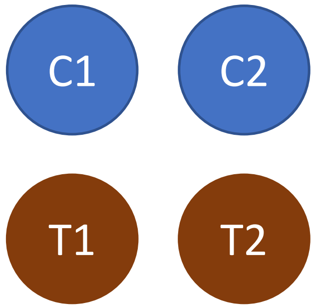
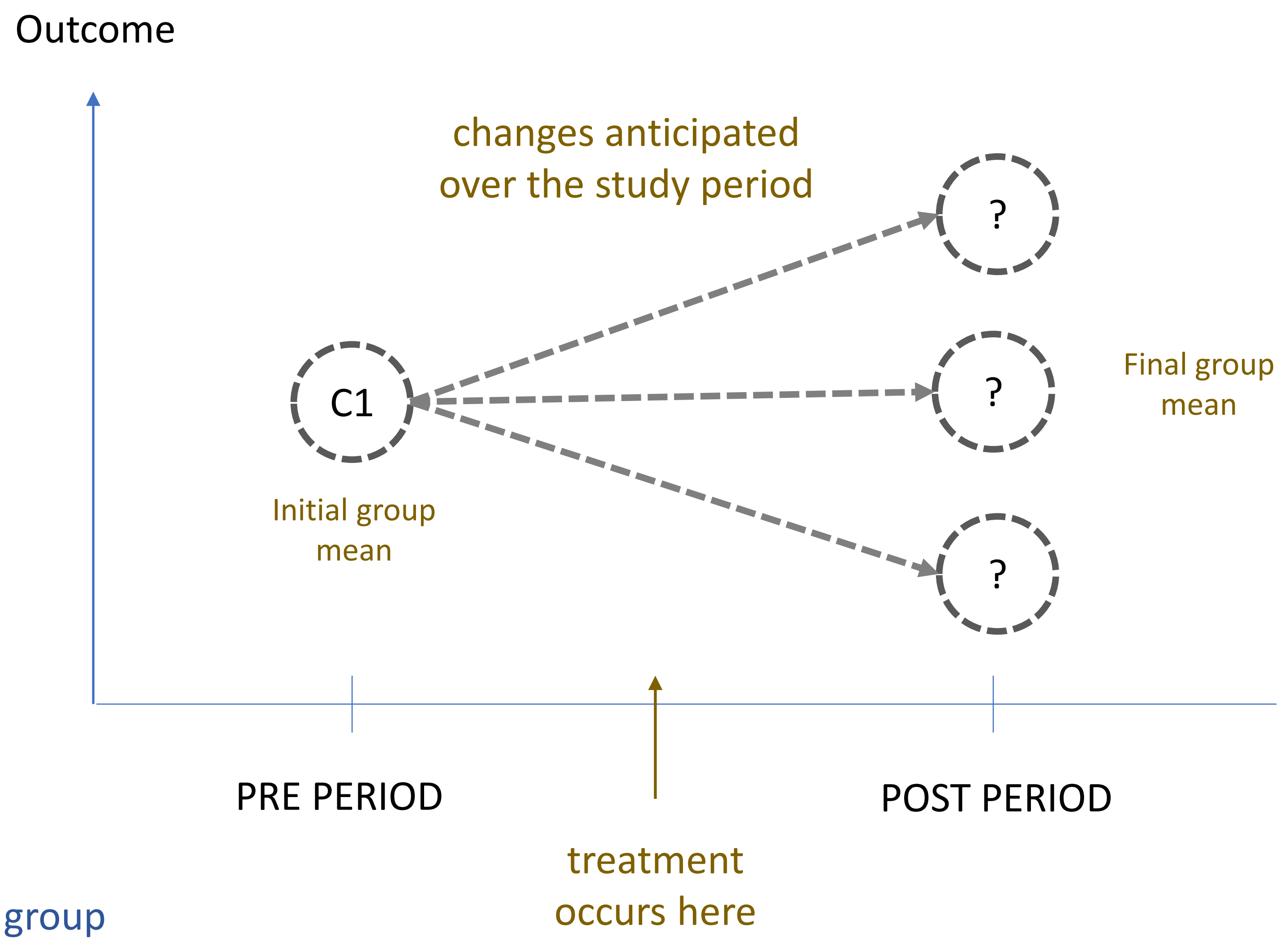
DIAGRAM YOUR RESEARCH DESIGN

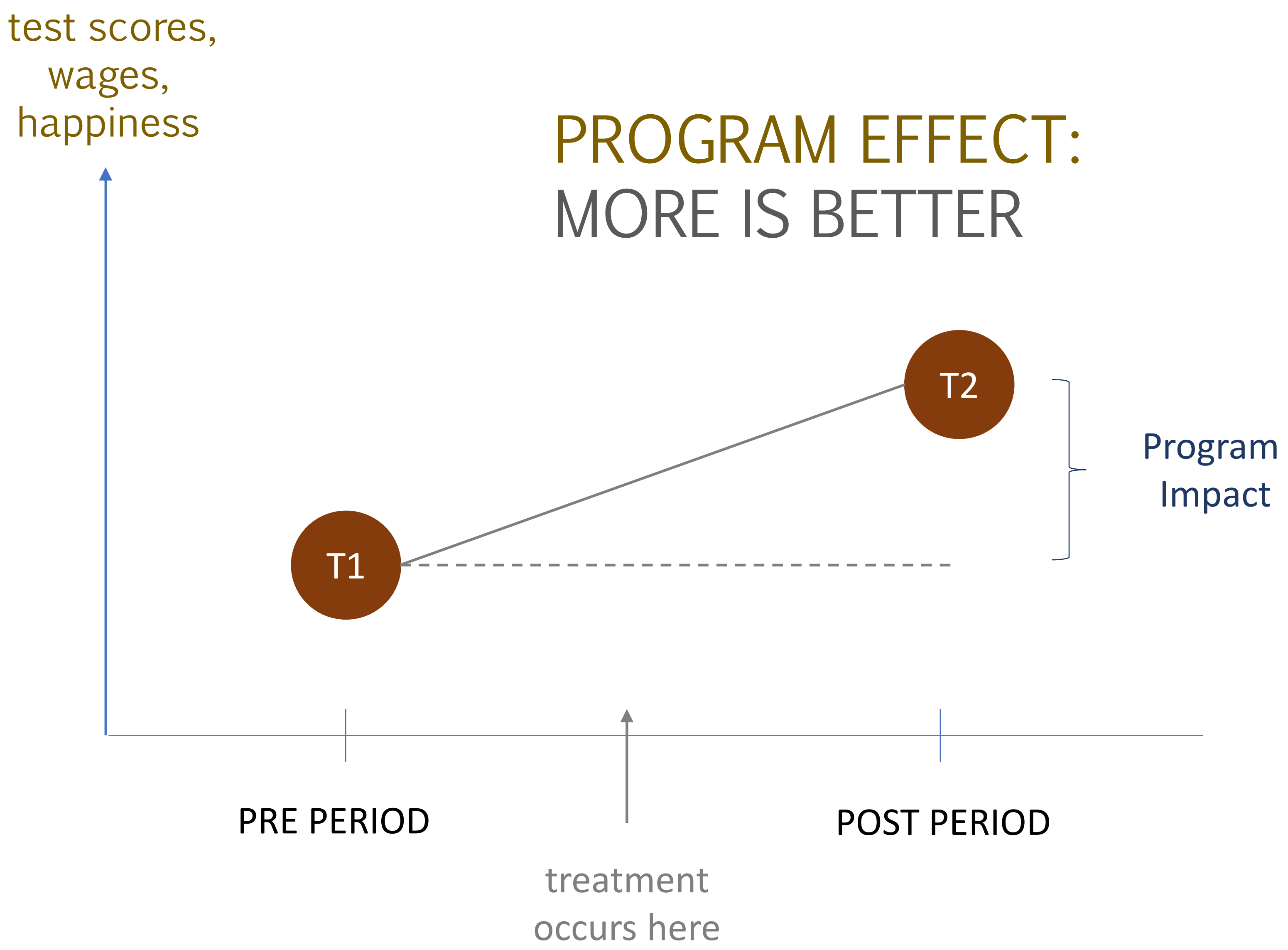
STUDY GROUP MEANS



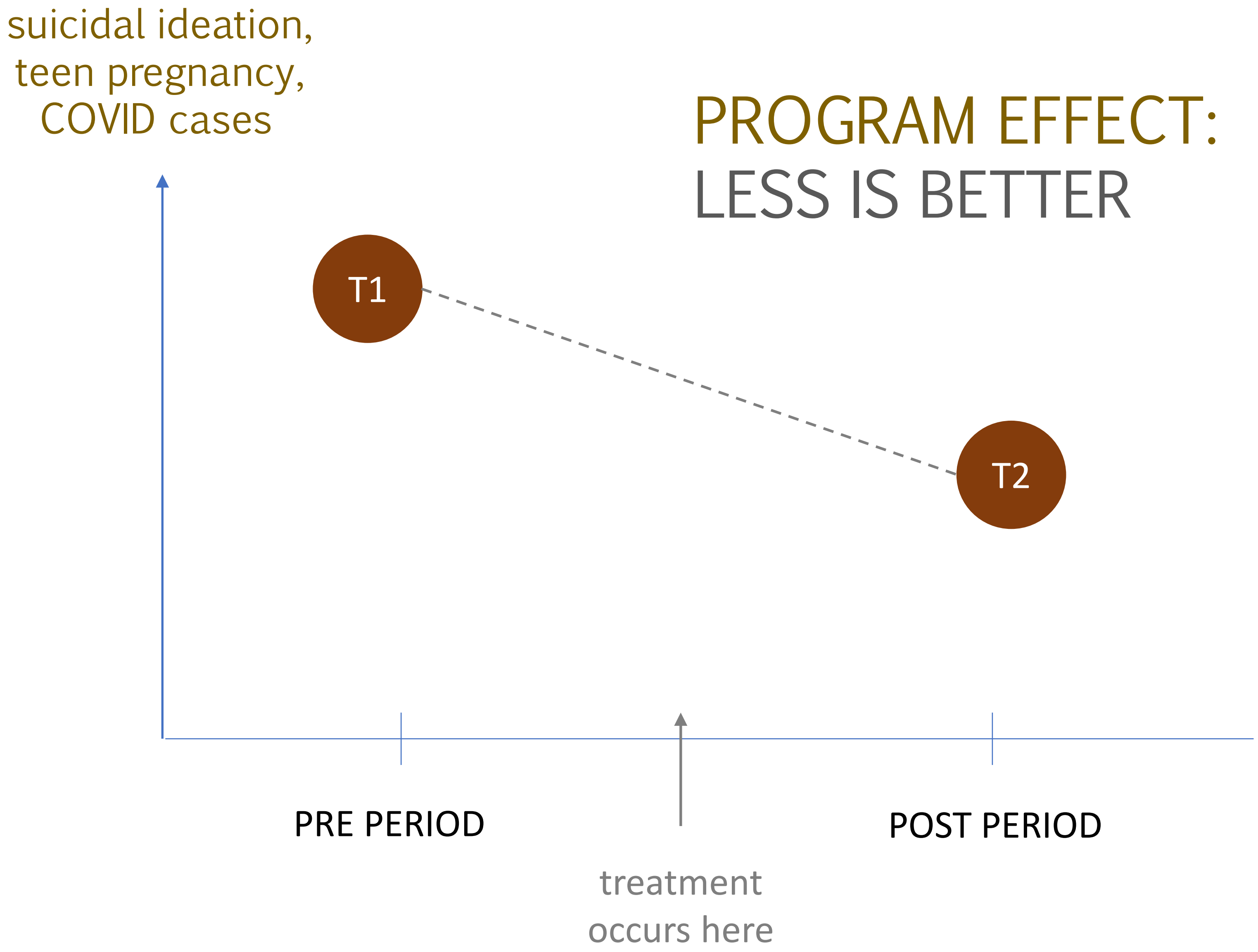


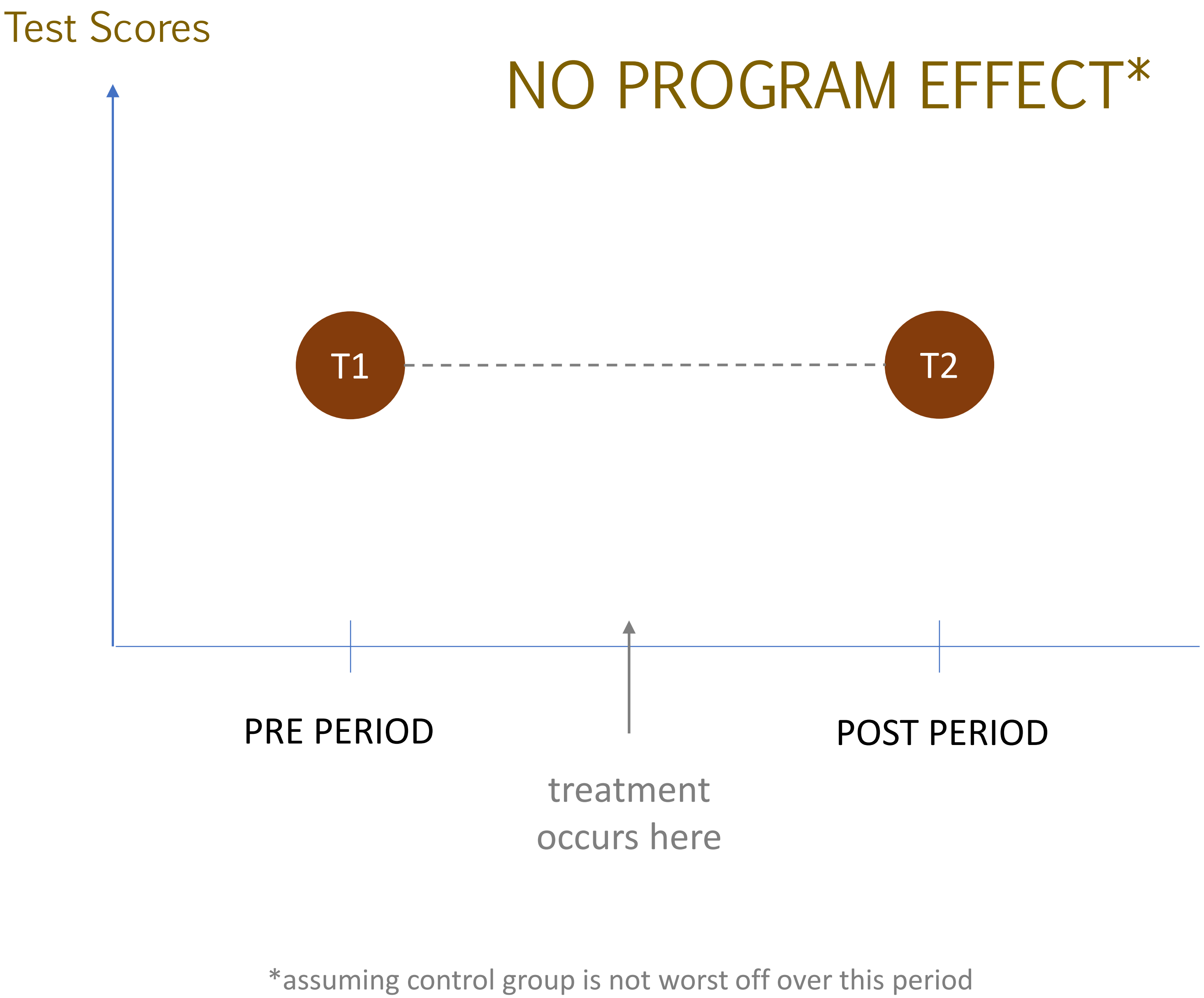
T = “treatment group”
C = “comparison” or “control” group
1 = pre-treatment period, 2=post

PROGRAM EFFECT: MORE IS BETTER



PROGRAM EFFECT:
LESS IS BETTER




Test Scores
NO PROGRAM EFFECT*
T1
T2
PRE PERIOD
POST PERIOD
treatment
occurs here
*assuming control group is not worst off over this period

On creating two study groups:

Conceptually you will have a “treatment” and “control” group. A true control group represents the counterfactual, and a comparison group is used to construct your counterfactual (typically by providing a measure of trend that is then used to predict where the treatment group would have ended up if it had not been for the treatment).

If everyone in your study receives different levels of the treatment (blood pressure medication, hours of tutoring each week) then for this exercise create two groups – a little treatment and a lot of the treatment – and use those to reason through your design.

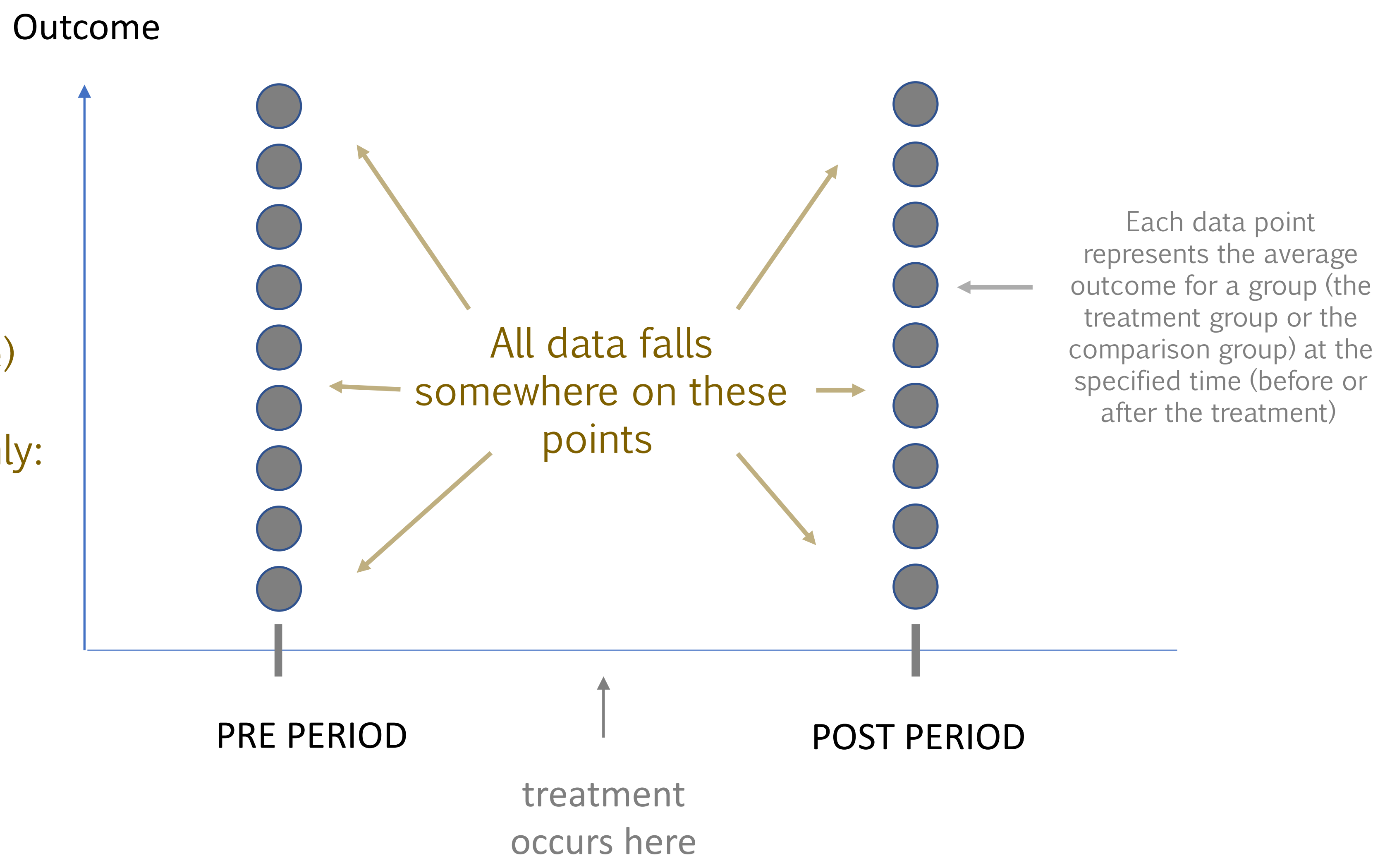
In general, it is helpful to think through your research design in terms of these pure groups, and make sure you understand the nature of your counterfactual in your specific context.

PRE-TREATMENT DIFFERENCES

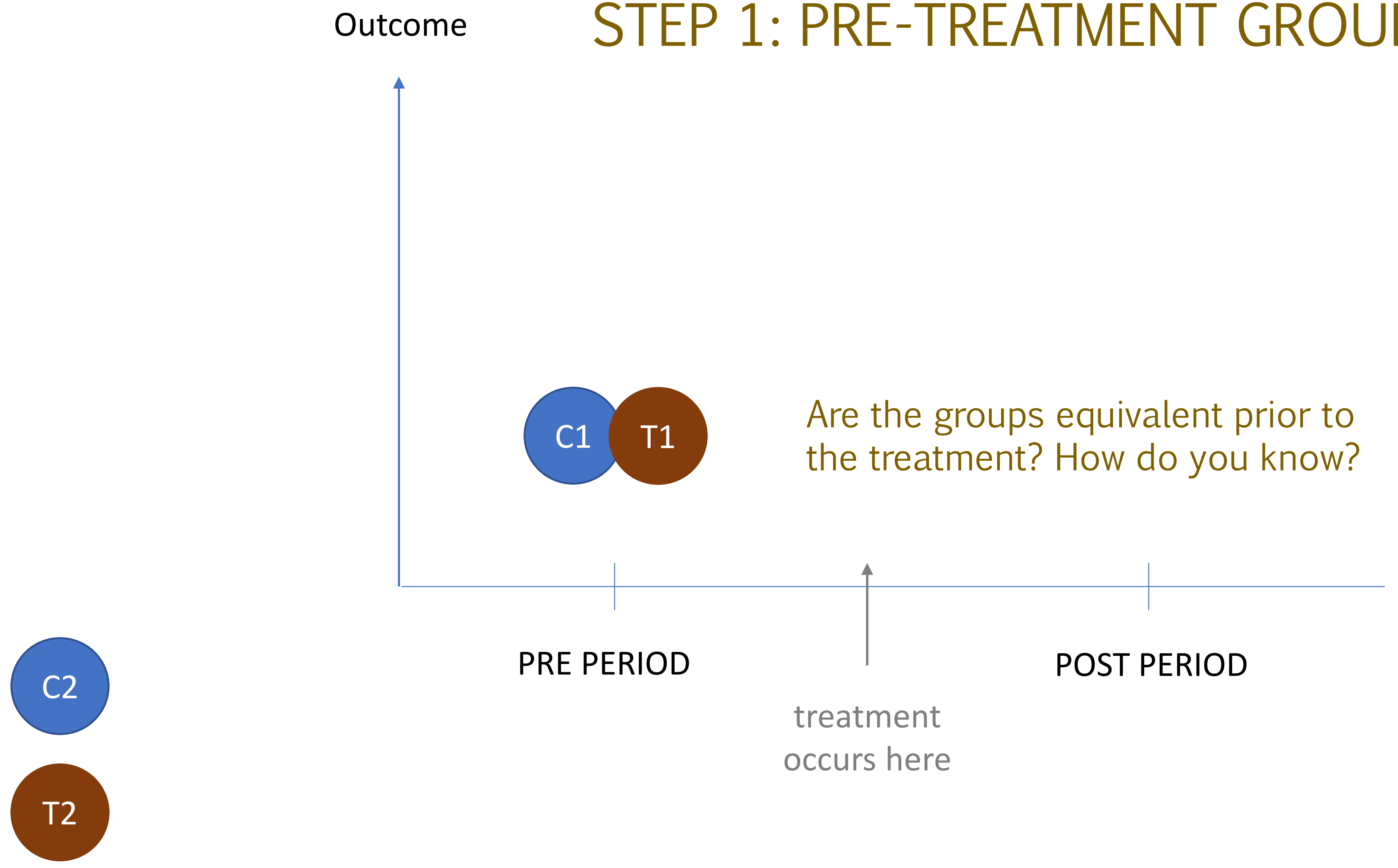
SET UP YOUR DIAGRAM:

Y-axis is continuous
(any position on axis possible)

X-axis is two discrete points only:
‘before’ and ‘after’ treatment

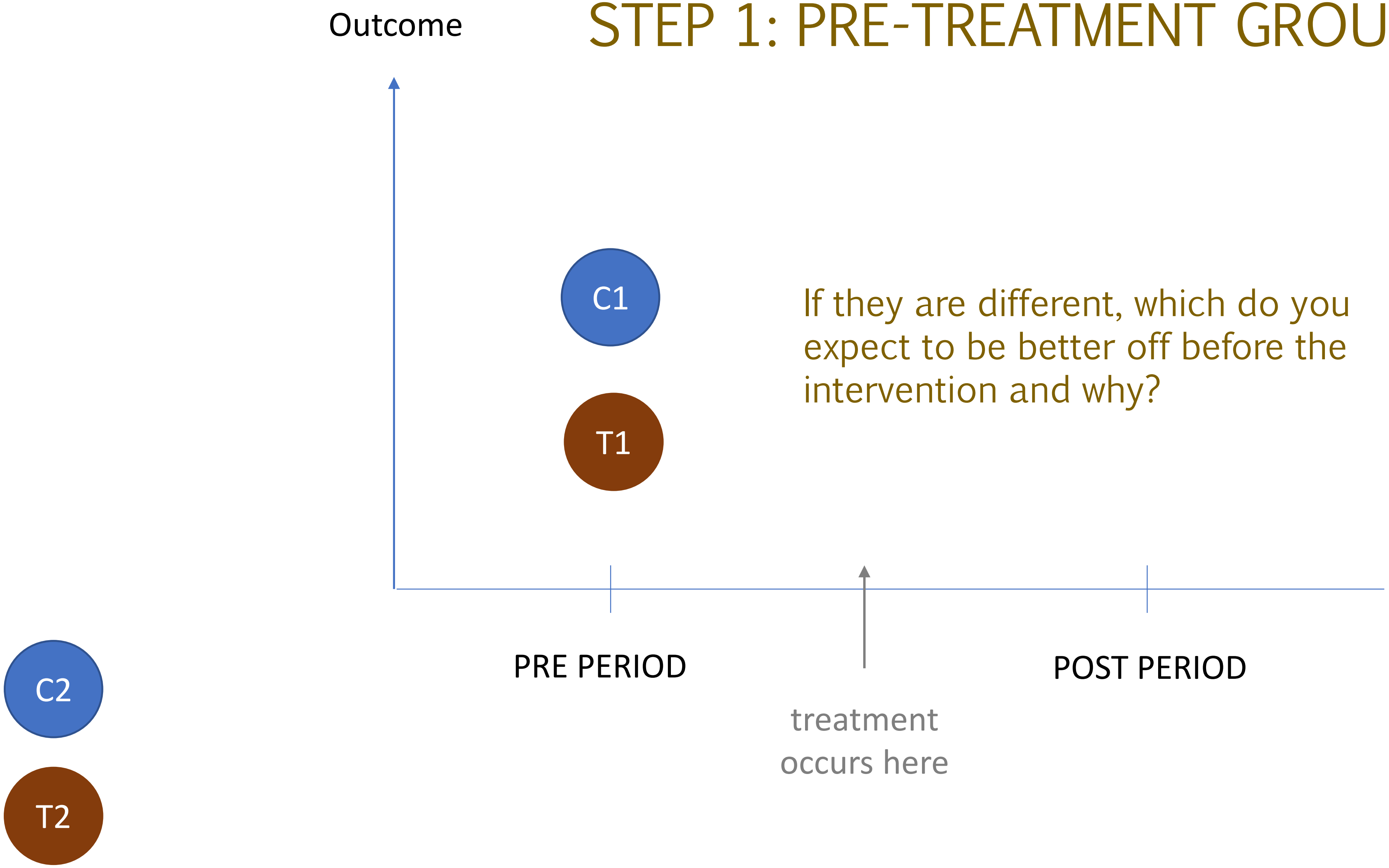


STEP 1: PRE-TREATMENT GROUPS



Are the groups equivalent prior to the treatment? How do you know?

STEP 1: PRE-TREATMENT GROUPS

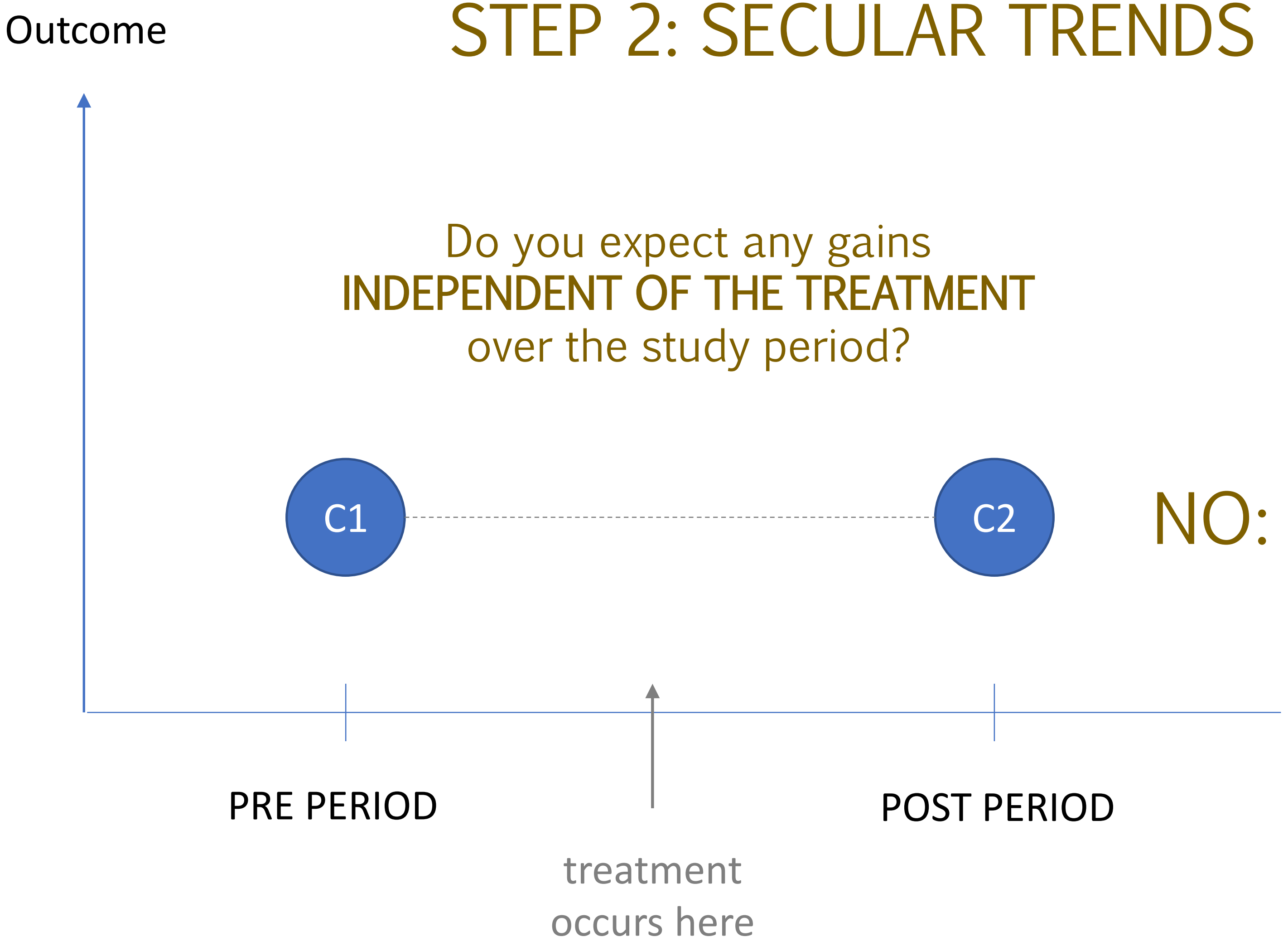


If they are different, which do you expect to be better off before the intervention and why?

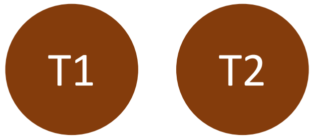
GAINS IN THE COMPARISON GROUP

STEP 2: SECULAR TRENDS

Do you expect any gains
INDEPENDENT OF THE TREATMENT
over the study period?



NO: C2-C1=0

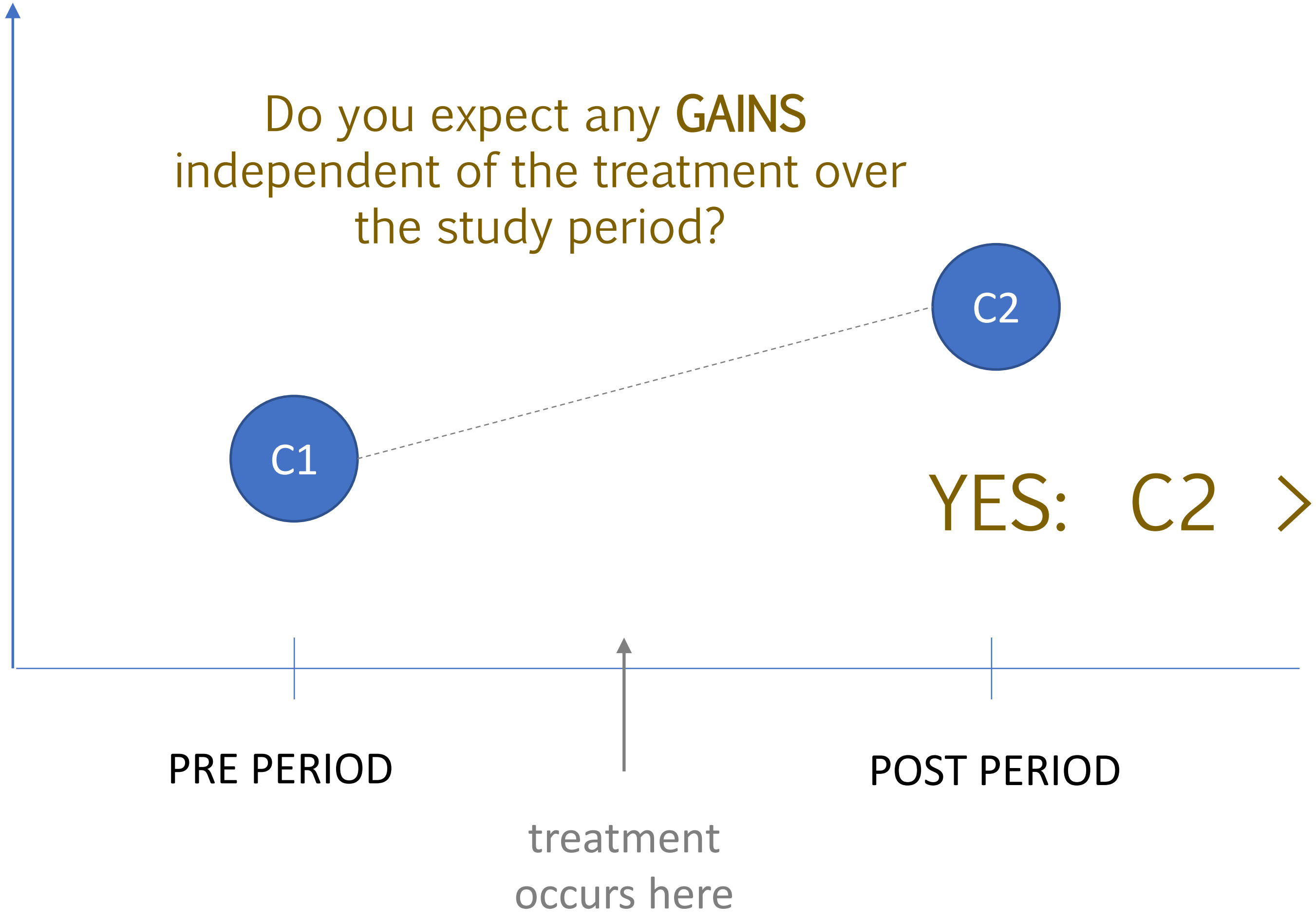


STEP 2: SECULAR TRENDS

EXAMPLE:

Vocabulary of elementary school children

Do you expect any **GAINS** independent of the treatment over the study period?



YES: C2 > C1

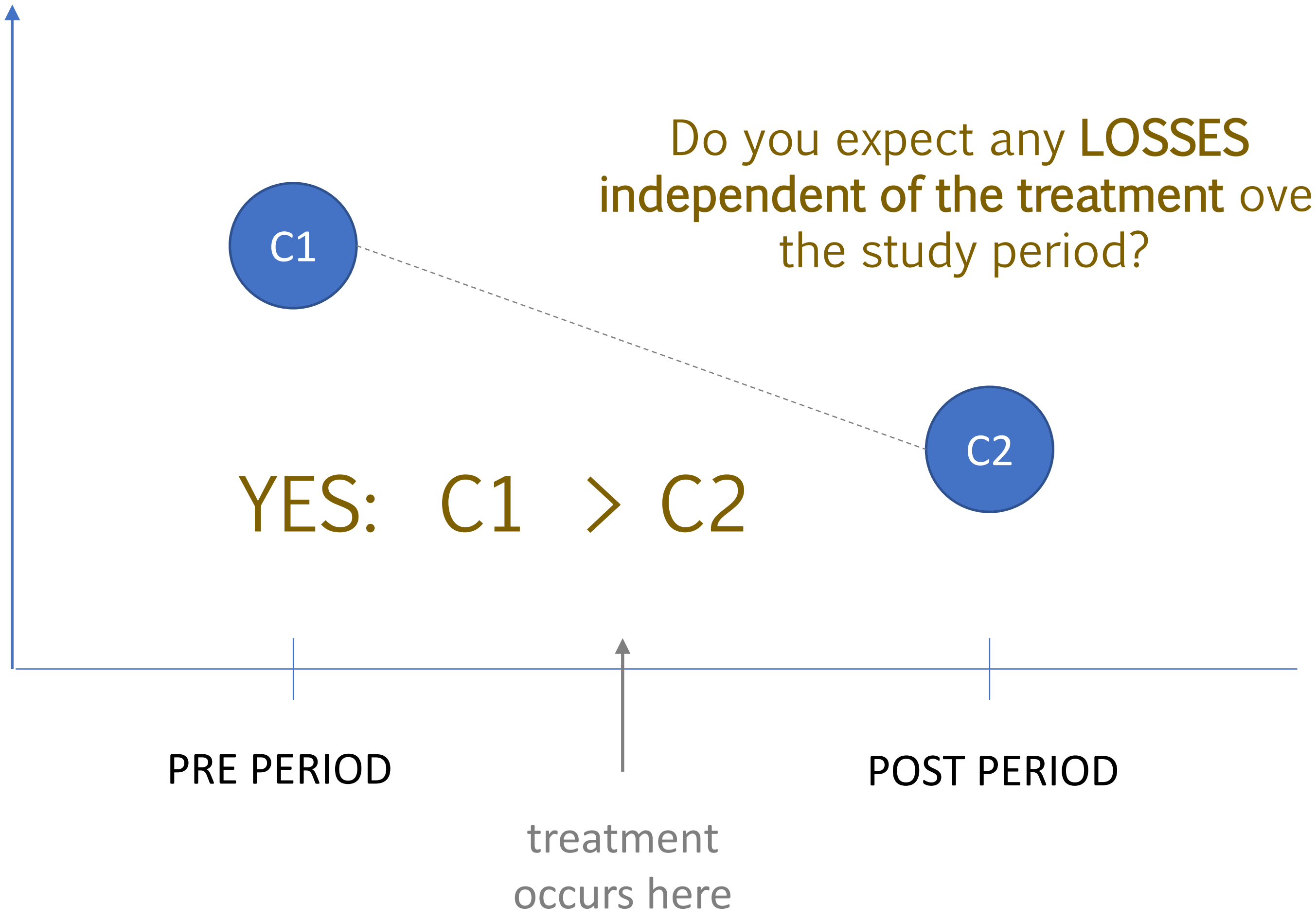
T1 T2

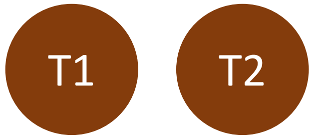
EXAMPLE:

Cognitive function of elderly in nursing homes

STEP 2: SECULAR TRENDS

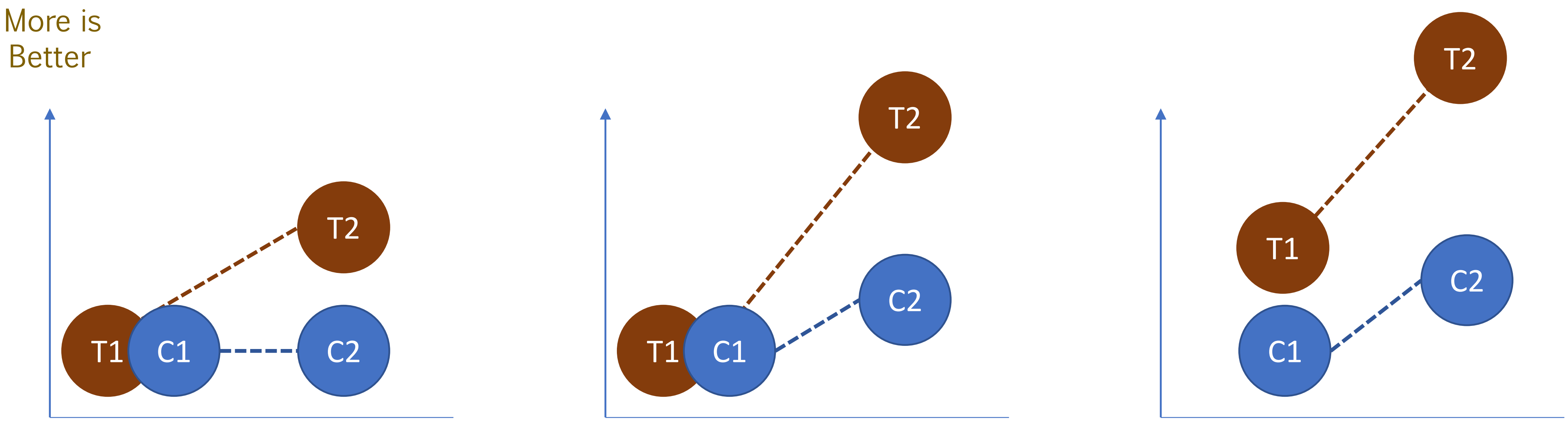
Do you expect any **LOSSES independent of the treatment** over the study period?





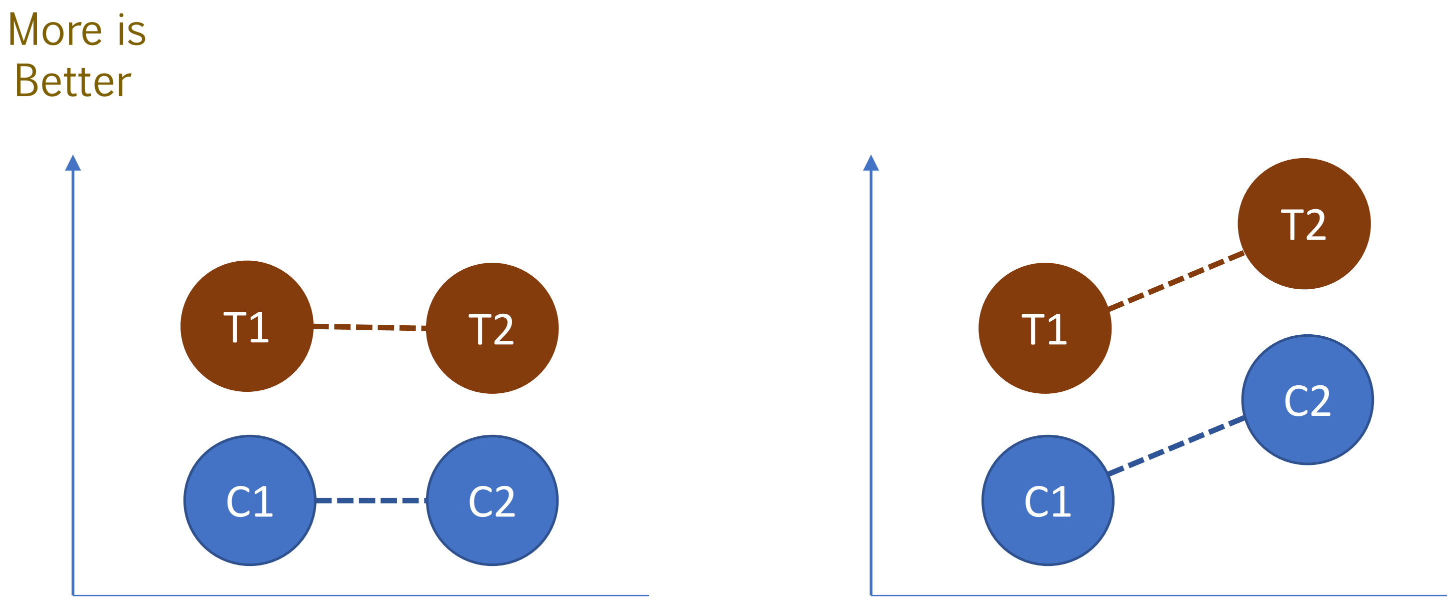
GAINS IN THE TREATMENT GROUP

STEP 3: DOES THE TREATMENT WORK?



YES: T2-T1 > C2-C1

STEP 3: DOES THE TREATMENT WORK?



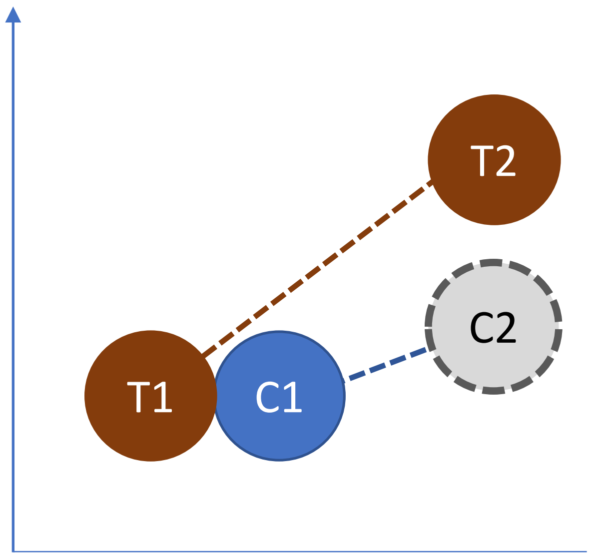
NO: T2-T1 = C2-C1

lines are parallel ::: gains are equal ::: no program impact

COUNTERFACTUALS AND EFFECTS

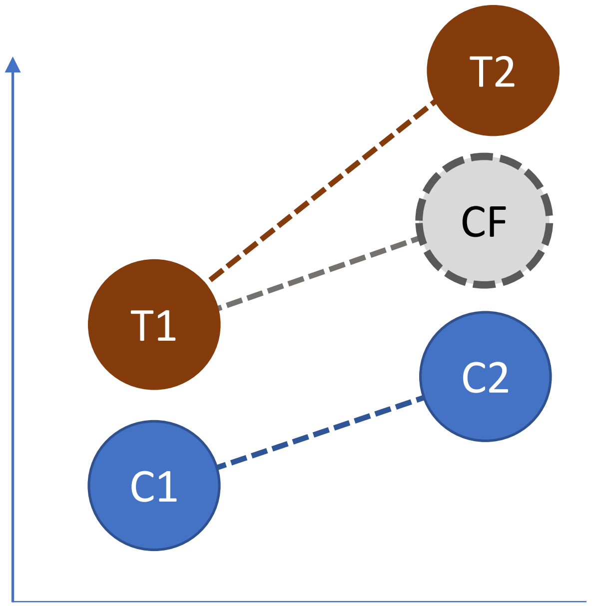
STEP 4: WHERE IS YOUR COUNTERFACTUAL?

What would your treatment group look like if it had not received the treatment?



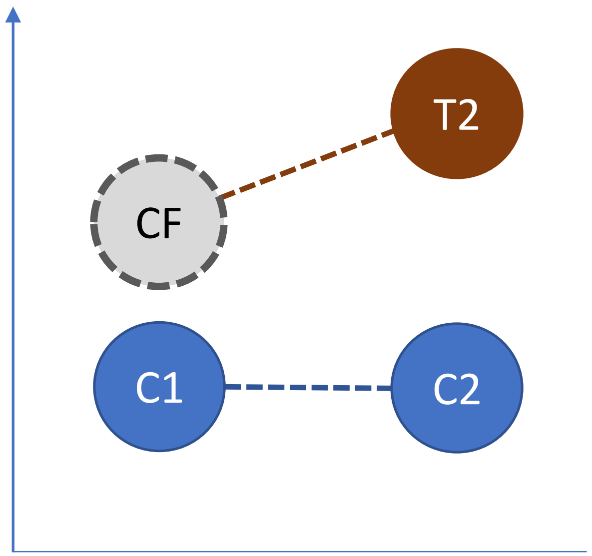
C1=T1

C2 is the counterfactual



C1≠T1

T1 + (C2-C1) is the counterfactual

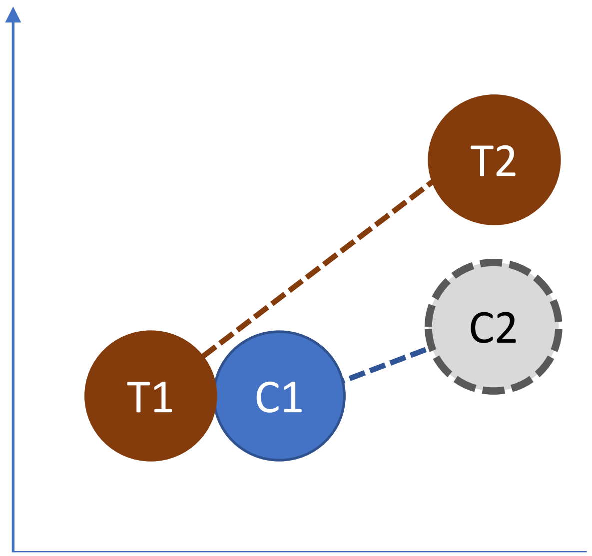


C1=C2

T1 is the counterfactual

STEP 4: WHERE IS YOUR COUNTERFACTUAL?

What would your treatment group look like if it had not received the treatment?



C1=T1

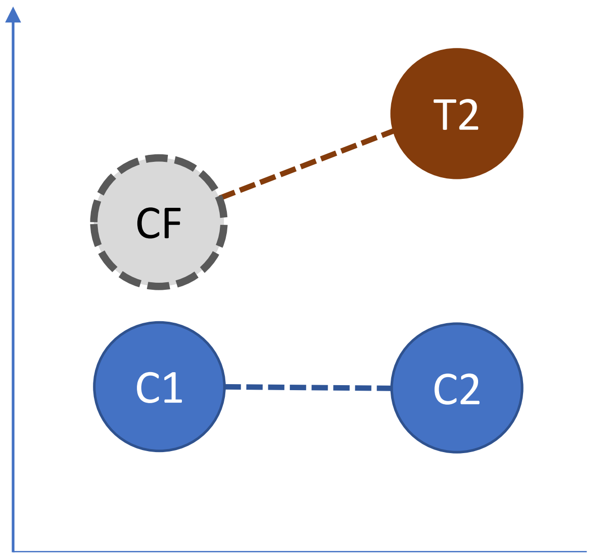
C2 is the counterfactual

T2-C2 captures program effects

STEP 4: WHERE IS YOUR COUNTERFACTUAL?

What would your treatment group look like if it had not received the treatment?

T2-T1 captures program effects

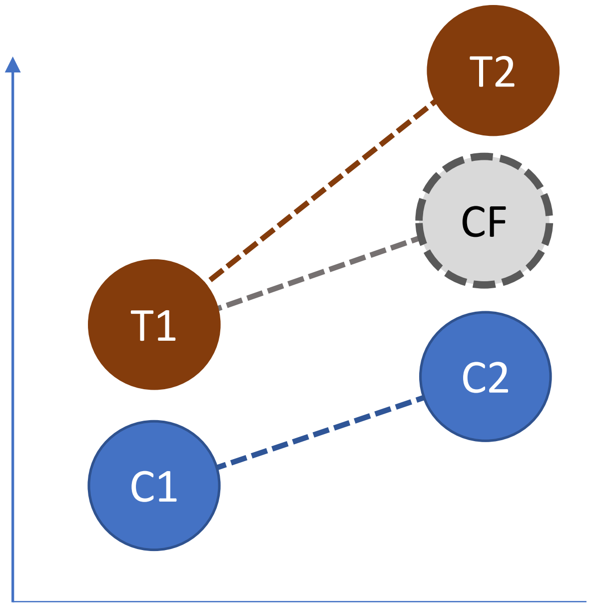


C1=C2

T1 is the counterfactual

STEP 4: WHERE IS YOUR COUNTERFACTUAL?

What would your treatment group look like if it had not received the treatment?



(T2-T1) – (C2-C1)

(total gain – trend)

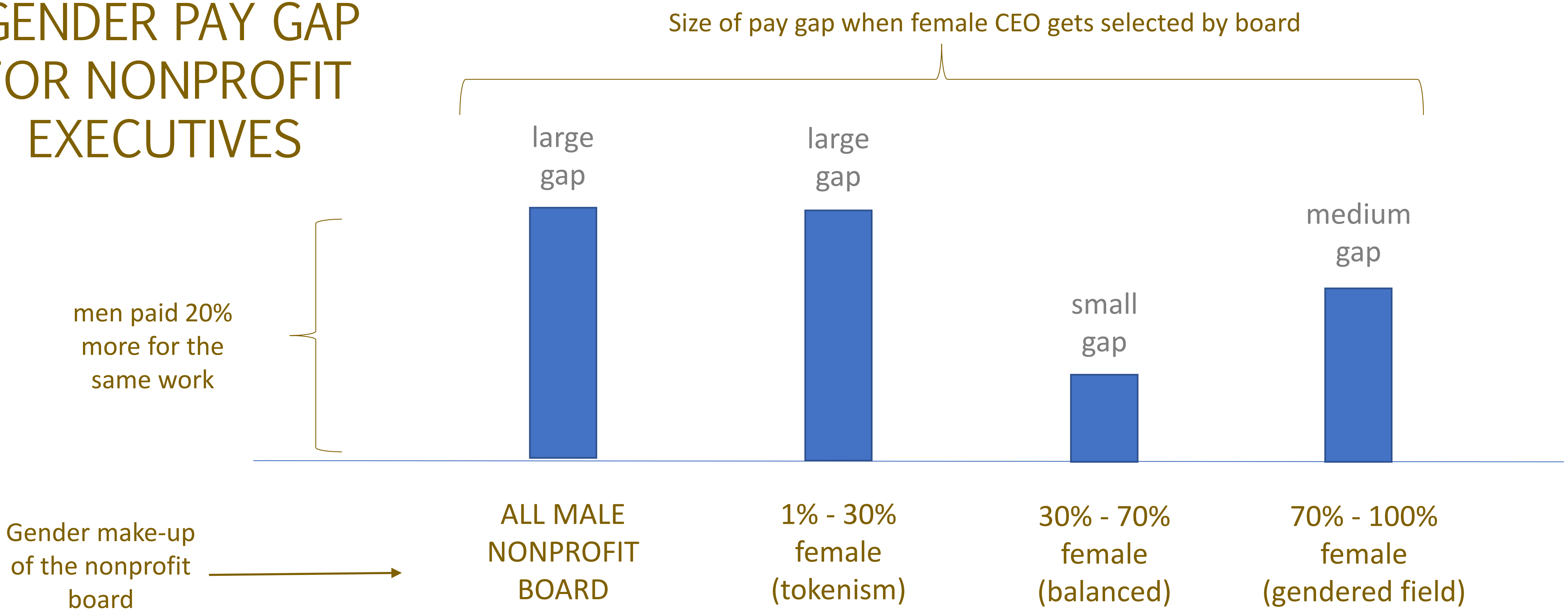
captures
program effects

<u>C1≠T1</u>

T1 + (C2-C1)
is the
counterfactual

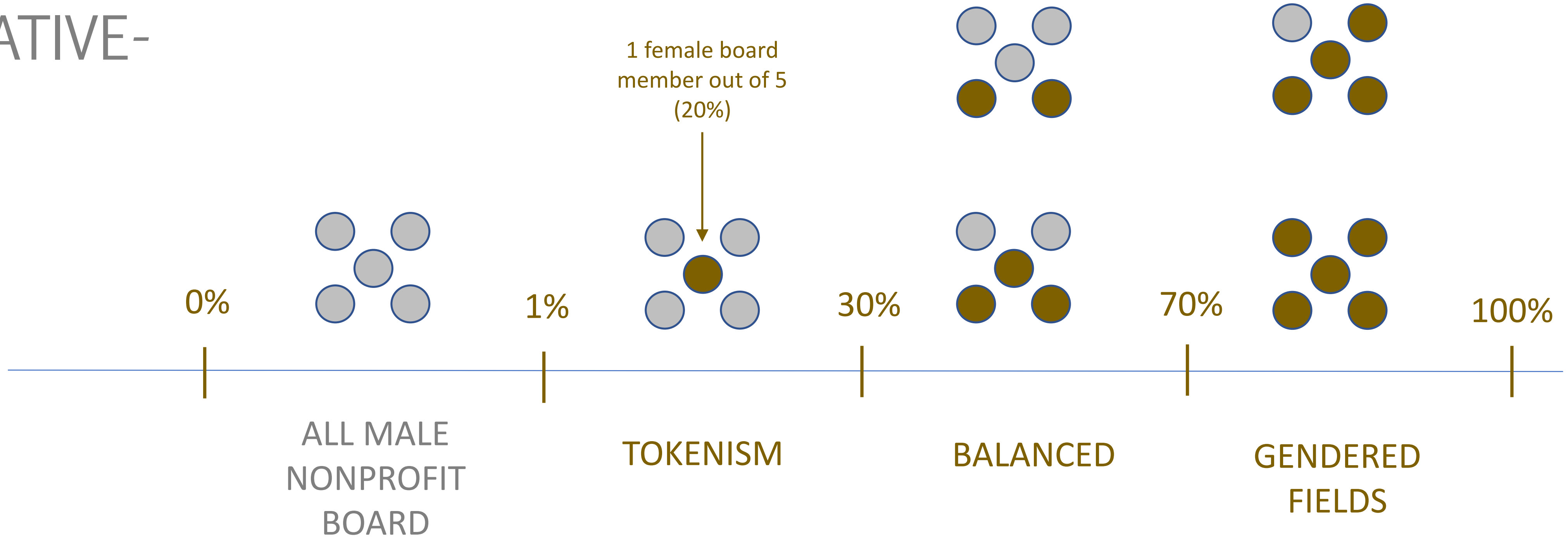
REAL-WORLD EXAMPLE

GENDER PAY GAP FOR NONPROFIT EXECUTIVES



Critical Mass Theory (CMT): CMT argues that there needs to be proportional representation (balance) for changes (e.g. – social issues) to be seen in organizations.

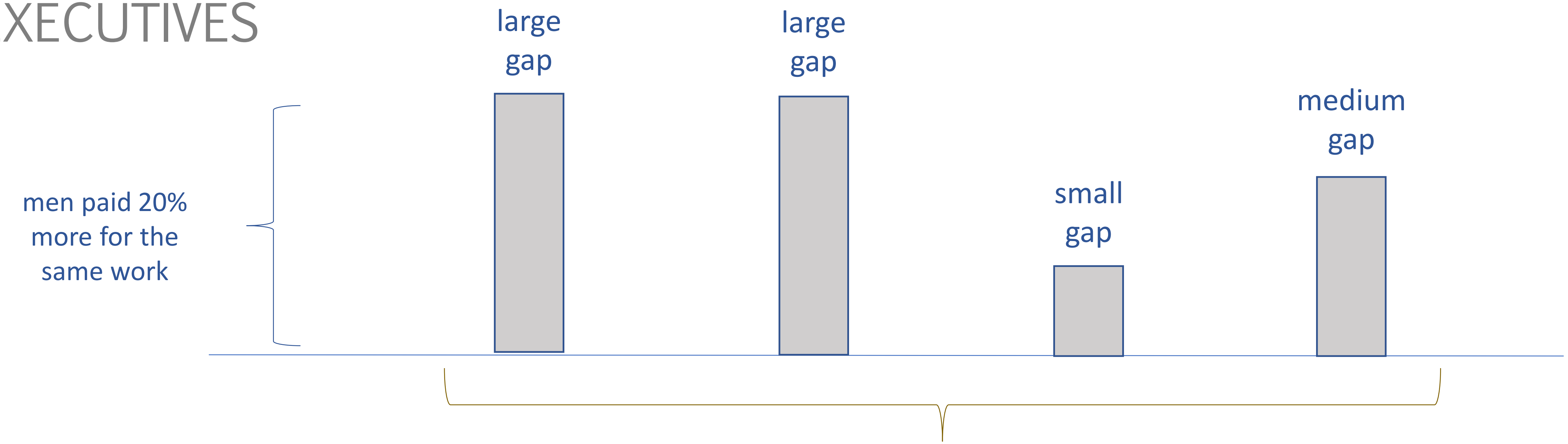
BOARD REPRESENTATIVE-NESS



- **TOKENISM** (less than proportional representation) may cause members to conform to the majority.
- **BALANCE** is expected to give female board members influence over the final salary.
- **MAJORITY FEMALE** boards tend to occur in gendered fields like reproductive health and thus pay is lower relative to other similar nonprofits (size, performance, etc.) doing generic public health.

GENDER PAY GAP FOR NONPROFIT EXECUTIVES

SIZE OF PAY GAP WHEN FEMALE CEO GETS SELECTED BY THE BOARD

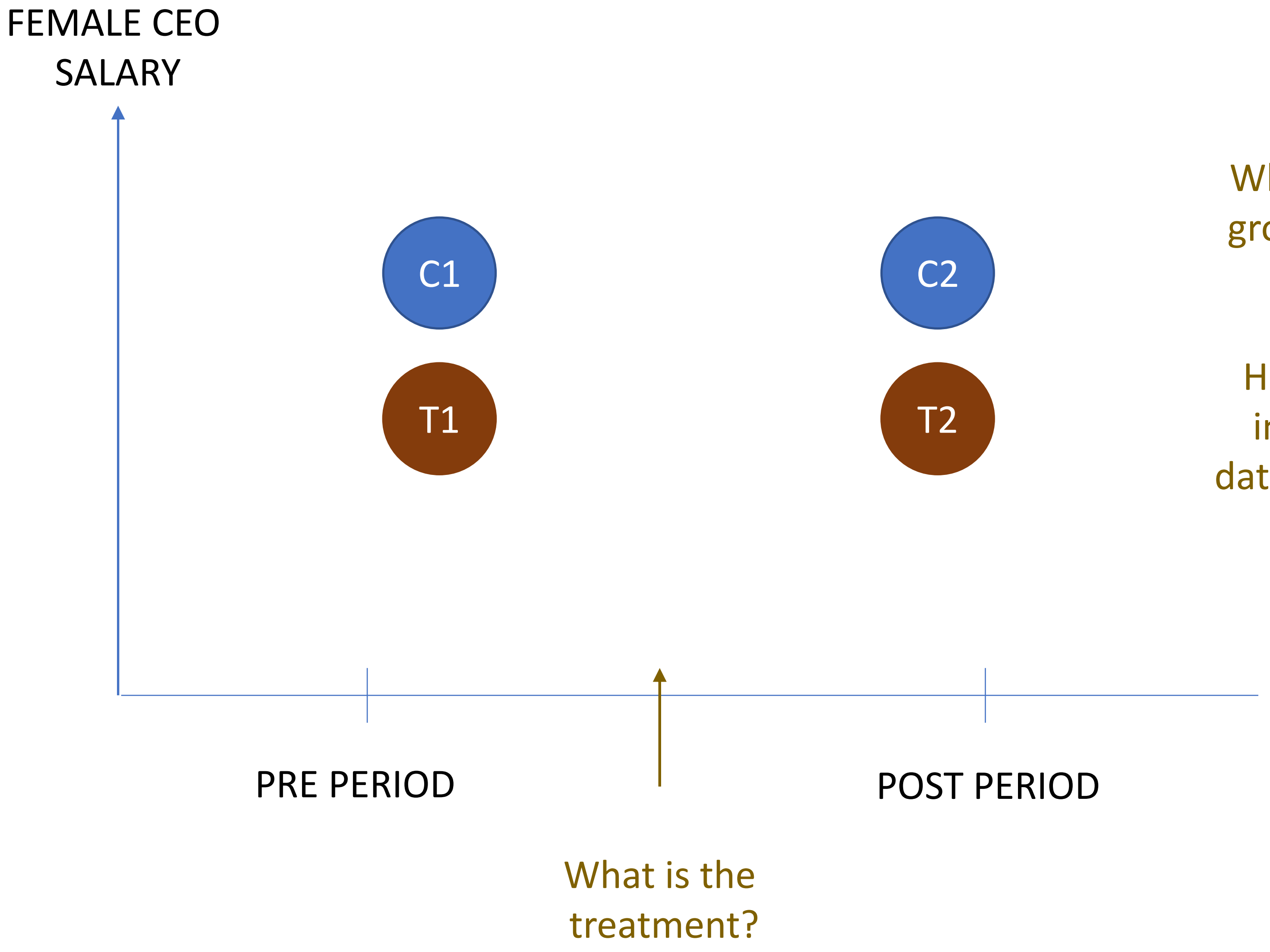


The **treatment** is the balance of the board when a female CEO is hired.
(are there 4 different treatments, or 3 treatments and a control?)

The **outcome** of interest is the gender pay gap.

The **counterfactual** would then be,
what would the salary be if the organization had a different board when they hired the current female CEO?

What **observable groups** capture the treatment in each case???



Who is in the treatment group, and who is in the control group?

How are you assigning individuals from your dataset to each category?

GENERATING HYPOTHESES: WHEN DO WE EXPECT TO SEE CHANGES IN THE DATA?

How can we use our data to construct meaningful thought experiments to test our theory?

The quasi-experimental approach (use thought experiments to create testable hypotheses).

According to the theory, what SHOULD BE TRUE IF:

- **board diversity increases between time=1 and time=2**?

- **board diversity decreases over the same time period**?

DIAGRAM YOUR DATA BASED UPON EXPECTATIONS INFORMED BY THEORY

DATA:

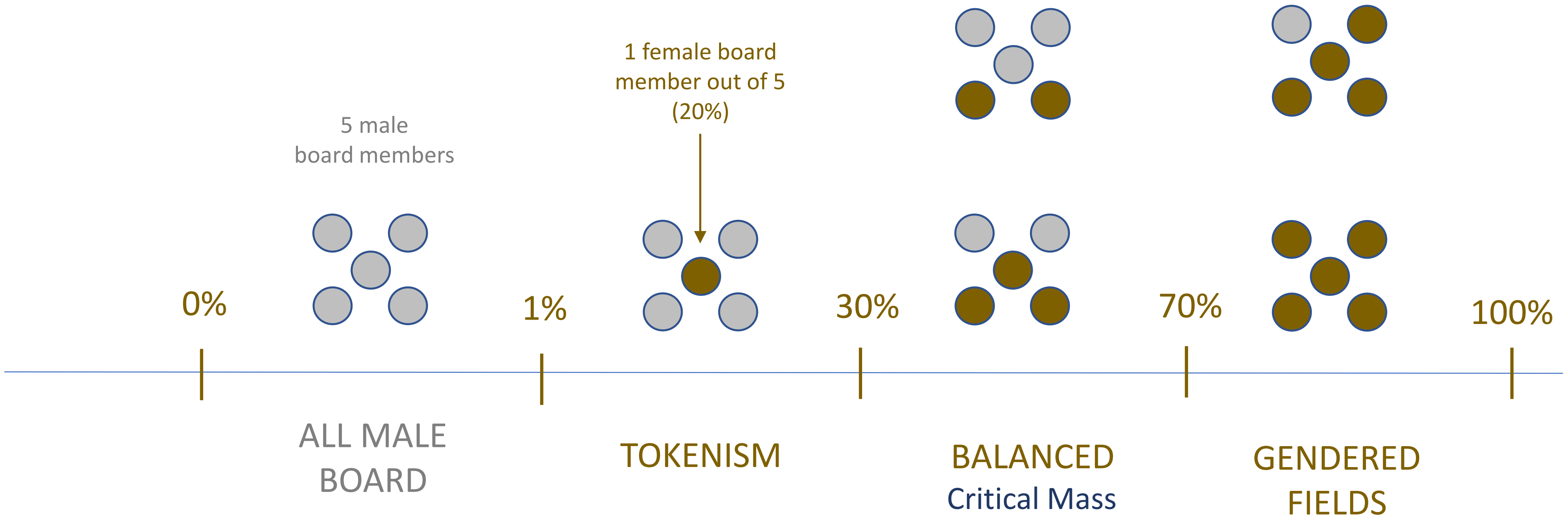
- 10-year panel on nonprofit CEO pay
- Data includes the gender of CEOs all board members in the nonprofits
- *Note that male/female binary is limited, but only metrics available in the data (future research could use different constructs for gender)*
- Limit the data to only organizations that HIRE TWO FEMALE CEOs within the 10-year study period so that we can see the starting salary of both
- Treatment – isolate baseline board diversity scenarios so all organizations are the same in the first time period, then look for cases that change in the second time period.
- Theory: The board sets CEO pay at the time of hire, so the diversity of the board during compensation discussions impacts whether there is implicit bias in .

For simplicity assume all boards have 5 members so that:

- 0 female board members = **no diversity**
- 1 female member (20%) = **token representation**
- 2 female members (40%) = **critical mass / balanced**

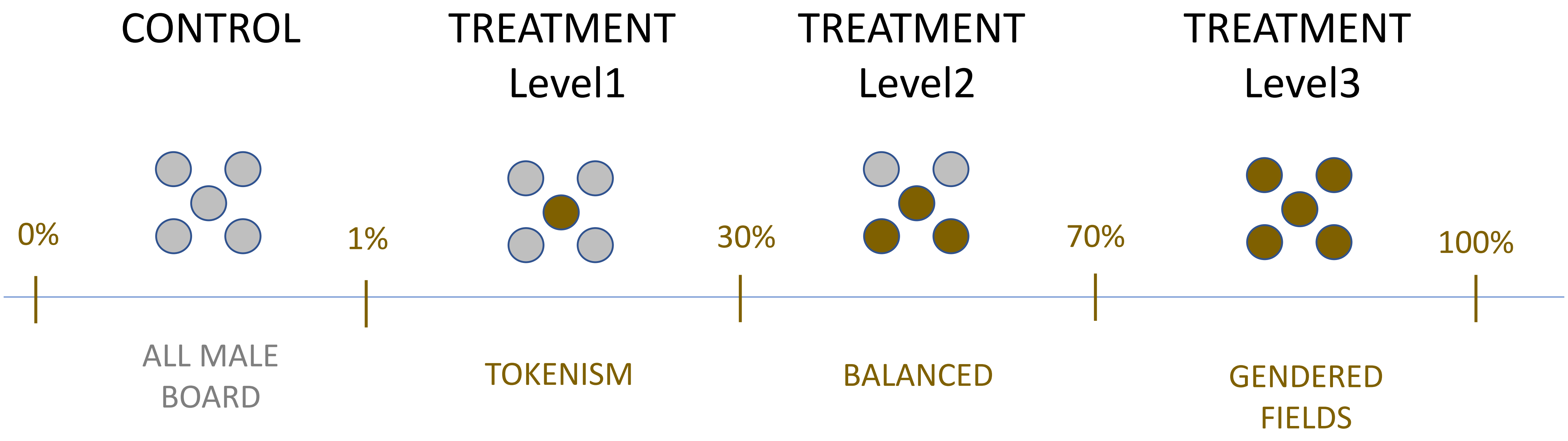
For simplicity assume all boards have 5 members so that:

- 0 female board members = **no diversity**
- 1 female member (20%) = **token representation**
- 2/3 female members (40/60%) = **critical mass / balanced**



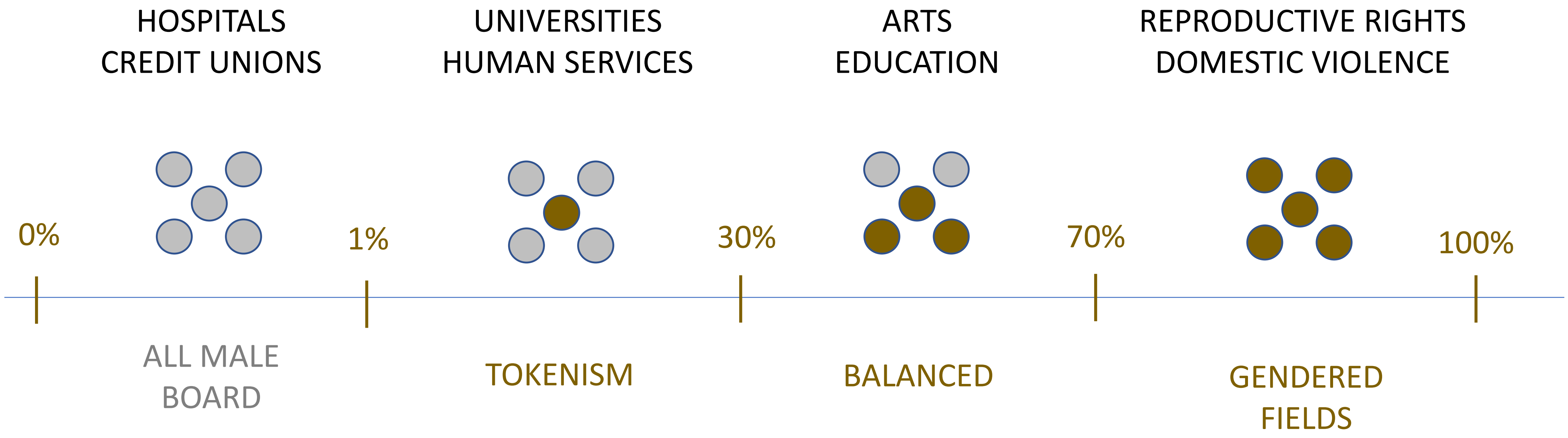
- **TOKENISM** (less than proportional representation) may cause members to conform to the majority.
- **BALANCE (CRITICAL MASS OF REPRESENTATION)** is expected to give female board members influence over the final salary.

THE CHALLENGE OF OPERATIONALIZING THE "TREATMENT GROUP"



Wage gap = b0 + b1(PropBoardFemale) + e

If we would like to know whether board diversity has an impact on the gender wage gap, what is the fundamental problem with this approach to research design?

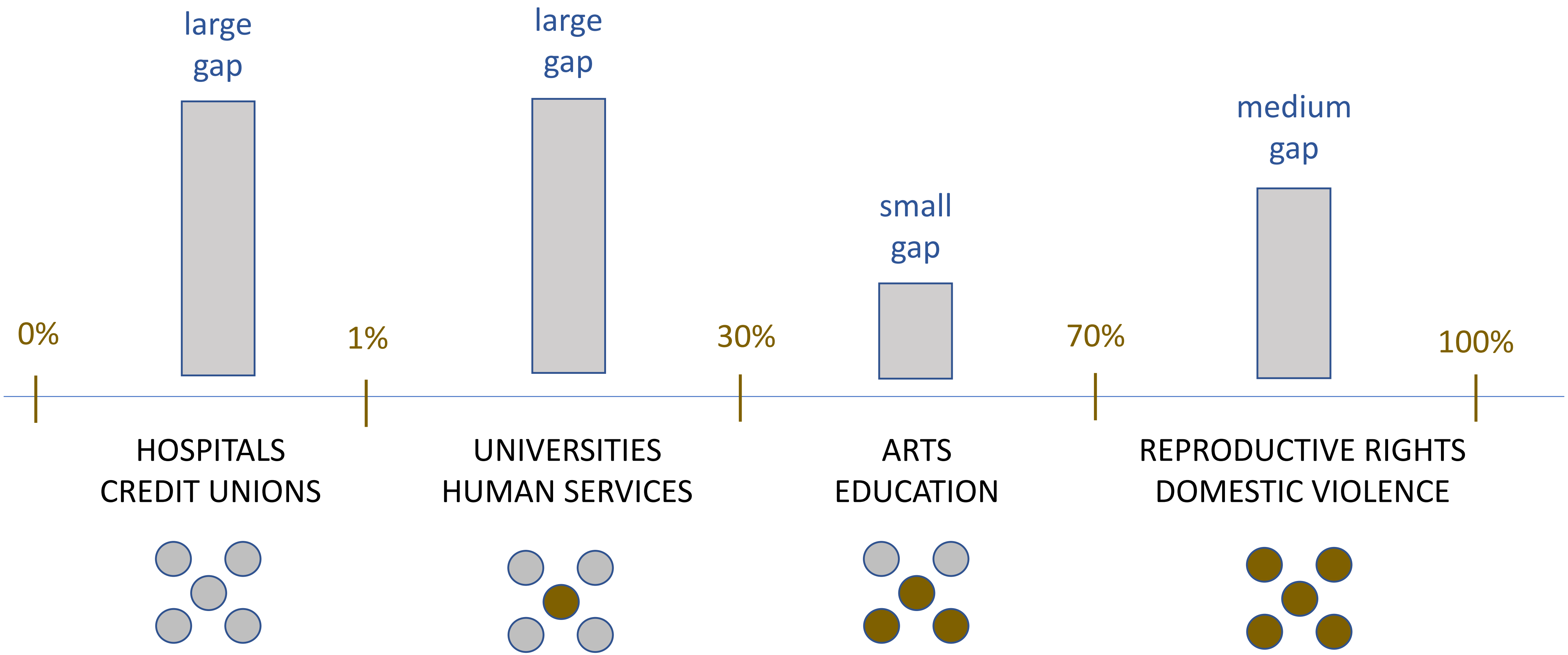


THE SELECTION PROBLEM

Board diversity will be correlated with the industry the nonprofit operates in. Some fields are either highly “gendered” (nursing and daycare) or gender toxic and discriminatory (economics, tech, finance).

Pay gaps vary by industry. Board diversity also varies by industry.

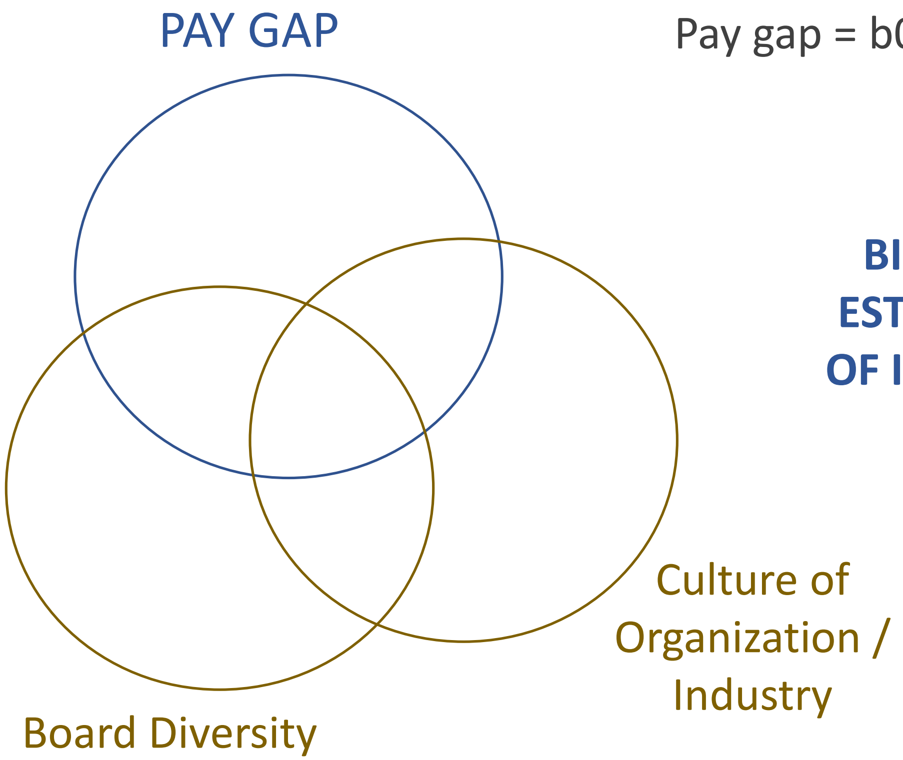
If we observe a pay gap as a function of board diversity, is it diversity that causes the pay gap, or selection into or out of an industry by men and women because of their work preferences and experiences?



THE SELECTION PROBLEM

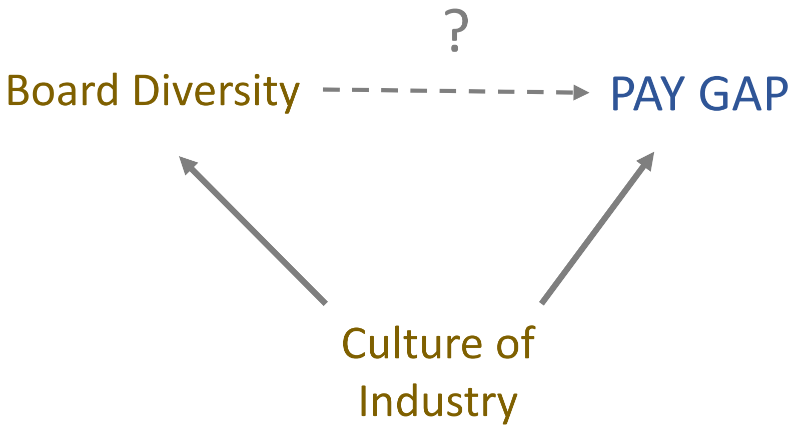
We have an omitted variable problem. We cannot isolate board diversity from other characteristics of nonprofits like their industry and size. As a result, if we estimate the impact of diversity on the pay gap we are building omitted variable bias into the analysis.

As a result, we **CANNOT USE CROSS-SECTIONAL DIFFERENCES** to estimate the impact of board diversity on pay balance.



Pay gap = b0 + **b1**(PropFemaleBoard) + e

BIASED ESTIMATE OF IMPACT

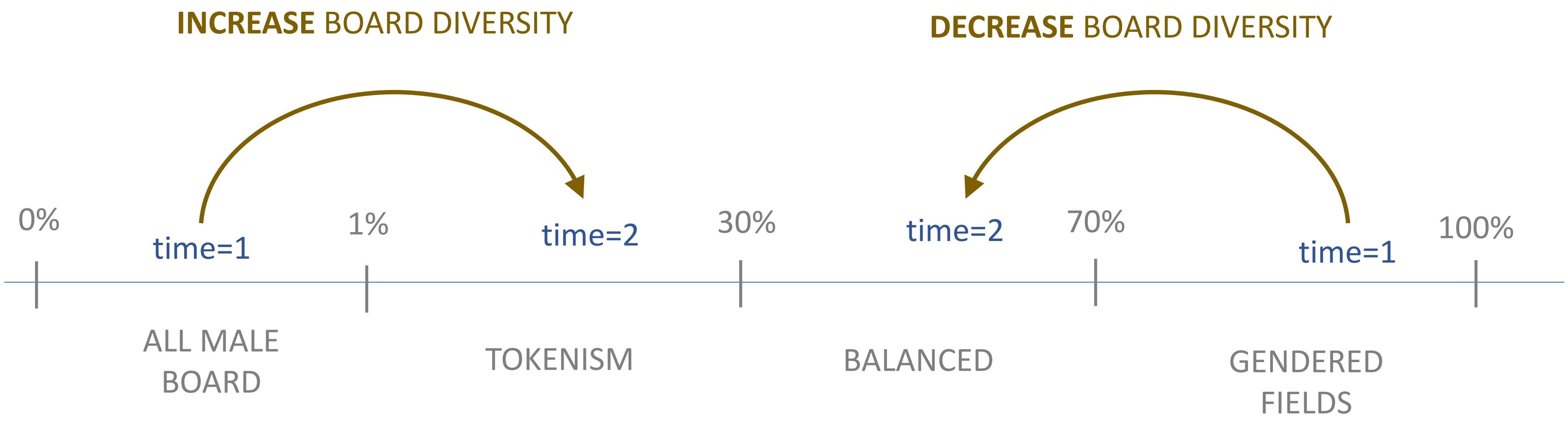


THE SELECTION PROBLEM

We have an omitted variable problem. We cannot isolate board diversity from other characteristics of nonprofits like their industry and size. As a result, if we estimate the impact of diversity on the pay gap we are building omitted variable bias into the analysis.

As a result, we **CANNOT USE CROSS-SECTIONAL DIFFERENCES** to estimate the impact of board diversity on pay balance.

What is our treatment?



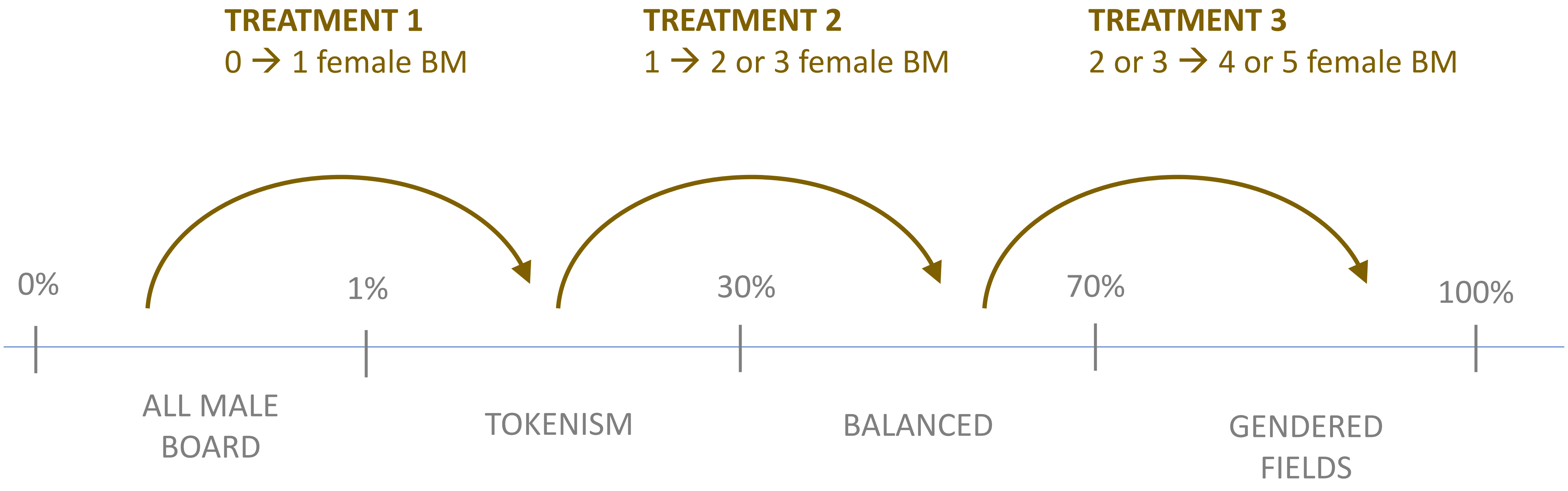
THE POWER OF PANEL DATA

We can simplify the design if we are more precise about our research question.

Initial research question: Does the level of board diversity impact the gender pay gap?

Improved: Does increasing board diversity reduce the gender pay gap?

We can now leverage **panel data** to look at how changes over time impact the outcome.

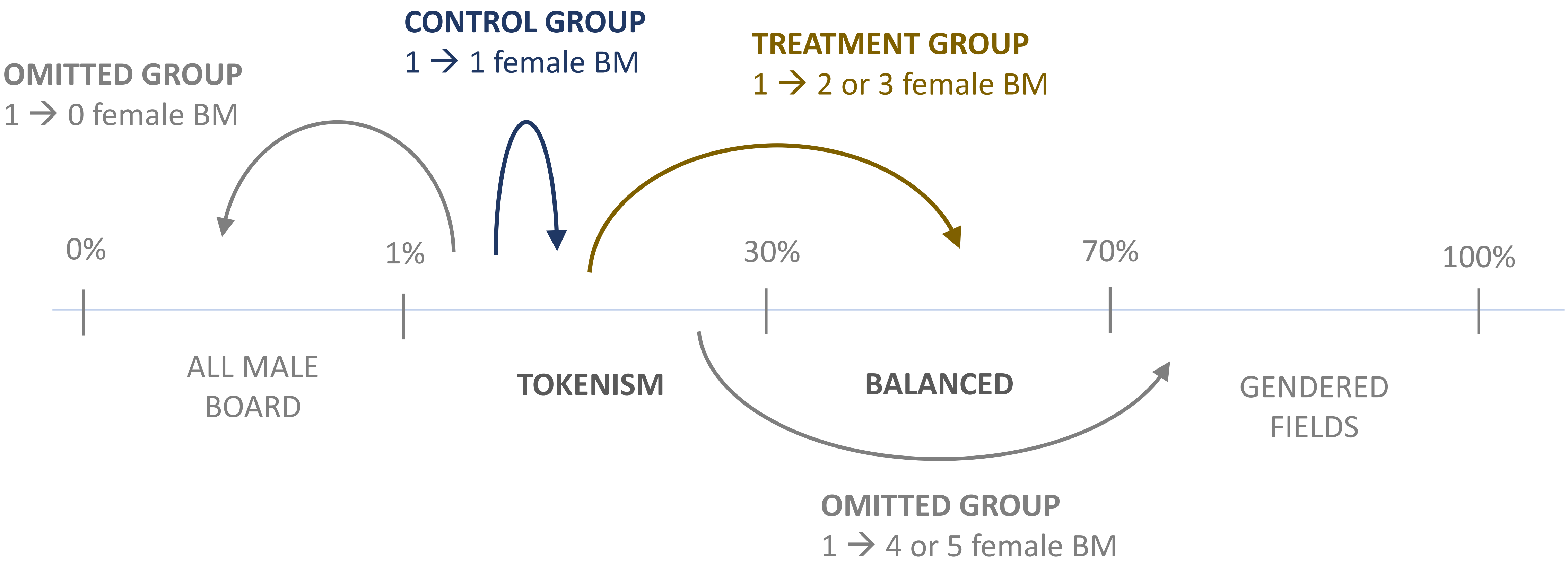


DEFINING TREATMENT GROUPS

Board members turn over regularly. Leverage changes in board membership to isolate the impact of board diversity.

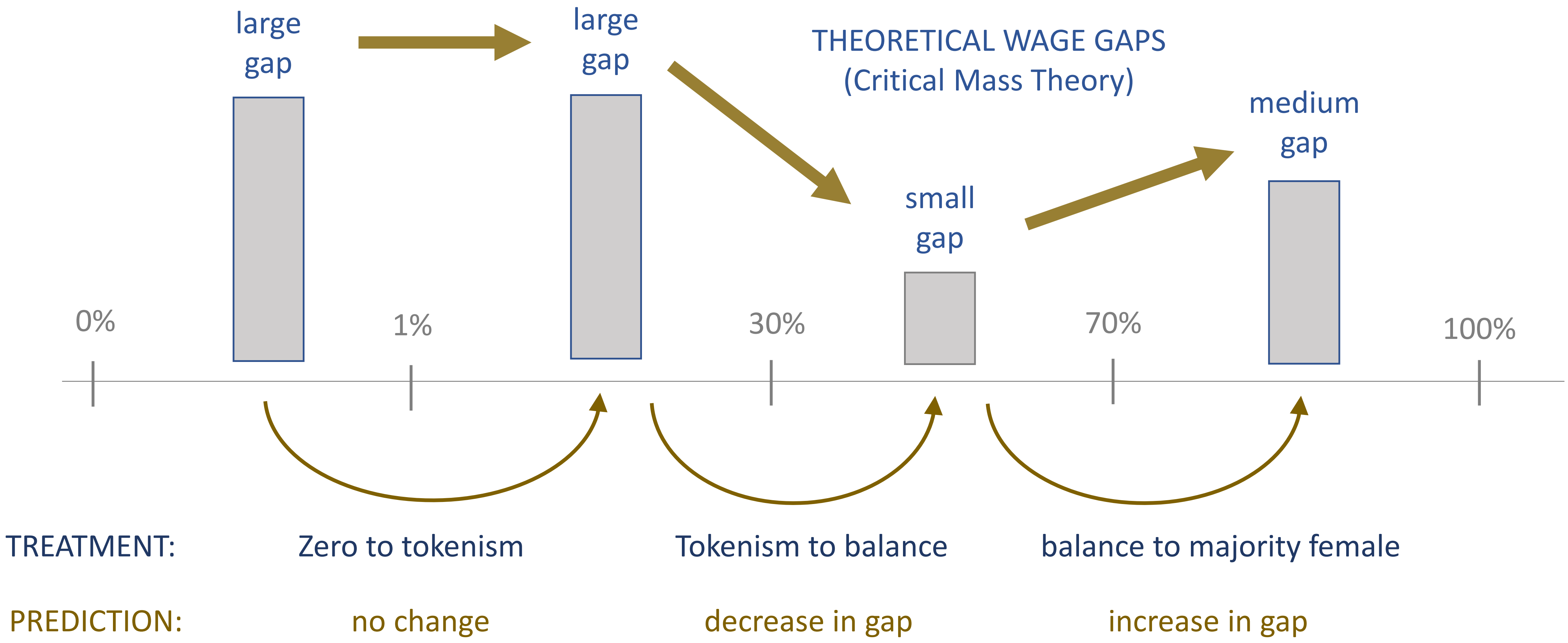
The “treatment” then is defined as a board moving from one level to another.

TREATMENT: TRANSITION FROM TOKEN LEVEL OF DIVERSITY TO BALANCED DIVERSITY



CONSTRUCTING CONTROL GROUPS

To construct a reasonable counterfactual we choose a comparison that has two traits:
(1) Both the treatment and control group start out in the “tokenism” category in the first time period
(2) Ignore cases than lose diversity or make a large leap in diversity gains to ensure the “treatment” category is well-defined

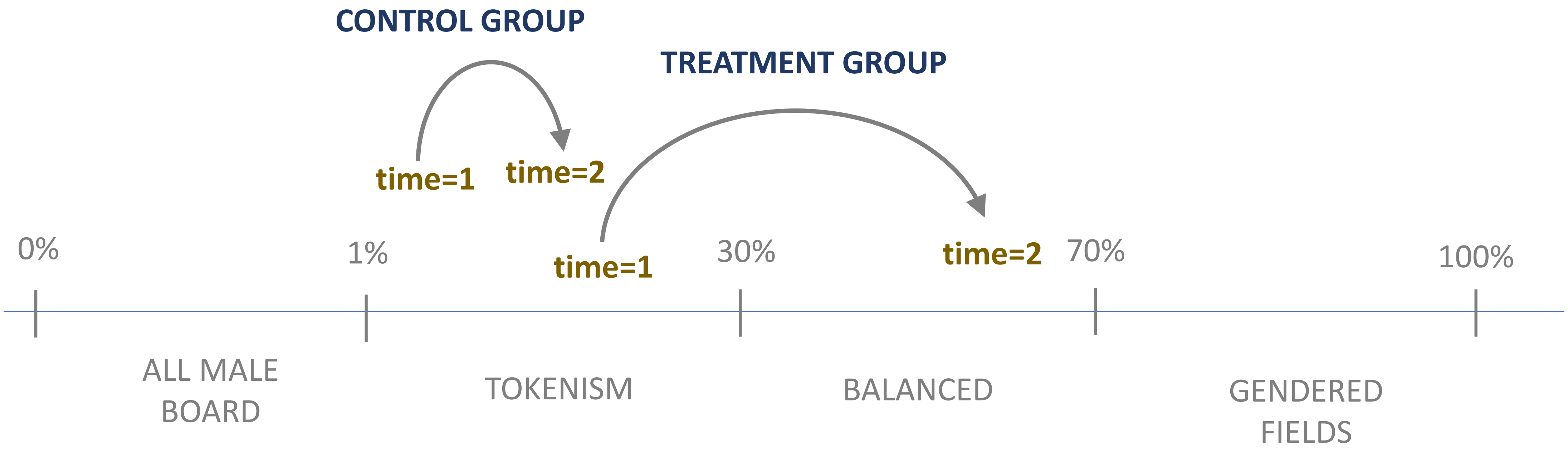


DIAGRAMMING EXPECTATIONS:

We have three distinct treatments in this case, so we need one diagram per "treatment".

The comparison group will start in the base category and experience no change in board diversity between time 1 and time 2.

The "treated' group will experience a large enough increase in diversity to transition categories defined by critical mass theory.



WHAT DO TIME PERIODS REPRESENT IN THE STUDY?

Research Question: Does increasing board diversity reduce the gender pay gap?

Time is a little tricky in this study because the salary of an executive director is set at the time of hire, and once set it would not be increase or decreased dramatically as a result of new board members joining.

This is important because it means board diversity really only matters AT THE TIME OF HIRE. So time periods here do not represent specific years. Rather, they are **time periods in which new executive directors are hired**.

We need to find organizations that have hired two executive directors within our panel so that we can observe board diversity at the time of hire. Time does not always represent calendar units. It can also represent events.

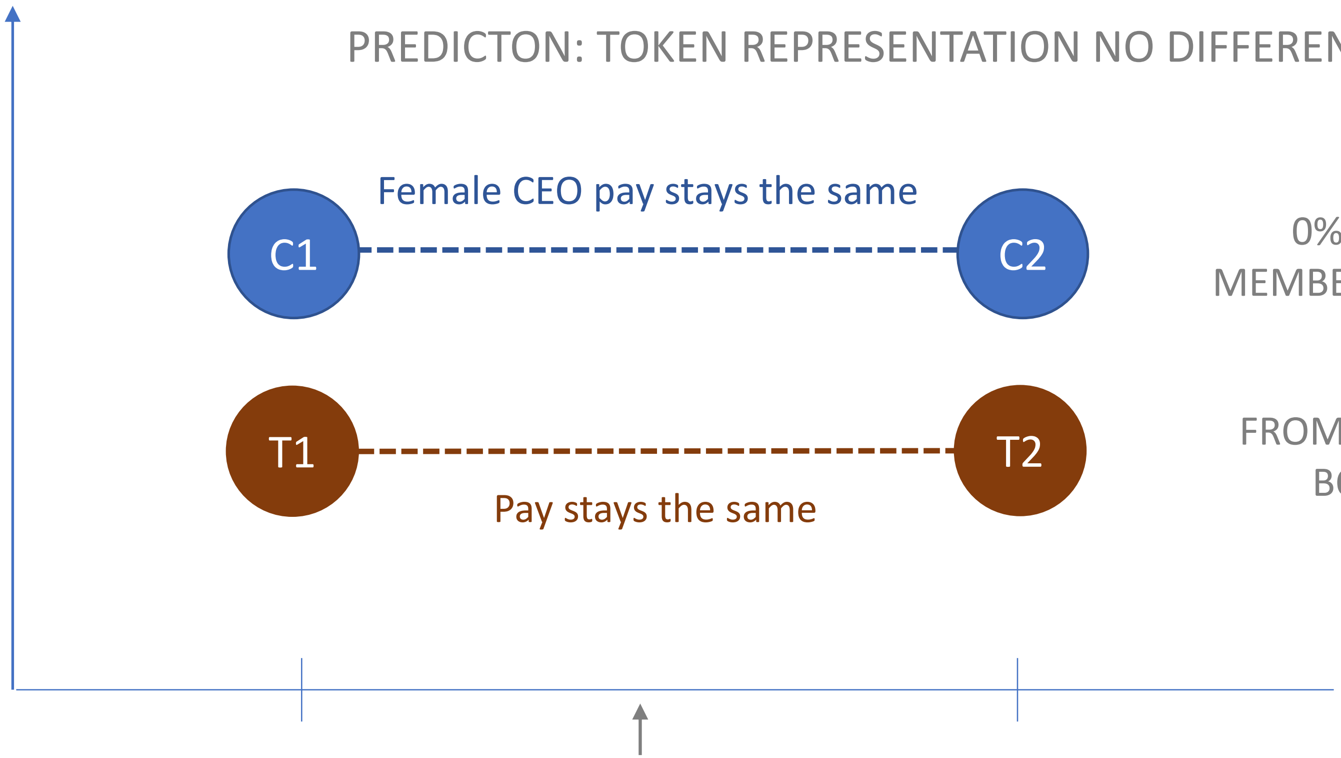
SALARY OF
FEMALE CEOS

TREATMENT TYPE 1: INCREASE DIVERSITY (TOKENISM)

PREDICTON: TOKEN REPRESENTATION NO DIFFERENT THAN ZERO DIVERSITY

Y

Outcomes: Ave salary of newly-hired female CEOs inflation-adjusted for the year of hire



0% FEMALE BOARD MEMBERS IN BOTH PERIODS

FROM 0% TO 20% FEMALE BOARD MEMBERS

PRE PERIOD

TREATMENT:
from “zero” to “token” board diversity

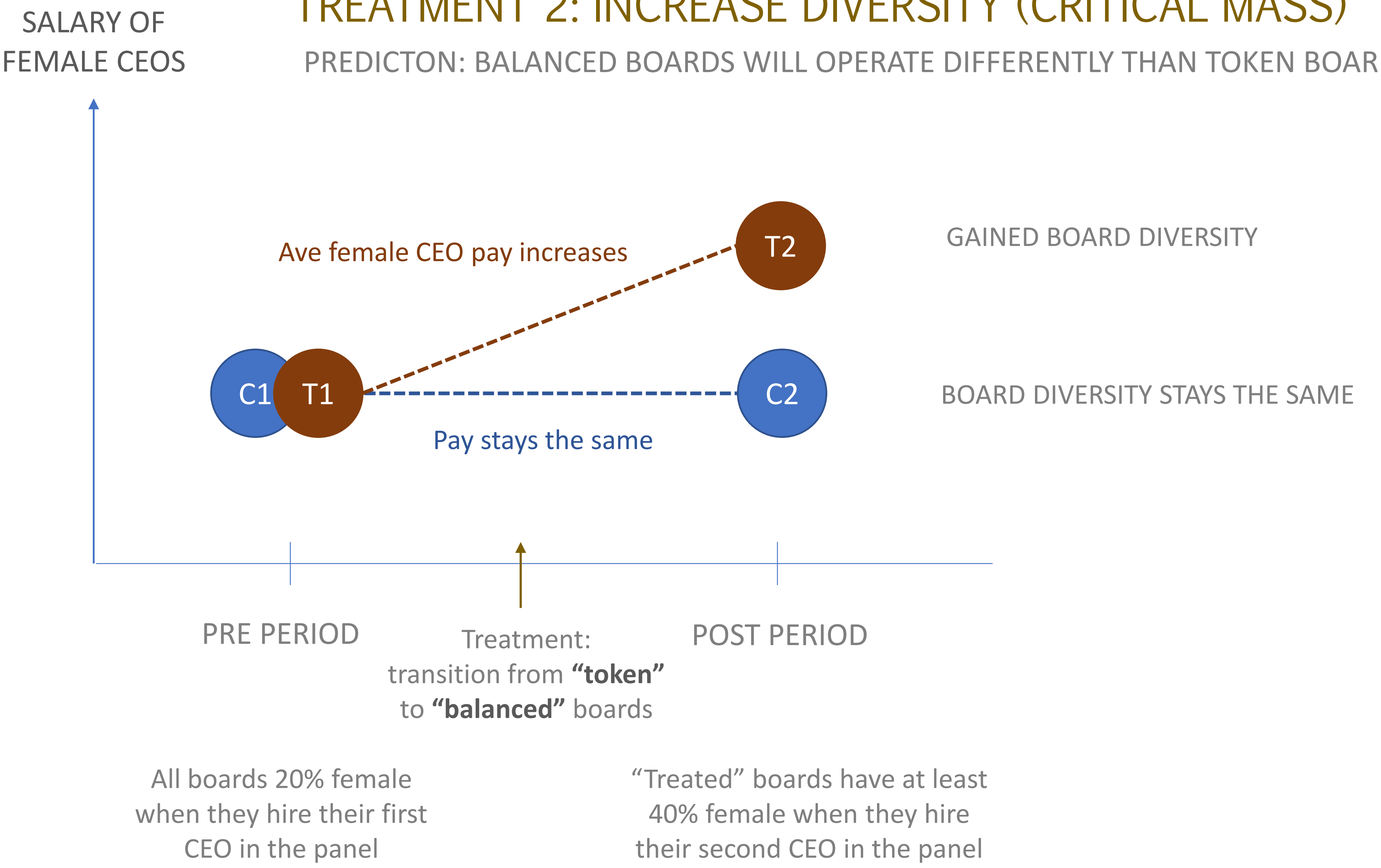
POST PERIOD

All boards have zero female member out of 5 total (0% female) when they hire their first CEO

“Treated” boards have at least 1 female member (20% female) when they hire their second CEO

TREATMENT 2: INCREASE DIVERSITY (CRITICAL MASS)

PREDICTON: BALANCED BOARDS WILL OPERATE DIFFERENTLY THAN TOKEN BOARDS



ASSUMPTIONS:

- C1 and C2 have the same set of nonprofits;
- T1 and T2 have the same set of nonprofits

  Holding nonprofits constant controls for differences in pay across organizational types

- Since they are small boards, all female board members participate in the compensation discussions (some large boards have committees that do not include all board members)

ASSUMPTIONS:

- Is there a **lurking variable** that could explain both increase in board diversity (more women) and increase in female CEO pay?

  Or stated differently, do we believe that changes to board diversity are "stochastic" or somewhat random (every time old board member leaves they recruit a new board member and are open to male or female replacements, so decision is driven by availability and not preference).

  If so, can a marginal increase in board diversity be considered **"exogenous"** making this a reasonable **"identification strategy"** for the effects of board diversity in a quasi-experimental world.

  Quasi-experimental design means looking for "natural experiments" within observational data instead of having full control over deciding who receives the treatment and who does not.

GENERALIZATION (external validity):

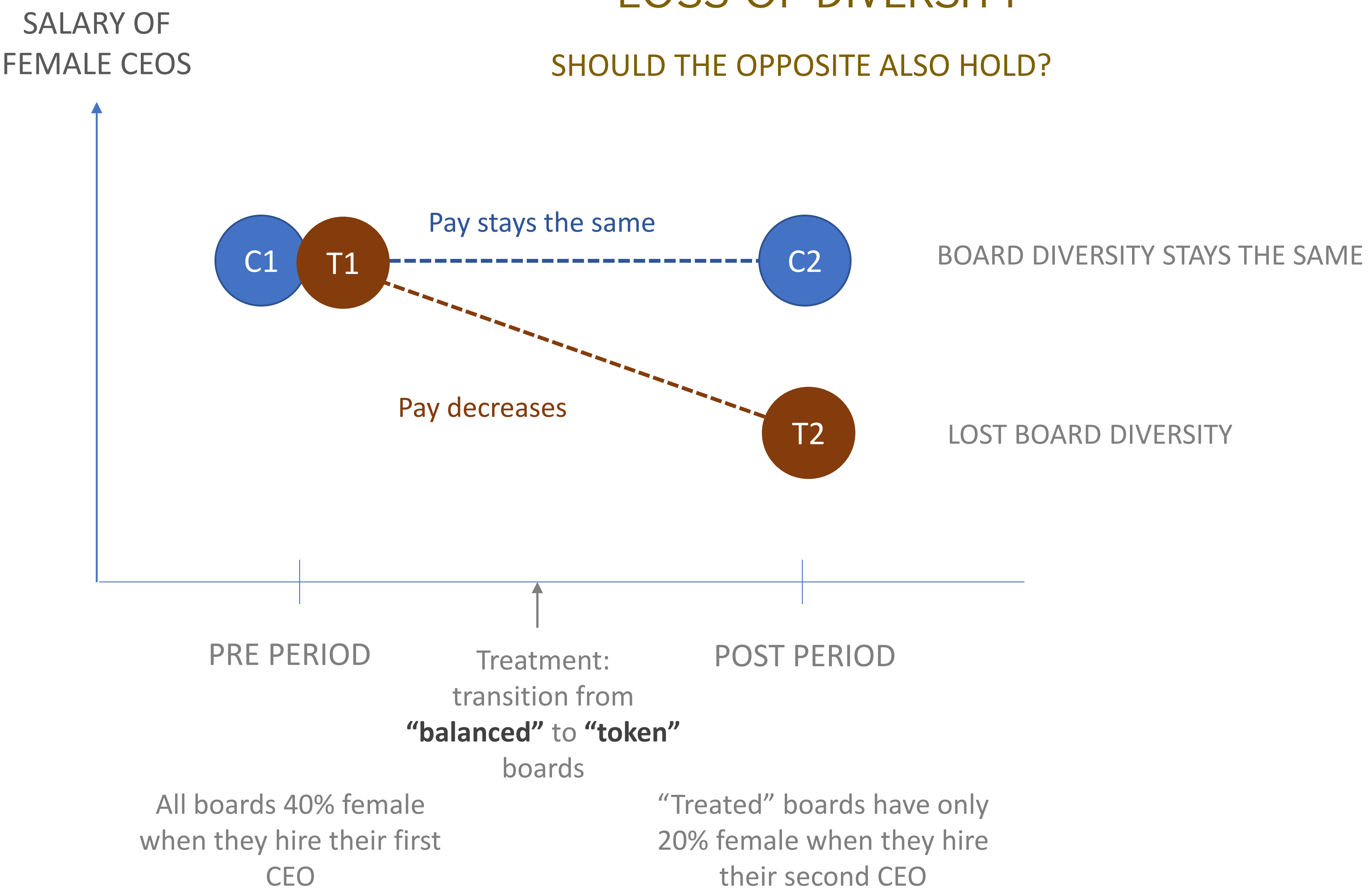
- Are organizations that hire two new female CEOs different from other nonprofits in the population?

  Would the results generalize to the entire nonprofit sector? Or do they only apply to those with female CEOs?

- Can we think of other ways to design the study that capture gender wage gaps but do not limit the sample to those that hire two female CEOs in the study period?

LOSS OF DIVERSITY

SHOULD THE OPPOSITE ALSO HOLD?



MEASURING THE GENDER WAGE GAP

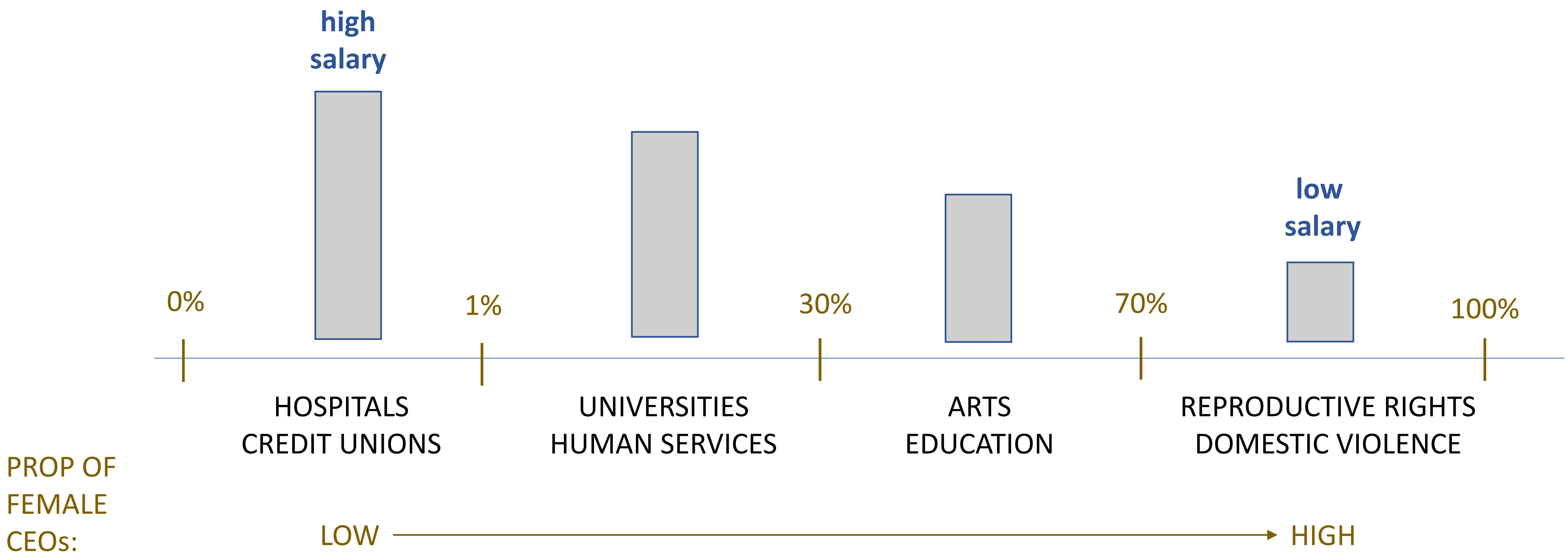
EXERCISE:

- What is the problem with running the model:

  salary = b0 + b1(male_dummy) + e

  To identify the wage gap for nonprofit executives?

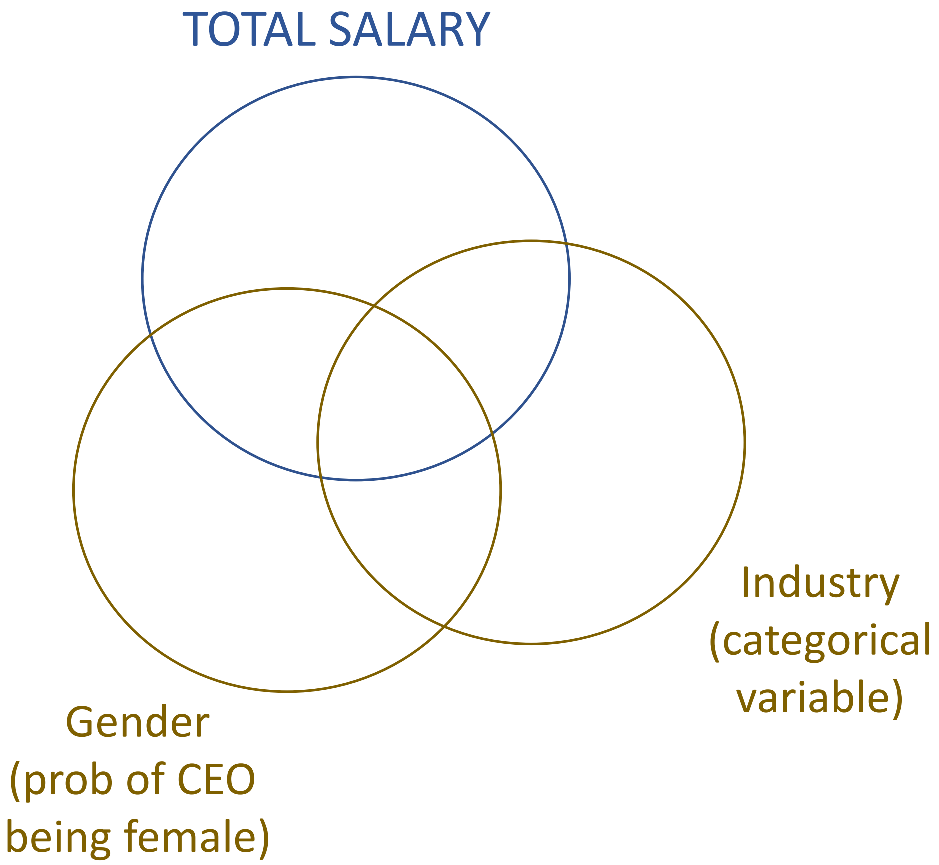
- What would this model tell us? Interpret b1 in this context.



THE SELECTION PROBLEM (revisited)

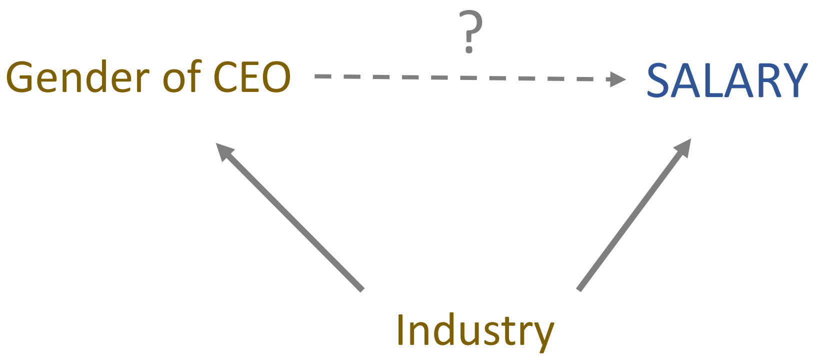
We have an omitted variable problem. Men are more likely to be CEOs in high-paying industries. Women are more likely to be CEOs in industries with modest pay. Thus the regression salary ~ male_dummy will capture **industry preference** more than **wage bias**. Wage bias is defined as inequal pay for identical work. Thus it cannot be calculated by comparing CEO salaries across industries since the qualifications and demands of the job will vary by industry.

As a result, we **CANNOT USE CROSS-SECTIONAL DIFFERENCES** to measure the pay gap.



Salary = b0 + **b1**(male_dummy) + e

Accurate measure of pay difference (purely descriptive), biased estimate of wage discrimination (causal explanation for the gap)

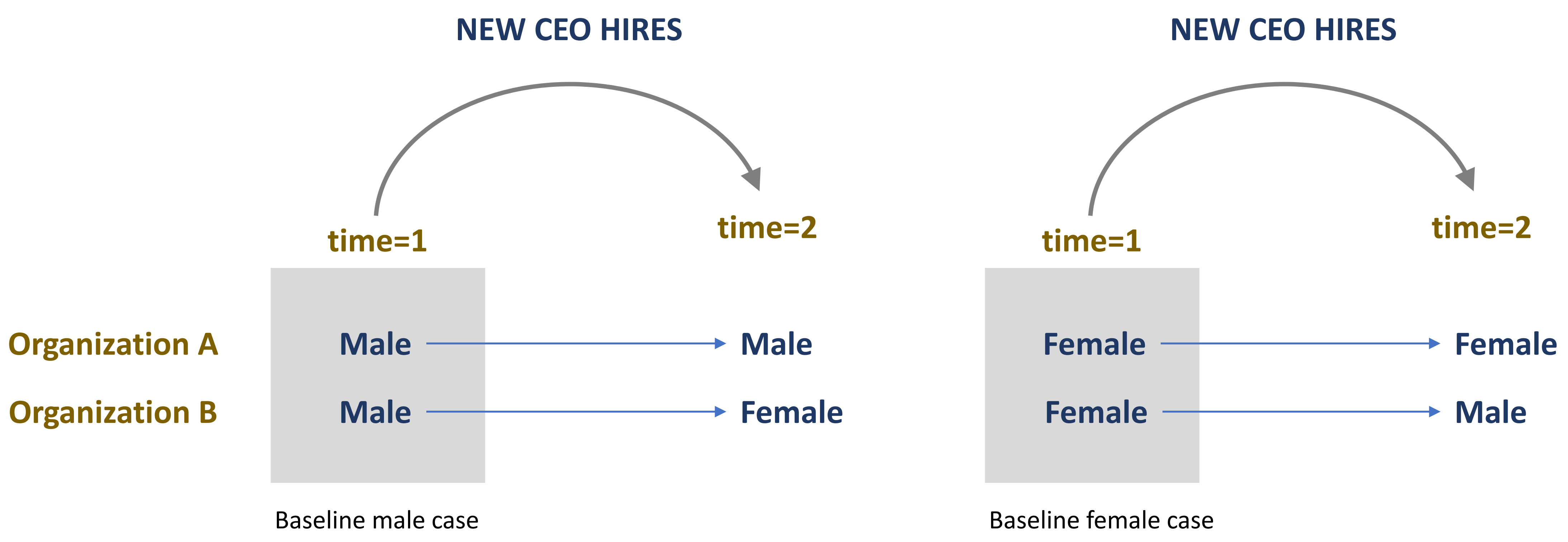


THE SELECTION PROBLEM

To place the problem back into the regression framework, if we try to estimate the pay gap without accounting for the selection problem our estimate will be biased. Selection here is that men and women hold CEO positions in different industries.

Instead of looking at differences across organizations in different industries, conceptualize the pay gap as a single nonprofit hiring a new CEO. There is one male and one female candidate, both with identical qualifications. The nonprofit offers different salaries to each.

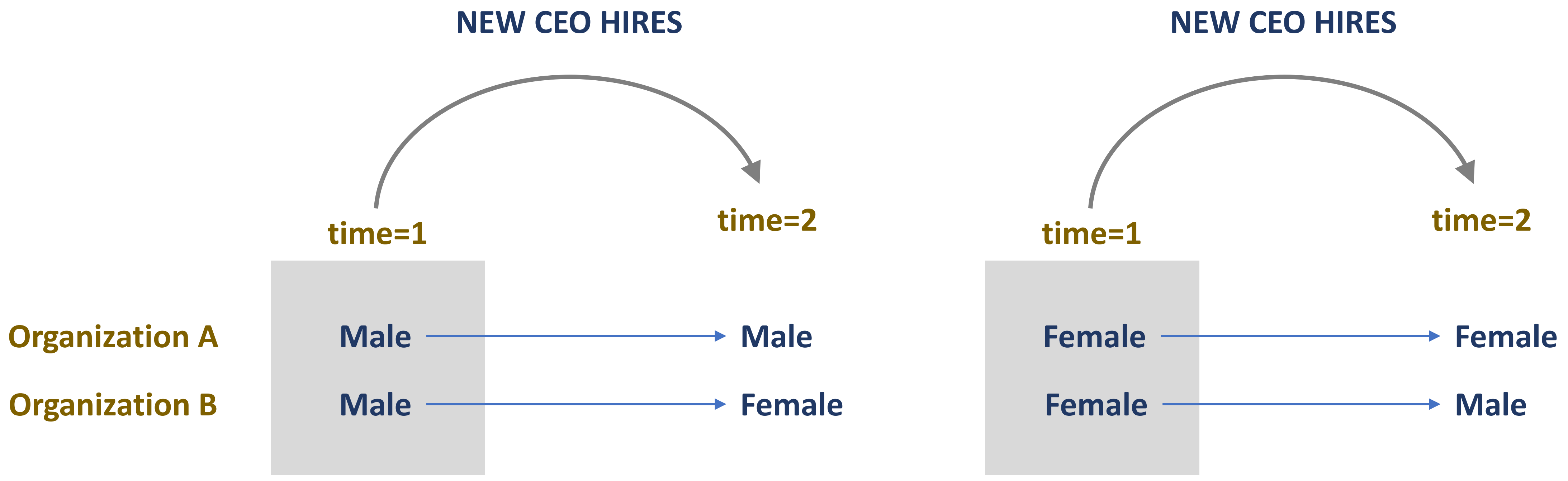
So how can we measure the gender pay gap?



GENDER PAY GAP MEASURE

We know that measuring across industries is problematic. Similarly, even measuring across organizations can be problematic because they will be different sizes, have different levels of financial stability, and different demand in the job. Thus pay differences across organizations (even in the same industry) might reflect the task differences, not gender pay discrimination.

Again, through the power of panel data we can control for environmental factors by looking at how THE SAME ORGANIZATION compensates male and female CEOs differently.



Baseline male case

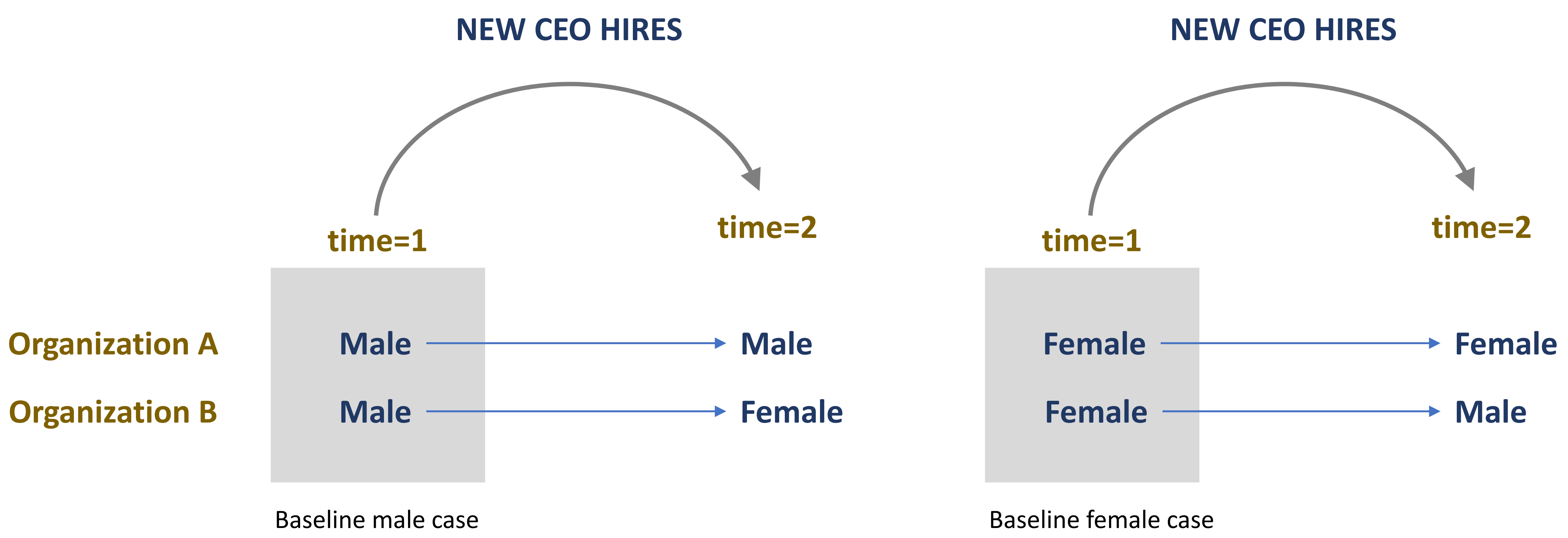
Baseline female case

GENDER PAY GAP MEASURE

If there is **gender pay discrimination** we would expect Organization A and Organization B to behave differently.

The salary of outgoing and income CEOs should be roughly equivalent for Organization A since the gender of the CEO is not changing.

If a gender pay gap exists, we would expect Organization B's CEO salary to fall in the baseline male case (replacing a male CEO with a female CEO), and increase in the baseline female case (replacing a female CEO with a male CEO).



GENDER PAY GAP MEASURE

If **gender pay discrimination** is a problem in the sector:

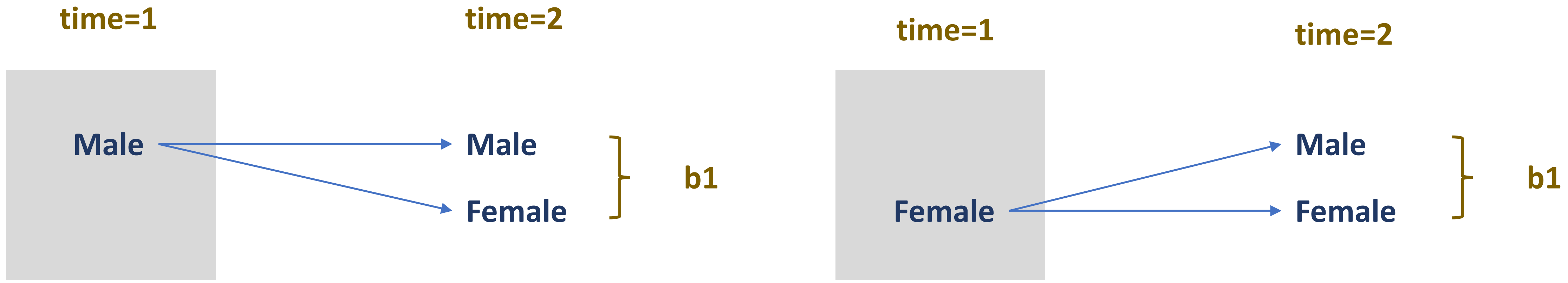
The salary should **stay approximately the same** when the gender of the outgoing and incoming CEO do not change.

Pay should **increase** when a female CEO is replaced by a male CEO.

Pay should **decrease** when a male CEO is replaced by a female CEO.

If we observe these patterns then we have strong evidence of a *discriminatory* gender pay gap (different pay for identical work).

$$(SALARY_{time=2} / SALARY_{time=1}) - 1 = b0 + b1(male\_at\_time=2) + e$$



Baseline male case

Baseline female case

ESTIMATING THE GENDER PAY GAP

The DV in this model is the change in salary with the new hire relative to the old CEO: (T2-T1)/T1 or equivalently T2/T1 – 1

The model now uses a dummy for the replacement CEO only.

Note that we are lowering female pay in the first scenario and raising male pay in the second, but b1 would capture the difference (the pay bump or lack of pay decrease for being male) in either case. But this specification does impose the restriction that the pay gap is the same for both baseline cases. If we want to test that assumption we can use a more nuanced model.

$$(SALARY_{time=2} \,/\, SALARY_{time=1}) - 1 = \mathbf{0.95} + (\mathbf{0.10})(male\_at\_time{=}2) + e$$

$100k

Male → **Male** ($100k), **Female** ($90k) — **10%**

Baseline male case

female: (90/100) = 0.90
male: (100/100) = 1.00

$90k

Female → **Male** ($99k), **Female** ($90k) — **10%**

Baseline female case

female: (90/90) = 1.00
male: (99/90) = 1.10

Ave female salary (proportionate to outgoing CEO) = (0.90+1.00)/2 = 0.95 ⟶ **b0 = 0.95** (reference group)
Ave male salary (proportionate to outgoing CEO) = (1.00+1.10)/2 = 1.05 ⟶ **b0 + b1 → b1 = 0.10** (pay gap)

INTERPRETING COEFFICIENTS

What does b0 represent in this case? b0=0.95 is NOT saying that female salaries are 95% of male salaries. It is a composite measure of two baseline cases that is not that meaningful other than as a reference point to calculate b1.

NOTE, this math assumes that we have an equal proportion of observations in each baseline case!

REGRESSION-BASED APPROACHES

DATA:

- 10-year panel on nonprofit CEO pay
- Data includes the gender of CEOs all board members in the nonprofits
- Limit the data to organizations that hire at least 1 new CEO during the panel, and that have a MALE CEO prior to the hire
- When organizations replace a male CEO with a female CEO, do they pay them less?
- *You could conversely limit the data to all organizations that have female CEOs prior to the new hires, and see how pay changes when they hire new male CEOs.*

ESTIMATE PAY GAP:

$$(salary_{NEW} / salary_{OLD} - 1) = b_0 + b_1(D_{male}) + b_2(D_{male} \times D_{treat}) + b_3(X_1) + b_4(X_2) + time + e$$

Y = increase or decrease in the salary of the new CEO relative to the outgoing CEO (in %)

D_{male}= gender dummy: 1 if the new CEO is male, 0 otherwise

D_{treat}= gender dummy: 1 if nonprofit increased board diversity, 0 otherwise

Where X_i = control variables for organization (size, industry, financial health, etc)

Time = time fixed effects to control for economic conditions at time of the new CEO hire

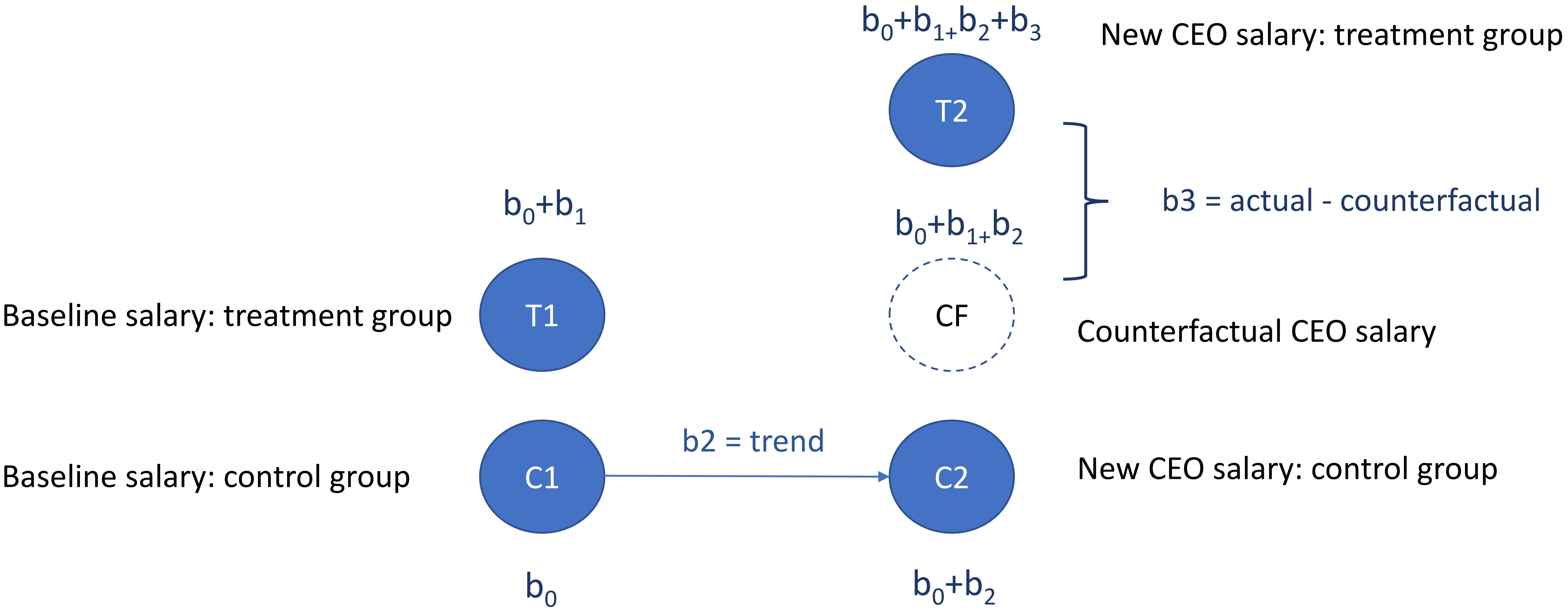
Then b_1 would capture the pay gap – the expected pay premiums the new hire would get for being a male

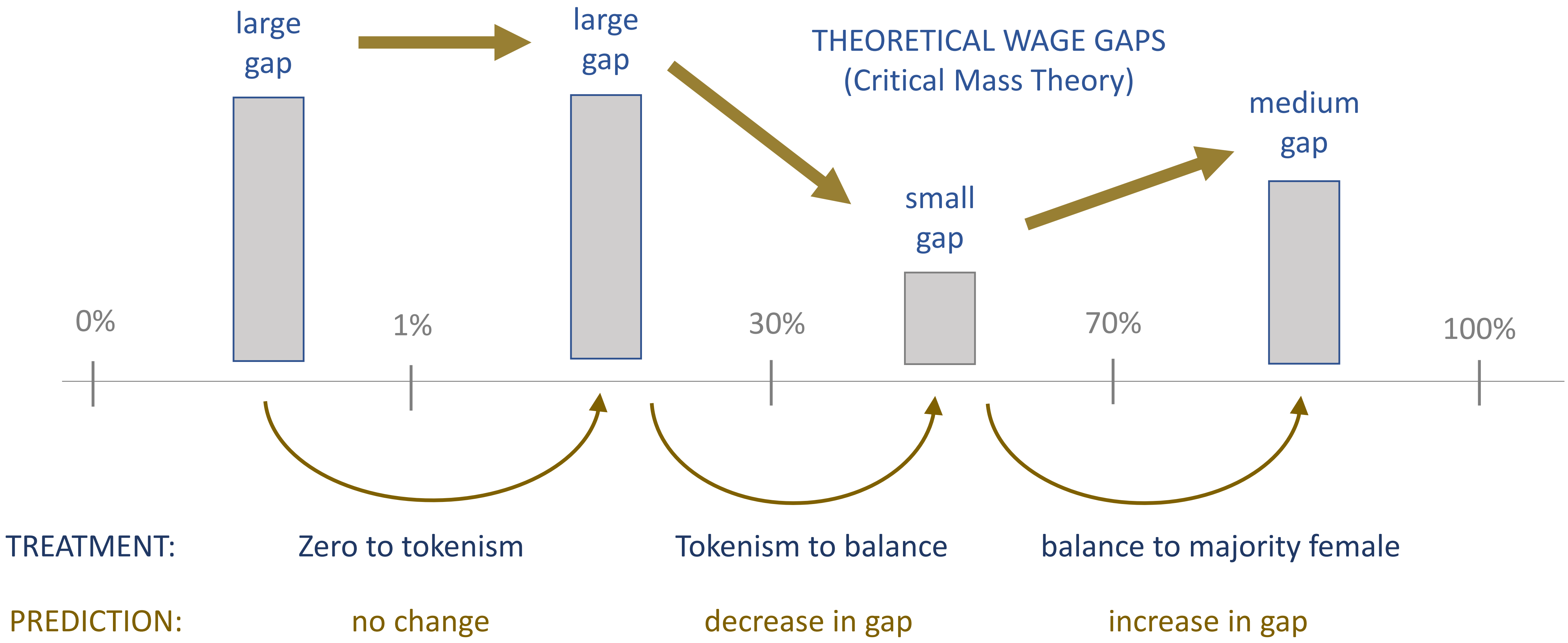
Then b_2 would capture the pay gap reduction achieved by increasing board diversity (the treatment in this study)

Board diversity increase captures changes in board diversity from when the first CEO was hired to when the 2nd CEO was hired in the panel, and whether they transitioned to a new critical mass category.

DIFFERENCE IN DIFFERENCE APPROACH:

$$\text{SALARY} = b_0 + b_1(D_{\text{baseline\_male}}) + b_2(D_{\text{time=2}}) + b_3(D_{\text{baseline\_male}} \times D_{\text{time=2}}) + e$$





DIAGRAMMING EXPECTATIONS:

We have three distinct treatments in this case, so we need one diagram per “treatment”.

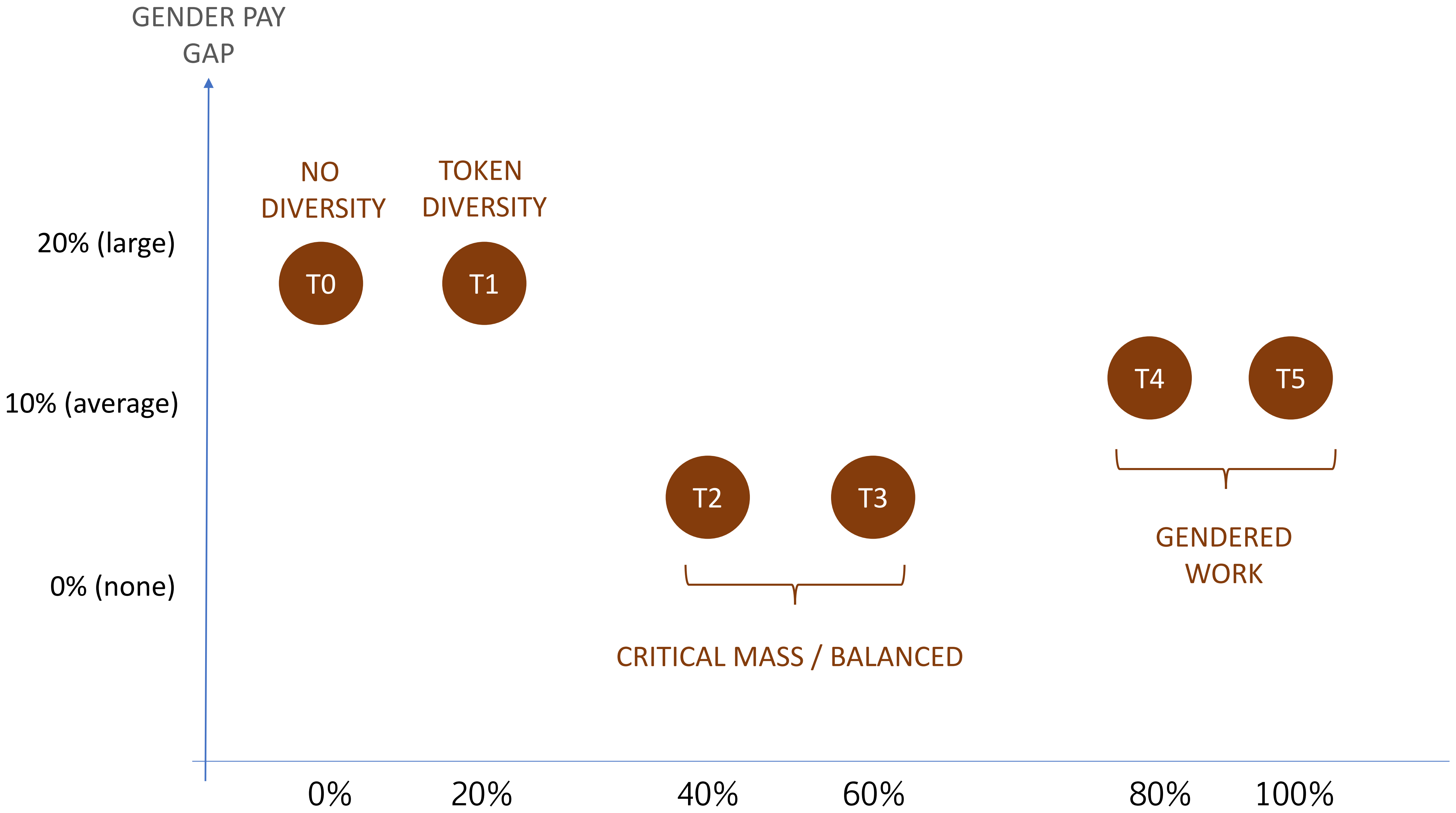
The comparison group will start in the base category and experience no change in board diversity between time 1 and time 2.

The “treated’ group will experience a large enough increase in diversity to transition categories defined by critical mass theory.

POOLING DATA FOR TESTS:

Critical Mass Theory makes several predictions. Explain how you would pool data (e.g. combine T1+T2 to form a new group) to test the following:

- Token diversity will behave no differently than no diversity so we can combine T0 + T1
- Nonprofits with balanced boards will have lower pay gaps: T2 + T3
- Over the critical mass threshold, additional board diversity does not matter: T4 + T5


GENDER PAY GAP
NO DIVERSITY
TOKEN DIVERSITY
20% (large)
T0
T1
10% (average)
T4
T5
T2
T3
GENDERED WORK
0% (none)
CRITICAL MASS / BALANCED
0%
20%
40%
60%
80%
100%