University of Glasgow | Institute of Health & Wellbeing

University of Glasgow

College of Medical, Veterinary & Life Sciences

GLASGOW MOLECULAR PATHOLOGY NODE

EPSRC Engineering and Physical Sciences Research Council

NHS Greater Glasgow and Clyde

MRC Medical Research Council

Advanced Epidemiology: Methods designed to accommodate unobserved confounding

Janet Bouttell, February 2020

Health Economics and Health Technology Assessment (HEHTA)

Intended learning outcomes:

- To understand that there are methods which attempt to deal with unmeasured confounding
- To be familiar with the names and general approaches of some of these methods
- To have an appreciation of the strengths and limitations of these methods

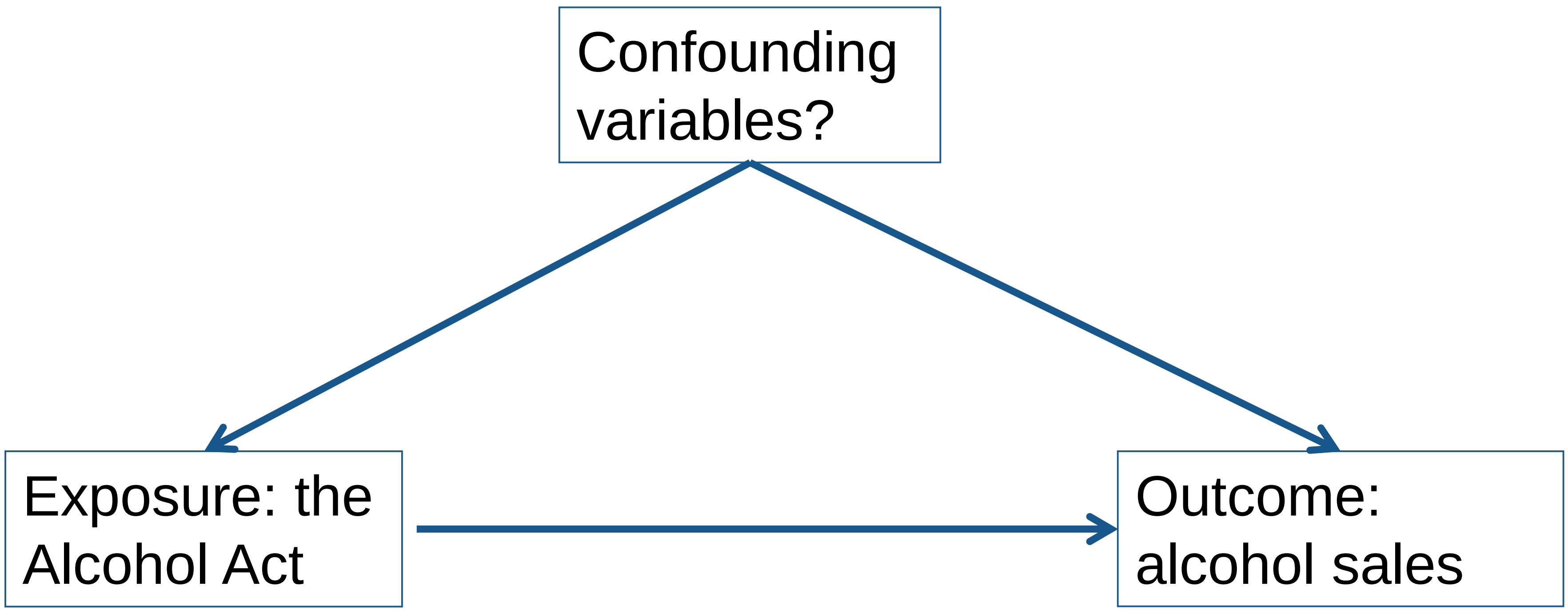
University of Glasgow | Institute of Health & Wellbeing

- **What is confounding? – (**variable) that influences both the dependent variable and independent variable causing a spurious association
- **What is observed confounding?**
- **What is unmeasured confounding?**
- **What is unmeasured time-varying confounding?**

University of Glasgow | Institute of Health & Wellbeing

Confounding - what's the problem?

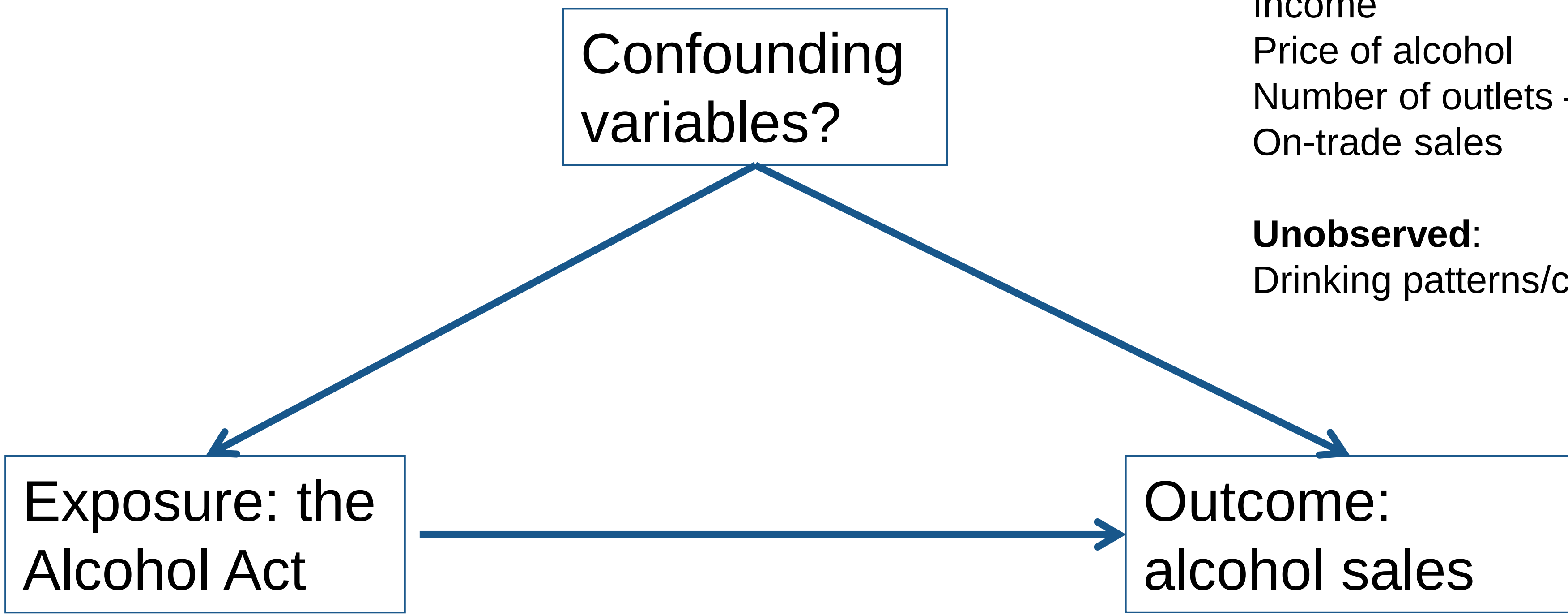
Hypothesis – the Alcohol Act (2011) reduced alcohol consumption in Scotland



Robinson et al (2014)

Confounding - what's the problem?

Hypothesis – the Alcohol Act (2011) reduced alcohol consumption in Scotland



Observed:
Income
Price of alcohol
Number of outlets – on and off
On-trade sales

Unobserved:
Drinking patterns/culture?

University of Glasgow | Institute of Health & Wellbeing

Methods designed to accommodate unobserved confounding

1. Difference-in-difference
2. Interrupted time series
3. Synthetic controls
4. Regression discontinuity
5. Instrumental Variables

University of Glasgow | Institute of Health & Wellbeing

Difference in difference

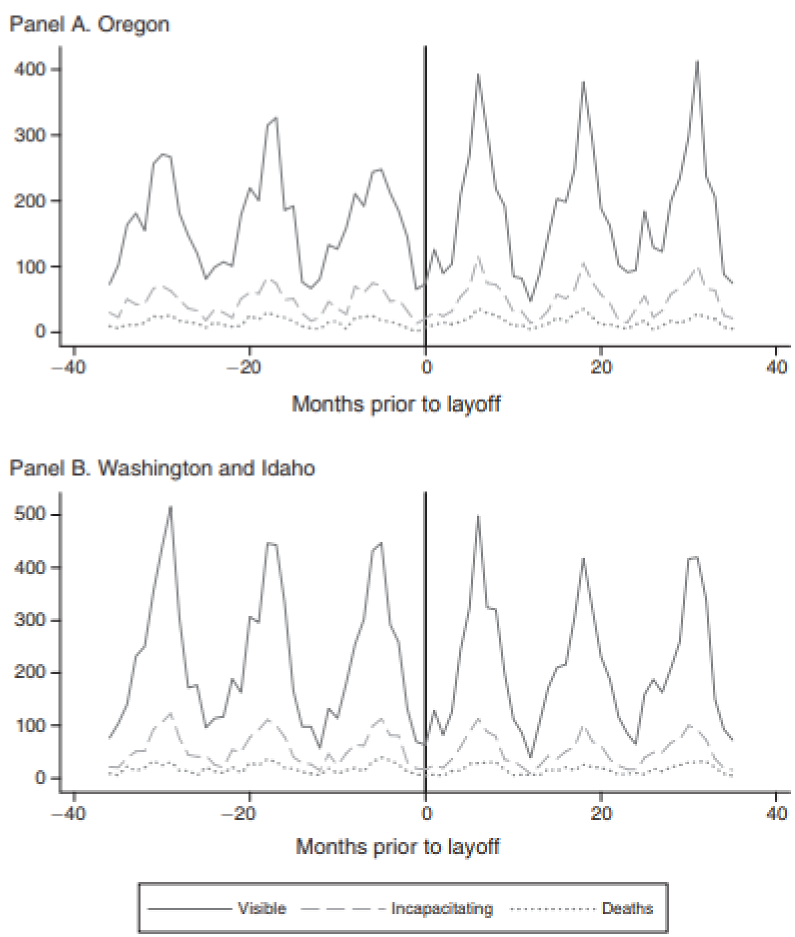


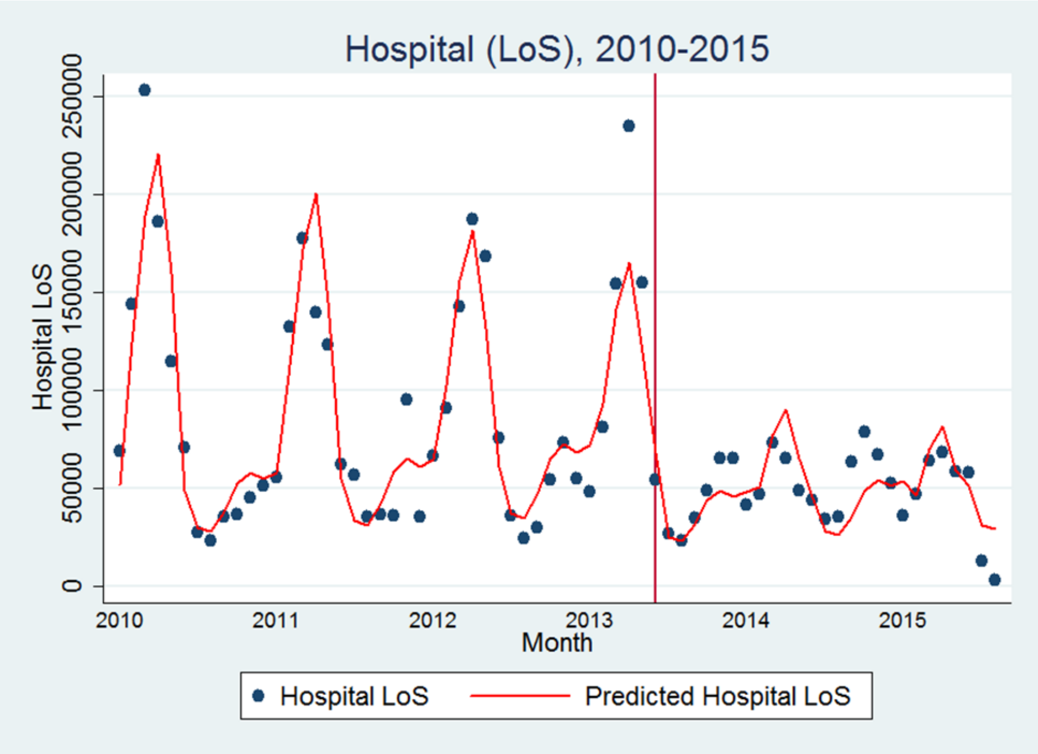
FIGURE 4. INJURIES ON HIGHWAYS IN OREGON, WASHINGTON, AND IDAHO OUTSIDE CITY LIMITS—DRY WEATHER CONDITIONS

- Change in outcome variable in intervention/control areas before and after the intervention
- Difference between those two differences
- E.g +50 Oregon -10 Washington and Idaho – DiD = 60 injuries
- Sometimes uses just one time point before and after
- How does it account for unobserved confounding?
- Limitations – parallel trends (similarity of control area)

DeAngelo and Hansen: Police enforcement and traffic fatalities (2014)

University of Glasgow | Institute of Health & Wellbeing

Interrupted time series



- Trend in outcome compared pre and post intervention
- What is the control/counterfactual here?
- Model accounts for auto-correlation, seasonality and underlying trend
- Can also add in a control – overlap with DiD
- Limitations – power/data, similarity of control area (if used)

Heggie et al, Hospital costs and rotavirus innoculation (2018)

Assess the impact of proposition 99

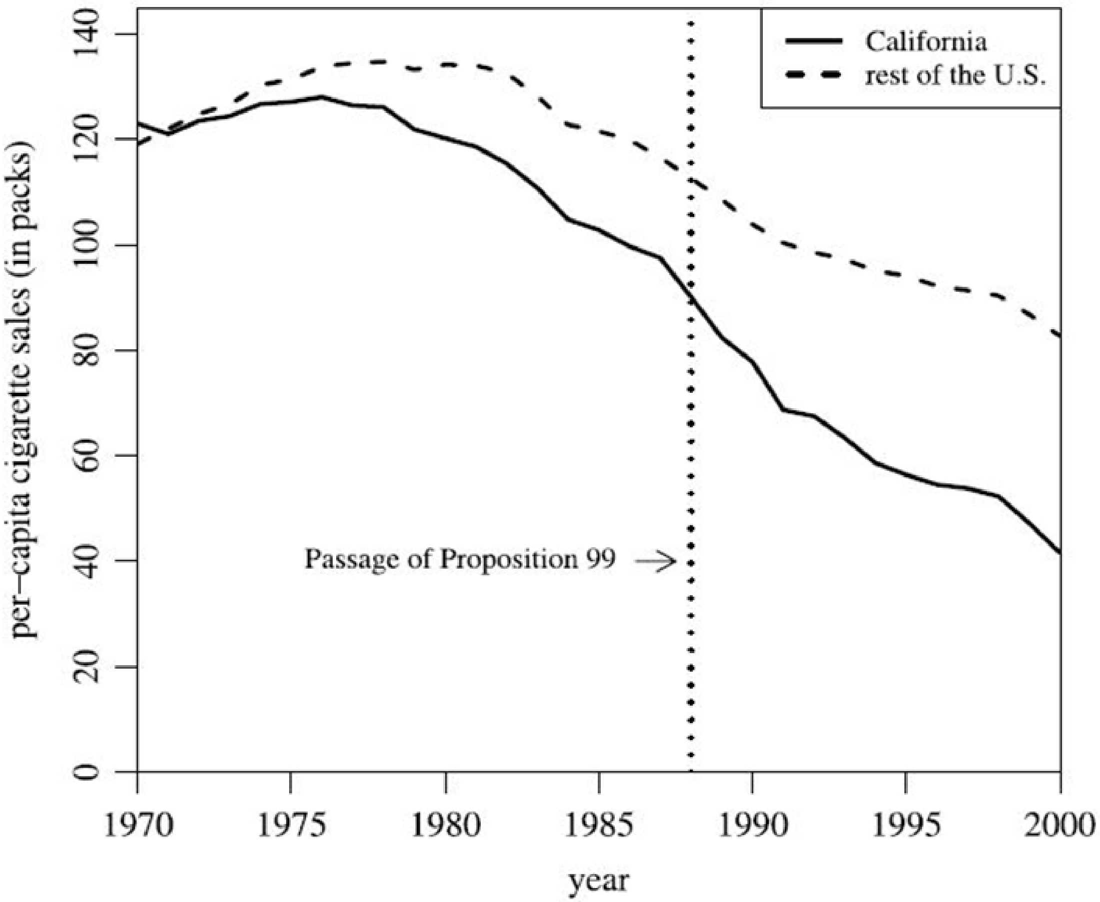


Figure 1. Trends in per-capita cigarette sales: California vs. the rest of the United States.

- How would you assess the impact of Proposition 99?
- Could you use an interrupted time series or a difference in difference approach?
- What would be the limitations of that?
- Could you overcome the problem?

Abadie et al (2010)

Synthetic controls

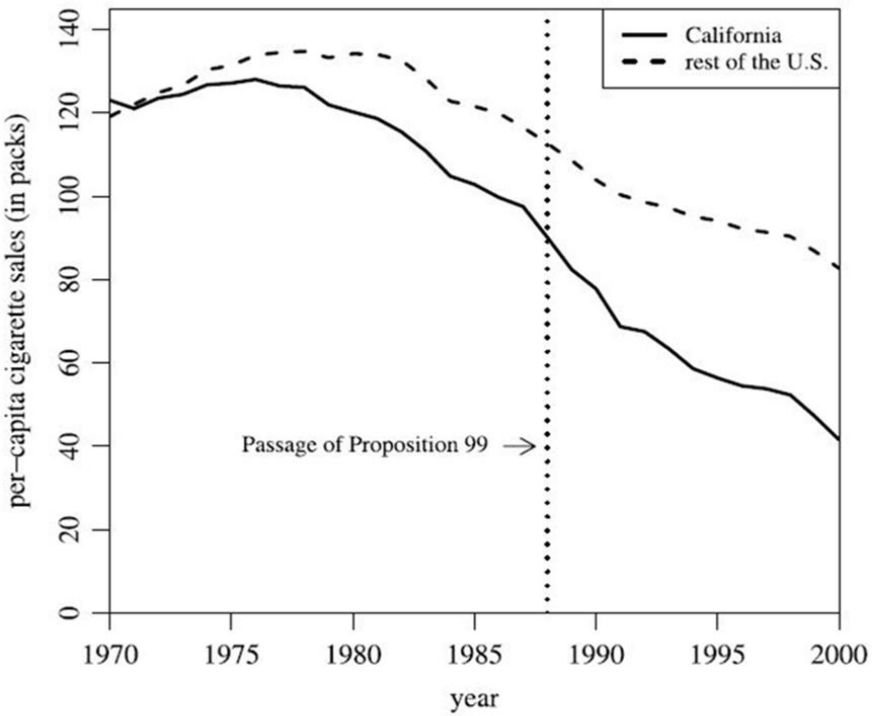


Figure 1. Trends in per-capita cigarette sales: California vs. the rest of the United States.

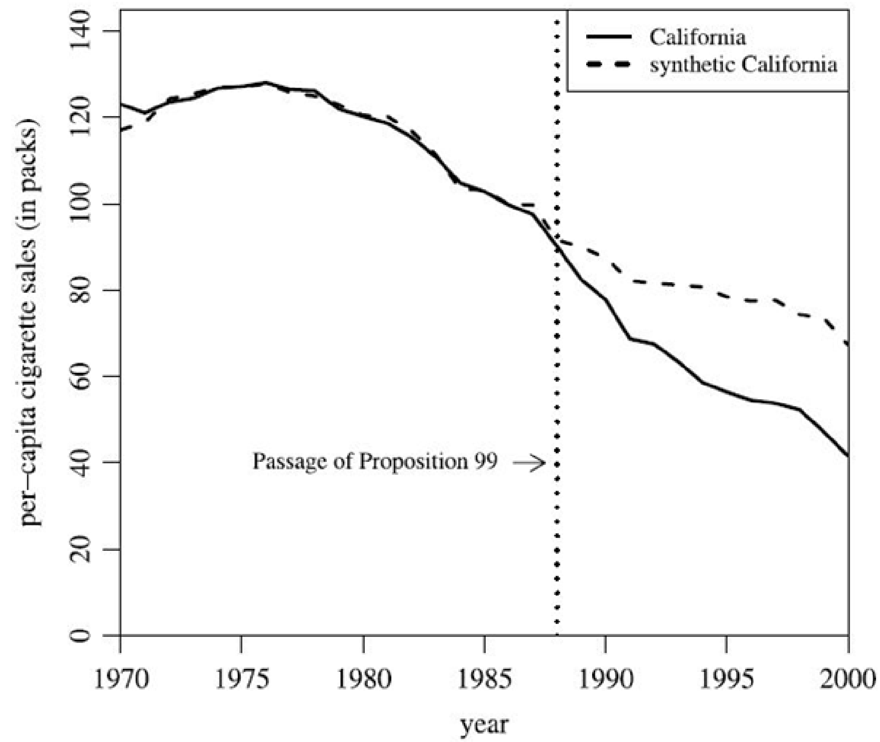


Figure 2. Trends in per-capita cigarette sales: California vs. synthetic California.

University of Glasgow | Institute of Health & Wellbeing

Synthetic controls

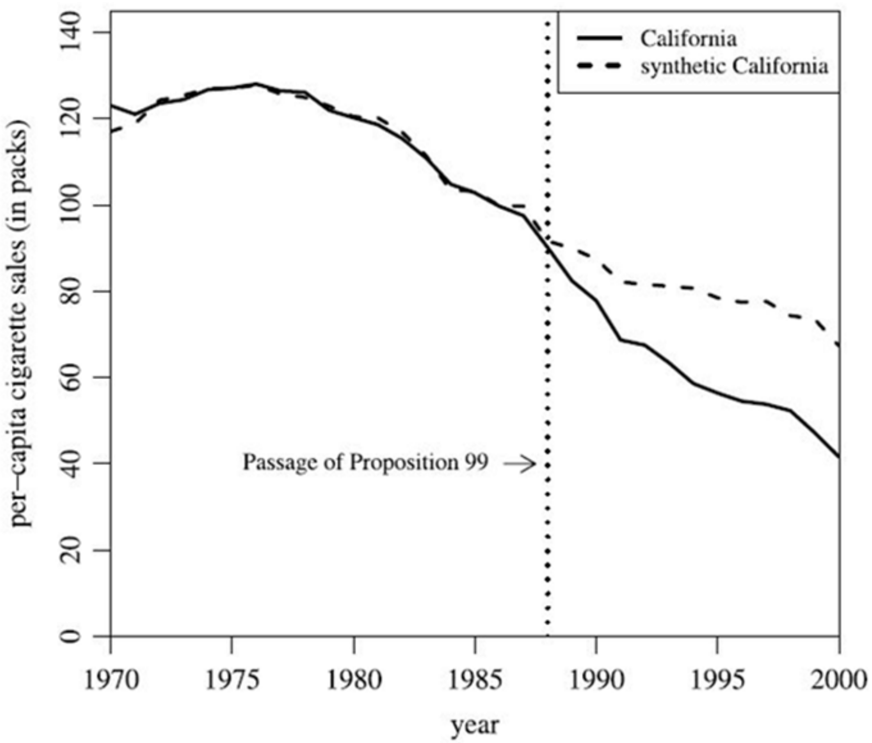
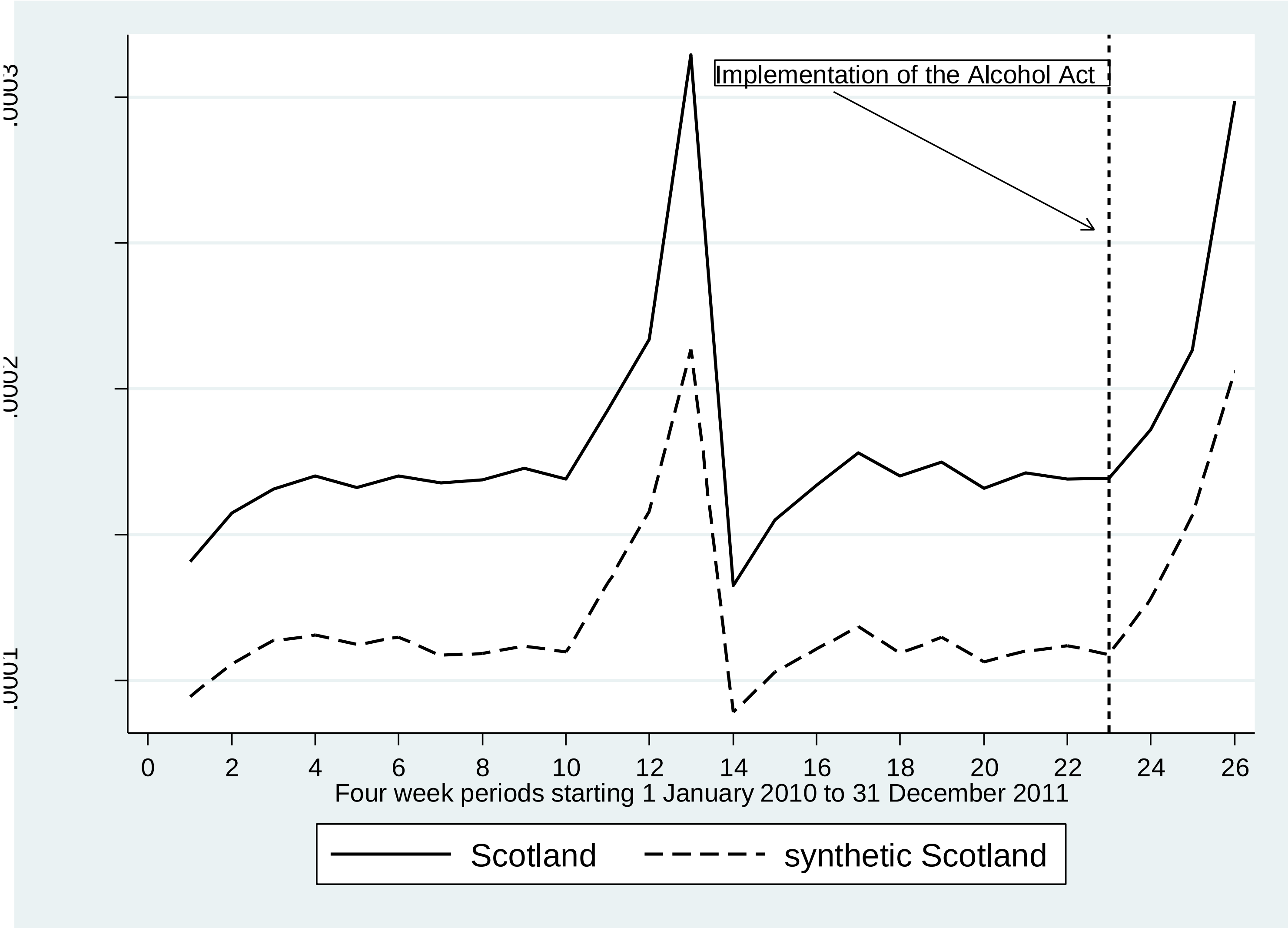


Figure 2. Trends in per-capita cigarette sales: California vs. synthetic California.

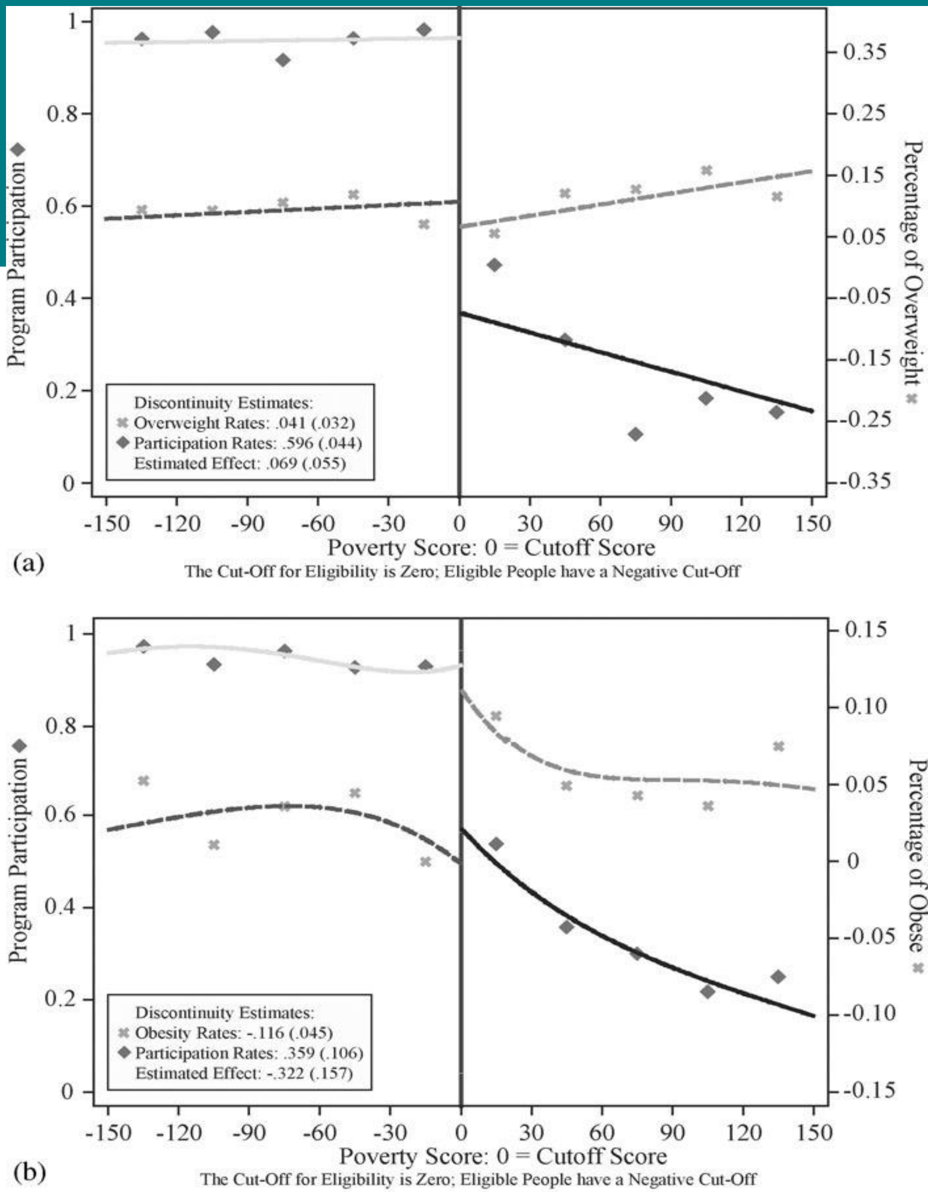
- Trend in the outcome in intervention compared to synthetic control area
- Synthetic control is weighted average from pool of potential controls
- How does it overcome unmeasured confounding?
- Why can this method cope with time-varying unmeasured confounding where DiD and ITS can’t?
- Limitations – data and outliers

Synthetic controls – outlier problem – sales of spirits in Scotland



Oportunidades to reduce overweight and obesity in Mexico?

Regression discontinuity



Average program participation (marked with ◆) and overweight/obesity rates (marked with ✖) are plotted as a function of five categories of the povery score on each side of the eligibility cut-off. The lines are conditional expectations of specification as in equations (3) and (4) in the text. They are estimated using the whole range of data with poverty scores of +/- 500. Based on Schwarz (1978), the preferred specifications for men include first order polynomial terms on both sides of the cut-off and fifth order for women.

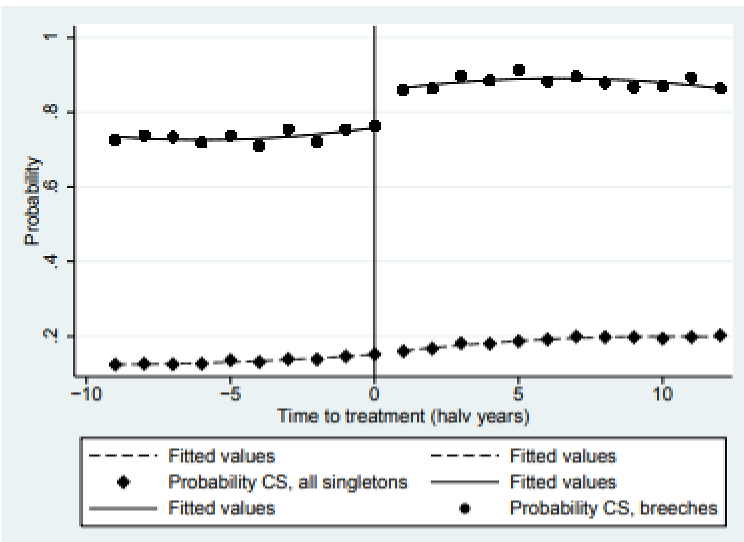
- Outcomes compared in units defined by scores above and below a cut-off in a continuous variable that determines exposure
- Units either side of the cut-off should be similar
- Trade-off between power (want numbers) and minimising confounding
- Limitations – only appropriate where you have a cut-off for entitlement

Andalon (2011)

Health Economics, Volume: 20, Issue: S1, Pages: 1-18, First published: 01 August 2011, DOI: (10.1002/hec.1773)

Regression discontinuity

Figure 1: CS rate for all non-breech and breech pregnancies, 1996-2006

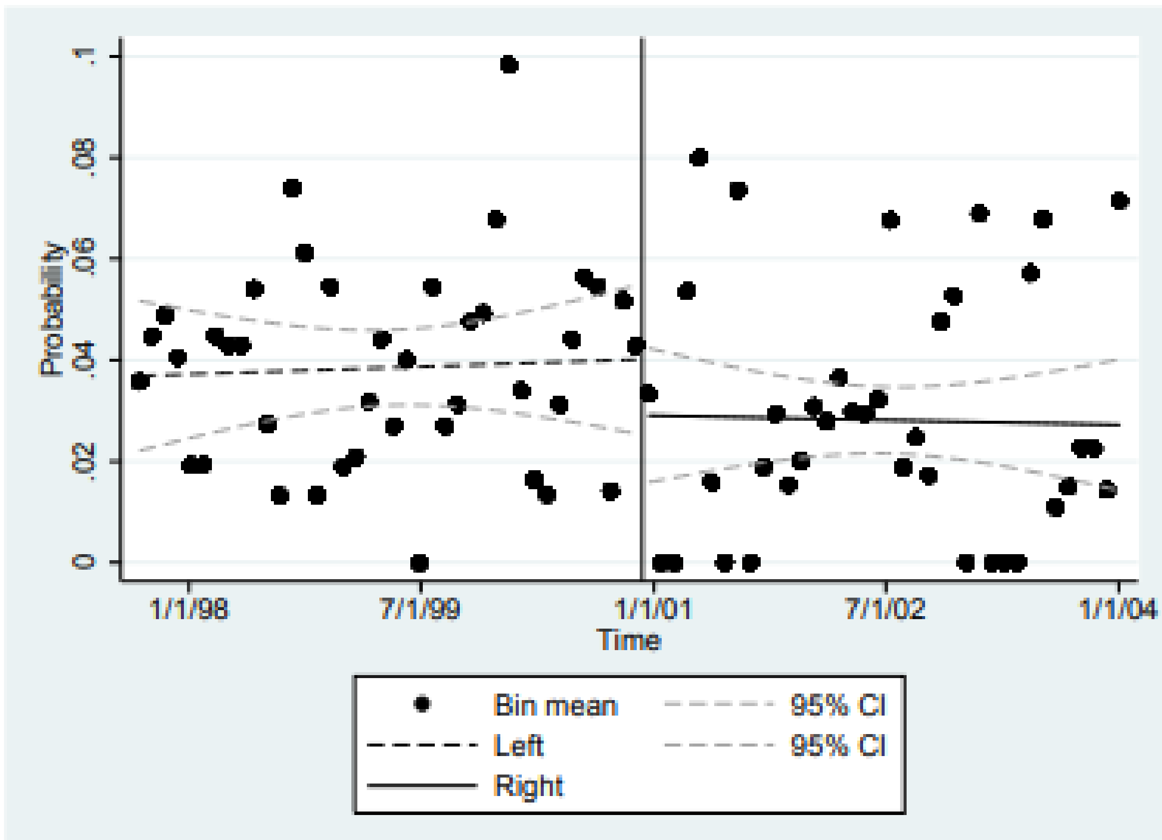


Notes: The plot shows the average probability of a CS per half-year. The vertical line is the date for the Danish dissemination of the TBT results. The sample includes all singleton births irrespective of parity.

- Second example – Jensen and Wust (2015)
- RCT in 2000 suggested that caesarean section safer for mothers and babies when baby breech.
- Major change in practice – very quickly
- Time discontinuity allowed study of women and babies on the margin

Regression discontinuity

Figure 8.7: Probability of APGAR score<=7 at 1 min for breech babies at term with parity>1



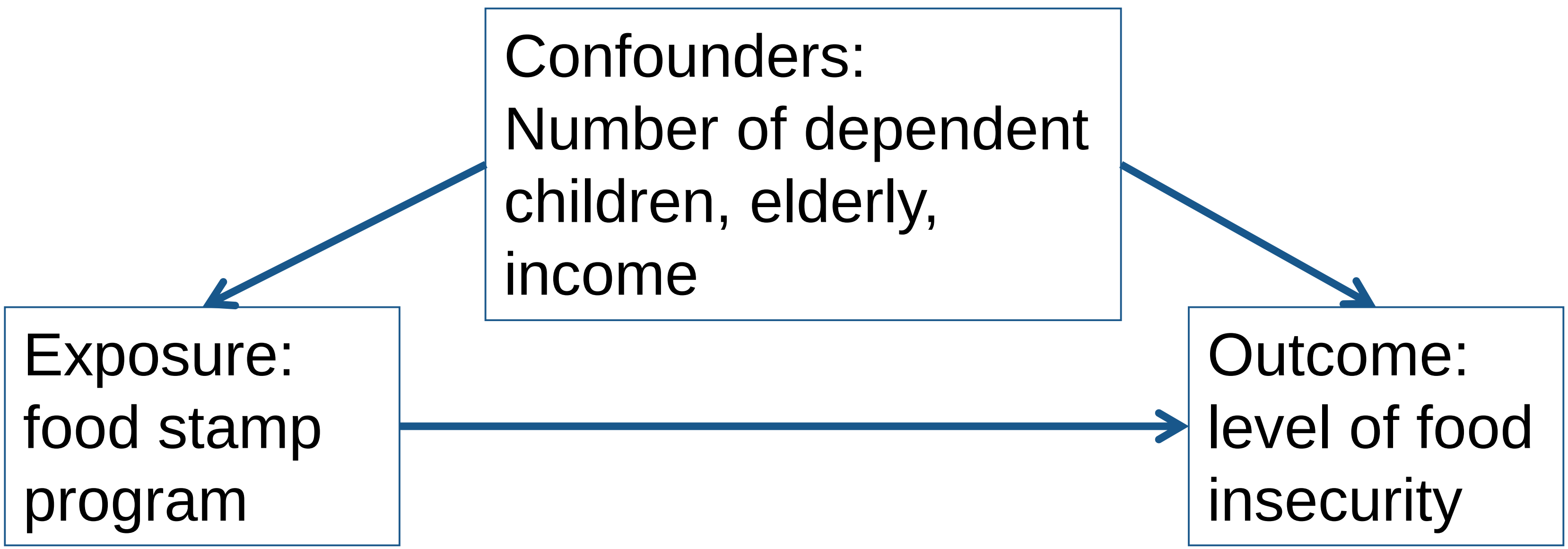
Hypothesis – food stamp program participation reduces food insecurity

Exposure: food stamp program $\longrightarrow$ Outcome: level of food insecurity

(Yen et al, 2008)

Instrumental Variables

Hypothesis – food stamp program participation reduces food insecurity

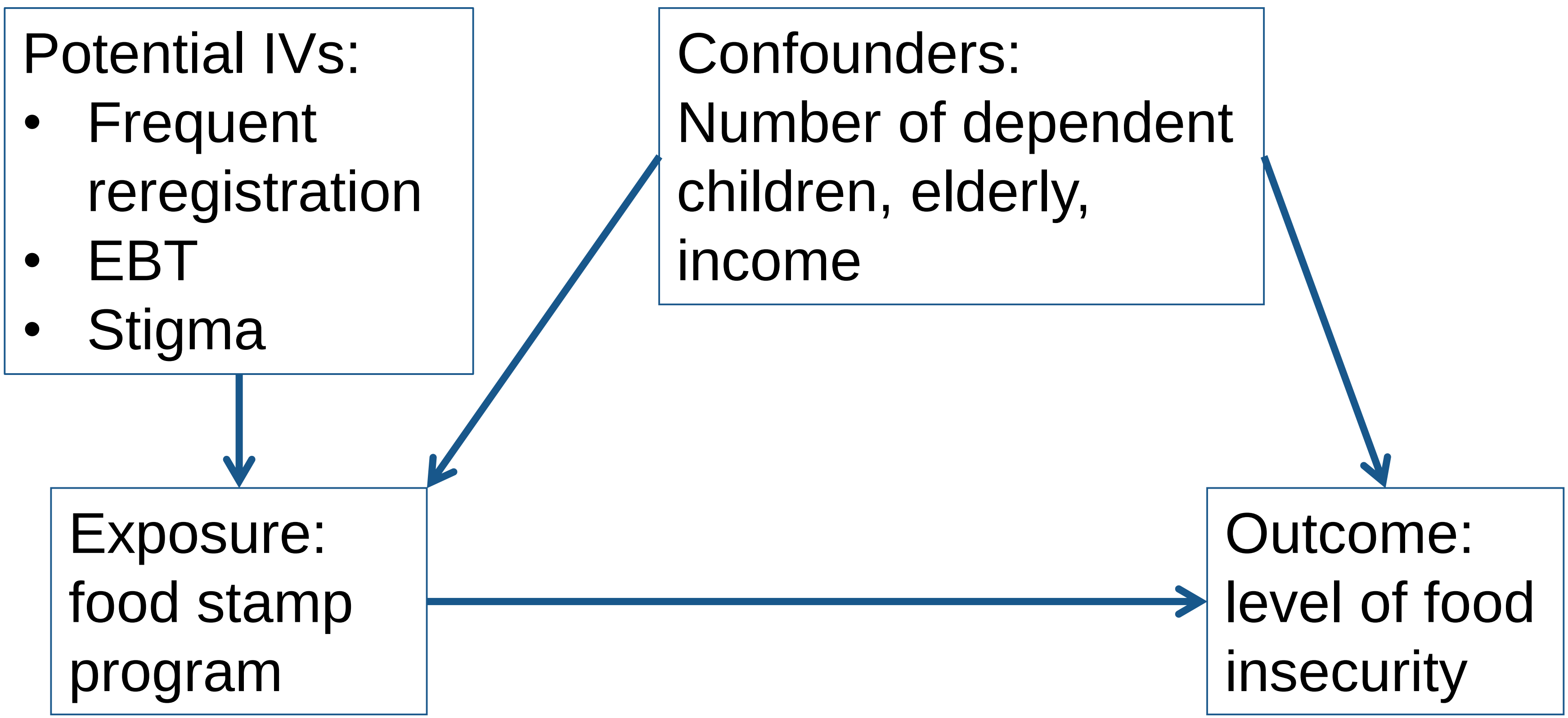


(Yen et al, 2008)

University of Glasgow | Institute of Health & Wellbeing

Instrumental Variables

Hypothesis – food stamp program participation reduces food insecurity (Yen et al, 2008)



Instrumental variables - summary

3 conditions

1. The IV must be correlated with the exposure – the stronger the better
2. The IV must not be associated with the outcome
3. The IV should not be associated with the confounding variable (which influences the outcome as well as the exposure)

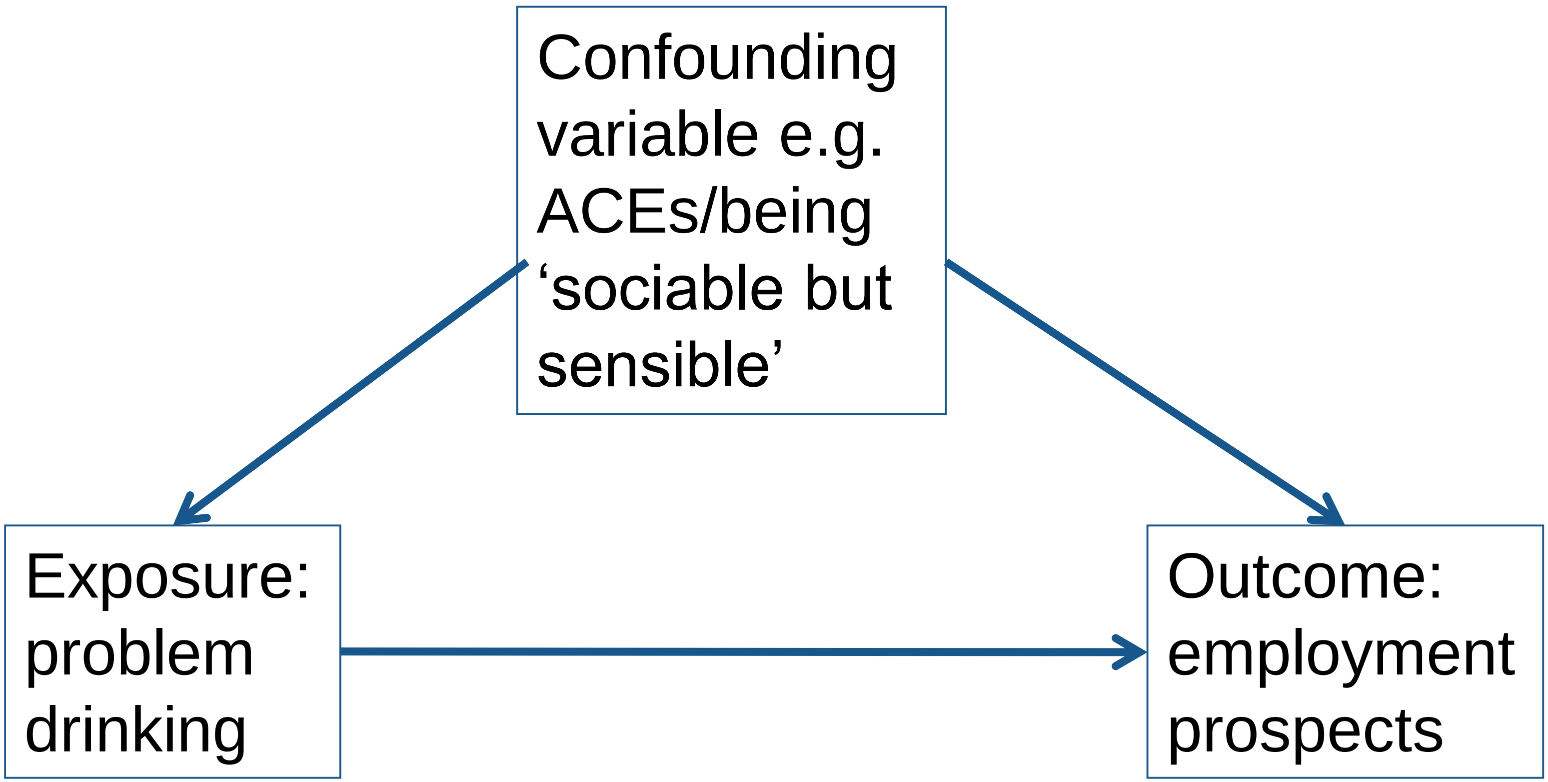
Limitations

1. Finding suitable instruments
2. Demonstrating that the assumptions are met

Hypothesis – problem drinking impacts employment prospects

Macdonald et al (2001)

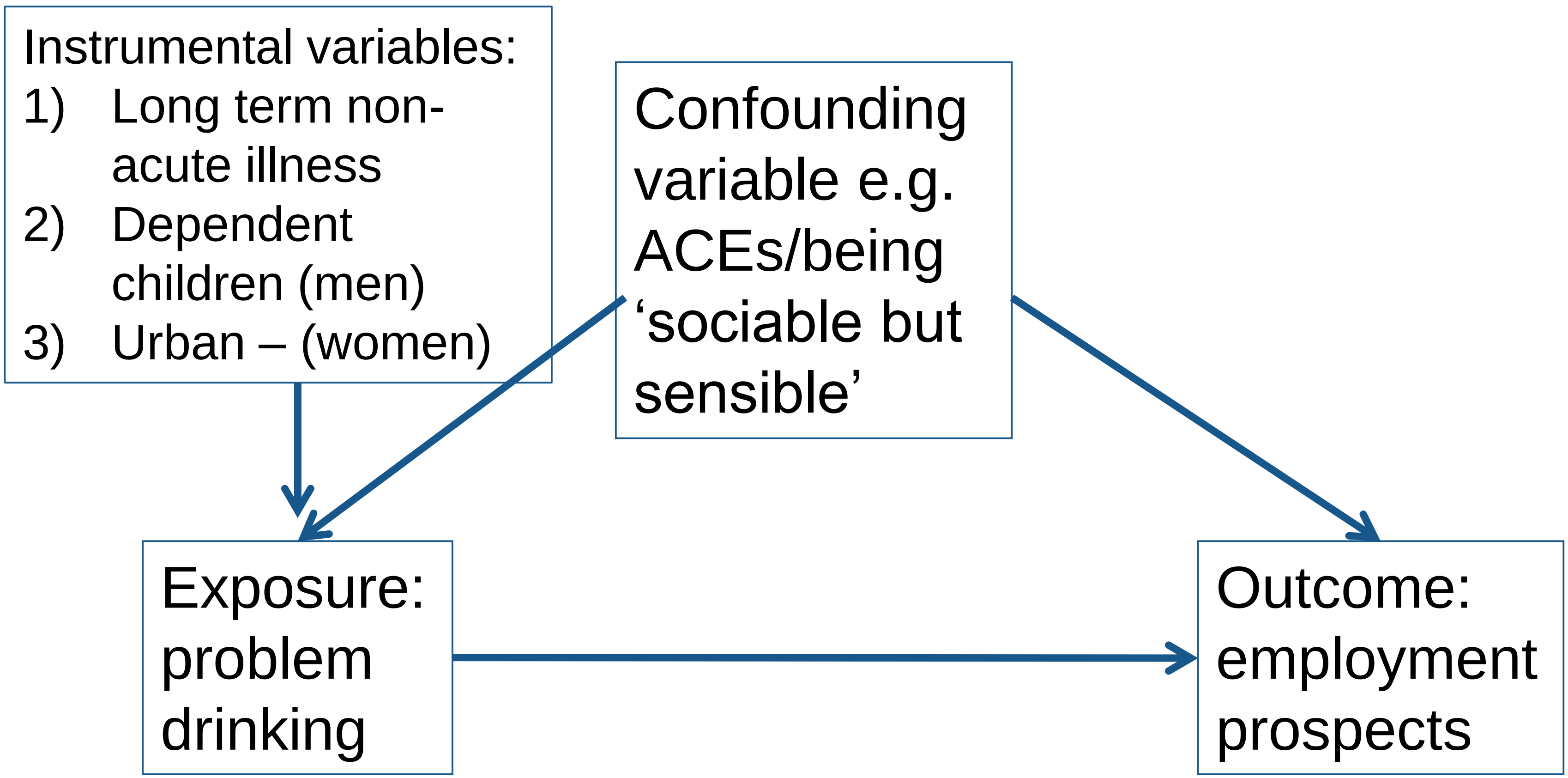
Hypothesis – problem drinking impacts employment prospects



Macdonald et al (2001)

Instrumental Variables

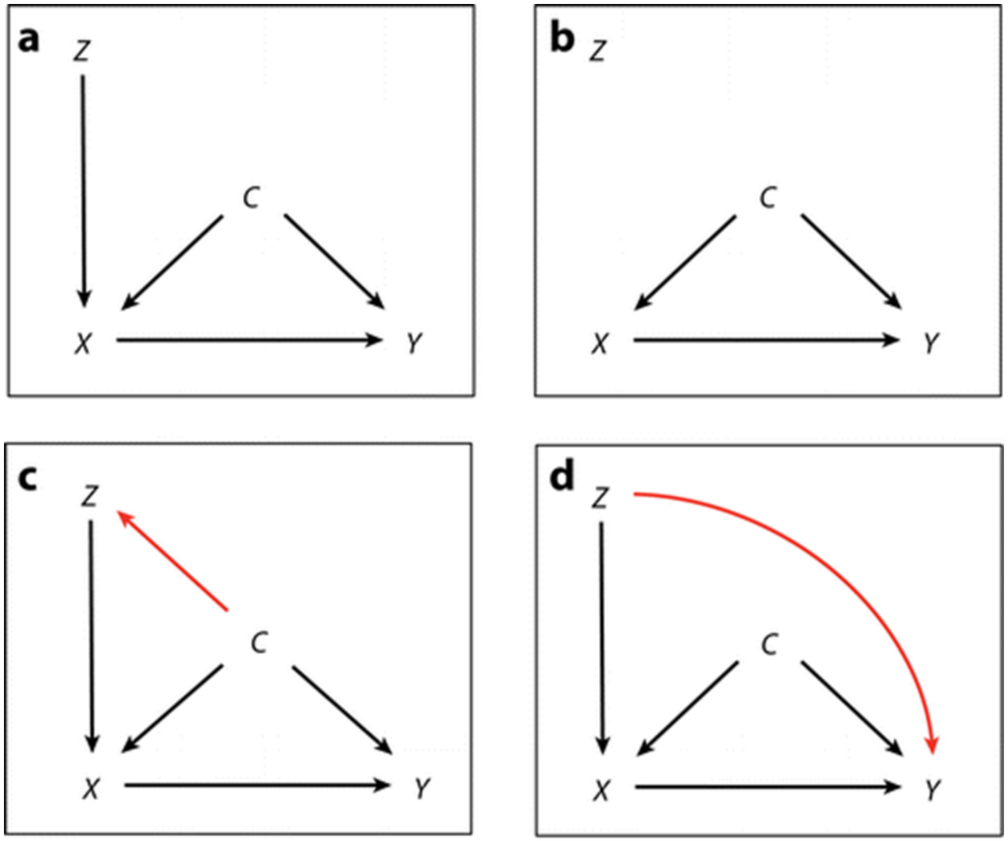
Hypothesis – problem drinking impacts employment prospects



Macdonald et al (2001)

University of Glasgow | Institute of Health & Wellbeing

Instrumental variables - quiz



X = exposure of interest
Y = outcome of interest
C = confounder
Z = instrumental variable

In which of a-d would Z be an effective instrumental variable? Why?

Craig P, et al. 2017.
Annu. Rev. Public Health. 38:39–56

Practical exercise

1. **Three case studies – brief detail given (please make and note any assumptions you need)**
2. **Complete the outcome and interventions row and the confounders row**
3. **Decide which methods would be possible in each scenario - see description of the methods (Table 2 in Craig et al, 2017) overleaf**
4. **Select your preferred method then identify your counterfactual, data required and limitations**

Mendellian randomisation: a form of instrumental variable

https://www.youtube.com/watch?v=LoTgfGotaQ4

University of Glasgow | Institute of Health & Wellbeing

References (1 of 2)

Overview article

Natural Experiments: An Overview of Methods, Approaches, and Contributions to Public Health Intervention Research

Peter Craig, Srinivasa Vittal Katikireddi, Alastair Leyland, and Frank Popham. Annual Review of Public Health 2017 38:1, 39-56 [https://www.annualreviews.org/doi/10.1146/annurev-publhealth-031816-044327]

Difference in difference

DeAngelo, G. and Hansen, B., 2014. Life and death in the fast lane: Police enforcement and traffic fatalities. *American Economic Journal: Economic Policy*, *6*(2), pp.231-57.

Interrupted time series

Bernal, J.L., Cummins, S. and Gasparrini, A., 2017. Interrupted time series regression for the evaluation of public health interventions: a tutorial. *International journal of epidemiology*, *46*(1), pp.348-355. [https://academic.oup.com/ije/article/46/1/348/2622842]

Heggie, R., Murdoch, H., McIntosh, E. , Cameron, C., Smith-Palmer, A. and Bouttell, J. (2018) Cost-impact study of rotavirus vaccination programme in Scotland. *Human Vaccines and Immunotherapeutics*,(doi:10.1080/21645515.2018.1543522)(PMID:30395774) (Early Online Publication)

Robinson M, Geue C, Lewsey J, Mackay D, McCartney G, Curnock E, Beeston C. Evaluating the impact of the alcohol act on off-trade alcohol sales: a natural experiment in Scotland. Addiction. 2014 Dec;109(12):2035-43.

References (2 of 2)

Synthetic controls

Abadie, A., Diamond, A. and Hainmueller, J., 2010. Synthetic control methods for comparative case studies: Estimating the effect of California’s tobacco control program. *Journal of the American statistical Association*, *105*(490), pp.493-505.

Bouttell, J., Craig, P. , Lewsey, J. , Robinson, M. and Popham, F. (2018)Synthetic control methodology as a tool for evaluating population-level health interventions. *Journal of Epidemiology and Community Health*, 72(8), pp. 673-678. (doi:10.1136/jech-2017-210106)(PMID:29653993) (PMCID:PMC6204967)

Regression discontinuity

Andalón M. Oportunidades to reduce overweight and obesity in Mexico?. Health economics. 2011 Sep;20(S1):1-8.

Moscoe E, Bor J, Barnighausen T. 2015. Regression discontinuity designs are under-used in medicine, epidemiology and public health: a review of current and best practice. *J. Clin. Epidemiol.* 68:132–43 [https://www-sciencedirect-com.ezproxy.lib.gla.ac.uk/science/article/pii/S0895435614003990]

Callaghan, R.C., Sanches, M., Gatley, J.M. and Stockwell, T., 2014. Impacts of drinking-age laws on mortality in Canada, 1980–2009. *Drug and alcohol dependence*, *138*, pp.137-145.

Jensen VM, Wüst M. Can Caesarean section improve child and maternal health? The case of breech babies. Journal of health economics. 2015 Jan 1;39:289-302.

Instrumental variables

Ichida, Y., Hirai, H., Kondo, K., Kawachi, I., Takeda, T. and Endo, H., 2013. Does social participation improve self-rated health in the older population? A quasi-experimental intervention study. *Social science & medicine*, *94*, pp.83-90.

MacDonald, Z. and Shields, M.A., 2001. The impact of alcohol consumption on occupational attainment in England. *Economica*, *68*(271), pp.427-453.

Yen ST, Andrews M, Chen Z, Eastwood DB. Food Stamp Program participation and food insecurity: an instrumental variables approach. American Journal of Agricultural Economics. 2008 Feb 1;90(1):117-32.