Making Causal Critiques

Day 2 - Fundamental Critiques

Jonathan Phillips

January 27, 2019

What do political scientists **know**?

- Door-to-door political campaigning works
- Proportional Representation electoral systems have more parties
- Democracies do not go to war with each other
- Development helps democracies endure
- ...And that's about it

What do political scientists **know**?

- Thousands of books and papers have *not* generated any knowledge about what explains political processes
  - Many add **descriptive** knowledge
  - Many investigate **specific** events, not generalizable variables
  - Many highlight **correlations** between variables

Learning from Data

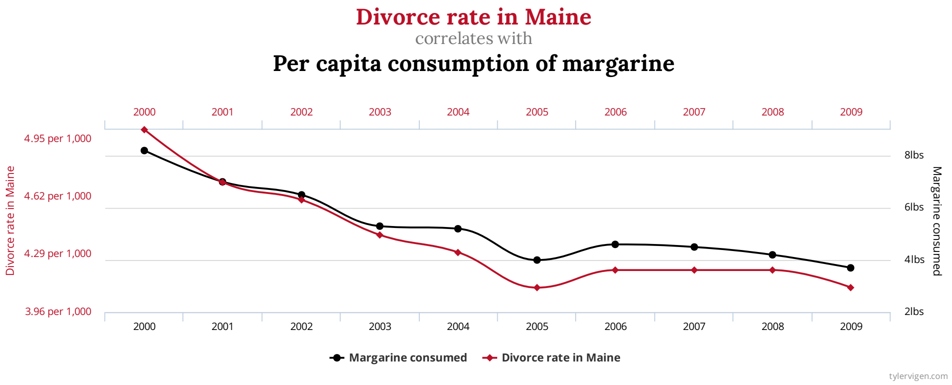
- Why aren't case studies enough?
  - If we want to know why some countries are more successful democracies than others, surely we have to examine the successful countries in detail?
  - Yes! But that's not *sufficient*
- The problem is that there are many variables that *could* explain success
- And detailed case studies can help us identify plausible hypotheses
- But the only way to *confirm* the hypothesis is to verify that:
  1. In other cases, the presence of the condition also produces the same outcome (if not, the explanation is not sufficient)
  2. The absence of the condition does not produce the same outcome (if not, the explanation is not necessary)

Learning from Data

- For example, we could look at India and conclude large Asian countries produce successful democracies
  - But...China
  - But...Costa Rica
- Only by looking at other cases, particularly 'control' cases (small non-Asian countries) can we understand if this explanation is plausible

Learning from Data

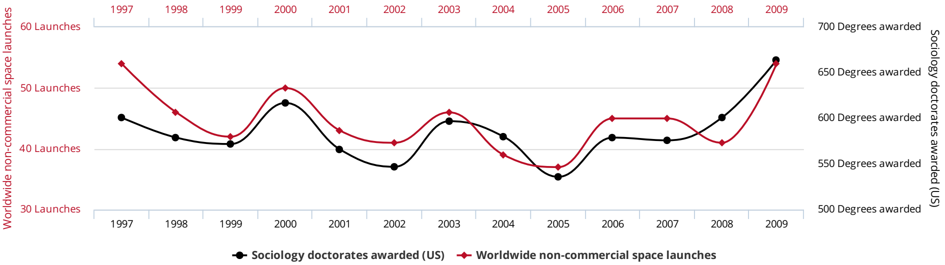
- Even when we compare multiple cases:
- **Correlation is not causation**
  - If we look hard enough we can always find correlations
  - By chance...
  - Due to complex social patterns...
  - But we cannot conclude that there is a causal effect of x on y
- *More* data will not help
- The problem is the *type* of data; it does not allow us to answer causal question


Divorce rate in Maine
correlates with
Per capita consumption of margarine
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
4.95 per 1,000
4.62 per 1,000
4.29 per 1,000
3.96 per 1,000
Divorce rate in Maine
8lbs
6lbs
4lbs
2lbs
Margarine consumed
Margarine consumed
Divorce rate in Maine
tylervigen.com

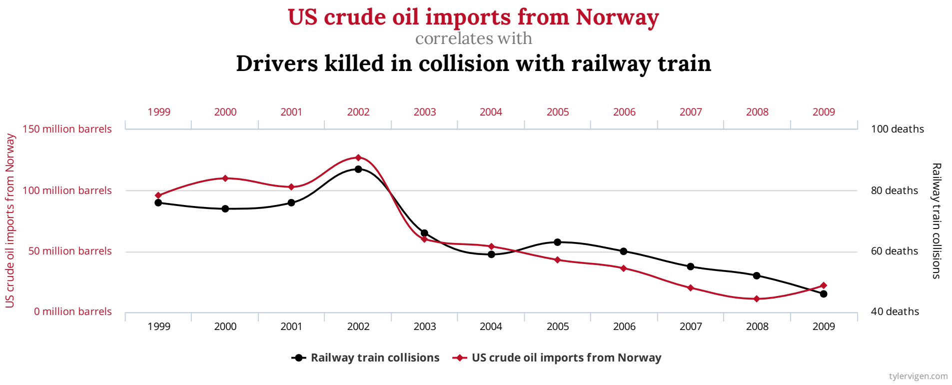
Worldwide non-commercial space launches

correlates with

Sociology doctorates awarded (US)



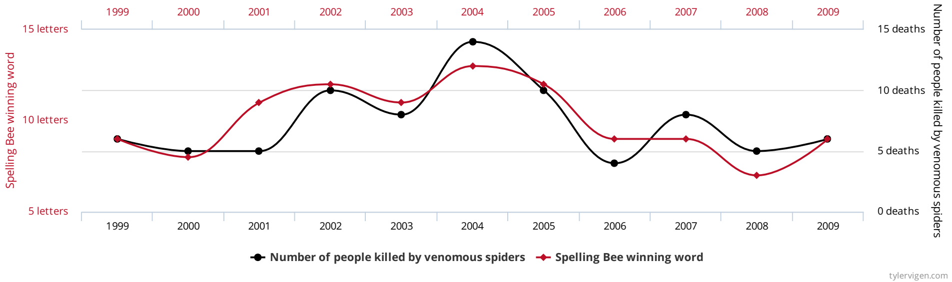
tylervigen.com


US crude oil imports from Norway
correlates with
Drivers killed in collision with railway train
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
150 million barrels
100 million barrels
50 million barrels
0 million barrels
US crude oil imports from Norway
100 deaths
80 deaths
60 deaths
40 deaths
Railway train collisions
Railway train collisions
US crude oil imports from Norway
tylervigen.com

Letters in Winning Word of Scripps National Spelling Bee

correlates with

Number of people killed by venomous spiders



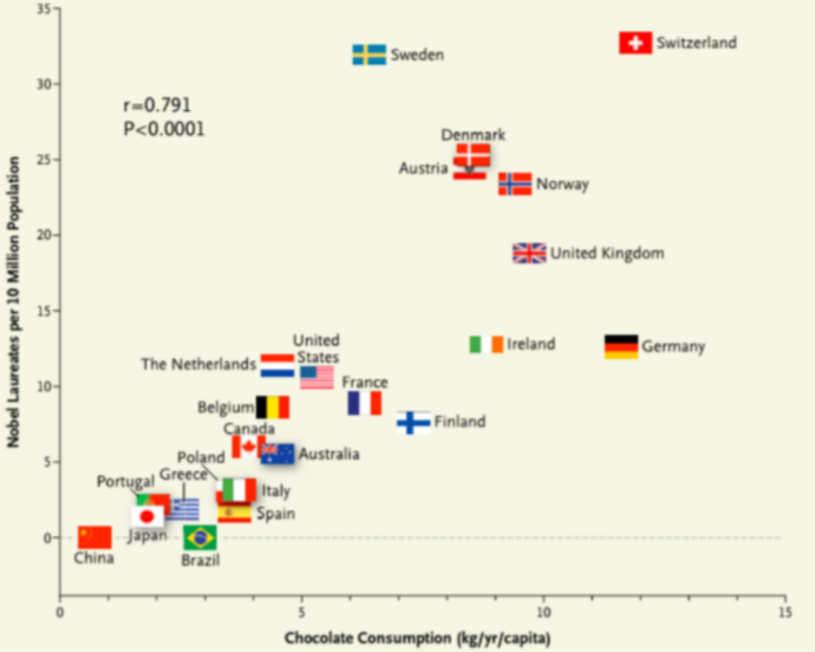


Figure 1. Correlation between Countries' Annual Per Capita Chocolate Consumption and the Number of Nobel Laureates per 10 Million Population.

Learning from Data

- Why isn't correlation enough?
  - For *prediction*, correlation is fine: If we know a country has income of US$50,000 per capita we can confidently predict it is perceived as being less corrupt
  - But for *intervention*, correlation does not help: investing to boost the economy does nothing on its own to reduce corruption
- So if we want to provide policy-relevant advice, we need to know more than just correlation

Learning from Data

- Why isn't correlation enough?
  - For *explanation*, correlation also fails - it is no *explanation* to say that you are a student because you are in your 20's
  - Explanation means identifying the direct causal effects

Learning from Data

- Why isn't correlation enough?
  - People are strategic, so their behaviour changes
- **The Lucas Critique**: Relationships fall apart when we intervene with policy
  - The data shows no-one lies on their tax forms
  - So let's abandon tax checks; the government wants to save money
  - But reducing checks reduces the chances of getting caught
  - Citizens start to lie on their tax forms
- That means we need to understand what *causes* people to lie on tax forms, so we can better understand their behaviour

Learning from Data

- To accumulate knowledge, we have to ask specific types of questions:
  - Specifically, about the **effects of causes**

Causes of Effects	**Effects of Causes**
What caused Y?	Does X cause Y?
Why did the United States grow faster than Bolivia in the twentieth century?	Did the more permanent colonial settlement of the United States compared to Bolivia affect their subsequent growth rates?

Causal Inference

- A focus on a single explanatory variable X requires us to clearly define this 'treatment'
- AND to clearly define a control
  - What is the opposite of investing \$1bn in education?
  - No investment, or investing it elsewhere?
- Define treatment:

$$D_i = \begin{cases} 1, \text{ if treated} \\ 0, \text{ if not treated} \end{cases}$$

Causal Inference

- Defining our outcome is also crucial:
  - Can we measure our outcome of interest?
  - Is that outcome the end of the causal chain?
  - Tempting to look at many outcomes, but the risk of cherry-picking
    - All outcomes are probabilistic
    - If we study 20 outcomes, on average one will show a significant effect even with no real causal effect

Causal Inference

- We want to know how some variable affects another variable
- eg. how a proportional representation electoral system affects investment in education
    - The **treatment** is a change to a PR electoral system (vs FPTP)
    - The **outcome** is the level of (public?) investment in education

Causal Inference

- So we need a precise framework to analyze causation
- The causal effect of treatment is how the unit's outcome differs when it is treated and not treated
- These are the **potential outcomes** for unit *i*:

$$Y_{Di} = \begin{cases} Y_{1i} \text{ Potential Outcome if unit i treated} \\ Y_{0i} \text{ Potential Outcome if unit i not treated} \end{cases}$$

- Treatment Effect $= Y_{1i} - Y_{0i}$

Causal Inference

- We are relying on **counterfactuals**
  - What would have happened to the same unit if the treatment had not happened?
  - Would World War I still have happened if Archduke Franz Ferdinand had not been assassinated in 1914?
  - Would people have voted for Brexit if the campaign had been better regulated?
  - Would Brazil have won the 2014 World Cup if Neymar had not been injured?
- To explain a class of events - not a single event - we need multiple counterfactual comparisons

Causal Inference

Potential Outcomes Example

	Investment in Education if PR system	Investment in Education if FPTP system	
	Y_1	Y_0	Treatment Effect
Brasil	8	4	4
Argentina	10	7	3
Bolivia	2	4	-2
Colombia	11	11	0
Peru	6	2	4

Causal Inference

- **The Fundamental Problem of Causal Inference**
  - No units can receive **both** treatment and control
  - So we can never observe both Y_1 and Y_0 for the same unit

Causal Inference

Potential Outcomes Example

	PR System?	Investment in Education if PR system	Investment in Education if FPTP system	
	D_i	Y_1	Y_0	Treatment Effect
Brasil	1	8	?	?
Argentina	1	10	?	?
Bolivia	0	?	4	?
Colombia	0	?	11	?
Peru	0	?	2	?

Causal Inference

- We can't even look at the change in countries that switch to a PR system
  - What if **all** countries had started to invest more in education at the same time, for different reasons?
  - The potential outcome for Country X in time 1 is different to at time 2

Causal Inference

- So we need to consider the **counterfactual** - what would have happened if the country had **not** switched to a PR system?
  - This is **impossible** to know
  - We can only estimate the effect by comparing **across** units in some way
  - That is why we are doing causal **inference**, not causal proof

Causal Inference

- Which comparisons to make?
- Control units can never be perfect substitutes
- Causal Inference is all about identifying a **plausible counterfactual**
    - Plausible means that the potential outcomes of the control unit are likely to be the same as those of the treated unit

Causal Inference

- The comparability of treatment and control units depends on *how* they got to be treated
  - On the **Treatment Assignment Mechanism**
- If we 'treated' an outlier like the Galapagos Islands, could we find a comparable control unit?
- Comparisons are easier where the **Treatment Assignment Mechanism is independent of potential outcomes**
  - This makes it more likely that potential outcomes are 'balanced' and comparable

Causal Inference

- Types of Research Design:

Add caption

	Researcher controls the treatment assignment	Treatme mechan ate com outcome
Controlled Experiments	Yes	Yes
Natural Experiments	No	Yes
Observable Studies	No	No

Problems with Observational Data

- Observational Studies
  - Household surveys
  - Simple regression on secondary data
  - Interviews of a random sample

Problems with Observational Data

- We do not know what the treatment assignment mechanism was
  - Which units were treated and why?
- Treatment assignment is unlikely to create comparable potential outcomes
  - Which units might be appropriate counterfactuals?

Problems with Observational Data

- With complete information on potential outcomes, calculating treatment effects is trivial

Calculating Treatment Effects

	D	Y_1	Y_0	Y_i	Real Effect, $Y_1 - Y_0$
A	1	7	4	7	3
B	0	9	5	5	4
C	0	4	4	4	0
D	1	4	3	4	1

Problems with Observational Data

- With complete information on potential outcomes, calculating treatment effects is trivial

Calculating Treatment Effects

	D	Y_1	Y_0	Y_i	Real Effect, $Y_1 - Y_0$
A	1	7	4	7	3
B	0	9	5	5	4
C	0	4	4	4	0
D	1	4	3	4	1
$E(Y_1) =$		6			
$E(Y_0) =$			4		

- $ATE = E(Y_1 - Y_0) = 8/4 = 2$
- $ATE = E(Y_1) - E(Y_0) = 6 - 4 = 2$

Problems with Observational Data

- From observed outcomes can we calculate an Average Treatment Effect?

Calculating Treatment Effects

	D	Y_1	Y_0	Y_i	Real Effect, $Y_1 - Y_0$
A	1	7	?	7	?
B	0	?	5	5	?
C	0	?	4	4	?
D	1	4	?	4	?

Problems with Observational Data

- From observed outcomes can we calculate an Average Treatment Effect?

Calculating Treatment Effects

	D	Y_1	Y_0	Y_i	Real Effect, $Y_1 - Y_0$
A	1	7	?	7	?
B	0	?	5	5	?
C	0	?	4	4	?
D	1	4	?	4	?
$E(Y_1) =$		5.5			
$E(Y_0) =$			4.5		

Problems with Observational Data

- If we use the control units as counterfactuals...
- Average Treatment Effect:

$$
\begin{aligned}
ATE &= E(Y_1) - E(Y_0) && (1)\\
&= 5.5 - 4.5 && (2)\\
&= 1 && (3)
\end{aligned}
$$

- Half the true treatment effect

Problems with Observational Data

- If we use the control units as counterfactuals...
- Average Treatment Effect:

$$
\begin{aligned}
ATE &= E(Y_1) - E(Y_0) && (1)\\
&= 5.5 - 4.5 && (2)\\
&= 1 && (3)
\end{aligned}
$$

- Half the true treatment effect
- Why?

Problems with Observational Data

- If we use the control units as counterfactuals...
- Average Treatment Effect:

$$
\begin{aligned}
ATE = & \quad E(Y_1) - E(Y_0) \qquad (1)\\
= & \quad 5.5 - 4.5 \qquad (2)\\
= & \quad 1 \qquad (3)
\end{aligned}
$$

- Half the true treatment effect
- Why?
  - The units that got treated had lower Y_1
  - The units that were controls had higher Y_0

Problems with Observational Data

- If we use the control units as counterfactuals...
- Average Treatment Effect:

$$
\begin{aligned}
ATE = &\quad E(Y_1) - E(Y_0) && (1)\\
= &\quad 5.5 - 4.5 && (2)\\
= &\quad 1 && (3)
\end{aligned}
$$

- Half the true treatment effect
- Why?
  - The units that got treated had lower Y_1
  - The units that were controls had higher Y_0
  - The ’stand-in’ counterfactuals were wrong

Problems with Observational Data

- The bias in units' potential outcomes depends on which units get treated and which ones don't
- In observational studies, we have very little protection against causal critiques
    1. Reverse Causation
    2. Omitted variable bias (confounding)
    3. Selection bias

Omitted Variable Bias

- Wealthier countries are more likely to be democracies
    - But wealthier countries are more likely to be European
    - And democracies are more likely to be European
- Maybe the correlation just reflects the fact that European countries are ’different’?

Omitted Variable Bias

```
## Error in create_graph() %>%
add_global_graph_attrs("graph", "rankdir", :  could
not find function "%>%"
## Error in render_graph(graph):  could not find
function "render_graph"
```

Omitted Variable Bias

```
## Error in create_graph() %>%
add_global_graph_attrs("graph", "rankdir", :  could
not find function "%>%"
## Error in render_graph(graph):  could not find
function "render_graph"
```

Omitted Variable Bias

- Imagine a treatment assignment mechanism where all women get treated

Treatment Assignment by Covariate

	X	D	Y_1	Y_0	Y_i	Real Effect
A	Man	0	7	4	4	3
B	Man	0	9	5	5	4
C	Woman	1	4	4	4	0
D	Woman	1	4	3	4	1

Omitted Variable Bias

- Imagine a treatment assignment mechanism where all women get treated

Treatment Assignment by Covariate

	X	D	Y_1	Y_0	Y_i	Real Effect
A	Man	0	7	4	4	3
B	Man	0	9	5	5	4
C	Woman	1	4	4	4	0
D	Woman	1	4	3	4	1
$E(Y_1) =$			4			
$E(Y_0) =$				4.5		

Omitted Variable Bias

- Imagine a treatment assignment mechanism where all women get treated

Treatment Assignment by Covariate

	X	D	Y_1	Y_0	Y_i	Real Effect
A	Man	0	7	4	4	3
B	Man	0	9	5	5	4
C	Woman	1	4	4	4	0
D	Woman	1	4	3	4	1
$E(Y_1) =$			4			
$E(Y_0) =$				4.5		

- ATE = 4 - 4.5 = -0.5
- This is **confounding** or an **omitted variable** - another variable affects both treatment and potential outcomes

Self-Selecion Bias

- Wealthier countries are more likely to be democracies
  - But wealthy autocracies and poor democracies do not like to report data
  - So we cannot compare them
  - Only wealthy democracies ’self-select’ into our sample

Self-Selection Bias

```
## Error in create_graph() %>%
add_global_graph_attrs("graph", "rankdir", :  could
not find function "%>%"
## Error in render_graph(graph):  could not find
function "render_graph"
```

Self-Selection Bias

```
## Error in create_graph() %>%
add_global_graph_attrs("graph", "rankdir", :  could
not find function "%>%"
## Error in render_graph(graph):  could not find
function "render_graph"
```

Self-Selection Bias

- Imagine a treatment assignment mechanism where people get to *choose* their treatment

Treatment Assignment by Self-Selection

	D	Y_1	Y_0	Y_i	Real Effect
A	1	7	4	7	3
B	1	9	5	9	4
C	0	4	4	4	0
D	0	4	3	3	1

Self-Selection Bias

- Imagine a treatment assignment mechanism where people get to *choose* their treatment

Treatment Assignment by Self-Selection

	D	Y_1	Y_0	Y_i	Real Effect
A	1	7	4	7	3
B	1	9	5	9	4
C	0	4	4	4	0
D	0	4	3	3	1
$E(Y_1) =$		8			
$E(Y_0) =$			3.5		

Self-Selection Bias

- Imagine a treatment assignment mechanism where people get to *choose* their treatment

Treatment Assignment by Self-Selection

	D	Y_1	Y_0	Y_i	Real Effect
A	1	7	4	7	3
B	1	9	5	9	4
C	0	4	4	4	0
D	0	4	3	3	1
$E(Y_1) =$		8			
$E(Y_0) =$			3.5		

- ATE = 8 - 3.5 = 4.5
- This is **self-selection bias** - treatment is affected by potential outcomes

Problems with Observational Data

- We can identify the source of these biases in potential outcomes:

Problems with Observational Data

- We can identify the source of these biases in potential outcomes:

$$\underbrace{E(Y_i|D=1) - E(Y_i|D=0)}_{\text{Observed Effect}} \quad (4)$$

Problems with Observational Data

- We can identify the source of these biases in potential outcomes:

$$\underbrace{E(Y_i|D=1) - E(Y_i|D=0)}_{\text{Observed Effect}} = \underbrace{E(Y_{1i} - Y_{0i})}_{\text{Real ATE}}$$
$$\underbrace{+\frac{1}{2}\Big[E(Y_{1i}|D=1) - E(Y_{1i}|D=0)\Big]}_{\text{Imbalance on } Y_1} + \underbrace{\frac{1}{2}\Big[E(Y_{0i}|D=1) - E(Y_{0i}|D=0)\Big]}_{\text{Imbalance on } Y_0} \quad (5)$$

NB: For equal-sized treatment and control groups

Problems with Observational Data

- Disaggregating the Self-Selection Bias:

$$\frac{(7+9-4-3)}{2} = \frac{(7+9+4+4-4-5-4-3)}{4} + \frac{1}{2}\Big[\frac{(7+9)}{2} - \frac{(4+4)}{2}\Big] + \frac{1}{2}\Big[\frac{(4+5)}{2} - \frac{(4+3)}{2}\Big]$$

$$4.5 = {\color{red}2} + 2 + \frac{1}{2} \qquad (6)$$

Problems with Observational Data

- Depending on the treatment assignment mechanism we get a range of Average Treatment Effects:

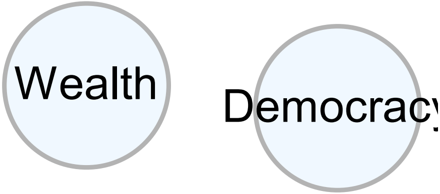
Comparing Average Treatment Effects

Treated Units	**ATE**
Real Effect for all units	2
A & D	1
Omitted Variable Bias (Women)	-0.5
Self-selection	4.5

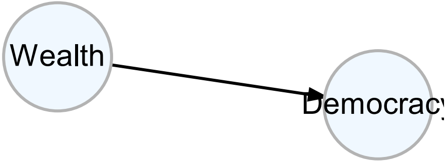
Reverse Causation

- Wealthier countries are more likely to be democracies
  - But does wealth create democracy?
  - Or democracy create wealth?
- We cannot tell from the correlation alone
- Both may be true

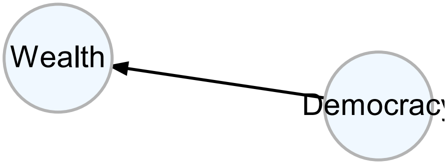
Reverse Causation



Reverse Causation



Reverse Causation



Reverse Causation

- Where treatment has *no* effect

Treatment Assignment by Covariate

	D	Y_1	Y_0	Y_i	Real Effect
A	0	7	7	7	0
B	0	9	9	9	0
C	1	4	4	4	0
D	1	4	4	4	0

Reverse Causation

- Where treatment has *no* effect

Treatment Assignment by Covariate

	D	Y_1	Y_0	Y_i	Real Effect
A	0	7	7	7	0
B	0	9	9	9	0
C	1	4	4	4	0
D	1	4	4	4	0
$E(Y_1) =$		4			
$E(Y_0) =$			4		

Reverse Causation

- Where treatment has *no* effect

Treatment Assignment by Covariate

	D	Y_1	Y_0	Y_i	Real Effect
A	0	7	7	7	0
B	0	9	9	9	0
C	1	4	4	4	0
D	1	4	4	4	0
$E(Y_1) =$		4			
$E(Y_0) =$			4		

- ATE = 4 - 4 = 0. There is no effect.
- The (negative) correlation between D and Y is because Y **causes** D

Exercise

- Does fruit make you happier?
  - Write down on a piece of paper a number between 0 and 10 representing how happy you would be if I gave you an apple now.
  - Label this number Y_1.
  - Then write down a second number between 0 and 10 representing how happy you would be if I did NOT give you an apple now.
  - Label this number Y_0.
- These are your **potential outcomes**.

Exercise

- Now we will consider how estimates of the average effect of fruit on happiness vary depending on how treatment (apples) are assigned.
  1. All the female participants are given an apple.
  2. The tallest half are given an apple.
  3. You are free to choose yourself to take an apple or not.
  4. Apples are distributed randomly