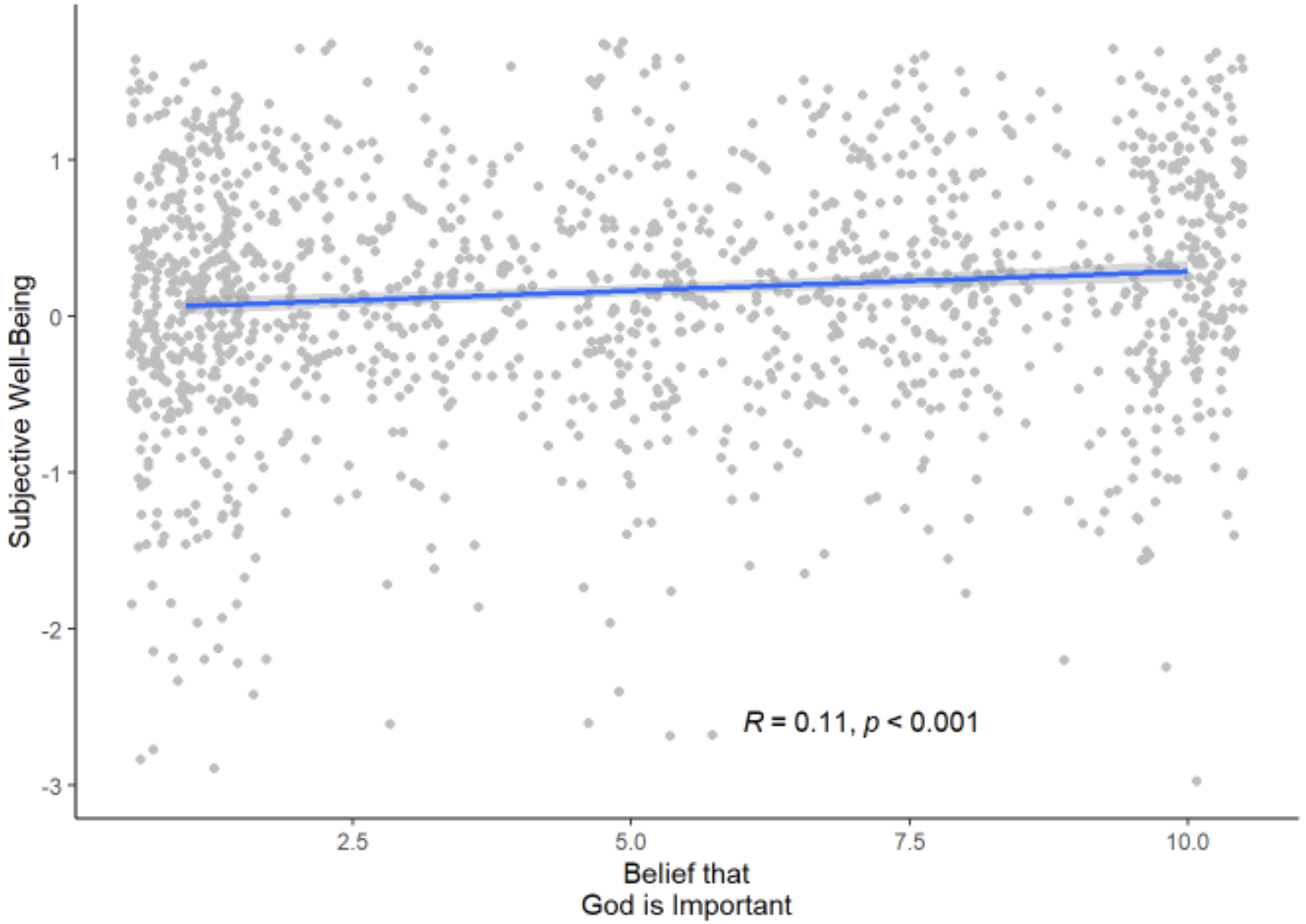


Unadjusted Linear Association



Importance of God and Happiness?

University of Bremen



socium
Research Center on
Inequality and Social Policy

Can Money Buy Happiness?

Example 3. NHST – Linear Association, adjusted

- One option is a **partial correlation**
- Another option is a **linear regression**
  - Advantage is that it can adjust for many variables
  - Computer does the work
  - Statistical software provides results/p-values

Can Money Buy Happiness?

Example 3. NHST – Linear Association, adjusted

- Germany – 3 regressions

	Model 1	Model 2	Model 3
Income	0.14 ***	0.15 ***	0.15 ***
	(0.01)	(0.01)	(0.01)
Education		-0.02	-0.02
		(0.01)	(0.01)
God Importance			0.03 ***
			(0.01)
N	1445	1445	1445
R2	0.11	0.11	0.13

*** p < 0.001; ** p < 0.01; * p < 0.05.

socium
Research Center on
Inequality and Social Policy

University of Bremen

Can Money Buy Happiness?

University of Bremen

Example 3. NHST – Linear Association, adjusted

- Compare:

Ethiopia

	Model 1	Model 2	Model 3
Income	0.05 ***	0.05 ***	0.05 ***
	(0.01)	(0.01)	(0.01)
Education		-0.02	-0.02
		(0.01)	(0.01)
God Importance			0.01
			(0.03)
N	1222	1222	1222
R2	0.01	0.02	0.02

*** p < 0.001; ** p < 0.01; * p < 0.05.

Vietnam

	Model 1	Model 2	Model 3
Income	0.06 ***	0.07 ***	0.07 ***
	(0.01)	(0.01)	(0.01)
Education		-0.06 ***	-0.06 ***
		(0.01)	(0.01)
God Importance			0.01
			(0.01)
N	1200	1200	1200
R2	0.02	0.04	0.04

*** p < 0.001; ** p < 0.01; * p < 0.05.

socium
Research Center on Inequality and Social Policy