Lecture 18

Randomized experiments

Julian Reif

Fall 2025

Reminder: get started on final project

- Goal: predict values of properties listed in `predict_property_data.csv`
- Use `historic_property_data.csv` to train your model
- Challenges: variable selection, missing values, and more
- Special bonus given to group with the most accurate valuations!

Recall: designing a causal analysis

- A causal inference analysis includes the following components:
  - **Outcome**: a measure we are interested in studying
  - **Treatment**: a variable measuring the causal relation of interest
  - **Treatment group**: subjects receiving the treatment
  - **Control group**: subjects not receiving the treatment

- Estimated effect of treatment: Difference in Group Means of outcome
  - Avg[Treatment group] - Avg[Control group]

- Good control group describes fate of the treated group, if they had not been treated

- If control group is bad, then estimate suffers from selection bias

Today: more on treatment effects and selection bias

- Explore the problem, using effect of health insurance on health as an example
- Learn how randomized experiments ("A/B testing") solves the selection bias problem
- Study real-world example that used random assignment

Effect of health insurance on health

- Design of our causal analysis:
  - Outcome, Y_i: measure of health (e.g., health index) for person i
  - Treatment, T_i: has health insurance (yes/no) (1/0)
  - Treatment group: people with health insurance $(T_i = 1)$
  - Control group: people without insurance $(T_i = 0)$

- Estimate: $\overline{Y}_{treatment}$ - $\overline{Y}_{control}$

- Will this estimate suffer from selection bias?

Recall: notation for potential outcomes

- Actual outcome for a specific individual is Y_i, the value recorded in the data

- To talk about potential outcomes, we add a 0/1 subscript
  - Y_{0i} ("y-zero-i"): the outcome for person i with no treatment
  - Y_{1i} ("y-one-i"): the outcome for person i with treatment

Suppose there are two types of people

Tabby 1

"Fit"

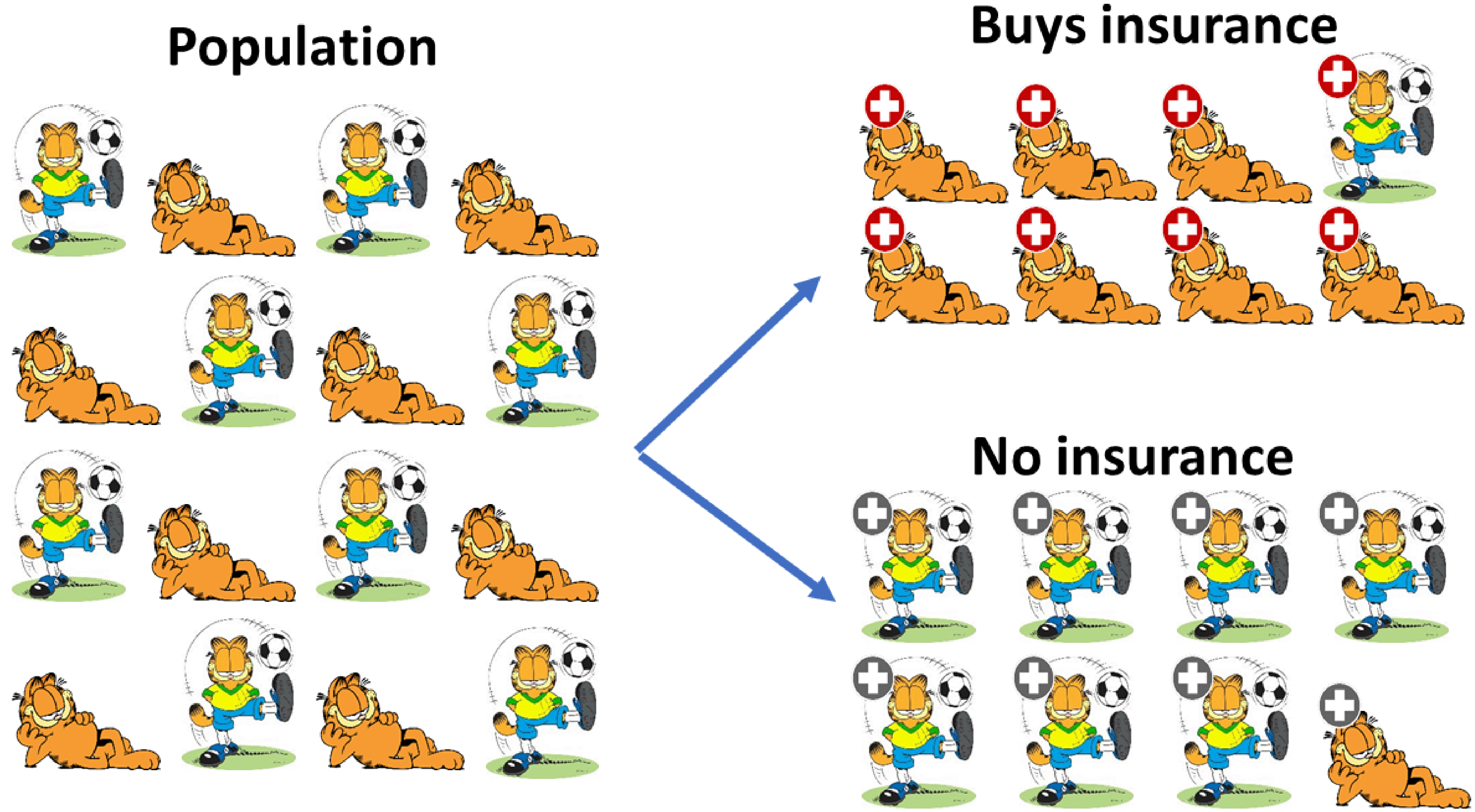
Tabby 2

"Fun"

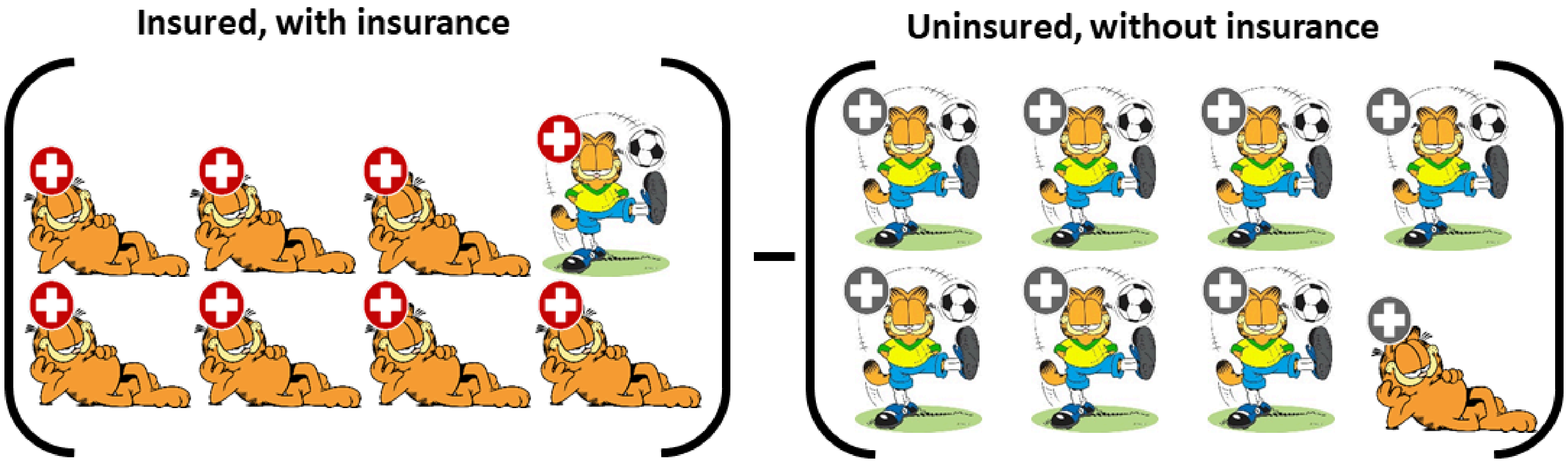
We have two potential outcomes for each type

Potential Outcomes	Tabby 1		Tabby 2	
Uninsured (Y_{0i})		Health = fit		Health = fun
Insured (Y_{1i})		Health = fit + β		Health = fun + β

Who buys health insurance? It's not random...



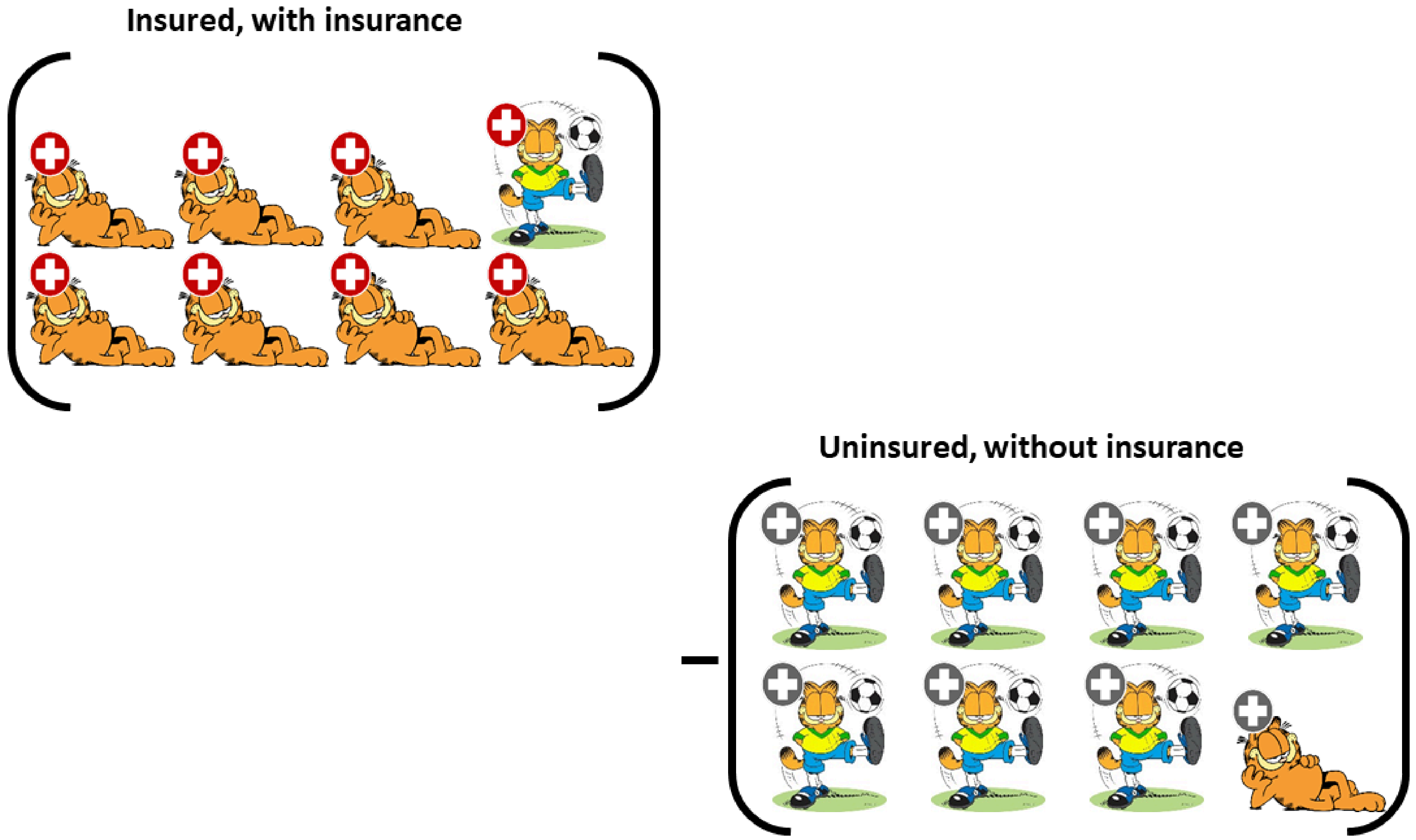
What happens if we compare insured to the uninsured?



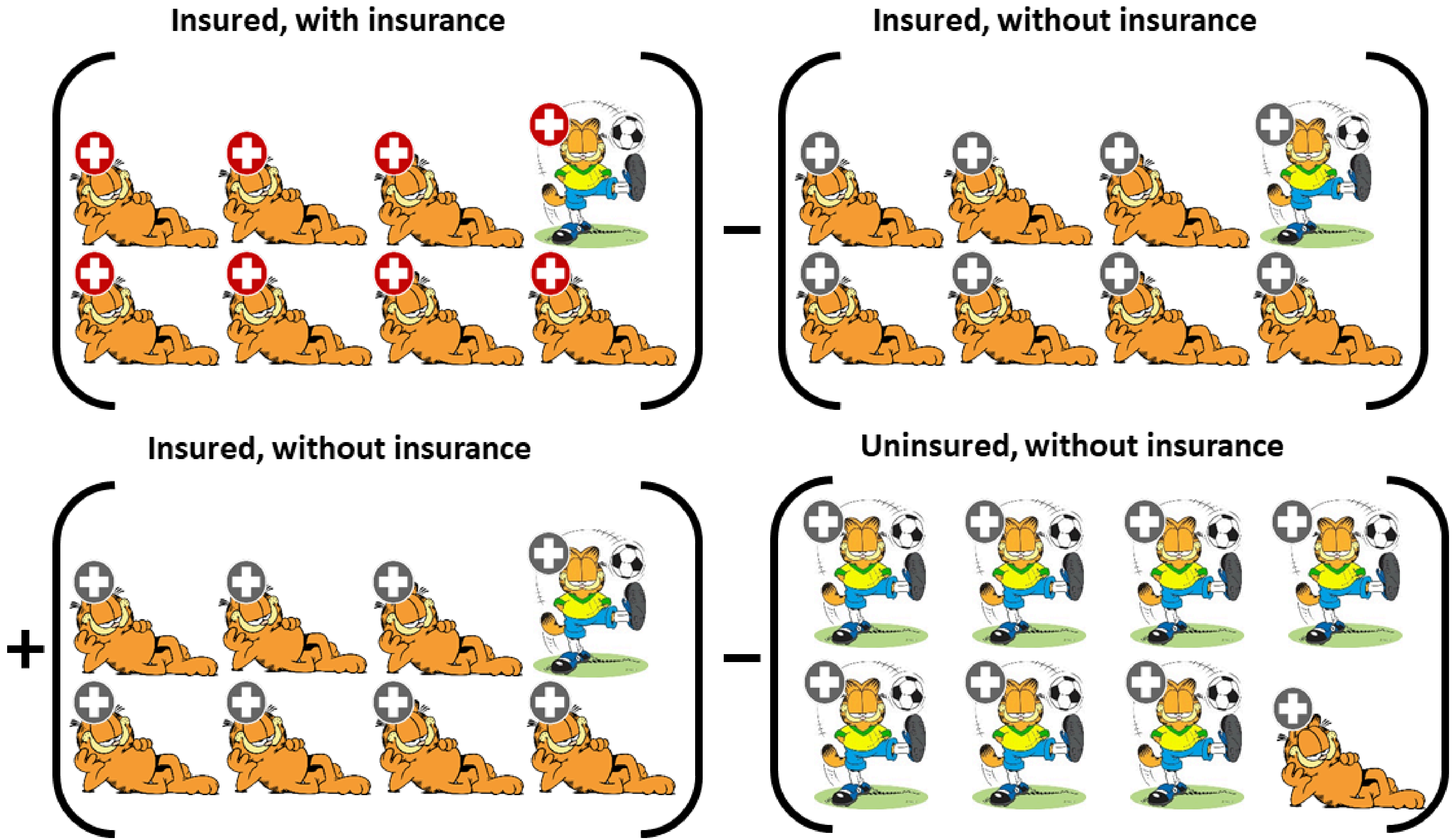
Calculate difference in means between treatment and control groups

Will this estimate equal the causal effect of health insurance?

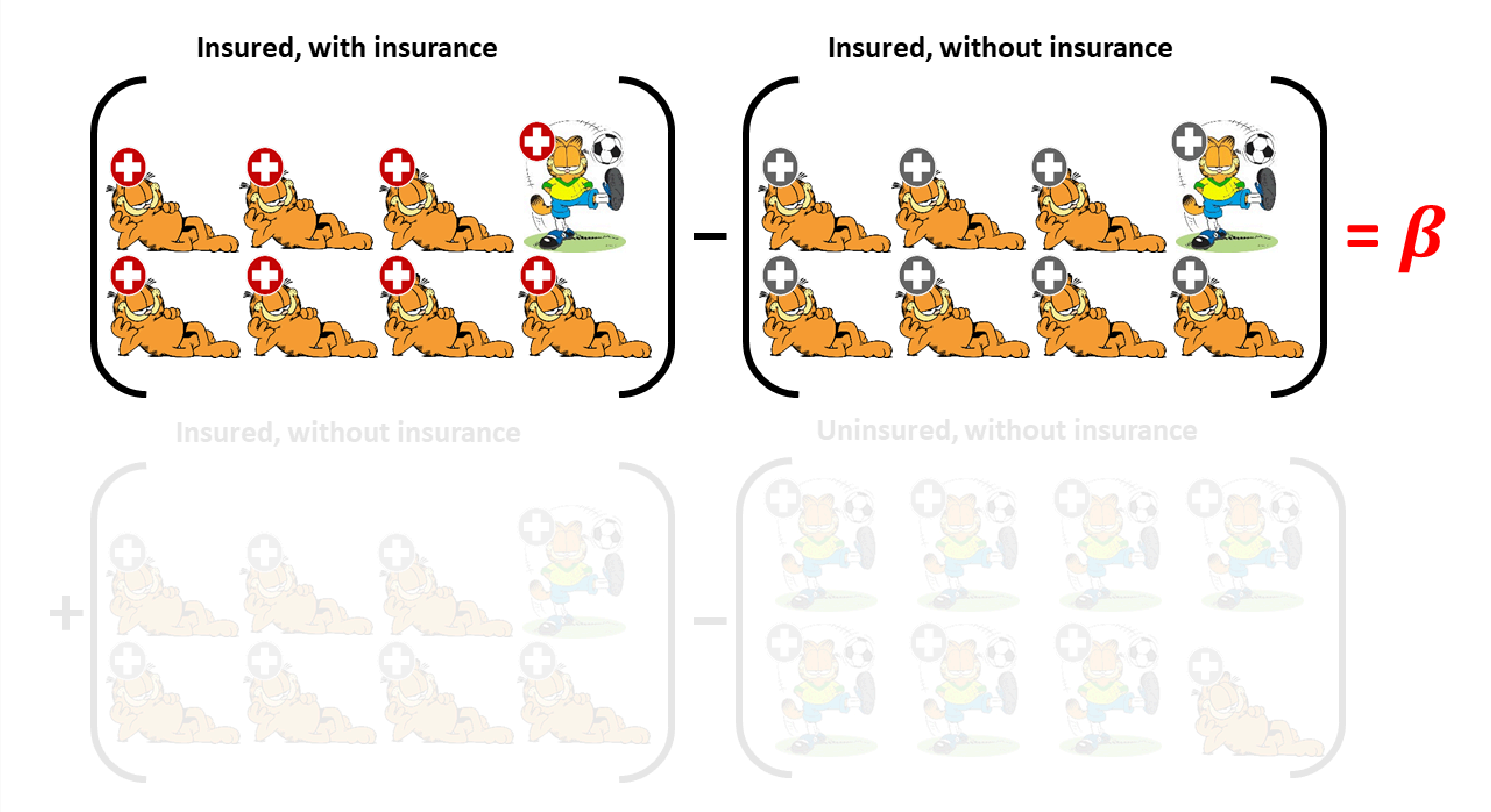
What happens if we compare insured to the uninsured?



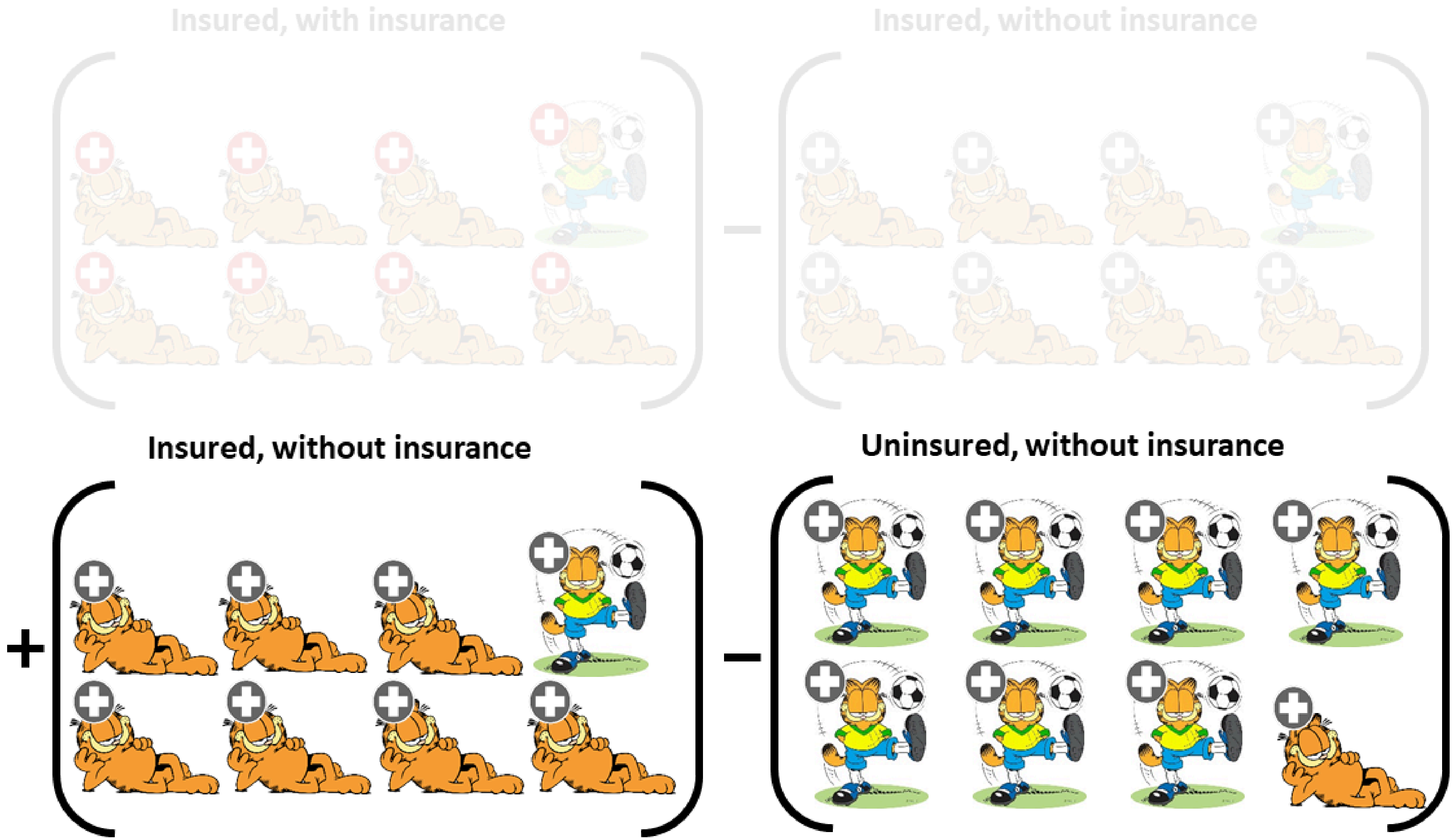
What happens if we compare insured to the uninsured?



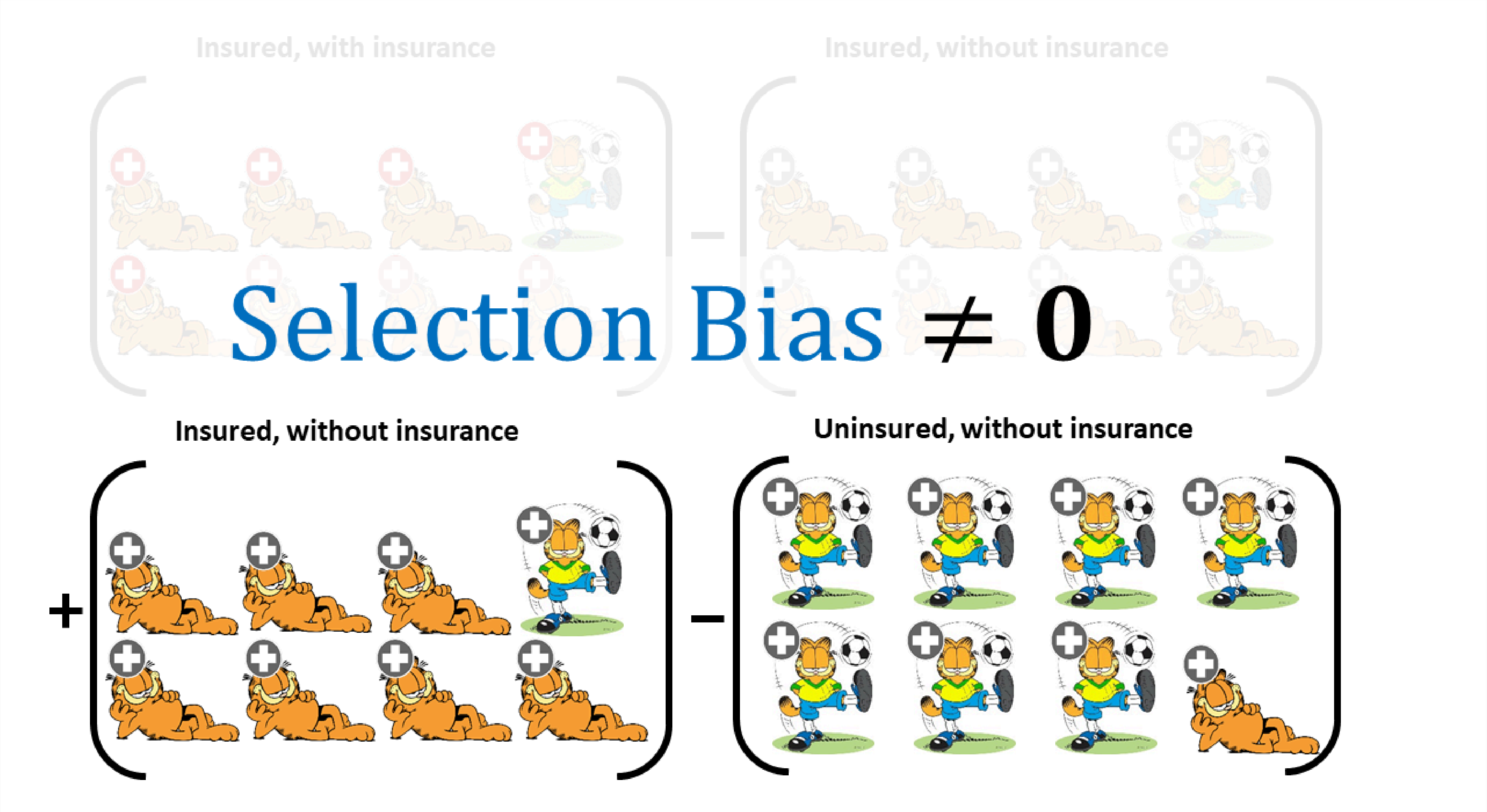
This part is the causal treatment effect



This part is selection bias



This part is selection bias



Selection bias arises when the control group is bad

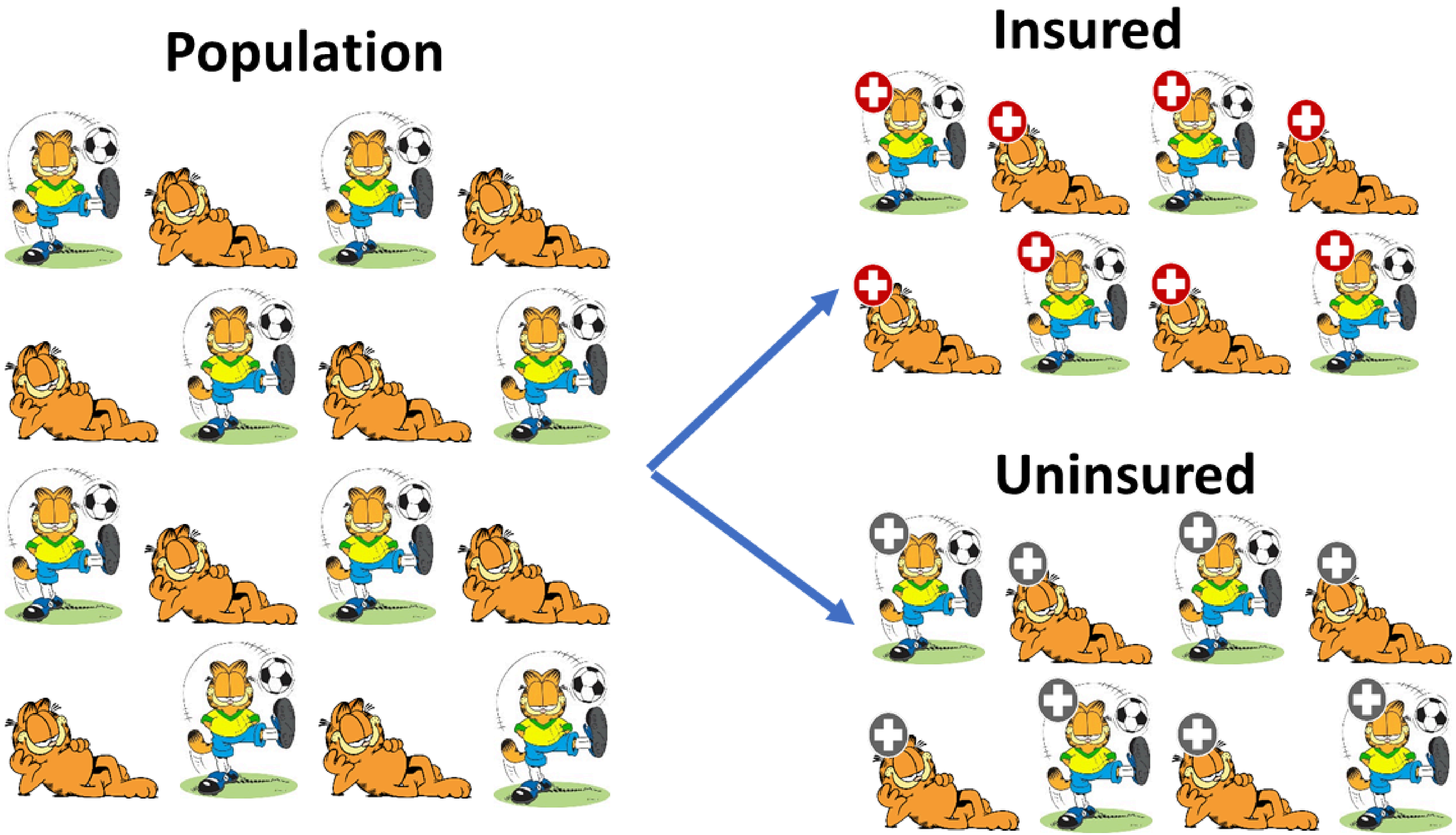
- Average in the treatment group minus average in the control group:

$$
\begin{aligned}
Avg[Y_{1i}|T_i = 1] - Avg[Y_{0i}|T_i = 0] &= \color{red}{(Avg[Y_{1i}|T_i = 1] - Avg[Y_{0i}|T_i = 1])} + \\
&\quad \color{blue}{(Avg[Y_{0i}|T_i = 1] - Avg[Y_{0i}|T_i = 0])} \\
&= \color{red}{\text{Causal Treatment Effect}} + \\
&\quad \color{blue}{\text{Selection Bias}}
\end{aligned}
$$

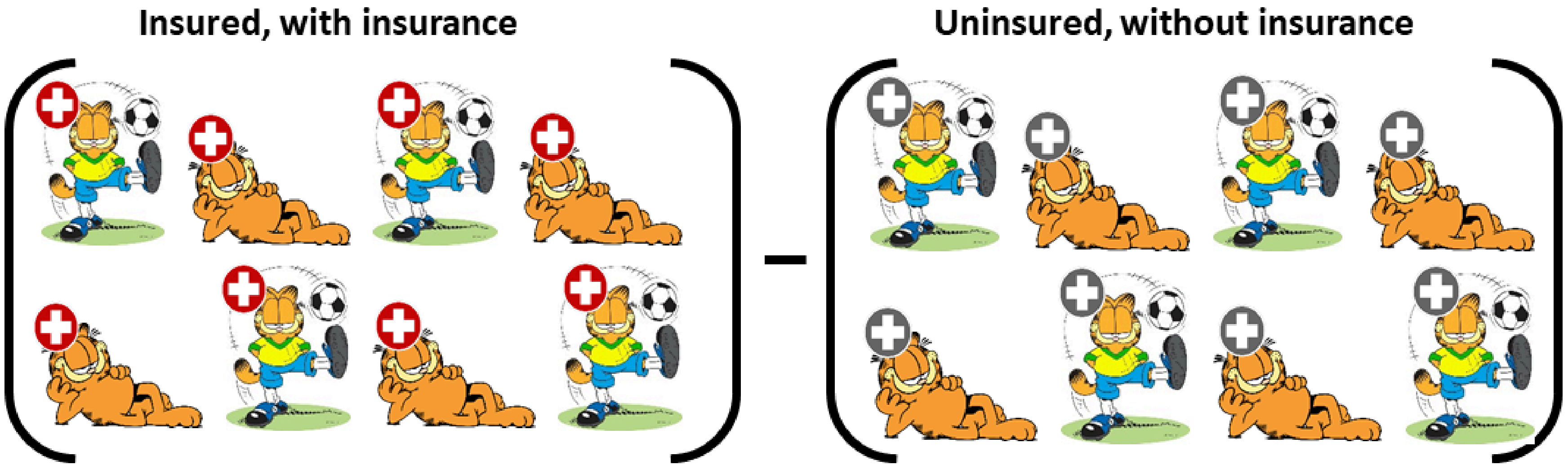
- Note: having "bigger" (more) data does not solve this problem!

- What can solve it? Randomization!

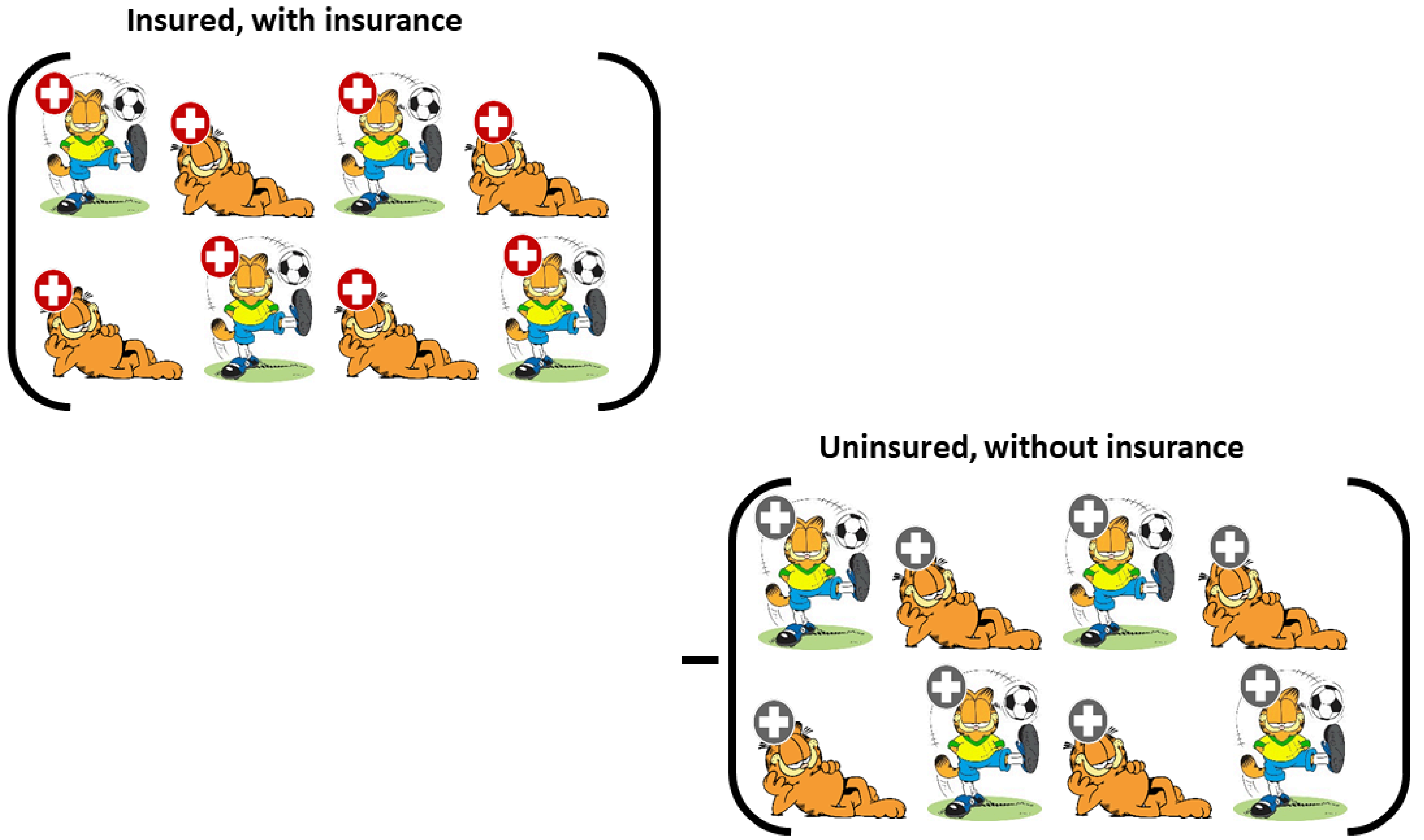
Randomly assign people to insurance



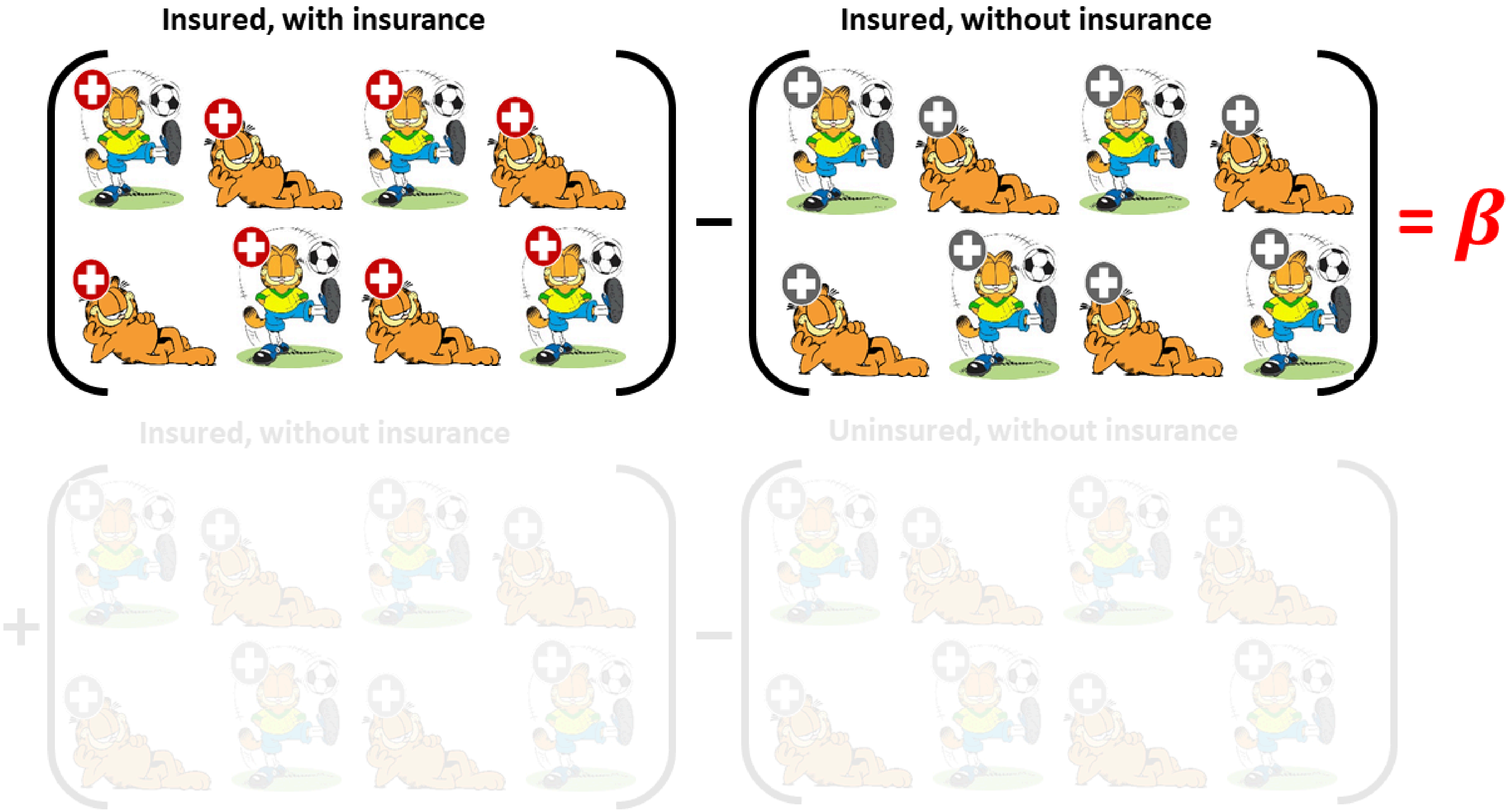
What happens if we compare insured to the uninsured?



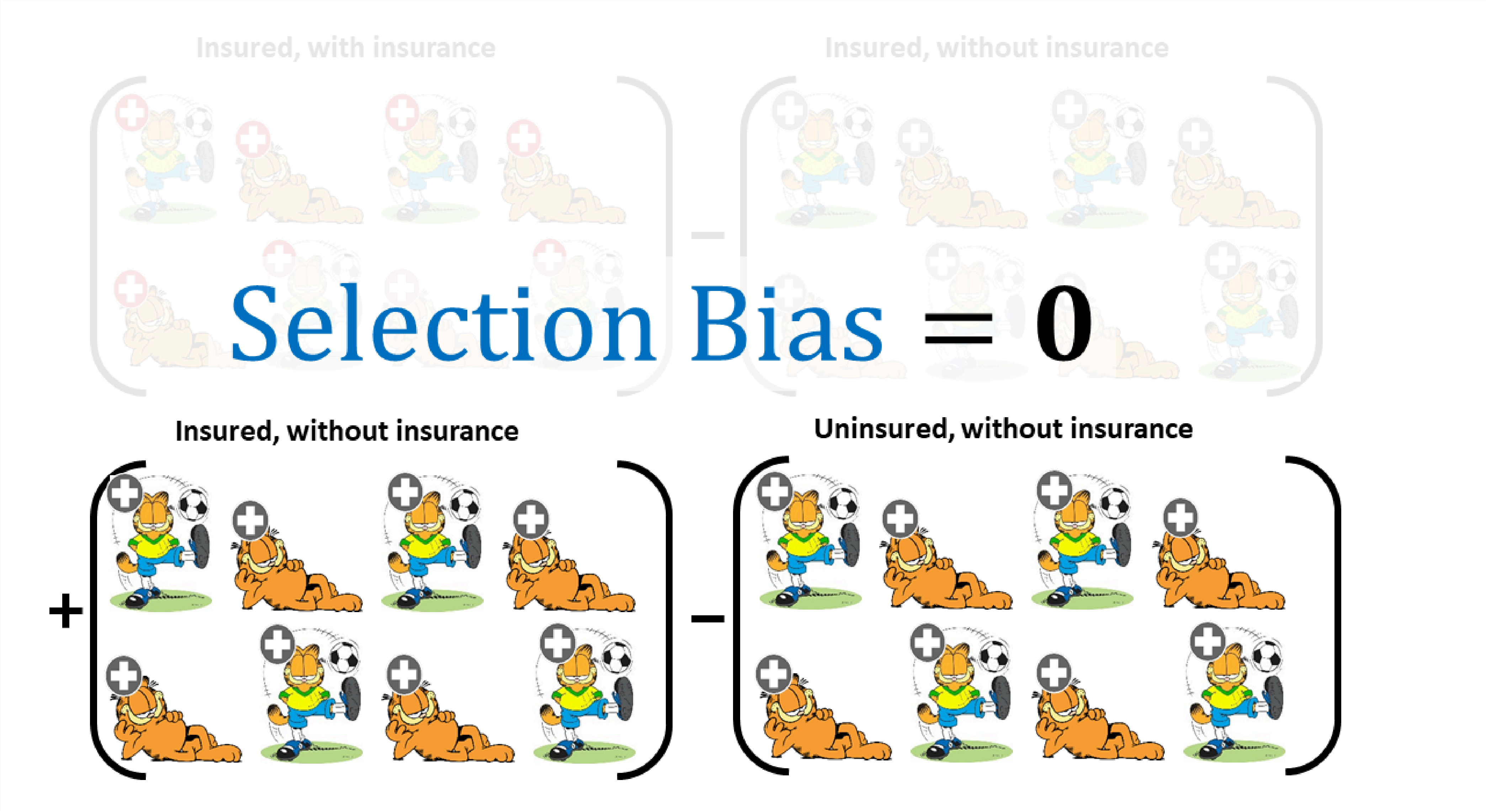
What happens if we compare insured to the uninsured?



Causal treatment effect



There is no selection bias



A good control group shows what would have happened to the treated group, if they had not been treated

The power of random assignment

$$\text{Selection bias} \;=\; \underbrace{Avg[Y_{0i}|T_i = 1]}_{\text{Treatment group (unobserved)}} \;-\; \underbrace{Avg[Y_{0i}|T_i = 0]}_{\text{Control group (observed)}}$$

- By law of large numbers, and since treatment is assigned randomly:

$$Avg[Y_{0i}|T_i = 1] \approx Avg[Y_{0i}|T_i = 0] \approx Avg[Y_{0i}]$$
$$\rightarrow \text{Selection bias} \;\approx 0$$

- When treatment is random and samples are sufficiently large, selection bias disappears

- Under random assignment, **difference in group means describes average causal effects**

Running a randomized experiment

1. Randomly assign individuals to treatment or control group

2. Check that treatment and control groups are balanced

3. Let the treatment (intervention) run its course

4. Measure the outcomes for treatment and control groups
    - Causal effect is $Avg[Treatment] - Avg[Control]$

- Note: Step 2 (balance tests) can also be done at the end

Balance tests

- Balance tests are standard way to check whether randomization worked

- Compare treatment to control group using outcomes measured **before treatment**
  - E.g., age, income, education

- If randomization worked, then the means of pre-treatment outcomes should be the same for the control and treatment groups
  - Key assumption: law of large numbers
  - Small sample sizes can lead to imbalance

Real-world health insurance experiments

- RAND Health Insurance Experiment
  - Ran from 1974-1982
  - Randomized individuals into different insurance plans

- Oregon Health Plan Experiment
  - Oregon expanded its Medicaid insurance program in 2008
  - Limited slots available, so access was determined by lottery

- Today: study results from RAND

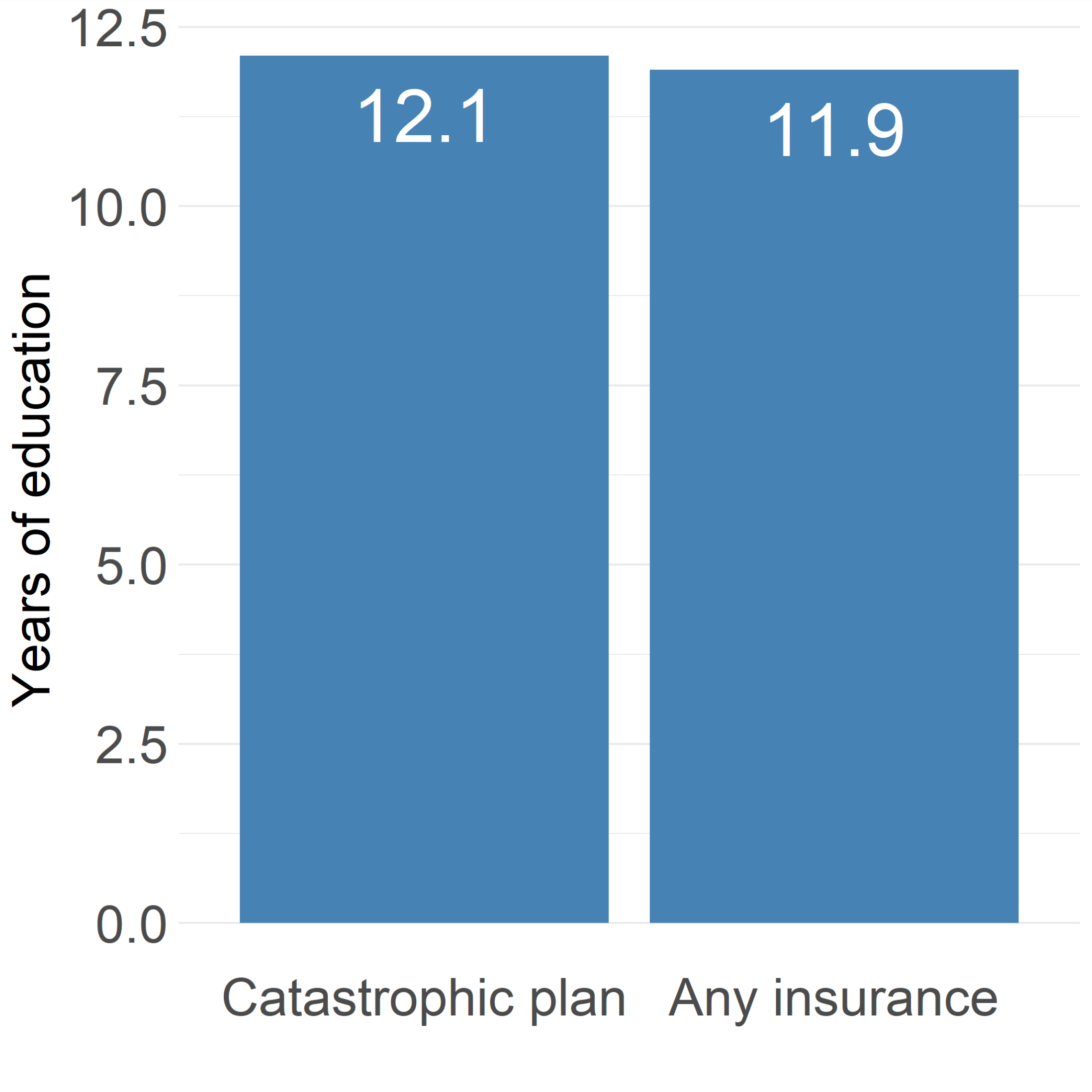
RAND Health Insurance experiment (1974-1982)

- Individuals were randomized into different insurance plans
  - Several types of generous plans (e.g., free care plan, coinsurance plan)
  - One stingy plan ("catastrophic plan" with high cost sharing)

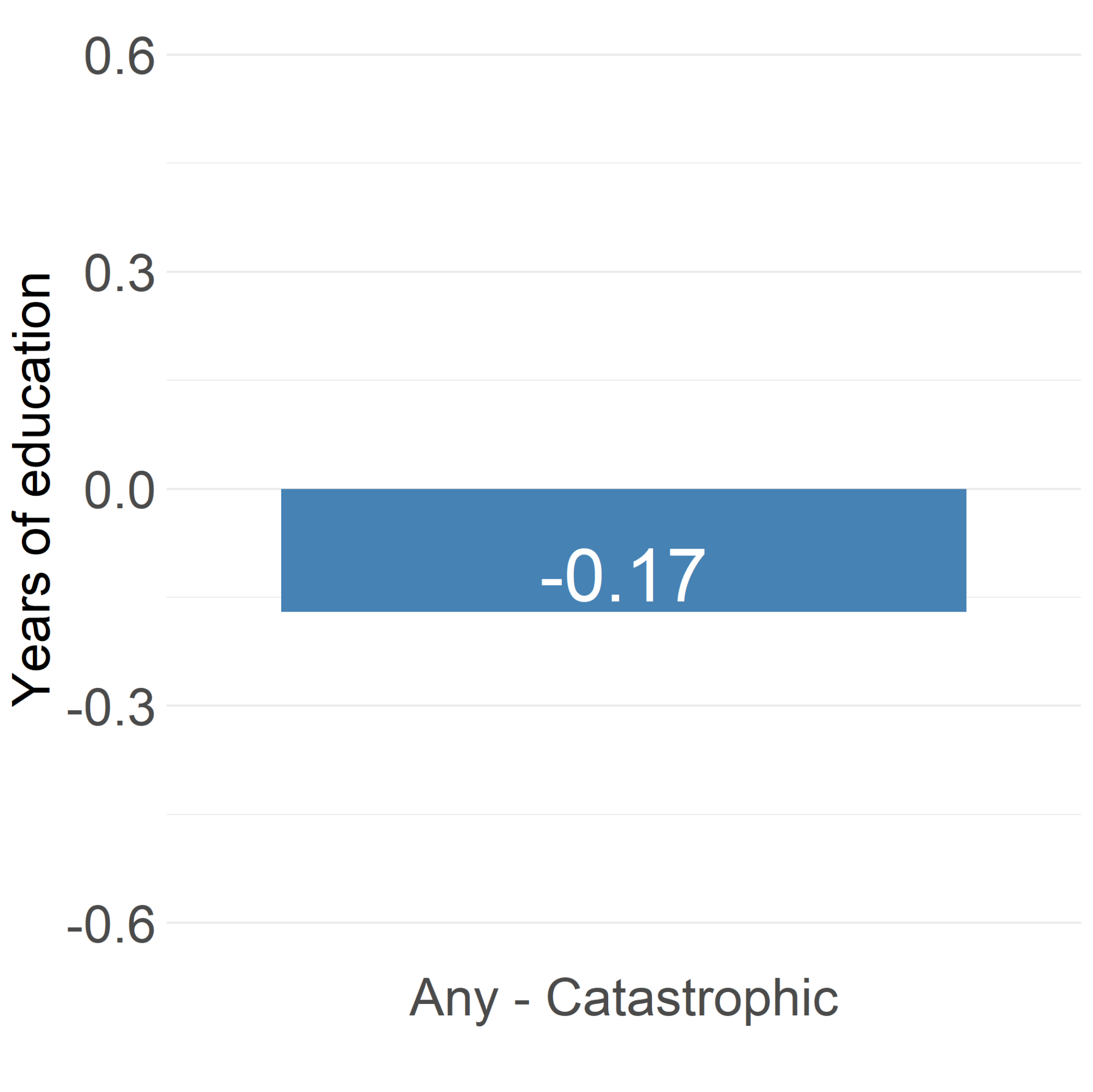
- We need to define treatment and control groups
  - Control group: catastrophic plan (minimal insurance)
  - Treatment group: all other plans ("any insurance")

- Research questions:
  - How does insurance affect use of healthcare services?
  - How does insurance affect quality of care?
  - What are health consequences of insurance?

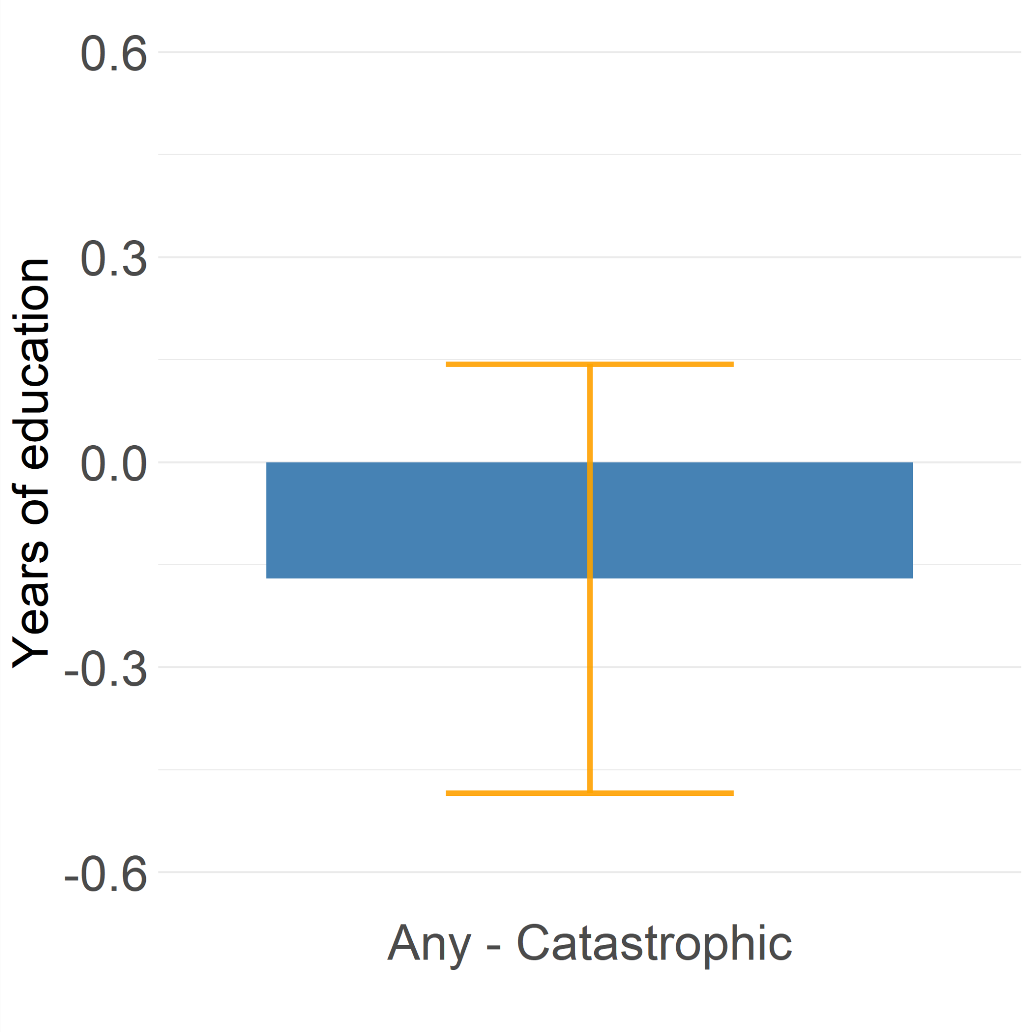
RAND balance test: years of education



RAND balance test: years of education (treat - control)



RAND balance test: years of education (treat - control)



- Bars show 95% confidence intervals
- Is the difference statistically significant?

RAND balance table

	Mean control	Treatment - Control
Female	0.56	-0.030 (0.013)
Nonwhite	0.172	-0.025 (0.022)
Age	32.4	0.64 (0.54)
Education	12.1	-0.17 (0.16)
Family income	31,603	-654 (1,181)
Hospitalized last year	0.115	0.001 (0.013)

- Standard errors given in parentheses
- Can you approximate the 95% confidence interval for education? How about others?
- Are any of these differences statistically significant?

Causal effect of insurance on utilization

	Mean control	Treatment - Control
Face-to-face visits	2.78	0.90 (0.20)
Outpatient expenses	248	101 (17)
Hospital admissions	0.099	0.017 (0.009)
Inpatient expenses	388	97 (53)
Total expenses	636	198 (63)

- Standard errors given in parentheses
- Are any of these differences statistically significant at the 95% level?

Causal effect of insurance on health

	Mean control	Treatment - Control
General health index	68.5	-0.36 (0.77)
Cholesterol (mg/dl)	203	-1.32 (2.08)
Systolic blood pressure (mm Hg)	122	-0.36 (0.85)
Mental health index	75.5	0.64 (0.75)

- Standard errors given in parentheses
- Are any of these differences statistically significant at the 95% level?

Summary

- Difference in Group Means = Average Causal Effect + Selection Bias
    - Selection bias happens when the treatment and control groups are not similar
    - That is, average Y_{i0} of the two groups is different

- When treatment is random and sample size is large, selection bias ≈ 0
    - Difference in Group Means = Average Causal Effect

- Balance tests are used to assess whether randomization was successful