A Framework for Causal Analysis

Tamer Çetin

Causal Analysis

- define subjects, interventions, and potential outcomes for a causal question;
- define individual treatment effects and the average treatment effect in the framework of potential outcomes;
- understand effects as ceteris paribus comparisons;
- assess potential sources of variation in a causal variable and whether they are exogenous or endogenous;
- understand why and how random assignment helps identify the average treatment effect;
- understand what kinds of variables we need to condition on and what kinds of variables we should not condition on, when our aim is to uncover an average effect using observational data;
- create and understand causal maps that visualize the analyst's assumptions about the causal relationships between variables;
- assess the internal and external validity of the results of a causal analysis.

Intervention, Treatment, Subjects, and Outcomes

- What is the intervention/treatment?
- Who/what are the subjects?
  - Treated/untreaded subjects/units
- What's the outcome (y) variable?
- What's the causal (x) variable that represents the intervention?
  - Is it binary or quantitative?
- What is/are the potential mechanisms?

Potential Outcomes

- Interventions:
  - binary causal variable
  - non-binary causal variables
- The outcome may be anything, including binary or quantitative variables.
- Potential outcomes framework for each subject:
  - potential treated outcome: what would their outcomes be if they were treated?
  - potential untreated outcome: what would their outcomes be if they were untreated?

Potential Outcomes: Setup

- y_i^1 = potential treated outcome
- y_i^0 = potential untreated outcome

- y_i = observable outcome
- $y_i = y_i^1$ = for subjects that end up being treated
- $y_i = y_i^0$ = for subjects that end up being not treated

Potential Outcomes: Setup

$y_i = y_i^1$
$y_i = y_i^0$

- **Counterfactual outcome** is the other potential outcome of a subject:
  - the one that remains unobserved.
- Each subject is either assigned to be treated or assigned to be untreated.
- For each subject, only one of the potential outcomes is observed, the other potential outcome is counterfactual.
- If a subject is assigned to be treated, their observable outcome is their potential treated outcome, and their counterfactual outcome is their potential untreated outcome.
- If a subject is assigned to be untreated, their observable outcome is their potential untreated outcome, and their counterfactual outcome is their potential treated outcome.

Potential Outcomes: Setup

$$y_i = y_i^1$$
$$y_i = y_i^0$$

- Counterfactual outcomes involve comparing the ***observed*** outcome (what actually happened) with the ***hypothetical*** outcome (what would have happened under a different scenario).
  - The difference between the ***observed*** outcome and the ***counterfactual*** outcome.
- So a counterfactual outcome typically refers to an unobserved outcome because it represents what would have happened under a different treatment or intervention but a counterfactual outcome is not always an ***unobserved*** outcome (randomized controlled trials or natural experiments).
- Counterfactual outcomes can refer to both observed and unobserved outcomes, ***depending on*** the context and the specific question being asked.
- We cannot estimate a causal effect by considering the difference between the ***unobserved*** outcome and the ***counterfactual*** outcome alone.

Individual Treatment Effect

- The individual treatment effect (te_i for subject i) is the difference between the potential treated outcome and the potential untreated outcome.

$$te_i = y_i^1 - y_i^0$$

Heterogeneous Treatment Effects

- Individual treatment effects are likely heterogeneous:
  - the same intervention likely has a different treatment effect for different subjects.

ATE: Average Treatment Effect

- ATE is;
  - the average of the differences between potential treated outcomes and potential untreated outcomes across all subjects,
$$ATE = E[te_i] = E[y_i^1 - y_i^0]$$
  - or the difference between the average of potential treated outcomes and the average of potential untreated outcomes:
$$ATE = E[y_i^1] - E[y_i^0]$$
- ATE refers to "the effect" of an intervention.
- The same ATE may hide the different distributions of individual treatment effects.

Average Effects in Subgroups and ATET

- The average of individual treatment effects can be defined for any subgroup of subjects.
- The most important of such averages is ATET:
  - the average treatment effect on the treated group/subjects.
- Policy relevance, targeted intervention, and/or program evaluation

Quantitative Causal Variables

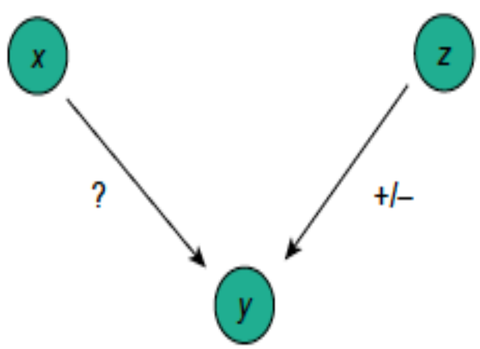
- With a quantitative causal variable, there are multiple potential outcomes, corresponding to the many values of the causal variable.
- Correspondingly, there is a series of individual treatment effects – e.g., by considering unit differences in the causal variable.
- Average individual effects are the expected increase in the potential outcome for a unit increase in the causal variable, for each subject.
- With a quantitative treatment variable, ATE is the average, across subjects, of those average individual effects.

Ceteris Paribus: Other Things Being the Same

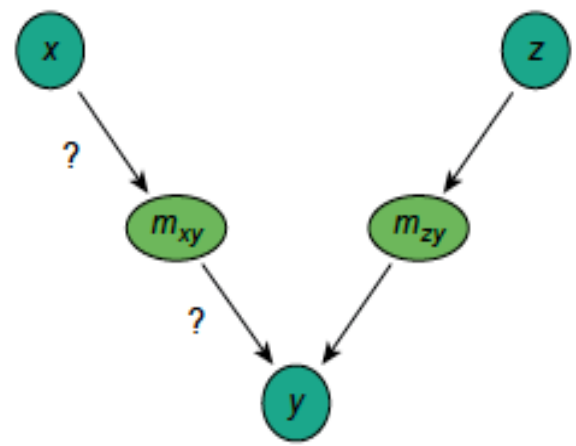
- The ceteris paribus definition of the individual treatment effect is about comparing the outcomes of a treated subject and an untreated subject.
- It's the difference in the outcomes of the two subjects if there are no other differences between them.
- Or, more generally that allows for a quantitative causal variable, it's the difference in outcome if the causal variable is different but everything else is the same.
- To make a ceteris paribus comparison, we want the two groups to be similar in all other factors that affect y independently .

Causal Maps

- We are interested in the effect of x on y (ATE).



A simple causal map



A simple causal map with mechanism variables

Random Assignment

- We can uncover the average treatment effect by comparing treated and untreated subjects if two conditions are satisfied:
  - the average outcome among treated subjects is close enough to what the average potential treated outcome would be among all subjects,
  - and the average outcome among untreated subjects is close enough to what the average potential untreated outcome would be among all subjects.
- The condition that ensures that those two conditions are satisfied is called **random assignment**.

Random Assignment

- The random assignment condition means that assignment is independent of **potential outcomes:**
  - whichever subject ends up being treated or untreated is independent of their potential outcomes.
- Random assignment here means independence of potential outcomes.
- With random assignment;
  - treated and untreated subjects are similar in their average potential outcomes
  - ATE and ATET are the same
  - the difference in average y between treated and untreated subjects is a good estimate of ATE and ATET.

Sources of Variation in the Causal Variable

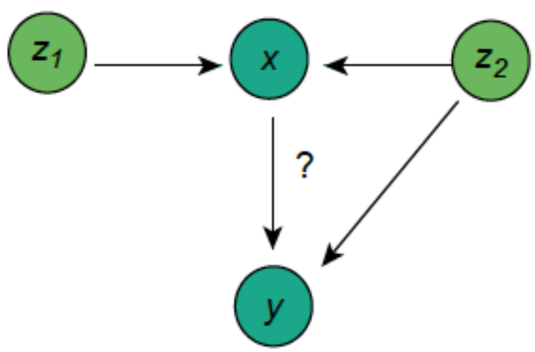
- The **sources of variation** in a variable are the reasons why different observations have different values.
- The sources of variation in the causal variable determine whether we can assume random assignment and what we can do when we can't assume random assignment.
- Assessing the sources of variation in x is a thinking process.
- The most important question about the sources of variation in x is whether and how they are related to y.
- A particular source of variation that is also related to y is an **endogenous source of variation** in the causal variable.

Sources of Variation in the Causal Variable

- A source of variation that affects x but is independent of y is an exogenous **source of variation** in the causal variable.
- Exogenous sources of variation in x affect x but are not related to y.
- Endogenous sources of variation in x affect x and are also related to y.
- To estimate the effect of x on y, we need exogenous variation in x only.

Sources of Variation in the Causal Variable

- z_1 is an exogenous source:
  - it affects x, but it is unrelated to y.
- z_2 is an endogenous source:
  - it affects both x and y.

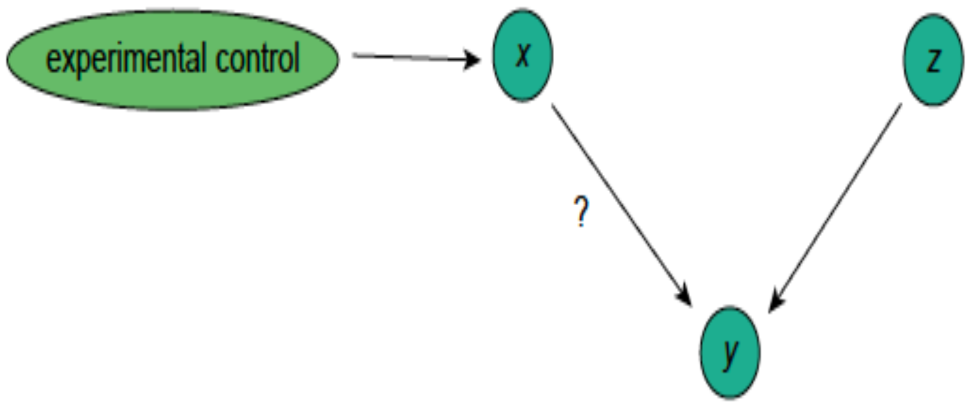


An exogenous and an endogenous source of variation in x

Experimenting versus Conditioning

- Two main approaches to uncover the average effect of an intervention are experiments and conditioning.
- The first approach uses **experimental data** controlling variation in the causal variable x.
  - Ideally, that makes all variation in x exogenous.
- The second approach uses **observational data**.
- In typical observational data, x has multiple sources of variation, some of which are likely endogenous.
- With observational data, most often, we try to condition on the endogenous sources of variation in x so that they compare observations that are different in x.

Experimenting versus Conditioning



Question: effect of x on y.
Experimental control is the only source of variation in x.
Other variables, summarized by z, may affect y but are unrelated to x.

Controlled experimental variation in x

Confounders in Observational Data

- We need to condition on other variables if we want to uncover the average effect of an intervention using observational data.
- Those other variables are endogenous sources of variation in the causal variable (x):
  - they affect x and are related to y at the same time.

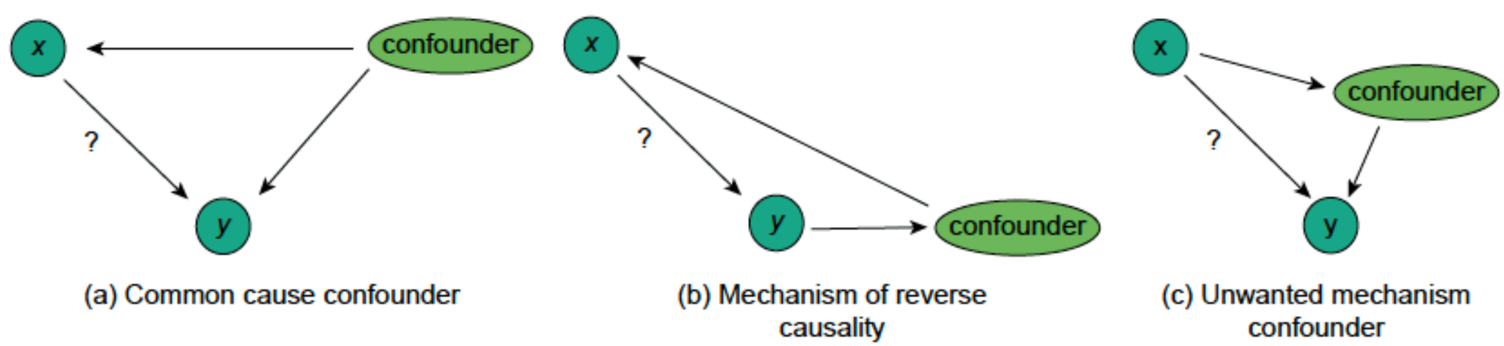


Figure 19.6 Three types of confounders

Confounders in Observational Data

- In observational data, we need to condition on all endogenous sources of variation in x.
- That conditioning requires measuring all confounders:
  - all variables that are endogenous sources of variation.
- We use some specific statistical techniques such as regression analysis and instrumental variables;
  - To control for the effects of other variables and ensure that the analysis accurately captures the relationship between the variable of interest (x) and the outcome variable.

CASE STUDY – Food and Health

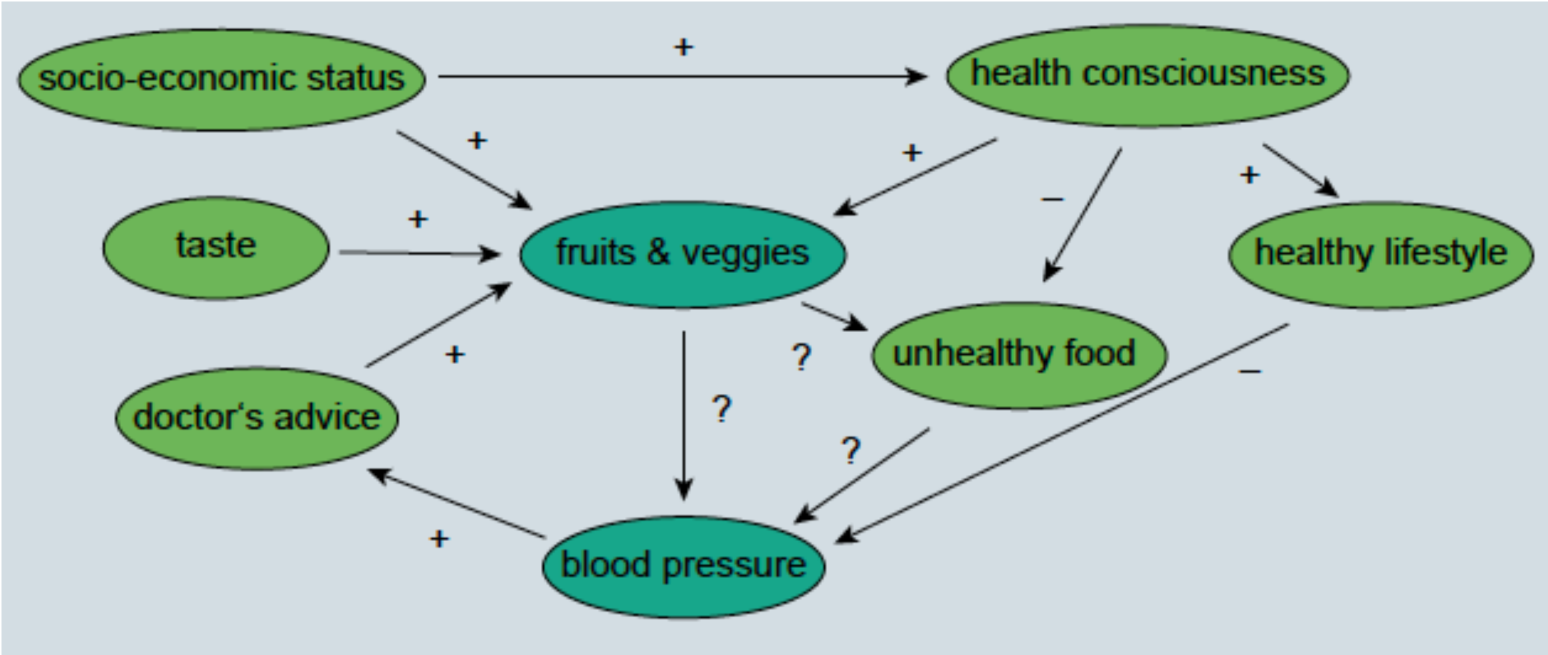


Figure 19.4 Causal map for the effect of consuming fruit and vegetables on blood pressure

Question: Effect of consuming fruit and vegetables on blood pressure.

CASE STUDY – Food and Health

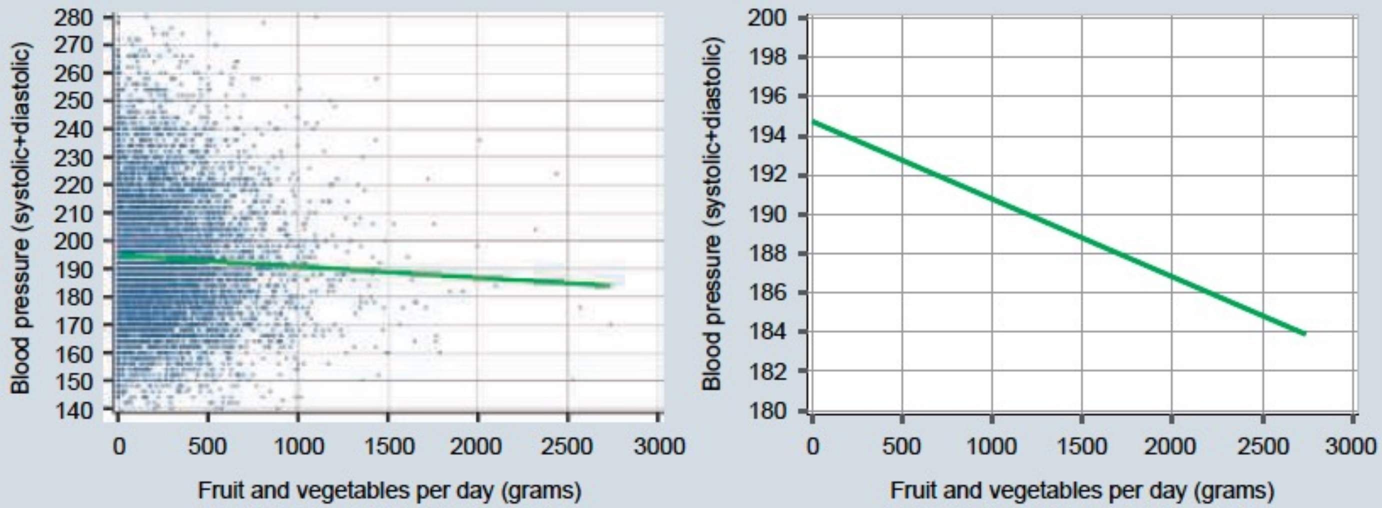
Table 19.1 Food and health – descriptive statistics

	Mean	Median	Std.Dev.	Minimum	Maximum	Observations
Blood pressure (systolic+diastolic)	194	192	24	129	300	7358
Fruit and vegetables per day, grams	260	188	274	0	2740	7358

Source: `food-health` dataset, USA, ages 30 to 59, 2009–2013.

CASE STUDY – Food and Health

- Variables to condition on



(a) Scatterplot and regression line

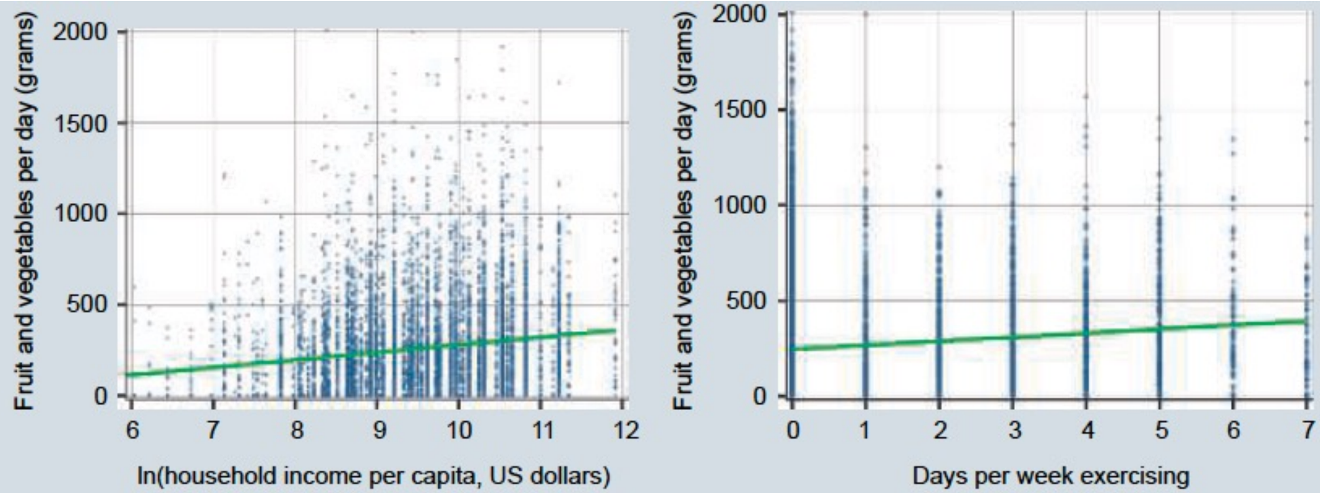
(b) Regression line only

Figure 19.8 The amount of fruit and vegetables consumed per day and blood pressure

Note: Blood pressure is sum of systolic and diastolic measures. Fruit and vegetables is the amount consumed per day (g)
Source: `food-health` dataset, USA, ages 30–59, 2009–2013. N=7358.

CASE STUDY – Food and Health

- Variables to condition on



(a) Log household income and the amount of fruit and vegetables

(b) Days per week of exercising and the amount of fruit and vegetables

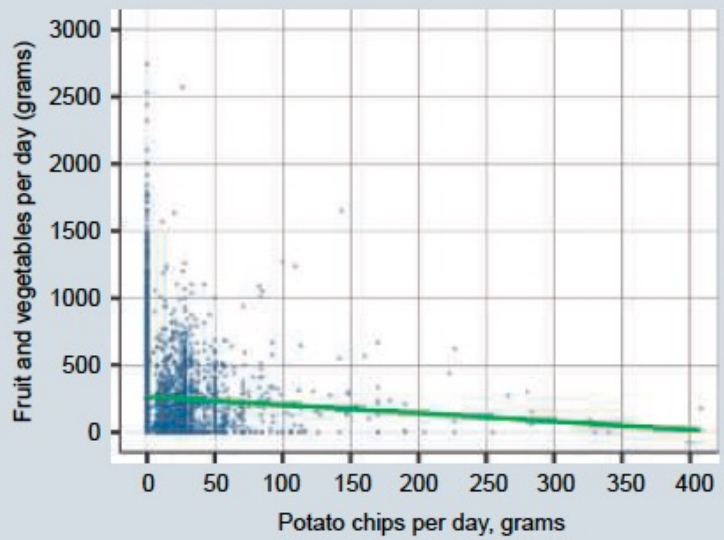
Figure 19.9 Two sources of variation in eating fruit and vegatables

Note: Scatterplots and the estimated regression lines.

Source: `food-health` dataset. USA, ages 30–59, 2009–2013. N=7358.

CASE STUDY – Food and Health

- Variables to condition on



(a) The amount of potato chips and fruit and vegetables

Figure 19.10 Consumption of an unhealthy food item: a potential bad conditioning variable

Note: Scatterplots and the estimated regression lines.
Source: `food-health` dataset. USA, ages 30–59, 2009–2013. N=7358.

Designing and Analyzing Experiments

- The goal of causal analysis is to uncover the average effect of an intervention.
- We have seen that, with **random assignment** of the causal variable x, the difference in average outcomes y between observations with different values of x gives a good estimate of the average effect of x on y.
- Random assignment means that the causal variable x is independent of potential outcomes.

Field Experiments, A/B Testing, Survey Experiments

- Field experiments make the experimental situation as close to the "natural environment" of the interventions as possible.
- A/B tests are specific field experiments:
  - To evaluate the design choices of products or online ads or websites, by presenting alternatives online.
- Survey experiments present different content to different respondents and record their answers to survey questions.

Experimental Setup

- Random assignment means that treatment is independent of potential outcomes.
- Random assignment is achieved by an assignment rule.
- The main idea is to make sure assignment is independent of everything that may affect the outcome.
- With a binary intervention, subjects are assigned either to be treated or not to be treated:
  - The **treatment group** or the **treated group**
  - The **non-treatment/untreated/control/comparison group**

Experimental Setup

- Random numbers generated by computers!
- Deciding on the number of subjects to include and the proportion of the treatment group both affect the precision of the effect estimate.
- Assigning 50 percent of subjects to treatment leads to the highest precision.
- Choosing the number of subjects balances the need for precision with the costs of the experiment.

Random Assignment and Covariate Balance

- Covariate balance means that the covariates that may affect y have the same distribution in the treatment group and in the non-treatment group.
- Covariate balance is important because it helps ensure that any observed differences in the outcome variable can be attributed to the treatment or intervention itself rather than confounding factors.
- By achieving covariate balance, we can reduce the potential for bias and improve the validity of causal inferences or treatment effects estimation.
- It allows us to isolate the impact of the treatment variable on the outcome variable, controlling for other covariates that may influence the outcome.

Random Assignment and Covariate Balance

- In practice, we check whether the average of covariates is similar in the two groups.
- The results of checking covariate balance can be:
  1. all covariates are balanced, no evidence against random assignment: we can go ahead with the analysis;
  2. some covariates are unbalanced, weak evidence against random assignment: we should keep this in mind during the analysis;
  3. many covariates are unbalanced, strong evidence against random assignment: we should go back and check what went wrong, and if we can't undo it, we need to treat the data as observational as opposed to experimental.

Imperfect Compliance and Intent-to-Treat

- **Compliance is perfect** if assignment and actual treatment are the same:
  - all subjects assigned to treatment end up being treated and all subjects assigned to be untreated end up being untreated, as planned.
- **Compliance is imperfect** if there is some **non-compliance**:
  - some subjects assigned to treatment end up being untreated, and/or some subjects assigned to be untreated end up being treated.
- With imperfect compliance, we distinguish two kinds of average treatment effects.
- One is the average treatment effect (ATE), the effect of the treatment itself.
- The other one is the effect of being assigned to the treatment.
- The average effect of being assigned to treatment is called the **average intent-to-treat effect**, sometimes abbreviated as **AITTE**.

Estimation and Statistical Inference

- The average effect of an experiment is best estimated by the regression
$$y^E = \alpha + \beta x$$
- The estimate of β is the difference in the average outcome (y) between the treatment group ($x = 1$) and the non-treatment group ($x = 0$).
- β is a good approximation of ATE if assignment is random and compliance is perfect.
- β is a good approximation of AITTE if assignment is random and compliance is imperfect.

Including Covariates in a Regression

- When assignment is random, there is no need to include covariates in the regression to estimate ATE (or AITTE if compliance is imperfect).
- When assignment is random, including covariates in the regression may lead to smaller SE and thus narrower CI.
- When we aren't sure if assignment is random, because covariate balance is not perfect, we can check if those imbalances make a difference by comparing the estimated effect with or without including the covariates.

Spillovers

- When there are spillovers, the fact that a particular subject is in the treatment group has an effect on the outcomes of some of the other subjects.
- So, spillovers, or extraneous effects, affect other subjects when a particular subject is treated.
- With spillovers, the total effect of an intervention includes the effects on all subjects.
- When spillovers affect members of the comparison group, estimates of the average effect for the treated subjects are biased.
- When spillovers are a possibility, good experimental design takes them into account – for example, by randomized assignment across groups as opposed to individuals.

Regression and Matching with Observational Data

- Eventually, the goal of causal analysis will be to use exogenous sources in x.
- In observational data, one way to ensure that, at least in principle, is conditioning on all endogenous sources of variation – these are the confounders.

Thought Experiments

- Mostly, experimental data is not available in economics, so we rely on observational data.
- It is still good practice to think through a thought experiment when doing causal analysis on observational data.
  - experiment that is designed but not carried out.
- First, a thought experiment can clarify the details of the intervention:
  - the subjects, the treatment variable, the outcome we want to examine, and how it compares to the causal variable in the data.
- Second, it can help understand the mechanisms through which the causal variable may affect the outcome.
  - Understanding what mechanisms may play a role is important for various reasons, including identifying the variables that we should not condition on.

Thought Experiments

- Third, a thought experiment helps understand how the ideal situation compares to what we have:
  - How a hypothetical random assignment compares to the sources of variation in the causal variable in our data.

Variables to Condition on, Variables Not to Condition On

- After defining the causal variable and the outcome variable and describing an appropriate thought experiment, we should turn to the sources of variation in the causal variable.
- Exogenous sources are variables that are independent of potential outcomes, while endogenous sources are variables that are related to potential outcomes.
- Eventually, the goal of our analysis will be to use exogenous sources in x.
- In observational data, one way to ensure that, at least in principle, is conditioning on all endogenous sources of variation – these are the confounders.

Variables to Condition on, Variables Not to Condition On

- There are three types of variables that are endogenous sources of variation, which we should condition on.
- These confounder variables are:
  - Common cause: the variable affects x and y.
  - Mechanism of reverse causality: y affects x through this variable.
  - Unwanted mechanism: x affects y through this variable, but we don't want to consider it when estimating the effect of x on y.

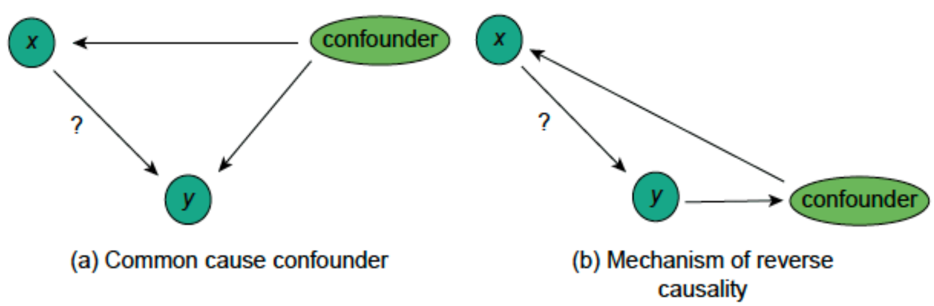


Figure 19.6 Three types of confounders

Note: Question: effect of x on y. Simple examples of confounder variables.

Variables to Condition on, Variables Not to Condition On

- There are three kinds of variables (**bad conditioners**) that we should not condition on:
  - An exogenous source of variation in x.
  - A mechanism that we want to include in the effect to be uncovered.
  - Common consequence: both x and y affect the variable.

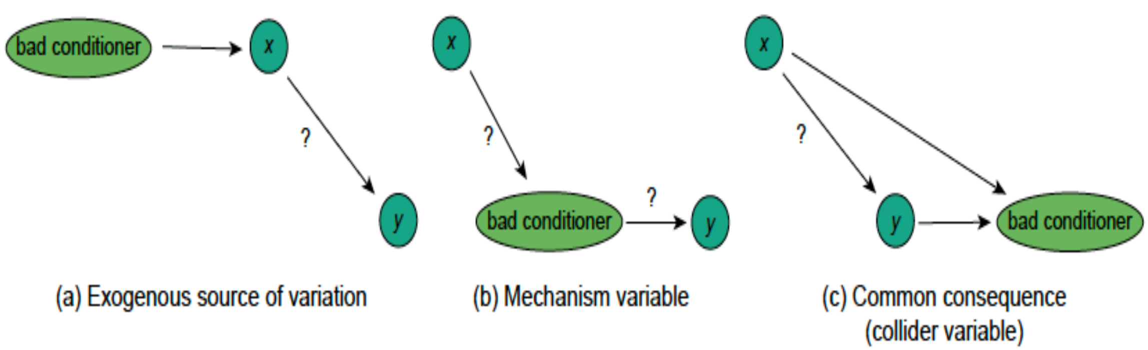


Figure 19.7 The three types of bad conditioning variables

Conditioning on Confounders by Regression

- In a causal analysis, we would like to use the multiple regression to estimate the effect of x on y, conditioning on observable confounder variables $(z_1, z_2, \ldots)$:
$$y^E = \beta_0 + \beta_1 x + \beta_2 z_1 + \beta_3 z_2 + \cdots$$
- As we know, here β_1 approximates the average difference in y between observations that are different in x but have the same values for z_1, z_2, . . .
- This interpretation of β_1 is always true regardless of whether those z variables capture all endogenous variation in x.

Conditioning on Confounders by Regression

- If the $z_1, z_2, \ldots$ variables capture all **endogenous sources of variation**, we say that x is **exogenous in the regression**.
- That means that, conditional on the $z_1, z_2, \ldots$ variables in the regression, the variation in x is exogenous.
- In this case, the OLS estimate of β_1 in the regression is a good estimate of the average effect of x on y.
- Unfortunately, as we know, that's very unlikely to happen in observational data.

Conditioning on Confounders by Regression

- Instead, the $z_1, z_2, \ldots$ variables tend to capture only some, but not all, of the endogenous sources of variation in x.
- x is endogenous in the regression.
- When x is endogenous in the regression, the OLS estimate of β_1 is a not good estimate of the average effect of x on y.
- Instead, it is a biased estimate of that effect.
- That bias is called the omitted variables bias because we omit some confounders.

Instrumental Variables

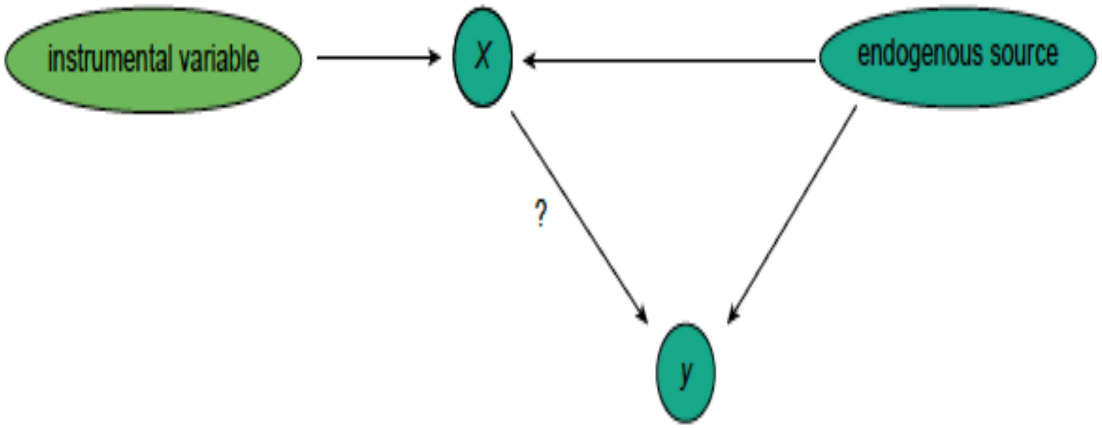


Figure 21.2 Instrumental variable that helps uncover the effect of *x* on *y*

Note: The question is the effect of *x* on *y*; there is an exogenous instrumental variable and another variable that is an endogenous source of variation in *x*.

The power of IV regression lies in the fact that we don't have to observe all, or any, variables that are endogenous sources.
But we have to observe at least one variable that is an exogenous source of variation in.

Instrumental Variables

- Two-stage least squares (2SLS)

$$y^E = \beta_0 + \beta_1 x + \beta_2 z$$

- S1: $\hat{x} = \mu_0 + \mu_1 z + \mu_2 IV$
- S2: $\hat{y} = \theta_0 + \theta_1 \hat{x}$

Reduced Form

- To obtain the reduced form of the instrumental variable (IV) regression for the equation $y^E = \beta_0 + \beta_1 x + \beta_2 z$, we can start by substituting the first-stage equation (S1) into the original equation:
- $y^E = \beta_0 + \beta_1 \hat{x} + \beta_2 z$
- Now, we can substitute the predicted value of x ($\hat{x}$) from the first stage into the equation:
- $y^E = \beta_0 + \beta_1(\mu_0 + \mu_1 z_1 + \mu_2 IV) + \beta_2 z$
- Simplifying the equation, we get:
- $y^E = \beta_0 + \beta_1 \mu_0 + \beta_1 \mu_1 z_1 + \beta_1 \mu_2 IV + \beta_2 z$

Reduced Form

- Now, let us define the reduced form coefficients:

$$\gamma_0 = \beta_0 + \beta_1 \mu_0$$
$$\gamma_1 = \beta_1 \mu_1$$
$$\gamma_2 = \beta_1 \mu_2$$
$$\lambda = \beta_2$$

- The reduced form equation becomes:

$$y^E = \gamma_0 + \gamma_1 z_1 + \gamma_2 IV + \lambda z$$

- In this reduced form equation, γ_0, γ_1, and γ_2 represent the coefficients of the exogenous variables z_1 and IV, while λ represents the coefficient of the endogenous variable z.
- Note that the reduced form equation provides the relationship between the endogenous variable y and the exogenous variables z and IV without explicitly involving the instrumental variable $\hat{x}$.