SOCI 40258

Causal Mediation Analysis

Week 1: Foundations of Causal Inference

Instructor

- Geoff Wodtke
- Office: SSRB Room 425
- Office Hours: Mon 5:30-6:30pm
- Email: wodtke@uchicago.edu

Teaching Assistant

- Simone Zhang
- Office Hours: by appointment
- Email: xinmengz@uchicago.edu

Course description

- Social scientists often seek to uncover not just whether but also *why* an exposure affects an outcome by quantifying the intermediate mechanisms through which causation operates

- This course covers methods for explaining causal effects

- Specifically, it focuses on:
  - Potential outcomes and graphical causal models
  - The natural effects decomposition
  - Interventional and path-specific effects
  - Non-, semi-, and fully parametric methods of estimation
  - (Qausi-)experimental designs tailored for analyses of mediation

Prerequisites

- Students should have a solid understanding of…
  - the foundations of causal inference
  - probability and statistical inference
  - multiple regression and general linear models

- Knowledge of linear algebra and calculus will be an asset but are not required

Readings/texts

- All readings are listed in the syllabus and linked on Canvas

- Required text:

  - Wodtke, Geoffrey T. and Xiang Zhou. Forthcoming. *Causal Mediation Analysis*. Cambridge University Press.

Software

- Statistical computing for this course will be done using Stata and R
- Stata is available through the UChicago Vlab, while R is available from CRAN
- All the Stata programs and R packages used for this course, complete with detailed documentation, are accessible at:
  - https://github.com/causalMedAnalysis

Homework

- Three problem sets
  - About 10 questions each
  - Double-spaced, in 12-point font, and accompanied by clearly demarcated replication files
  - Students may re-write incorrect answers for partial credit

Final research project

- This project may consist of…
  - an empirical study with a causal research question that employs one of the methods covered in this course
  - a methodological study related to causal inference (e.g., a simulation study that employs one of the methods covered in this course)

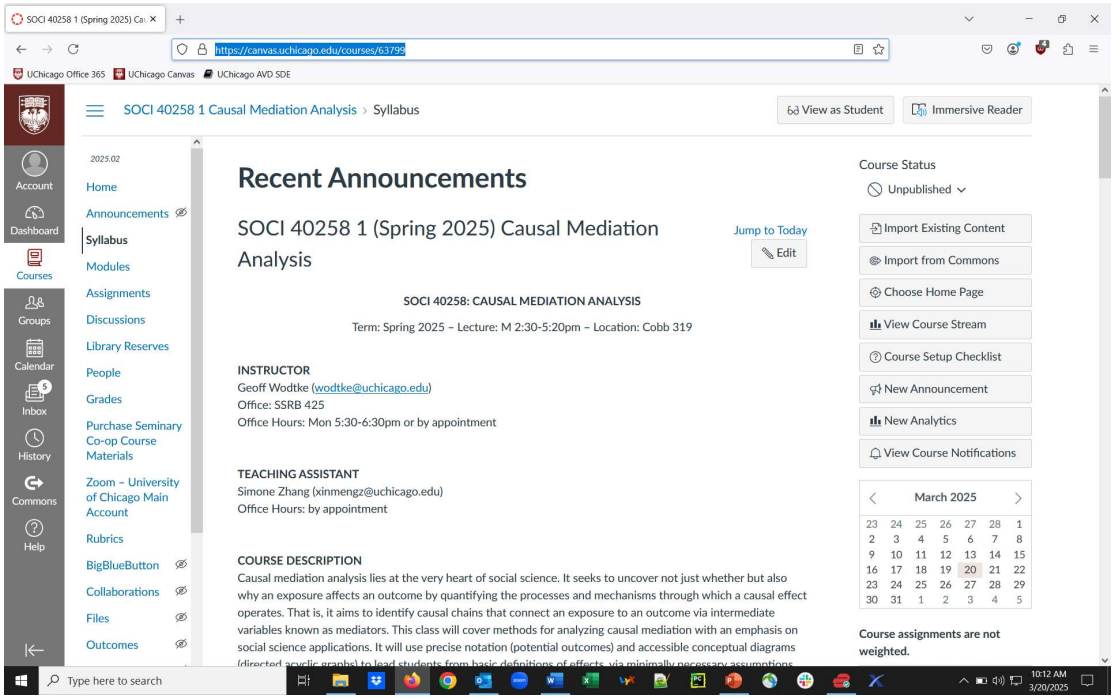
- The project should include:
  - a brief description of the background and motivation for the study
  - a well-defined causal question or methodological problem
  - a description of the data and methods
  - a summary of key findings
  - a full set of replication files

Grades

- Homework assignments: 60 percent
- Final research project: 40 percent

Canvas

- https://canvas.uchicago.edu/courses/63799



Questions?

Outline

- Measures of association versus causation
- Observed versus potential outcomes
- The fundamental problem of causal inference
- Solutions for the fundamental problem
- Identification and estimation of total effects

Measures of association

- Measures of association are determined by the joint probability distribution of observed random variables

- For example, the conditional probability of a response variable, Y_i, given a binary explanatory variable, D_i, $P(Y_i = y|D_i = d) = P(Y_i = y, D_i = d)/P(D_i = d)$, is a measure of association
  - It tells you how the probability that $Y_i = y$ differs when $D_i = 0$ rather than $D_i = 1$

- Similarly, the conditional expectation of a response variable, Y_i, given a binary explanatory variable, D_i, $E(Y_i|D_i = d) = \sum_y yP(Y_i = y|D_i = d)$, is also a measure of association
  - It tells you how the expected value of Y_i differs when $D_i = 0$ rather than $D_i = 1$

Measures of causation

- Measures of association do not provide a conceptual framework for defining measures of causation

- Causation involves "counterfactual" statements like…
  - "if this were to occur, then that would occur as a result"
  - "if this had occurred, then that would have occurred"
  - "if I were to do this, then that would happen"

- These types of counterfactual statements cannot be translated into measures of association, which involve only conditional statements like…
  - "if this occurs, then that also occurs"
  - "if this occurs, then that is more likely to occur"

Potential outcomes

- Potential outcomes provide one method for defining measures of causation

- A potential outcome, here denoted by $Y_i(d)$, is the value of the response variable for individual i that would occur if the explanatory variable, or treatment, for individual i were equal to d
  - With a binary treatment, for example…
    - $Y_i(1)$ is the value that would occur if individual i were to receive treatment
    - $Y_i(0)$ is the value that would occur if individual i were not to receive treatment

- Within this framework, each individual, or unit, i is conceived to have a set of potential outcomes corresponding to all possible levels of treatment

Individual causal effects

- The individual causal effect of treatment, then, is defined as the difference between potential outcomes for individual i:

$$ITE_i = Y_i(1) - Y_i(0)$$

- In other words, the ITE_i is the difference in the response variable that would occur if treatment were equal to 1 rather than 0 for individual i

- Note that the effect of treatment may differ across individuals—that is, ITE_i may not be the same for all i

Observed versus potential outcomes

- Each individual is conceived to have a set of potential outcomes corresponding to every different level of treatment, but we only ever observe a single element of this set

- Specifically, we only observe the one potential outcome that corresponds to the treatment actually received, while all the other potential outcomes are unobserved, or counterfactual

- For example, with a binary treatment, D_i, the observed outcome, Y_i, is formally related to the potential outcomes, $Y_i(d)$, via the following expression:

$$Y_i = D_i Y_i(1) + (1 - D_i)Y_i(0)$$

The fundamental problem

- We never observe the potential outcomes under treatment, $Y_i(1)$, for individuals who do not in fact receive treatment, that is, for whom $D_i = 0$
- And, we also never observe the potential outcomes under control, $Y_i(0)$, for individuals who do in fact receive treatment, that is, for whom $D_i = 1$
- As result, it is impossible to observe an individual causal effect, $Y_i(1) - Y_i(0)$, because we are always missing a key piece of data
- This dilemma is known as "the fundamental problem of causal inference"

The fundamental problem

- The "fundamental problem of causal inference" implies that causal effects cannot be easily *identified* from observed data

- *Identification* refers to whether a causal effect or set of causal effects of interest could be computed with data from an infinite sample
  - Identification is distinct from estimation, which refers to using data from a finite sample to compute an estimate of a causal effect that is subject to sampling error

- Causal analyses, therefore, should always proceed in two steps:
  - First, determine if the effect or effects of interest are identified from the observed data
  - Second, if identified, implement an appropriate estimator

The fundamental problem

- How do we "solve" the fundamental problem of causal inference? In other words, how do we identify causal effects from observed data?

- The fundamental problem is only "solved" by making unverifiable assumptions that allow one to substitute observed data in place of unobservable counterfactuals

- Broadly, there are two different "solutions" to the problem
  - The scientific solution
  - The statistical solution

The scientific solution

- Assume that different individuals, or units, are homogenous, except for the treatment received
  - That is, assume that $Y_i(d) = Y_j(d)$ for all individuals, or units, $i \neq j$, where $D_i = 1$ and $D_j = 0$

- Under this assumption, the $ITE_i = Y_i(1) - Y_i(0)$ can be computed as $Y_i - Y_j$ because…
  - $Y_i(1) = Y_i$ is observed
  - $Y_j(0) = Y_j$ is observed and can be substituted for $Y_i(0)$ because $Y_i(d) = Y_j(d)$

- This assumption is commonly employed in the physical and natural sciences
  - For example, when two different molecules in a vacuum are assumed to be homogenous, except for some treatment that was differentially applied to each

The scientific solution continued

- Alternatively, assume that the same individual, or unit, is identical at two different points in time, except for the treatment received
  - That is, assume that $Y_{it}(d) = Y_{it'}(d)$ for $t \neq t'$, where $D_{it} = 0$ and $D_{it'} = 1$

- Under this assumption, the $ITE_{it} = Y_{it}(1) - Y_{it}(0)$ at any single time t can be computed as $Y_{it'} - Y_{it}$ because…
  - $Y_{it}(0) = Y_{it}$ is observed
  - $Y_{it'}(1) = Y_{it'}$ is observed and can be substituted for $Y_{it}(0)$ because $Y_{it}(d) = Y_{it'}(d)$

- This assumption is also commonly employed in the physical sciences
  - For example, when the same molecule is assumed to be identical at two different points in time, except that a treatment was applied to it following an initial observation period

The scientific solution continued

- Are the assumptions involved with the “scientific solution” likely to ever be reasonable in the social sciences?

The statistical solution

- First, abandon the goal of computing individual causal effects, which requires assumptions that are unreasonable in the social sciences, and instead focus on computing *average* causal effects

- The average total effect of treatment is defined as

$$ATE = E(ITE_i) = E\big(Y_i(1) - Y_i(0)\big) = E\big(Y_i(1)\big) - E\big(Y_i(0)\big)$$

- The ATE is the difference in the mean of the outcome if all individuals in a population were exposed, rather than unexposed, to treatment

The statistical solution continued

- Second, assume that the value of $Y_i(1)$ will be the same no matter what mechanism is used to assign treatment and no matter what treatments are assigned to other individuals
  - This is sometimes called the stable unit treatment value assumption (SUTVA)
  - It requires that the potential outcomes of individuals are unaffected by changes in the treatments received by other individuals (i.e., no interference between units)
  - It also requires that there not be multiple versions of treatment

- Formally, suppose that $\mathbf{D}$ is a vector of treatment indicators for N individuals and that $Y_i(1; \mathbf{D})$ denotes a set of potential outcomes under treatment, which is now indexed by the vector of treatment variables
  - With this notation, SUTVA means that $Y_i(1; \mathbf{D}) = Y_i(1)$, or in other words, each individual's potential outcome under treatment does not depend on the configuration of treatments received by others

The statistical solution continued

- Third, assume that the potential outcomes, $\{Y_i(0), Y_i(1)\}$, are jointly independent of an individual's observed treatment status, D_i

- Formally, this independence assumption can be expressed as,

$$\{Y_i(0), Y_i(1)\} \perp D_i$$

where $\perp$ denotes statistical independence

- Recall that statistical independence between two random variables, X_i and Z_i, implies that

$$P(X_i, Z_i) = P(X_i)P(Z_i) \text{ and by extension, that } E(X_i|Z_i) = E(X_i)$$

The statistical solution continued

- Under the SUTVA and independence assumption, it's possible to substitute observed conditional means with fundamentally unobservable means of the potential outcomes

- Recall that we want to estimate $ATE = E\big(Y_i(1)\big) - E\big(Y_i(0)\big)$, but we only ever observe $E(Y_i(1)|D_i = 1)$ and $E(Y_i(0)|D_i = 0)$, which need not be equal to $E\big(Y_i(1)\big)$ and $E\big(Y_i(0)\big)$, respectively

- If $\{Y_i(0), Y_i(1)\} \perp D_i$, however, then…
  - $E\big(Y_i(1)\big) = E(Y_i(1)|D_i = 1) = E(Y_i|D_i = 1)$;
  - $E\big(Y_i(0)\big) = E(Y_i(0)|D_i = 0) = E(Y_i|D_i = 0)$; and by extension
  - $ATE = E\big(Y_i(1)\big) - E\big(Y_i(0)\big) = E(Y_i(1)|D_i = 1) - E(Y_i(0)|A_i = 0) = E(Y_i|D_i = 1) - E(Y_i|D_i = 0)$

The statistical solution continued

- In this table, the ATE is equal to

  $P(D_i = 0)(C - A) + P(D_i = 1)(F - B)$,

  but B and C are not observed

- Under the independence assumption, however, $A = B$ and $F = C$, and thus the ATE is equal to

  $$P(D_i = 0)(F - A) + P(D_i = 1)(F - A)$$
  $$= \big(P(D_i = 0) + P(D_i = 1)\big)(F - A)$$
  $$= F - A,$$

  which only involves the observed data

Observed and unobserved potential outcomes

Observed Treatment	n	Mean of the Potential Outcomes	
		$E(Y_i(0)\|D_i)$	$E(Y_i(1)\|D_i)$
$D_i=0$	20	A	C
$D_i=1$	80	B	F

Red font denotes unobserved pieces of information

The statistical solution continued

- In sum, the "statistical solution" to the fundamental problem of causal inference is to…
  - first, focus on estimating *average* causal effects
  - second, assume there is no interference between units or multiple versions of treatment
  - third, assume the potential outcomes are independent of treatment

Bias and inconsistency

- What happens when we use the difference between the observed conditional means, $E(Y_i|D_i = 1) - E(Y_i|D_i = 0)$, to compute the ATE, $E\big(Y_i(1)\big) - E\big(Y_i(0)\big)$, but the potential outcomes are not independent of observed treatment?
  - Bias and inconsistency

- It is extremely important to recognize that $E(Y_i|D_i = 1) = E(Y_i(1)|D_i = 1)$ need not equal $E\big(Y_i(1)\big)$ and $E(Y_i|D_i = 1) = E(Y_i(0)|D_i = 0)$ need not equal $E\big(Y_i(0)\big)$

- In fact, these equalities will not hold anytime…
  - the observed treatment groups differ on factors that affect the outcome
  - the effect of treatment differs across the observed treatment groups

Bias and inconsistency continued

- In general,

$$E(Y_i(1)|D_i = 1) - E(Y_i(0)|D_i = 0)$$

$$= ATE +$$

$$\big(E(Y_i(0)|D_i = 1) - E(Y_i(0)|D_i = 0)\big) +$$

$$\Big(\big(E(Y_i(1)|D_i = 1) - E(Y_i(0)|D_i = 1)\big) - \big(E(Y_i(1)|D_i = 0) - E(Y_i(0)|D_i = 0)\big)\Big) P(D_i = 0),$$

where the terms in colored font denote the asymptotic bias of $E(Y_i(1)|D_i = 1) - E(Y_i(0)|D_i = 0)$ for the ATE when the independence assumption is violated

Baseline bias

- The first term in the bias expression,

$$\big(E(Y_i(0)|D_i = 1) - E(Y_i(0)|D_i = 0)\big),$$

  is sometimes called the "baseline bias"

- The "baseline bias" is equal to the mean difference in the potential outcomes in the absence of treatment between the observed treated and untreated groups

- It will be nonzero when the observed treatment groups differ in their composition on factors that affect the outcome

Heterogeneity bias

- The second term in the bias expression

$$\Big(\big(E(Y_i(1)|D_i=1)-E(Y_i(0)|D_i=1)\big)-\big(E(Y_i(1)|D_i=0)-E(Y_i(0)|D_i=0)\big)\Big)P(D_i=0),$$

  is sometimes called "heterogeneity bias"

- The "heterogeneity bias" is equal to the difference in the average treatment effect between those in the observed treatment group and those in the observed control group

- It will be nonzero when the effect of treatment differs across the observed treatment groups because they differ in their composition on factors that make them more or less sensitive to treatment

An example

- What is the value of the ATE?
- What is the observed mean difference in the outcome between treatment and control groups?
- What is the baseline bias?
- What is the effect heterogeneity bias?

Observed and unobserved potential outcomes

Observed Treatment	n	Mean of the Potential Outcomes	
		$E(Y_i(0)\|D_i)$	$E(Y_i(1)\|D_i)$
$D_i=0$	20	10	20
$D_i=1$	80	5	10

Red font denotes unobserved pieces of information

Conditional average effects

- So far, we have been focusing on the ATE, but sometimes, this effect is not of immediate scientific interest, or other effects are of more immediate interest

- Two other *conditional* effects that are sometimes of scientific interest:
  - Average treatment effect on the treated, $ATT = E(Y_i(1) - Y_i(0)|D_i = 1)$
    - The ATT is the effect of treatment among those who actually received treatment
  - Average treatment effect on the untreated, $ATC = E(Y_i(1) - Y_i(0)|D_i = 0)$
    - The ATC is the effect of treatment among those who did not in fact receive treatment

- Not only are these effects often of direct scientific interest, they can also be identified from observed data under weaker independence assumptions

Identifying the ATT

- The ATT is equal to $(F - B)$, but B is not observed
- Under the assumption that $Y_i(0) \perp D_i$, however, $B = A$, and thus the ATT is equal to $(F - A)$, both of which are observed
- The assumption that $Y_i(0) \perp D_i$ is weaker than the joint independence assumption required to identify the ATE
- Note also that, by focusing on the ATT, any effect heterogeneity bias is avoided

Observed and unobserved potential outcomes

Observed Treatment	n	Mean of the Potential Outcomes	
		$E(Y_i(0)\|D_i)$	$E(Y_i(1)\|D_i)$
$D_i=0$	20	A	C
$D_i=1$	80	B	F

Red font denotes unobserved pieces of information

Identifying the ATC

- The ATC is equal to $(C - A)$, but C is not observed
- Under the assumption that $Y_i(1) \perp D_i$, however, $F = C$, and thus the ATC is equal to $(F - A)$, which, as noted previously, are both observed
- As before, the assumption that $Y_i(1) \perp D_i$ is weaker than the joint independence assumption required to identify the ATE
- And, by focusing on the ATT, any effect heterogeneity bias is also avoided

Observed and unobserved potential outcomes

Observed Treatment	n	Mean of the Potential Outcomes	
		$E(Y_i(0)\|D_i)$	$E(Y_i(1)\|D_i)$
$D_i=0$	20	A	C
$D_i=1$	80	B	F

Red font denotes unobserved pieces of information

Conditional independence

- The independence assumption, $\{Y_i(0), Y_i(1)\} \perp D_i$, is a strong condition that is not likely to be met in most social science applications, except for randomized experiments

- Fortunately, average causal effects can also be identified under a weaker *conditional* independence assumption, which can be expressed as

$$\{Y_i(0), Y_i(1)\} \perp D_i | C_i,$$

where C_i is a set of observed covariates

- Substantively, this assumption states that, within the strata defined by C_i, the potential outcomes are jointly independent of treatment

Conditional independence

- Under the conditional independence assumption, it's possible to substitute observed conditional means with unobservable *conditional* means of the potential outcomes

- That is, if $\{Y_i(0), Y_i(1)\} \perp D_i | C_i$, then…
  - $E(Y_i(1)|C_i) = E(Y_i(1)|D_i = 1, C_i) = E(Y_i|D_i = 1, C_i)$;
  - $E(Y_i(0)|C_i) = E(Y_i(0)|D_i = 0, C_i) = E(Y_i|D_i = 0, C_i)$; and by extension
  - $E(Y_i(1) - Y_i(0)|C_i) = E(Y_i(1)|D_i = 1, C_i) - E(Y_i(0)|D_i = 0, C_i) = E(Y_i|D_i = 1, C_i) - E(Y_i|D_i = 0, \mathrm{C}_i)$

Conditional independence

- Then, to compute the ATE, $E\big(Y_i(1) - Y_i(0)\big)$, we need only average the conditional effects of treatment, given C_i, over the distribution of C_i

- Specifically…

$$
\begin{aligned}
ATE &= E\big(Y_i(1) - Y_i(0)\big) \\
&= E\big(E(Y_i(1) - Y_i(0)|C_i)\big) \\
&= E\big(E(Y_i(1)|D_i = 1, C_i) - E(Y_i(0)|D_i = 0, C_i)\big) \\
&= E\big(E(Y_i|D_i = 1, C_i) - E(Y_i|D_i = 0, C_i)\big) \\
&= \textstyle\sum_c \big(E(Y_i|D_i = 1, C_i = c) - E(Y_i|D_i = 0, C_i = c)\big)P(C_i = c)
\end{aligned}
$$

Conditional independence

- Under the conditional independence assumption, the expression

$$ATE = \sum_c \big(E(Y_i|D_i = 1, C_i = c) - E(Y_i|D_i = 0, C_i = c)\big)P(C_i = c)$$

  is called the *nonparametric identification formula* for the ATE

- In words, this expression implies that, when the conditional independence assumption is satisfied, the ATE is equal to…
  - the sum of the mean differences between treated and control groups in each stratum defined by C_i…
  - …with each stratum-specific difference weighted by the relative size of the stratum

- Put differently, the ATE is equal to the mean difference in the outcome between treatment and control groups when they are standardized by the marginal distribution of C_i

Estimation and inference

- To nonparametrically estimate the ATE under the marginal independence assumption, when $ATE = E(Y_i|D_i = 1) - E(Y_i|D_i = 0)$, all we need is an unbiased estimator for $E(Y_i|D_i = 1) - E(Y_i|D_i = 0)$
  - For example, the sample mean difference between treatment and control groups in a randomized experiment on a representative sample of the target population

- To nonparametrically estimate the ATE under the conditional independence assumption, when $ATE = \sum_c \big(E(Y_i|D_i = 1, C_i = c) - E(Y_i|D_i = 0, C_i = c)\big)P(C_i = c)$, we need unbiased estimators for $E(Y_i|D_i = 1, C_i = c)$, $E(Y_i|D_i = 0, C_i = c)$ and $P(C_i = c)$
  - For example, the sample mean difference between treatment and control groups for each stratum defined by C_i and then the sample proportion in each stratum

Parametric estimation

- In practice, due to data limitations, it is often impossible or impractical to estimate causal effects nonparametrically, and thus we may need to rely on parametric models to estimate certain quantities in the identification formula

- In this situation, valid causal inferences are premised not only the accuracy of the identification assumptions outlined previously but also on the accuracy of the parametric constraints imposed by the models

- Thus, we should try to avoid strong parametric assumptions when estimating causal effects to whatever degree is possible

Regression estimation

- Use a parametric model to estimate $E(Y_i|D_i := 1, C_i)$ and $E(Y_i|D_i := 0, C_i)$, and then average these estimates over the empirical distribution of C_i

- Implementation:
  - Fit a parametric model for $E(Y_i|D_i, C_i)$, denoted by $E(Y_i|D_i, C_i) = h(D_i, C_i; \gamma)$
  - Estimate $\hat{E}(Y_i|D_i := 1, C_i)$ as $h(D_i := 1, C_i; \hat{\gamma})$
  - Estimate $\hat{E}(Y_i|D_i := 0, C_i)$ as $h(D_i := 0, C_i; \hat{\gamma})$
  - Compute $\widehat{ATE} = \frac{1}{n}\sum_i \big(h(D_i := 1, C_i; \hat{\gamma}) - h(D_i := 0, C_i; \hat{\gamma})\big)$

Regression estimation

- A special and perhaps familiar case:
  - When the ATE is nonparametrically identified, and $h(D_i, C_i; \gamma) = \gamma_0 + \gamma_1 D_i + \gamma_2 C_i$ is linear and additive, then the coefficient on D_i captures the effect of interest

- Proof:

$$\begin{aligned} ATE &= \sum_c \big(E(Y_i|D_i := 1, C_i = c) - E(Y_i|D_i := 0, C_i = c)\big) P(C_i = c) \\ &= \sum_c \big((\gamma_0 + \gamma_1 + \gamma_2 c) - (\gamma_0 + \gamma_2 c)\big) P(C_i = c) \\ &= \gamma_1 \sum_c P(C_i = c) \\ &= \gamma_1 \end{aligned}$$

Inverse probability weighting (IPW)

- In the nonparametric identification formula, $\sum_c E(Y_i|D_i = d, C_i = c)P(C_i = c)$ can be rewritten as follows:

$$\begin{aligned}
&\sum_c E(Y_i|D_i = d, C_i = c)P(C_i = c) \\
&= \sum_c \frac{1}{P(D_i=d|C_i=c)} E(Y_i|D_i = d, C_i = c)P(D_i = d|C_i = c)P(C_i = c) \\
&= E\left(\frac{I(D_i=d)Y_i}{P(D_i=d|C_i)}\right),
\end{aligned}$$

where $E\left(\frac{I(D_i=d)Y_i}{P(D_i=d|C_i)}\right)$ is a weighted mean of the outcome in treatment group d with weights equal to $\frac{1}{P(D_i=d|C_i)}$, the inverse probability that $D_i = d$ given C_i

Inverse probability weighting (IPW)

- Use a parametric model to estimate $P(D_i = 1|C_i)$ and $P(D_i = 0|C_i)$, and then use these estimates to construct weighted means of the outcome

- Implementation:
  - Fit a parametric model for $P(D_i = 1|C_i)$, denoted by $P(D_i = 1|C_i) = f(C_i; \alpha)$
  - Estimate $\hat{P}(D_i = 1|C_i)$ as $f(C_i; \hat{\alpha})$
  - Estimate $\hat{P}(D_i = 0|C_i)$ as $1 - f(C_i; \hat{\alpha})$
  - Compute $\widehat{ATE} = \left[\frac{1}{n_1}\sum_{i:D_i=1}\frac{Y_i}{\hat{P}(D_i=1|C_i)}\right] - \left[\frac{1}{n_0}\sum_{i:D_i=0}\frac{Y_i}{\hat{P}(D_i=0|C_i)}\right]$

Inverse probability weighting (IPW)

- Weighting the treated subsample by $\frac{1}{P(D_i=1|C_i)}$ and the control subsample by $\frac{1}{P(D_i=0|C_i)}$ transforms the distribution of C_i in both groups so that it resembles the marginal distribution of C_i

- In other words, IPW is just another form of standardization, like regression estimation

- In fact, these two approaches are asymptotically equivalent, assuming correct models for $P(D_i = 1|C_i)$ and $E(Y_i|D_i, C_i)$

Example: NLSY79

- Does attending college reduce depression later in adulthood?
- 1979 National Longitudinal Study of Youth
  - Exposure (D)
    - sample member attended college before age 22
  - Outcome (Y):
    - standardized scores on the CES-D at age 40
  - Covariates (C):
    - Race, gender, parental education, occupation, income, household size, AFQT scores

Example: NLSY79

- Regression estimation for *ATE* of college attendance on depression

```
# load/install libraries #
packages<-c("dplyr", "tidyr", "foreign", "margins")
#install.packages(packages)

for (package.i in packages) {
    suppressPackageStartupMessages(library(package.i, character.only=TRUE))
    }

# load data #
datadir <- "C:/Users/Geoffrey Wodtke/Dropbox/D/courses/2023-24_UOFCHICAGO/SOCI_40258
nlsy <- read.dta(paste(datadir, "nlsy79.dta", sep=""))

nlsy <- nlsy[complete.cases(nlsy[, c("cesd_age40", "ever_unemp_age3539", "att22",
    "female", "black", "hispan", "paredu", "parprof", "parinc_prank", "famsize", "af

nlsy$std_cesd_age40 <- (nlsy$cesd_age40-mean(nlsy$cesd_age40))/sd(nlsy$cesd_age40)

# linear, additive regression estimates #
Ymodel <- lm(std_cesd_age40 ~ att22 + female + black + hispan + paredu + parprof +
    parinc_prank + famsize + afqt3, data=nlsy)

summary(Ymodel)
```

```
Call:
lm(formula = std_cesd_age40 ~ att22 + female + black + hispan +
    paredu + parprof + parinc_prank + famsize + afqt3, data = nlsy)

Residuals:
    Min      1Q  Median      3Q     Max
-1.2704 -0.6747 -0.3005  0.3726  4.3075

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)   0.3095917  0.0961828   3.219  0.00130 **
att22        -0.0792489  0.0437107  -1.813  0.06991 .
female        0.2005493  0.0332191   6.037 1.73e-09 ***
black        -0.0638799  0.0468273  -1.364  0.17261
hispan       -0.1354808  0.0505413  -2.681  0.00738 **
paredu       -0.0024354  0.0069234  -0.352  0.72503
parprof       0.0698722  0.0624627   1.119  0.26338
parinc_prank -0.0005091  0.0007458  -0.683  0.49486
famsize      -0.0111977  0.0086396  -1.296  0.19503
afqt3        -0.0057056  0.0007983  -7.147 1.08e-12 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.98 on 3487 degrees of freedom
Multiple R-squared:  0.04212,   Adjusted R-squared:  0.03965
F-statistic: 17.04 on 9 and 3487 DF,  p-value: < 2.2e-16
```

Example: NLSY79

- Regression estimation for *ATE* of college attendance on depression

```
# linear, additive regression estimates #
Ymodel <- lm(std_cesd_age40 ~ att22 + female + black + hispan + paredu + parprof +
    parinc_prank + famsize + afqt3, data=nlsy)

summary(Ymodel)

# non-linear, non-additive regression estimates w/ margins #
Ymodel.x <- lm(std_cesd_age40 ~ att22*(female + black + hispan + paredu + parprof +
    parinc_prank*parinc_prank + famsize + afqt3), data=nlsy)

ATEhat <- margins(Ymodel.x, variables="att22", dydx="att22")

summary(ATEhat)

# non-linear, non-additive regression estimates 'by hand' #
gdata <- nlsy
gdata$att22 <- 1

yhat1 <- predict(Ymodel.x, newdata=gdata)

gdata$att22 <- 0
yhat0 <- predict(Ymodel.x, newdata=gdata)

mean(yhat1-yhat0)
```

```
> # non-linear, non-additive regression estimates w/ margins #
> Ymodel.x <- lm(std_cesd_age40 ~ att22*(female + black + hispan
>
> ATEhat <- margins(Ymodel.x, variables="att22", dydx="att22")
>
> summary(ATEhat)
 factor     AME     SE       z      p   lower   upper
  att22 -0.1287 0.0495 -2.6027 0.0093 -0.2257 -0.0318
>
> # non-linear, non-additive regression estimates 'by hand' #
> gdata <- nlsy
> gdata$att22 <- 1
>
> yhat1 <- predict(Ymodel.x, newdata=gdata)
>
> gdata$att22 <- 0
> yhat0 <- predict(Ymodel.x, newdata=gdata)
>
> mean(yhat1-yhat0)
[1] -0.1287357
```

Example: NLSY79

- IPW estimation for ATE of college attendance on depression

```r
# ipw estimates #
Dmodel <- glm(att22 ~ female + black + hispan + paredu + parprof +
    parinc_prank + famsize + afqt3, data=nlsy, family=binomial(link="probit"))

phatD1 <- predict(Dmodel, type = "response")

nlsy$ipw <- (nlsy$att22/phatD1) + (1-nlsy$att22)/(1-phatD1)

design <- svydesign(ids=~1, data=nlsy, weights=~ipw)

Ymodel.wtd <- svyglm(std_cesd_age40 ~ att22, design=design)

summary(Ymodel.wtd)

```

```
Call:
svyglm(formula = std_cesd_age40 ~ att22, design = design)

Survey design:
svydesign(ids = ~1, data = nlsy, weights = ~ipw)

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  0.04135    0.02275   1.818  0.06916 .
att22       -0.19578    0.05593  -3.501  0.00047 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for gaussian family taken to be 0.9327147)

Number of Fisher Scoring iterations: 2
```

Looking forward…

- Causal mediation analysis extends analyses of average total effects
- It aims to uncover why a treatment affects an outcome by quantifying the causal processes, or mechanisms, through which the effect operates
  - In other words, it aims to identify causal chains that connect an exposure to an outcome via intermediate variables known as mediators