Rubin's Causal Model
Endogeneity and Exogeneity

Tadao Hoshino （星野匡郎）

ver. 2018 Fall Semester

Introduction: Correlation and Causality

- For variables X and Y, we say they are correlated when

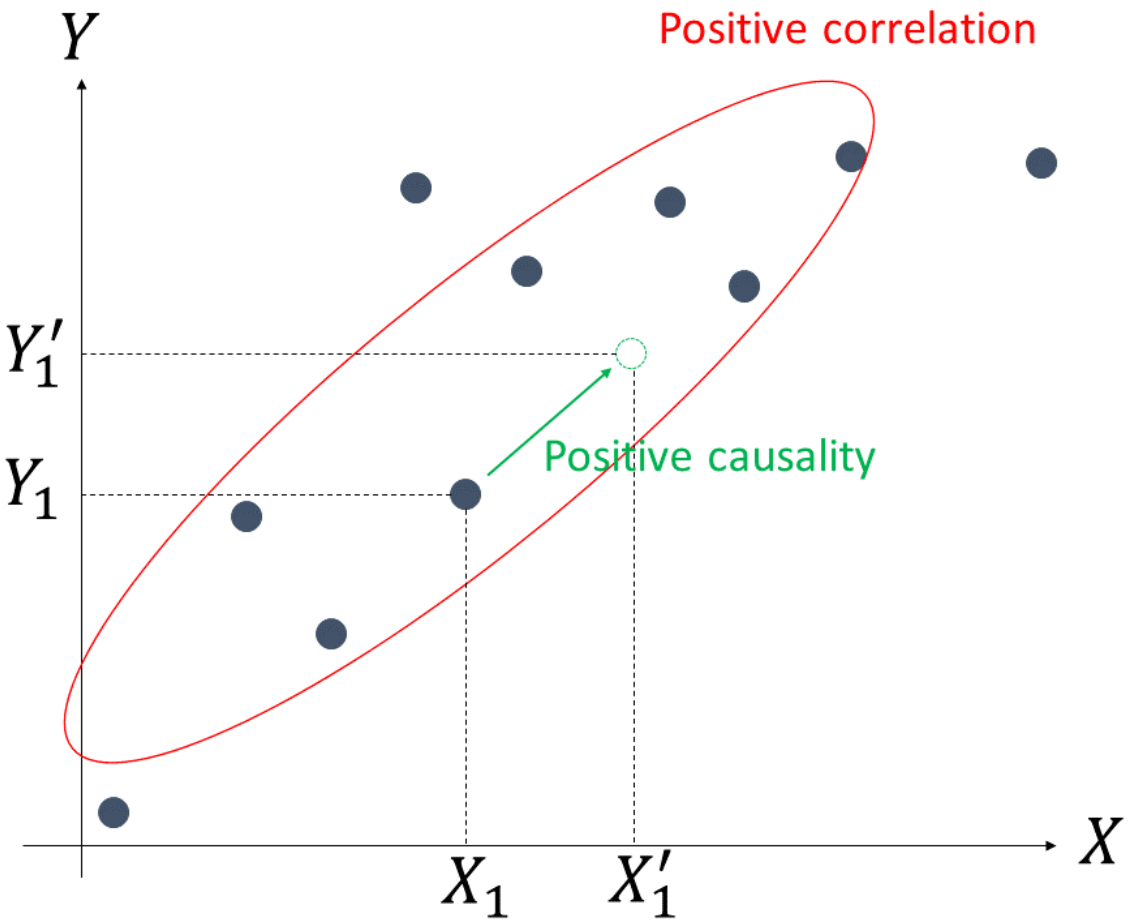
  the two variables vary together
  (Y tends to change as X changes)

  Correlation = overall tendency of the observed data
- On the other hand, when

  a change in X causes a change in Y
  (one can change X in order to change Y)

  we say there is a causal relationship from X to Y.

Introduction: Correlation and Causality



- For example, consider a case where X: weight and Y: height.
- Clearly, X and Y are correlated; but there is no causality from X to Y.

Introduction: Correlation and Causality

- In general, causality implies correlation, however correlation does not imply causality.
- We are often interested in quantifying the causality between variables, rather than a mere correlation.

Does raising minimum wage increase/decrease unemployment rate?

Even if we can observe a statistical correlation

"the unemp rate tends to be smaller in regions with lower mini wages"

this does *not* imply the causality

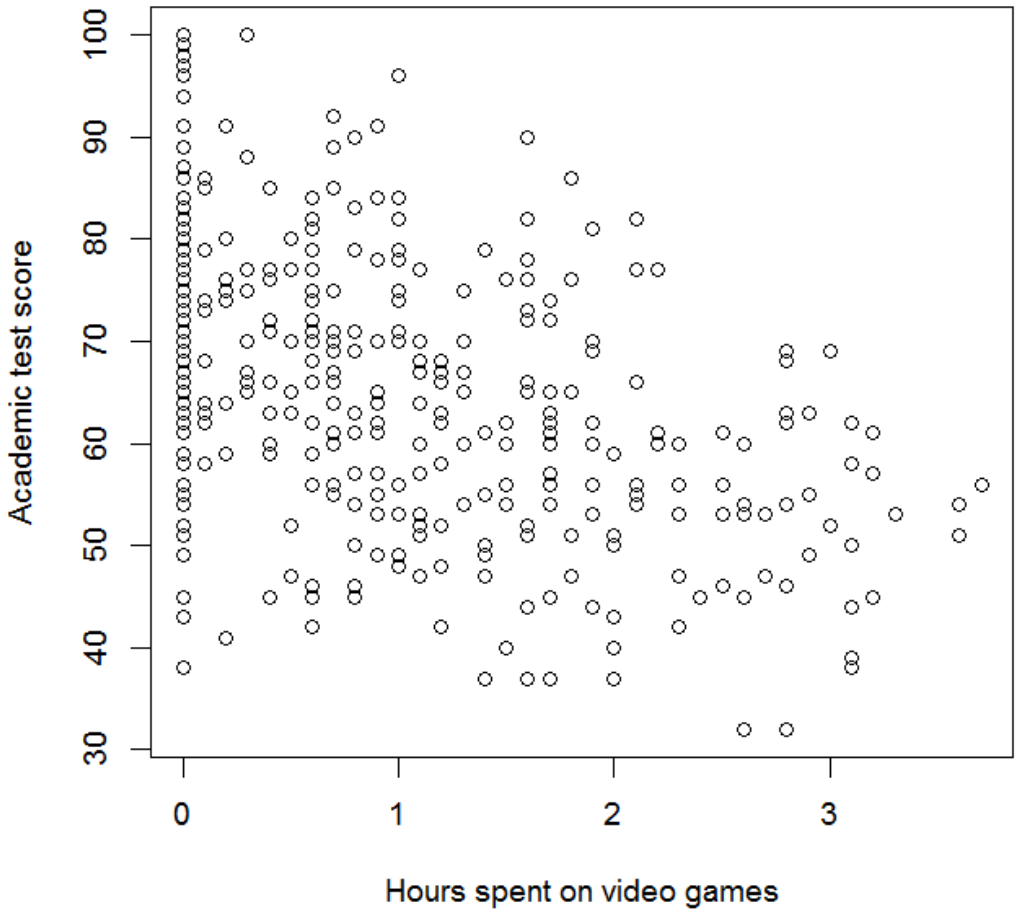
"the unemp rate can be reduced by lowering the mini wage"

- The above example highlights the importance of distinguishing between correlation and causality for policy analysis.
  $\Rightarrow$ Statistical Causal Inference

Correlation $\neq$ Causality

Spurious Correlation

Playing video games and academic performance[1]



[1] This is simulated data created by myself.

Spurious Correlation

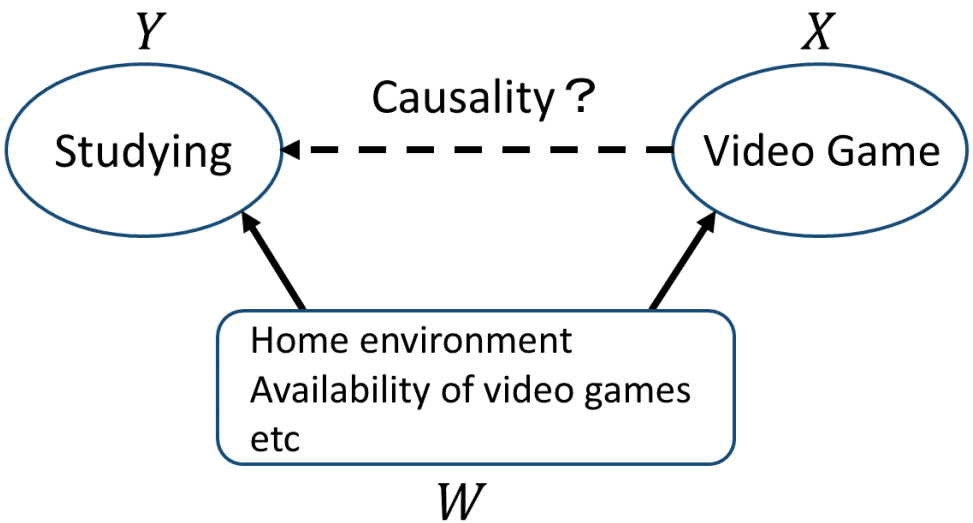
- We observe an overall negative correlation between video game and academic performance (correlation coef. : -0.49).
- From this result, we *cannot* conclude that "playing video games has a negative influence on student's academic performance".
- Regardless of whether playing video games negatively or even positively affects the academic performance, a negative correlation can exist between them. Why?

Spurious Correlation

A spurious correlation occurs when two variables (X and Y) that actually have no logical connection can be correlated due to the presence of an unobserved variable (W) that affects both X and Y.

Spurious Correlation

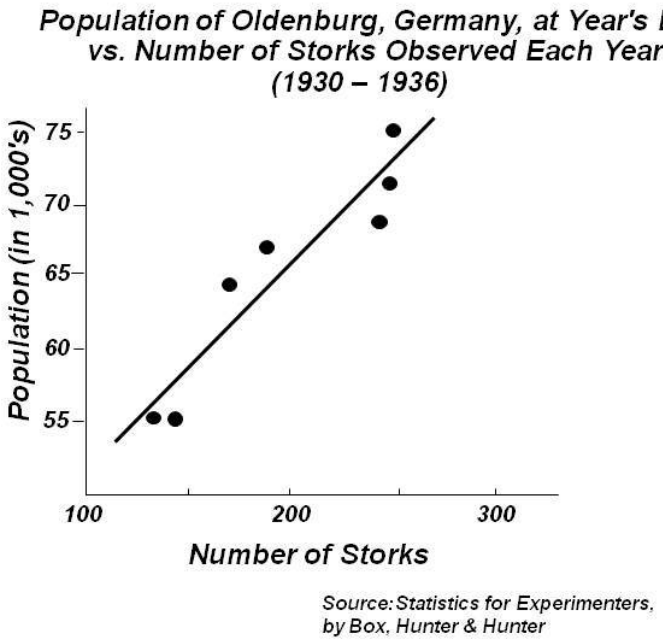
- Unobserved variables (W) that can affect both playing video games (X) and studying (Y):
  - home environment (e.g., both parents work late)
  - availability of video games in the child's own bedroom
  - etc
- In order to improve Y, improving the W's is much more important than merely decreasing X.



Spurious Correlation

Examples of spurious correlation:

- *Storks bring babies.*[2]

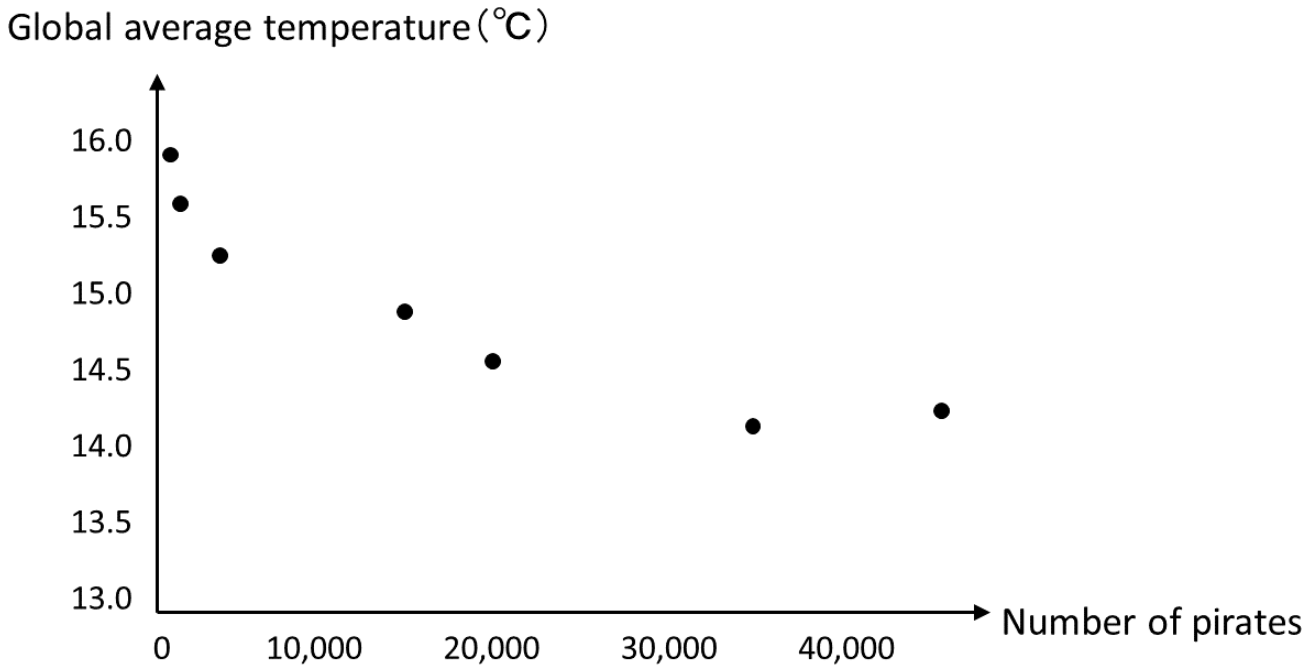


- Consumption of beers and the number of water accidents.
- etc

[2] storks = コウノトリ

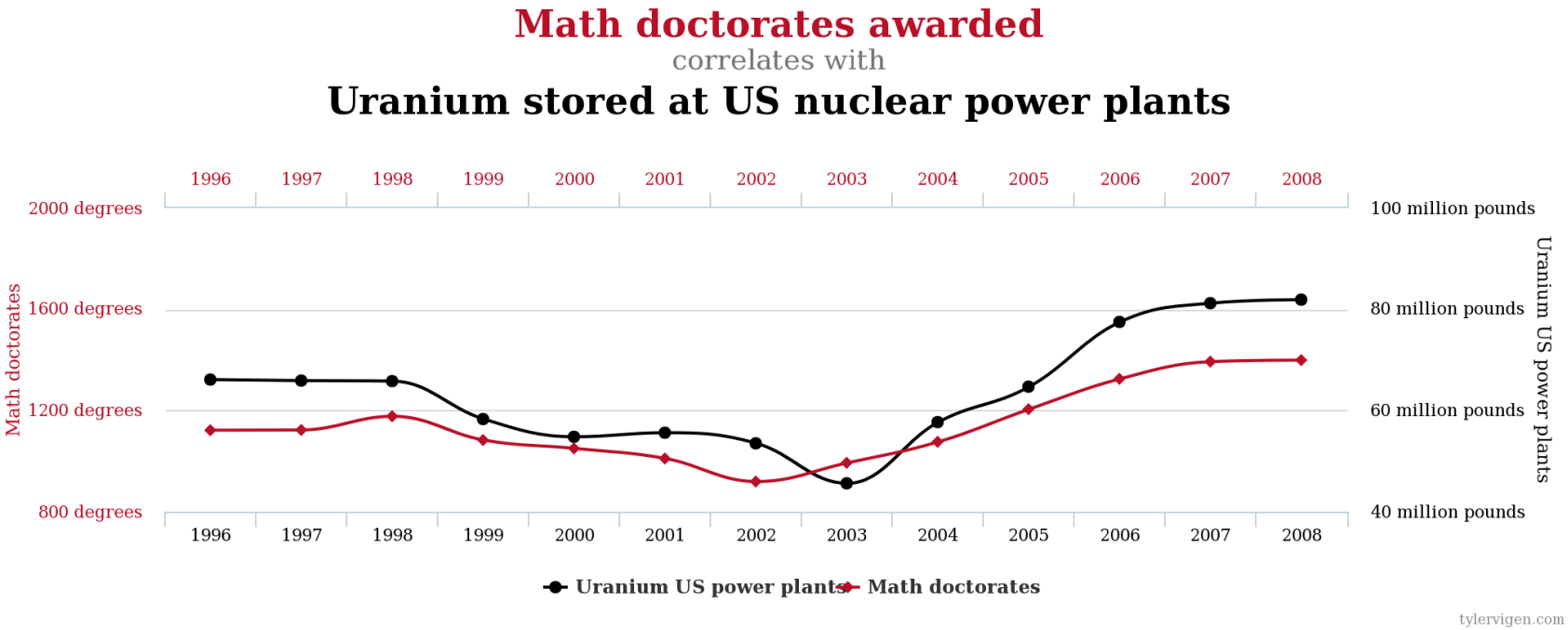
Correlation just by chance

Number of pirates and global warming



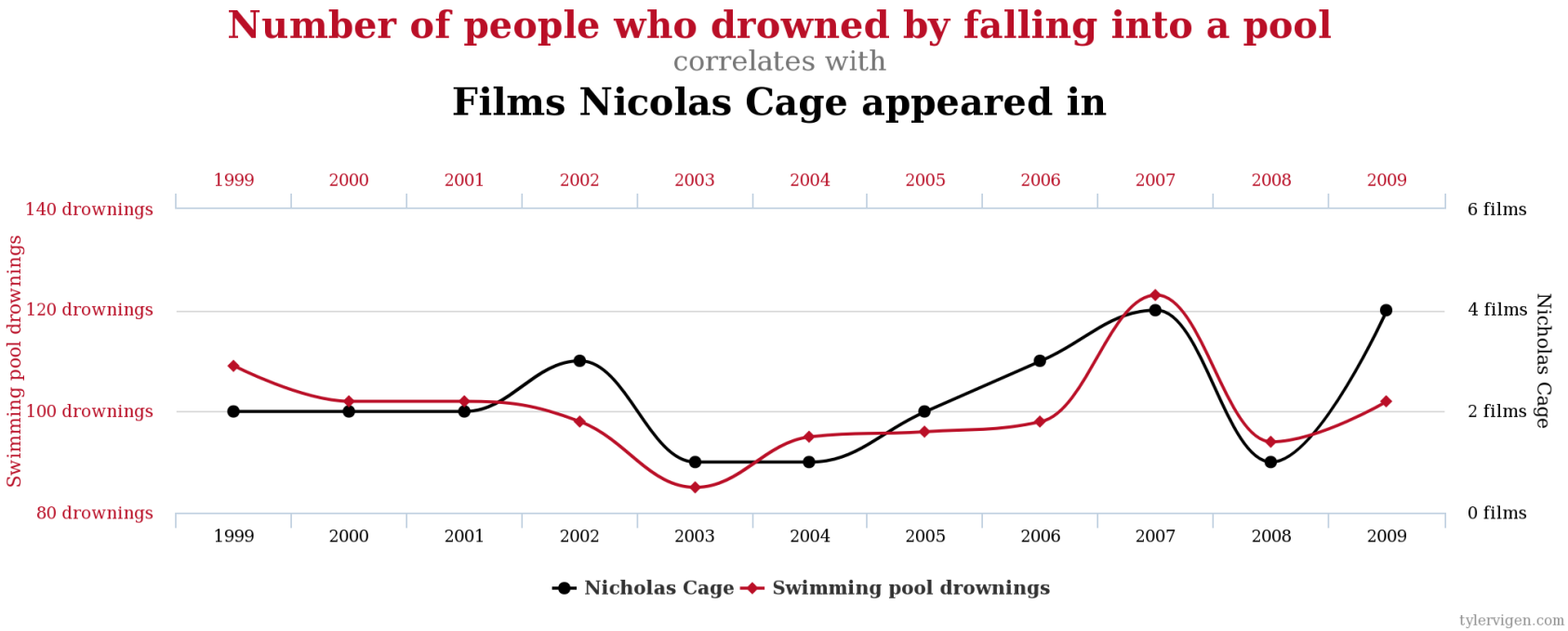
(ref.) http://sparrowism.soc.srcf.net/home/pirates.html

Correlation just by chance



(ref.) http://www.tylervigen.com/spurious-correlations

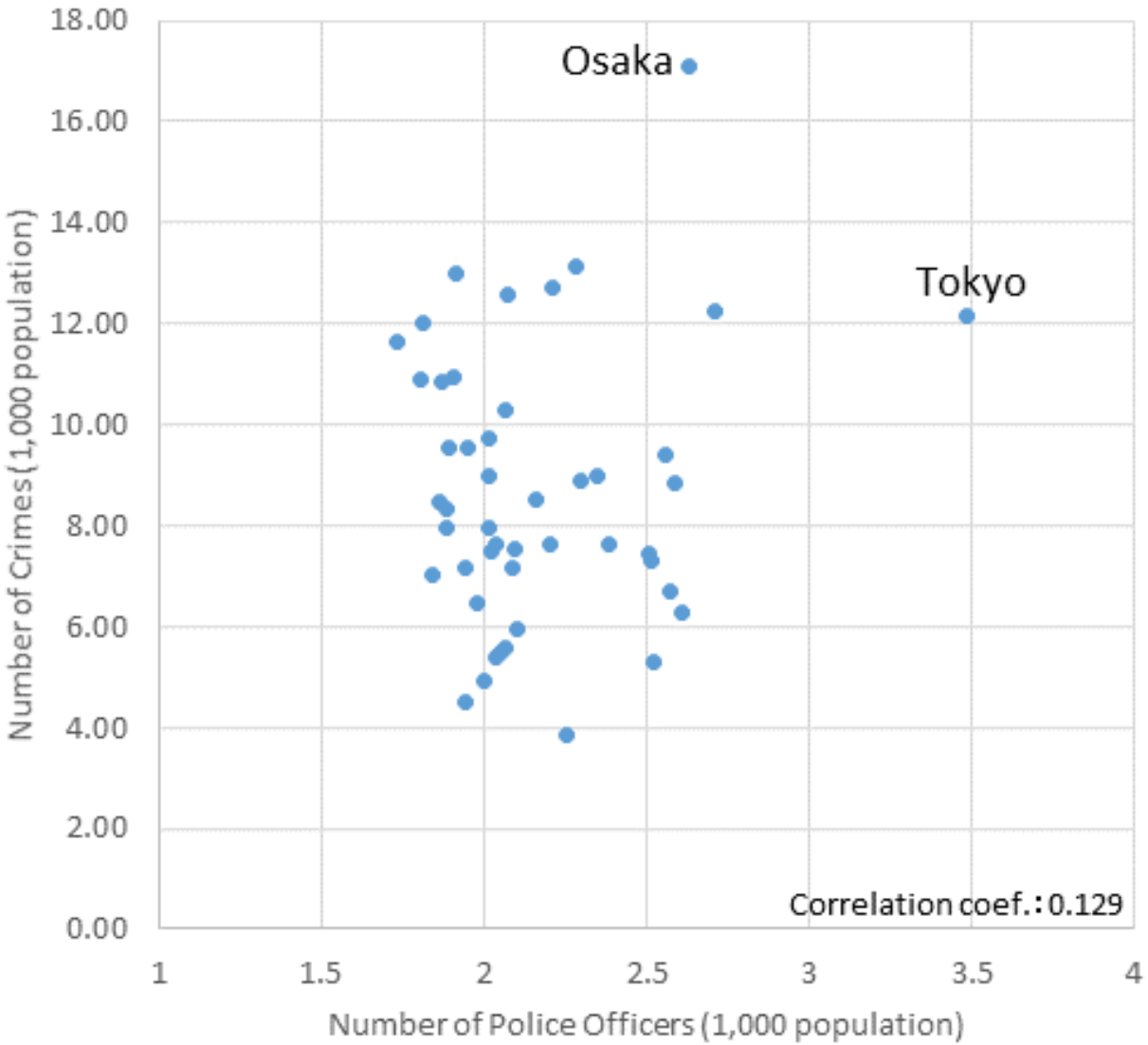
Correlation just by chance



(ref.) http://www.tylervigen.com/spurious-correlations

Simultaneity

Number of police officers and crimes per 1,000 population.



Simultaneity

- It is normally believed that increasing police reduces crimes.
  $=>$ We expect a negative causal relationship between police and crime.
- However, the observed correlation has the opposite sign to the expected causal effect. This phenomenon can be explained by the simultaneity of police levels and crime.

Simultaneity

For variables X and Y, simultaneity means that X is a cause of Y and at the same time Y is a cause of X.

- If the arrest rates are about the same across regions, this would generate a tendency "police $\uparrow$ - crime $\uparrow$" and "police $\downarrow$ - crime $\downarrow$", which results in positive correlation between police and crime.

Linear Regression $=$ Correlation Analysis $\not\Rightarrow$ Causality

Correlation Coefficient

Let $\{(X_1, Y_1), ..., (X_n, Y_n)\}$ be the sample data of (X, Y) of size n. The (sample) correlation coefficient r_{XY} is defined as

$$r_{XY} = \frac{\sum_{i=1}^{n}(X_i - \bar{X}_n)(Y_i - \bar{Y}_n)}{\sqrt{\sum_{i=1}^{n}(X_i - \bar{X}_n)^2}\sqrt{\sum_{i=1}^{n}(Y_i - \bar{Y}_n)^2}}$$

where $\bar{X}_n$ and $\bar{Y}_n$ are the sample averages of X and Y, respectively.

The correlation coefficient takes a value ranging from -1 to 1:

$r_{XY} > 0$ positive correlation ($r_{XY} = 1$: perfect positive correlation)
$r_{XY} = 0$ no correlation
$r_{XY} < 0$ negative correlation ($r_{XY} = -1$: perfect negative correlation)

Linear Regression = Correlation Analysis

- Consider a simple linear regression model:

$$Y_i = \beta_0 + X_i\beta_1 + \varepsilon_i, \quad i = 1, ..., n$$

- Recall that the OLS estimator of β_1 is given by

$$\hat{\beta}_{n1} = \frac{\sum_{i=1}^{n}(X_i - \bar{X}_n)(Y_i - \bar{Y}_n)}{\sum_{i=1}^{n}(X_i - \bar{X}_n)^2}$$

Linear Regression = Correlation Analysis

- Observe that

$$r_{XY} = \frac{\sum_{i=1}^{n}(X_i - \bar{X}_n)(Y_i - \bar{Y}_n)}{\sqrt{\sum_{i=1}^{n}(X_i - \bar{X}_n)^2}\sqrt{\sum_{i=1}^{n}(Y_i - \bar{Y}_n)^2}}$$

multiplying both the num and denom by $\sqrt{\sum_{i=1}^{n}(X_i - \bar{X}_n)^2}$

$$= \frac{\sum_{i=1}^{n}(X_i - \bar{X}_n)(Y_i - \bar{Y}_n)}{\sum_{i=1}^{n}(X_i - \bar{X}_n)^2} \cdot \frac{\sqrt{\sum_{i=1}^{n}(X_i - \bar{X}_n)^2}}{\sqrt{\sum_{i=1}^{n}(Y_i - \bar{Y}_n)^2}}$$

$$= \hat{\beta}_{n1} \cdot \underbrace{\frac{\sqrt{\sum_{i=1}^{n}(X_i - \bar{X}_n)^2}}{\sqrt{\sum_{i=1}^{n}(Y_i - \bar{Y}_n)^2}}}_{\text{positive constant}}$$

- Therefore, the regression coefficient $\hat{\beta}_{n1}$ and the correlation coefficient r_{XY} are proportionally related, i.e., $\hat{\beta}_{n1} = 0 \iff r_{XY} = 0$.

Linear Regression = Correlation Analysis

- As stated above, correlation does not imply causation. Therefore, for a regression model

$$Y_i = \beta_0 + X_i\beta_1 + \varepsilon_i, \quad i = 1, ..., n$$

when β_1 is estimated to be significantly positive, although we can say that

"Y tends to increase as X increases" – correlation,

we cannot say that

"an increase in X causes an increase in Y" – causality

in general.[3]

[3] As described later, if X has a property called *exogeneity*, the estimated β_1 can be interpreted as the causal effect of X on Y.

Rubin's Causal Model

Rubin's Causal Model

In order to statistically deal with "causality", first of all, there is a question:

How can we mathematically define the causal relationship btwn variables?

In the literature, there are several different approaches:

- Rubin's Causal Model (D. Rubin)
- Structural Causal Models (J. Pearl)
- etc.

Rubin's Causal Model

- Among them, the Rubin's Causal Model (RCM) is the most commonly used framework in econometrics.
- We consider causal relationship between two variables (Y, T).
  - Y is an outcome variable of interest, and T is a variable that potentially affects the value of Y.
  - T is called the treatment variable.
- In the following, we focus on the case where the treatment T is a dummy variable, i.e., $T \in \{0, 1\}$.

Examples of (Y, T):

Y	T	Causal effect of interest
Wage	Completion of college degree yes: $T = 1$, no: $T = 0$	Wage returns to a college degree
Academic grade	Participation in remedial courses yes: $T = 1$, no: $T = 0$	Effect of remedial education on academic performance
Emission of pollutants	Introduction of emission ctrl tech yes: $T = 1$, no: $T = 0$	Effect of the control technology on pollutants

Rubin's Causal Model

- For each individual i, suppose that we observe (Y_i, T_i).
- Each individual has two potential outcomes, Y_1 and Y_0, one for each value of the treatment T:
  - Y_{1i} = the value of Y_i when $T_i = 1$ is true,
  - Y_{0i} = the value of Y_i when $T_i = 0$ is true.
- Note that it is *impossible* to observe an individual under two different treatment statuses (T_i must be either one or zero).
- That is, when $T_i = 1$ is realized, we can observe an outcome $Y_i = Y_{1i}$, but Y_{0i} is an unobservable potential outcome (the so-called counterfactual outcome.)

Treatment status	Observed outcome	Counterfactual outcome
$T_i = 1$	$Y_i = Y_{1i}$	Y_{0i}
$T_i = 0$	$Y_i = Y_{0i}$	Y_{1i}

Rubin's Causal Model

- Thus, the observed outcome Y_i can be expressed as

$$Y_i = T_i \cdot Y_{1i} + (1 - T_i) \cdot Y_{0i}$$

- In this framework, the causal effect of T on Y for individual i is defined as

$$Y_{1i} - Y_{0i}.$$

  This quantity is referred to as the (individual) treatment effect for i.

- Namely, the causal effect (i.e., treatment effect) is defined as the difference between

the observed outcome when $T_i = 1$ (0) is actually selected
and
the counterfactual outcome when $T_i = 0$ (1) is hypothetically selected.

Rubin's Causal Model

- The treatment effect may be heterogeneous among individuals.
  - It is possible that the treatment is beneficial for some people but not for others.
- Unless there is someone completely identical to i, we cannot observe both potential outcomes Y_{0i} and Y_{1i} at the same time, and thus the individual treatment effect cannot be directly measured.
- However, it is still possible to estimate the treatment effect averaged over some subpopulation.

Average Treatment Effect (ATE)

Average Treatment Effect (ATE)

- It is impossible to directly estimate the individual treatment effects from data.
- Let us focus on, not individual specific treatment effect, but the treatment effect averaged over the whole population, the so-called average treatment effect (ATE):

$$E[Y_{1i}] - E[Y_{0i}].$$

- When a sample of n observations $\{(Y_i, T_i) : i = 1, ..., n\}$ is available, how can we estimate the ATE?

Average Treatment Effect (ATE)

- Treatment group $\{i : T_i = 1\}$: individuals who actually receive the treatment.
- Control group $\{i : T_i = 0\}$: individuals who do not receive the treatment.
  - The sample size of the treatment group and that of the control group are $\sum_{i=1}^{n} T_i$ and $\sum_{i=1}^{n}(1 - T_i)$, respectively.
- A naive estimator of ATE: $E[Y_{1i}] - E[Y_{0i}]$ is just the difference between the average outcome of the treatment group and that of the control group:

$$\underbrace{\frac{\sum_{i=1}^{n} T_i Y_i}{\sum_{i=1}^{n} T_i}}_{\text{average } Y \text{ in TG}} - \underbrace{\frac{\sum_{i=1}^{n} (1 - T_i) Y_i}{\sum_{i=1}^{n} (1 - T_i)}}_{\text{average } Y \text{ in CG}} \tag{1}$$

Average Treatment Effect (ATE)

- In fact, (1) is numerically equivalent to the OLS slope estimator $\hat{\beta}_{n1}$ of the following simple linear regression model:

$$Y_i = \beta_0 + T_i\beta_1 + \varepsilon_i, \quad i = 1, ..., n$$

that is,

$$\frac{\sum_{i=1}^{n} T_i Y_i}{\sum_{i=1}^{n} T_i} - \frac{\sum_{i=1}^{n} (1 - T_i) Y_i}{\sum_{i=1}^{n} (1 - T_i)} = \frac{\sum_{i=1}^{n} (T_i - \bar{T}_n)(Y_i - \bar{Y}_n)}{\sum_{i=1}^{n} (T_i - \bar{T}_n)^2}$$

holds. (This equality can be confirmed by direct calculation.)

- Thus, the simple difference estimator (1) only shows the correlation of T and Y, and therefore, it may not necessarily reflect the true causal effect of T on Y.

Summary of the Discussion So Far

- Correlation does not imply causality due to the presence of, e.g., spurious correlation and simultaneity.
- The linear least squares estimator of the regression coefficients is proportional to the correlation coefficient.
  $=>$ Linear regression is a correlational analysis.
- In the RCM framework, the causal effect of T on Y averaged over the population is called the average treatment effect (ATE):

$$\text{ATE: } E[Y_{1i} - Y_{0i}].$$

Summary of the Discussion So Far (cont')

- At first glance, it seems that the ATE can be estimated by the difference between the average outcome of the treatment group and that of the control group.
- However, since such simple difference estimator is equivalent to the OLS estimator of β_1 in a regression model

$$Y_i = \beta_0 + T_i\beta_1 + \varepsilon_i,$$

the resulting estimate only reflects the correlation between T and Y.

Q. Under what conditions does the simple difference estimator correctly recover the ATE?
A. A sufficient condition is that the treatment variable T has a property called "exogeneity" with respect to the outcome variable Y.

Endogeneity and Exogeneity

Endogeneity and Exogeneity

- Consider a simple regression model:

$$Y_i = \beta_0 + X_i\beta_1 + \varepsilon_i, \qquad i = 1, ..., n,$$

  where $E[\varepsilon_i] = 0$. The causal impact of a unit increase in X on Y is β_1.
- The variable X is said to be exogenous in the model if it is uncorrelated with the error term ε.
- That is, when

$$\underbrace{Cov\left(X_i, \varepsilon_i\right)}_{=E[X_i\varepsilon_i]} = 0, \qquad i = 1, ..., n$$

  holds, X is an exogenous variable.
- The uncorrelation with the error term is referred to as exogeneity.

Endogeneity and Exogeneity

- Taking the expectation of both sides of the simple regression model yields $\beta_0 = E[Y_i] - E[X_i]\beta_1$.
- Thus, the model can be re-written as

$$Y_i - E[Y_i] = (X_i - E[X_i])\,\beta_1 + \varepsilon_i, \qquad i = 1, ..., n. \tag{2}$$

- If X is an exogenous variable, multiplying X_i on both sides of (2) and taking the expectation of them lead to

$$\begin{aligned} Cov(X_i, Y_i) &= V(X_i)\beta_1 + Cov\,(X_i, \varepsilon_i) \\ &= V(X_i)\beta_1. \end{aligned}$$

  Hence, the slope parameter β_1 can be characterized by

$$\beta_1 = \frac{Cov(X_i, Y_i)}{V(X_i)}.$$

Recall: The OLS estimator of β_1 is obtained by replacing Cov and V with the sample covariance and variance, respectively.

Endogeneity and Exogeneity

- When X is not exogenous, i.e.,

$$\underbrace{Cov\left(X_i, \varepsilon_i\right)}_{=E[X_i\varepsilon_i]} \neq 0$$

  X is said to be an endogenous variable.
- The correlation of X with the error term ε is referred to as endogeneity of X.
- When X is an endogenous variable, we have

$$\frac{Cov(X_i, Y_i)}{V(X_i)} = \beta_1 + \underbrace{\frac{Cov\left(X_i, \varepsilon_i\right)}{V(X_i)}}_{\text{endogeneity bias}}$$

$$\neq \beta_1$$

Endogeneity and Exogeneity

- Therefore, in this case, if one tries to estimate β_1 based on the formula

$$Cov(X_i, Y_i) / V(X_i),$$

this entails an estimation bias (so-called endogeneity bias) that amounts to

$$Cov\left(X_i, \varepsilon_i\right) / V(X_i).$$

- Unfortunately, the magnitude and even the sign of the endogeneity bias are uncertain (because ε is unobservable).
- In the presence of endogeneity, the OLS estimate is uninformative about the causal impact of X on Y (it only tells us the correlation between X and Y), i.e., the OLS estimator is inconsistent.

Endogeneity and Exogeneity

Recall that when the following assumptions are met, the OLS estimator $\hat{\beta}_{n1}$ becomes consistent.

Assumption 1. The conditional expectation of Y given X is given by a linear function $\beta_0 + X\beta_1$. That is,

$$Y = \beta_0 + X\beta_1 + \varepsilon$$
$$E(Y|X) = \beta_0 + X\beta_1 \quad (\iff E(\varepsilon|X) = 0)$$

Assumption 2. $E(X^2)$ is finite.

Assumption 3. The observations $\{(Y_1, X_1), ..., (Y_n, X_n)\}$ are independent and sampled from the same population.

Assumption 4. The error term ε is independent of X, and its variance is given by $E(\varepsilon^2) = \sigma^2$.

Endogeneity and Exogeneity

- Under these conditions,

$$\text{OLS } \hat{\beta}_{n1} = \frac{\text{sample } Cov(X,Y)}{\text{sample } V(X)} \xrightarrow{p} \frac{\text{pop } Cov(X,Y)}{\text{pop } V(X)} = \beta_1.$$

- This implies that X is exogenous under these conditions.
- Indeed, by Assumption 1 and the law of iterated expectations, we have

$$E(X\varepsilon) = E[X \underbrace{E(\varepsilon|X)}_{0}] = 0.$$

- Endogeneity of X indicates the violation of Assumption 1.

Endogeneity and Exogeneity

ATE can be estimated by OLS under exogeneity.

- For simplicity, we assume that

$$Y_{1i} - Y_{0i} = \beta, \qquad i = 1, ..., n$$

  i.e., the treatment effect is homogeneous among individuals:[4]

- In this case, the ATE is simply

$$\text{ATE: } E[Y_{1i} - Y_{0i}] = \beta.$$

[4] This assumption is not necessary but is made for simplicity only.

Endogeneity and Exogeneity

ATE can be estimated by OLS under exogeneity (cont').

- Since $Y_i = T_i \cdot Y_{1i} + (1 - T_i) \cdot Y_{0i}$, we can write

$$\begin{aligned} Y_i &= Y_{0i} + T_i\beta \\ &= \underbrace{E(Y_{0i})}_{=\alpha} + T_i\beta + \underbrace{Y_{0i} - E(Y_{0i})}_{=\varepsilon_i} \\ &= \alpha + T_i\beta + \varepsilon_i \text{ such that } E(\varepsilon_i) = 0. \end{aligned}$$

- Thus, if T is exogenous such that $Cov(T_i, \varepsilon_i) = 0$, the treatment effect (ATE) β can be estimated by OLS.
- However, unfortunately the exogeneity of T cannot be tested empirically since ε is unobservable. We have to evaluate the plausibility of the exogeneity assumption on a case by case basis.

Sources of Endogeneity

Sources of Endogeneity: Omitted Variables

- Consider the following multiple regression model with two regressors X_1 and X_2:

$$Y_i = \beta_0 + X_{1i}\beta_1 + X_{2i}\beta_2 + \varepsilon_i, \qquad i = 1, ..., n$$

- Here, we assume that both X_1 and X_2 are exogenous:

$$E[X_{1i}\varepsilon_i] = 0, \;\; E[X_{2i}\varepsilon_i] = 0$$

- Then, when both X_1 and X_2 are observable variables, the slope parameters β_1 and β_2 can be correctly (consistently) estimated by multiple OLS regression.

Sources of Endogeneity: Omitted Variables

- Now consider a case in which X_2 is unobservable for some reason.
- Excluding unobservable X_2 from the model, we regress Y only on X_1 based on the following simple regression model:

$$Y_i = \alpha + X_{1i}\beta_1 + \eta_i, \qquad i = 1, ..., n$$

where $\alpha = \beta_0 + E[X_{2i}]\beta_2$, and η_i is a new error term defined by

$$\eta_i = (X_{2i} - E[X_{2i}])\,\beta_2 + \varepsilon_i \text{ such that } E(\eta_i) = 0.$$

- In this simple regression model, the explanatory variable X_2, which needs to be included in the model if $\beta_2 \neq 0$, is omitted from the analysis: omitted variables.

Sources of Endogeneity: Omitted Variables

- In order to correctly estimate β_1 in this model, X_1 needs to be an exogenous variable, in the sense that

$$E[X_{1i}\eta_i] = 0.$$

- However, by the definition of η,

$$\begin{aligned} E[X_{1i}\eta_i] &= E\left\{X_{1i}\left[\left(X_{2i} - E[X_{2i}]\right)\beta_2 + \varepsilon_i\right]\right\} \\ &= Cov\left(X_{1i}, X_{2i}\right)\beta_2. \end{aligned}$$

  Hence, unless either $Cov\left(X_{1i}, X_{2i}\right) = 0$ or $\beta_2 = 0$ (or both), X_1 is an endogenous variable: $E[X_{1i}\eta_i] \neq 0$.

Sources of Endogeneity: Omitted Variables

Interpretation as the spurious correlation problem:

- Let Y and X_1 denote the academic test score and the hours spent on video games, respectively.
- Consider estimating the causal effect of X_1 on Y using a simple regression model.
- In addition, as a factor affecting both X_1 and Y, there is a variable X_2, e.g., student's home environment.

$$\text{Home env } X_{2i} \text{ affects video games } X_{1i} \iff Cov(X_{1i}, X_{2i}) \neq 0$$

$$\text{Home env } X_{2i} \text{ affects academic test score } Y_i \iff \beta_2 \neq 0$$

Sources of Endogeneity: Omitted Variables

- Therefore, when the variable X_2 is omitted from the regression model, the video game variable X_1 is necessarily an endogenous variable, and thus the causal effect cannot be estimated by OLS.
- The amount of the estimation bias is

$$\frac{Cov\left(X_{1i}, \eta_i\right)}{V(X_{1i})} = \frac{Cov(X_{1i}, X_{2i})\beta_2}{V(X_{1i})}$$

- If X_2 is an index of home educational environment, it is expected that $Cov(X_{1i}, X_{2i}) < 0$ and $\beta_2 > 0$, i.e., the estimation bias is negative.
- Even when the causal impact of video game on studying is actually zero (i.e., $\beta_1 = 0$), the estimate of β_1 can be significantly negative.

The non-identifiability of causality due to spurious correlation and the endogeneity bias due to omitted variables are mathematically equivalent problems.

Sources of Endogeneity: Simultaneity

- As an example, we consider the estimation of the causal effect of police on crime.
- Let X_i be the number of police officers in a district i, and Y_i be the number of crime incidents in this district.
- Simultaneity: If crime rate increases, a larger police force is needed. At the same time, increasing police force can reduce crime.
- This relationship can be expressed as:

$$Y_i = \beta_0 + X_i\beta_1 + \varepsilon_i \quad (\beta_1\text{: \# of police officers} \rightarrow \text{\# of crimes})$$
$$X_i = \gamma_0 + Y_i\gamma_1 + u_i \quad (\gamma_1\text{: \# of crimes} \rightarrow \text{\# of police officers})$$

For simplicity we do not consider other factors that can influence Y and X.

Sources of Endogeneity: Simultaneity

- If X_i is an endogenous variable such that $E[X_i\varepsilon_i] \neq 0$, we cannot correctly estimate β_1.
- However, according to the second model,

$$\begin{aligned} E[X_i\varepsilon_i] &= E[(\gamma_0 + Y_i\gamma_1 + u_i)\varepsilon_i] \\ &= E[Y_i\varepsilon_i]\gamma_1 + E[u_i\varepsilon_i]. \end{aligned}$$

Further,

$$\begin{aligned} E[Y_i\varepsilon_i] &= E[(\beta_0 + X_i\beta_1 + \varepsilon_i)\varepsilon_i] \\ &= E[X_i\varepsilon_i]\beta_1 + \underbrace{E[\varepsilon_i^2]}_{>0}. \end{aligned}$$

- Thus, $E[X_i\varepsilon_i] = 0$ does not hold.

Econometric methods to circumvent endogeneity

Pure-experimental approaches

Randomized experiment

- Experimental study designs in which individuals are randomly assigned to a treatment group or a control group.
- The randomized experiment is the "gold-standard" to identify and test causal relationships. However, conducting a randomized experiment is often costly not only financially, but also ethically and politically.
- More details follow in the next lecture.

Quasi-experimental approaches

Matching method

- Matching method is a method to evaluate causal effects of a treatment by comparing individuals with similar characteristics (socio-economic background, ethnicity, gender, etc) but receiving different treatment assignment.
- Matching method can be used when the source of endogeneity for T is spurious correlation (omitted variables), and the third variable(s) W can be observed by researchers.
- This is the last topic of this course.

Instrumental variables

- Consider a simple regression model $Y = \beta_0 + X\beta_1 + \varepsilon$.
- Whatever the source of endogeneity, if we can find a variable Z that satisfies

  (A) Z is not a direct determinant of Y such that $E(Z\varepsilon) = 0$

  (B) Z is a direct determinant of X such that $Cov(Z, X) \neq 0$,

  multiplying Z on both sides of (2) and taking the expectation of them lead to

$$Cov(Z_i, Y_i) = Cov(Z_i, X_i)\beta_1 + \underbrace{Cov(Z_i, \varepsilon_i)}_{E(Z_i\varepsilon_i)=0 \text{ by (A)}}$$

$$\iff \quad \beta_1 = \frac{Cov(Z_i, Y_i)}{\underbrace{Cov(Z_i, X_i)}_{\neq 0 \text{ by (B)}}}.$$

Quasi-experimental approaches

Instrumental variables (cont')

- The last equality implies that we can estimate β_1 by replacing the population covariances with their sample counterparts.
- A random variable Z satisfying the conditions (A) and (B) is called the instrumental variable for X.
- Example: X: video games, Y: academic performance.
  - Z: proximity to an internet cafe.
- Technical details of instrumental variable method are beyond the scope of this course.

There are many other quasi-experimental approaches often used in empirical studies in economics: regression discontinuity design (RDD), difference in differences (DID), synthetic control method, etc.

Glossary I

Glossary II