Conference or Workshop Item:
Martin, A orcid.org/0000-0002-2559-6483 and Bojke, C From muddied waters to causal clarity? The use and limitations of Instrumental Variables in health economics. In: Royal Statistical Society (Leeds/Bradford Local Group), 25 Apr 2017, Leeds, UK.

Leeds Institute of
Health Sciences

UNIVERSITY OF LEEDS

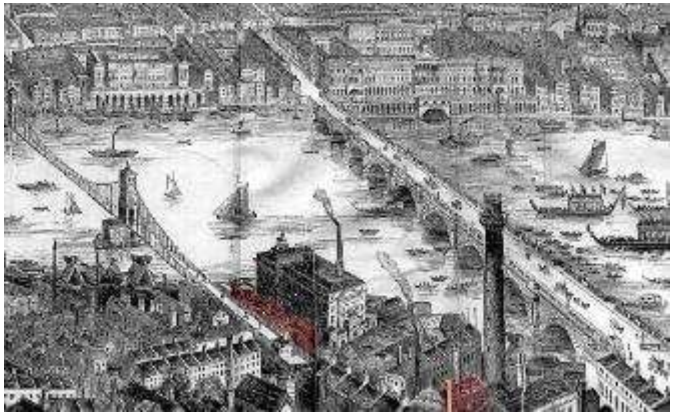
From muddied waters to causal clarity?

The use and limitations of instrumental variables in health economics

Dr Adam Martin

Professor Chris Bojke

Academic Unit of Health Economics



Outline

UNIVERSITY OF LEEDS

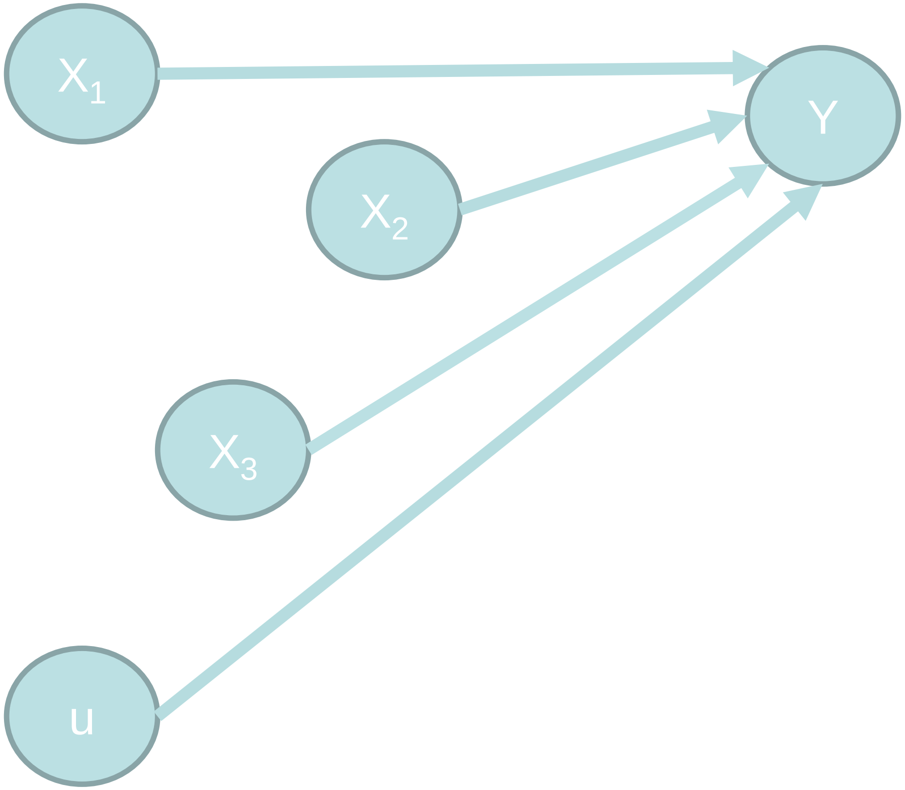
- Overview
  - Basic theory and key assumptions
- Brief history
- Review commonly used instruments in health economics
- What’s the future for IV in applied health econometrics?

Background

- Instrumental Variables (IV) is a method for estimating consistent parameters when assumptions of classical models are violated
- In modern day econometrics, the principal use of IV is to address the problem of omitted variables bias
  - "Causal inference has always been the name of the game in applied econometrics" (Angrist and Pischke, 2009)
  - Potentially powerful tool for analysing observational data when RCTs are not feasible, desirable, practical or ethical

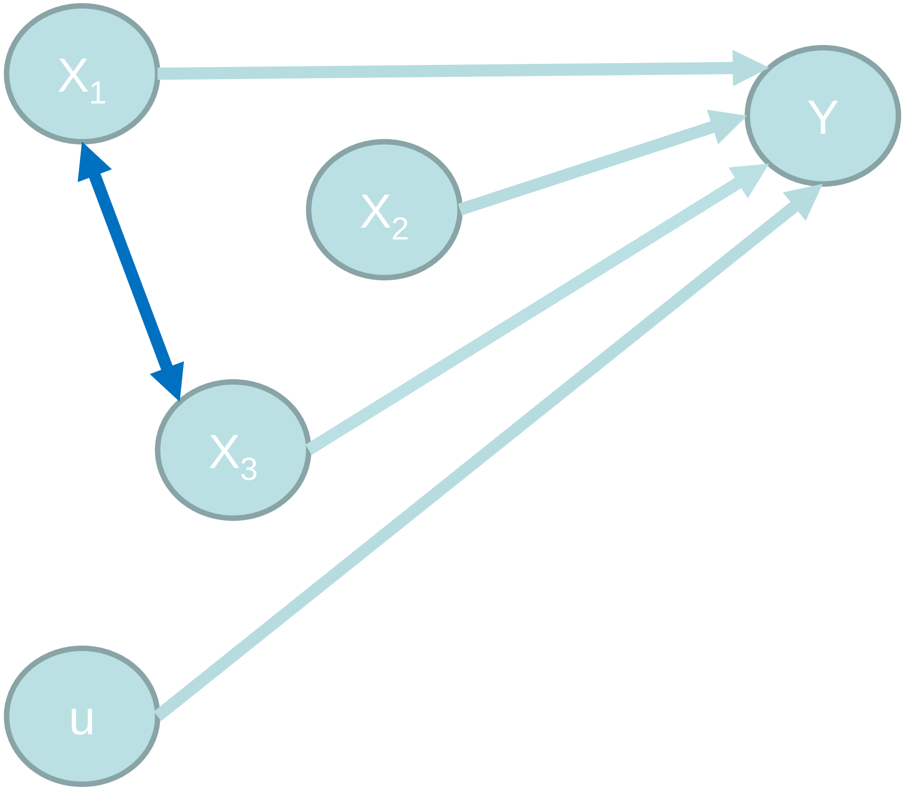
Standard regression

- No association between X_1 and u
- OLS estimates are consistent



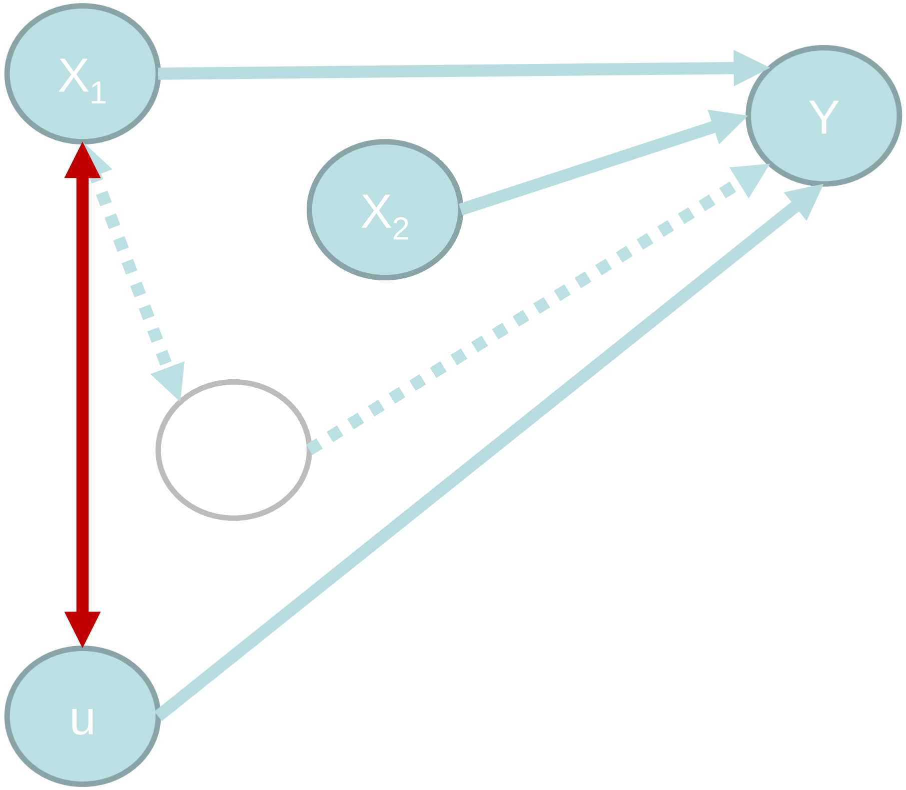
Standard regression

- No association between X_1 and u
- OLS estimates are consistent, even if association between X_1 and X_3



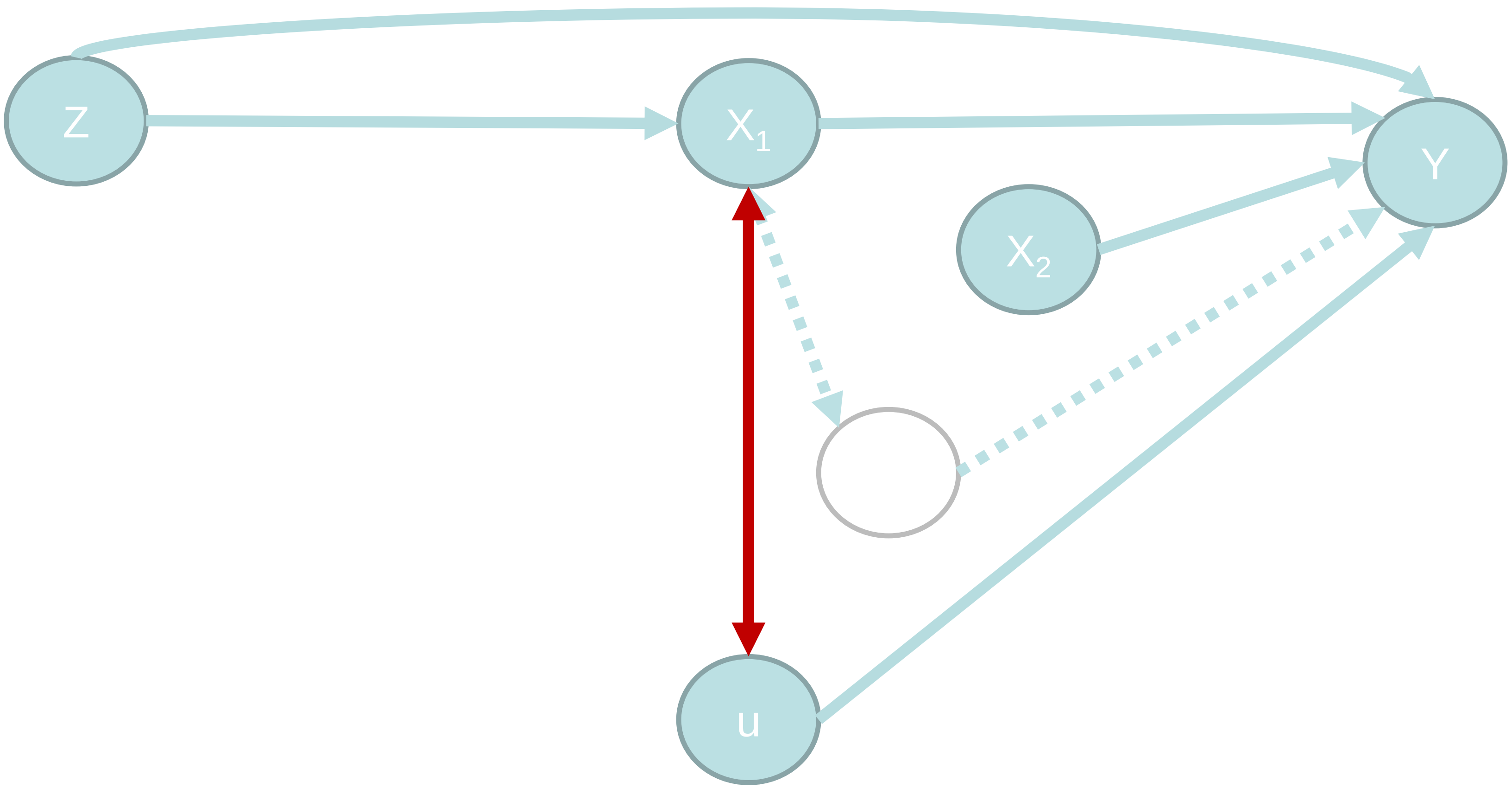
Endogeneity

- Association between X_1 and u if X_3 is unobserved
- Failure of the zero conditional mean assumption E[u|x] = 0
- OLS inconsistent



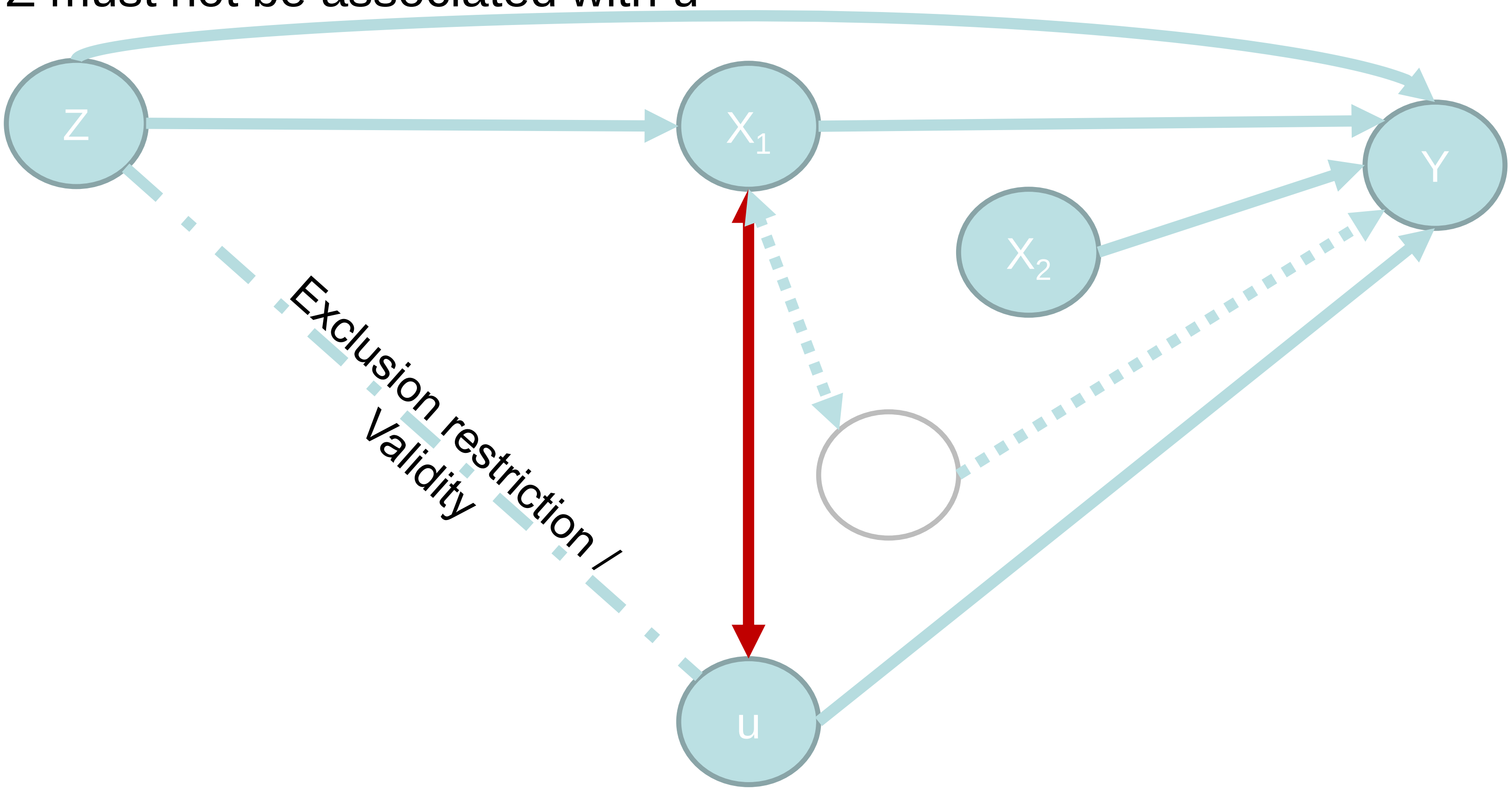
Instrumental variable

- Z must be associated with X_1



Instrumental variable

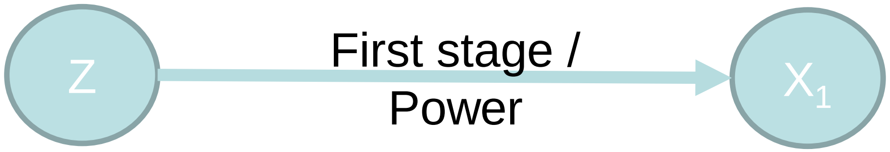
- Z must be associated with X_1
- Z must not be associated with u



First stage

UNIVERSITY OF LEEDS

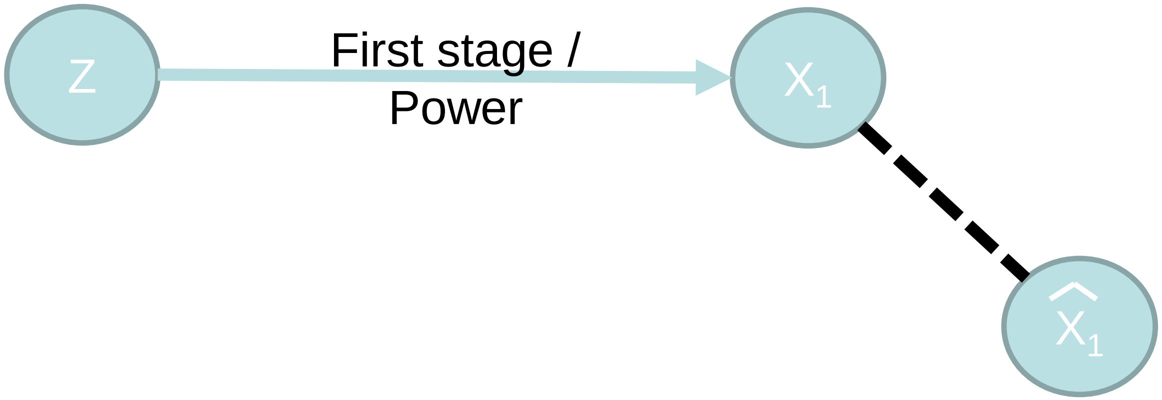
- Regress (endogenous X1) on Z



First stage

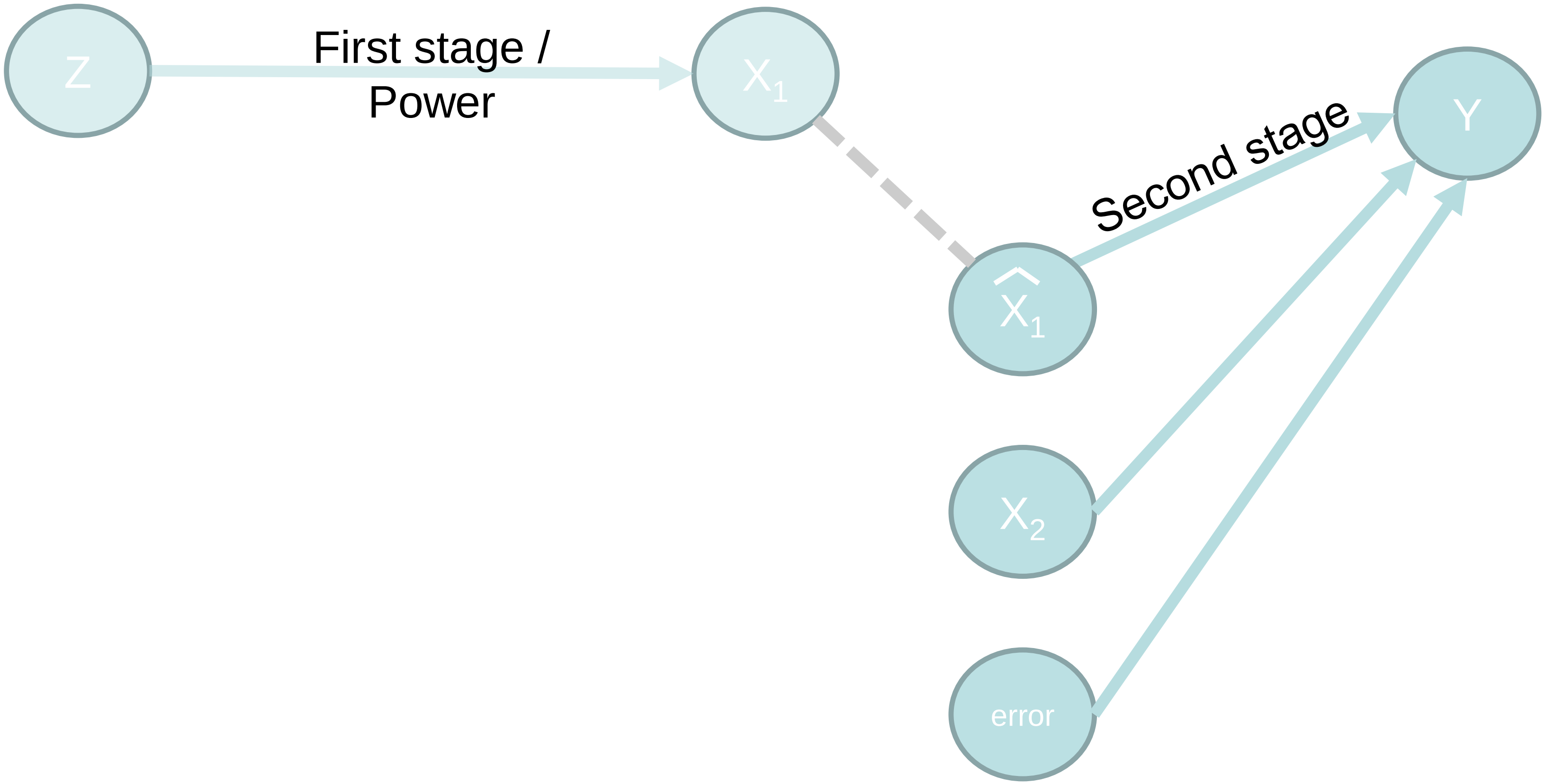
UNIVERSITY OF LEEDS

- Regress (endogenous X1) on Z

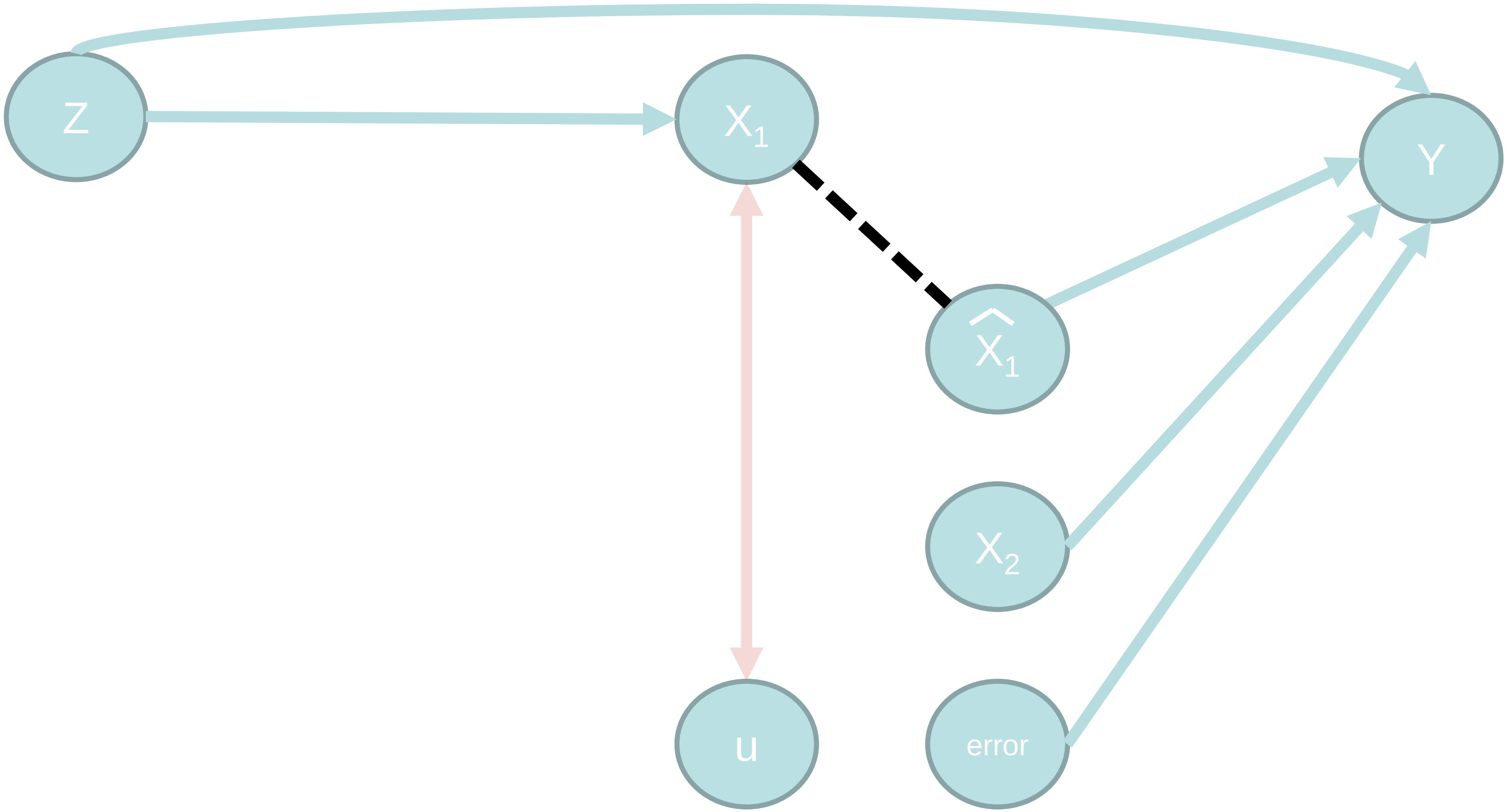


Second stage

- Use the predicted values of X_1 to estimate Y

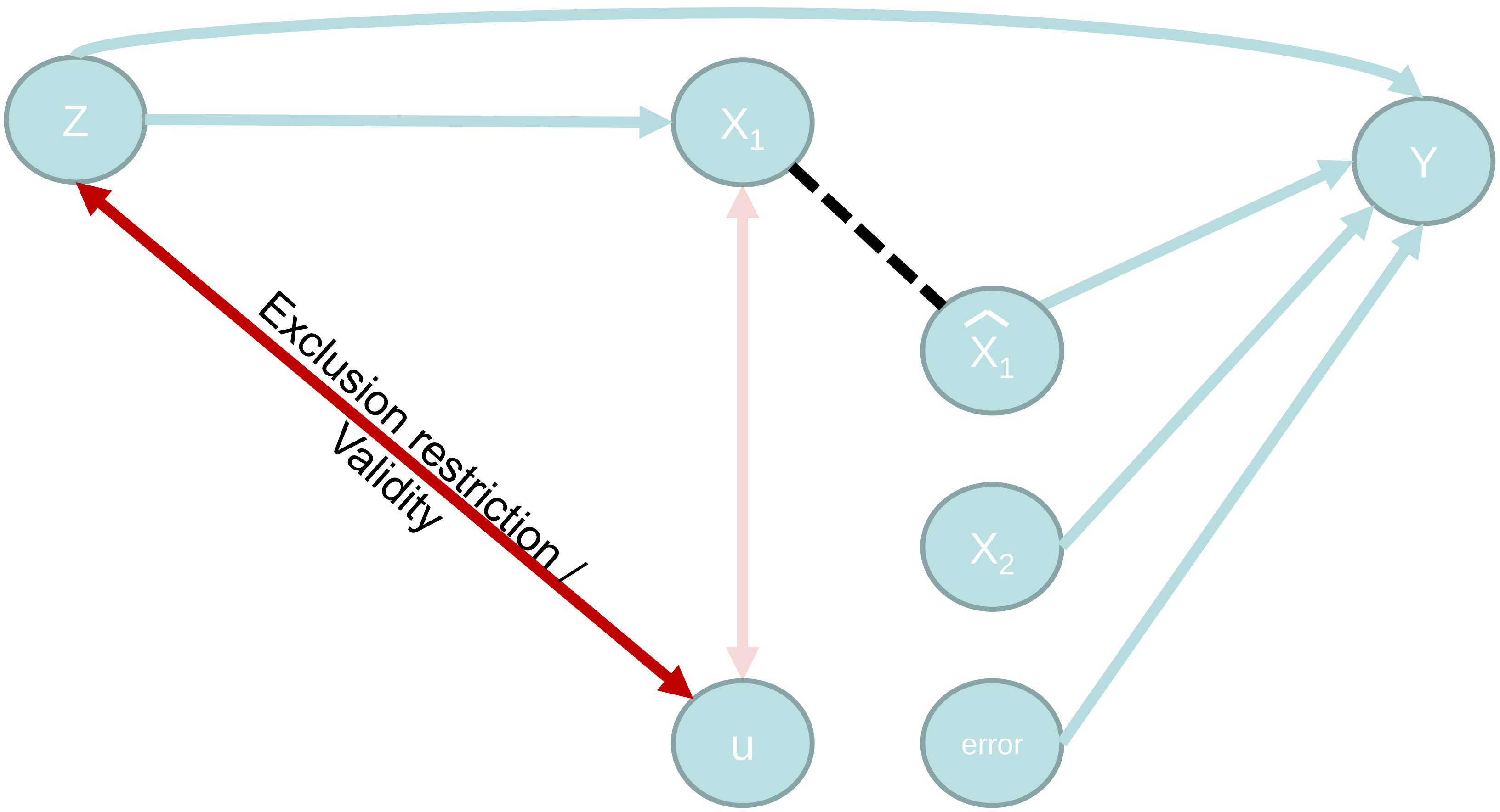


Two critical assumptions



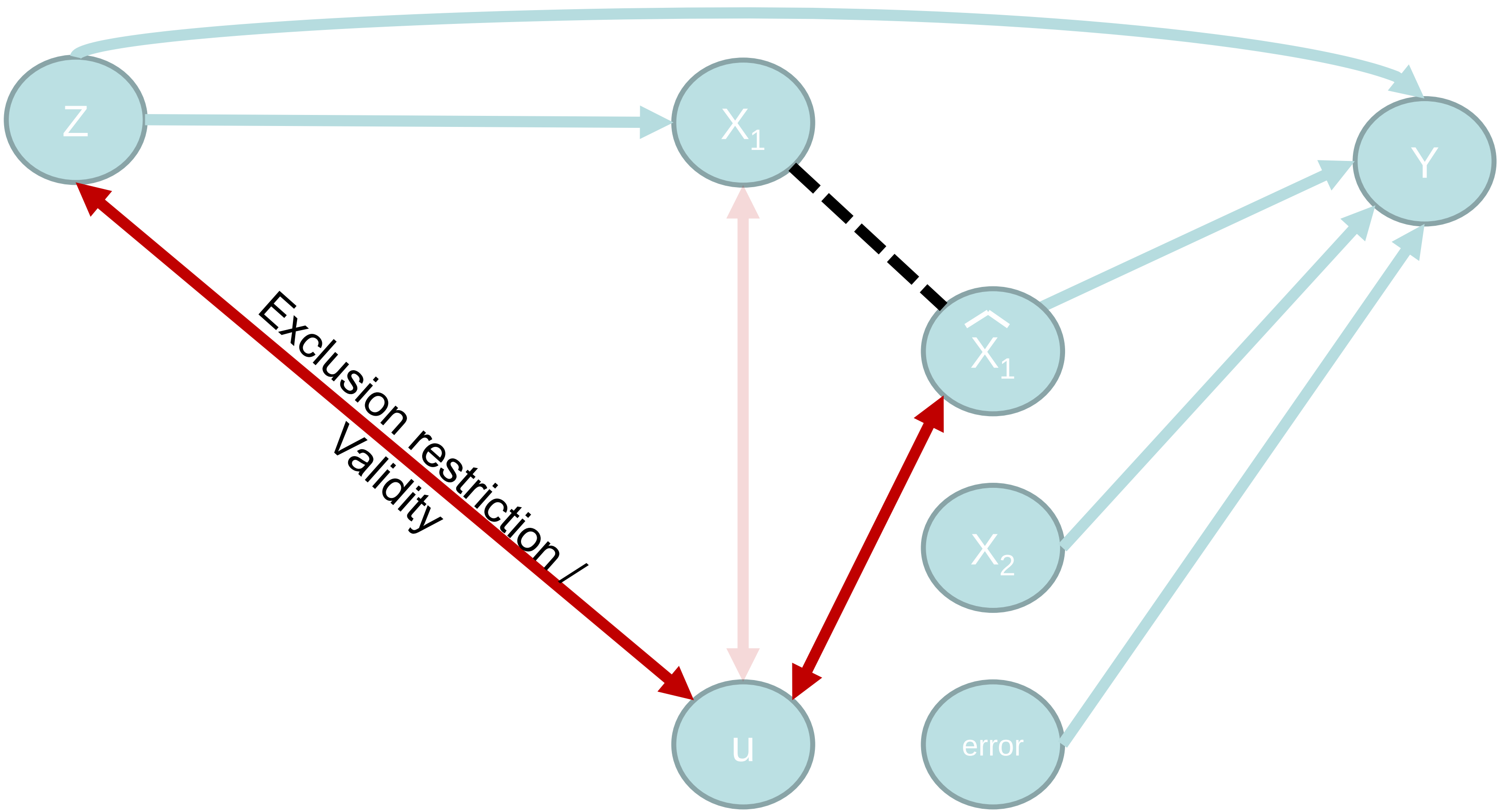
Two critical assumptions

- **There must be no association between Z and u**



Two critical assumptions

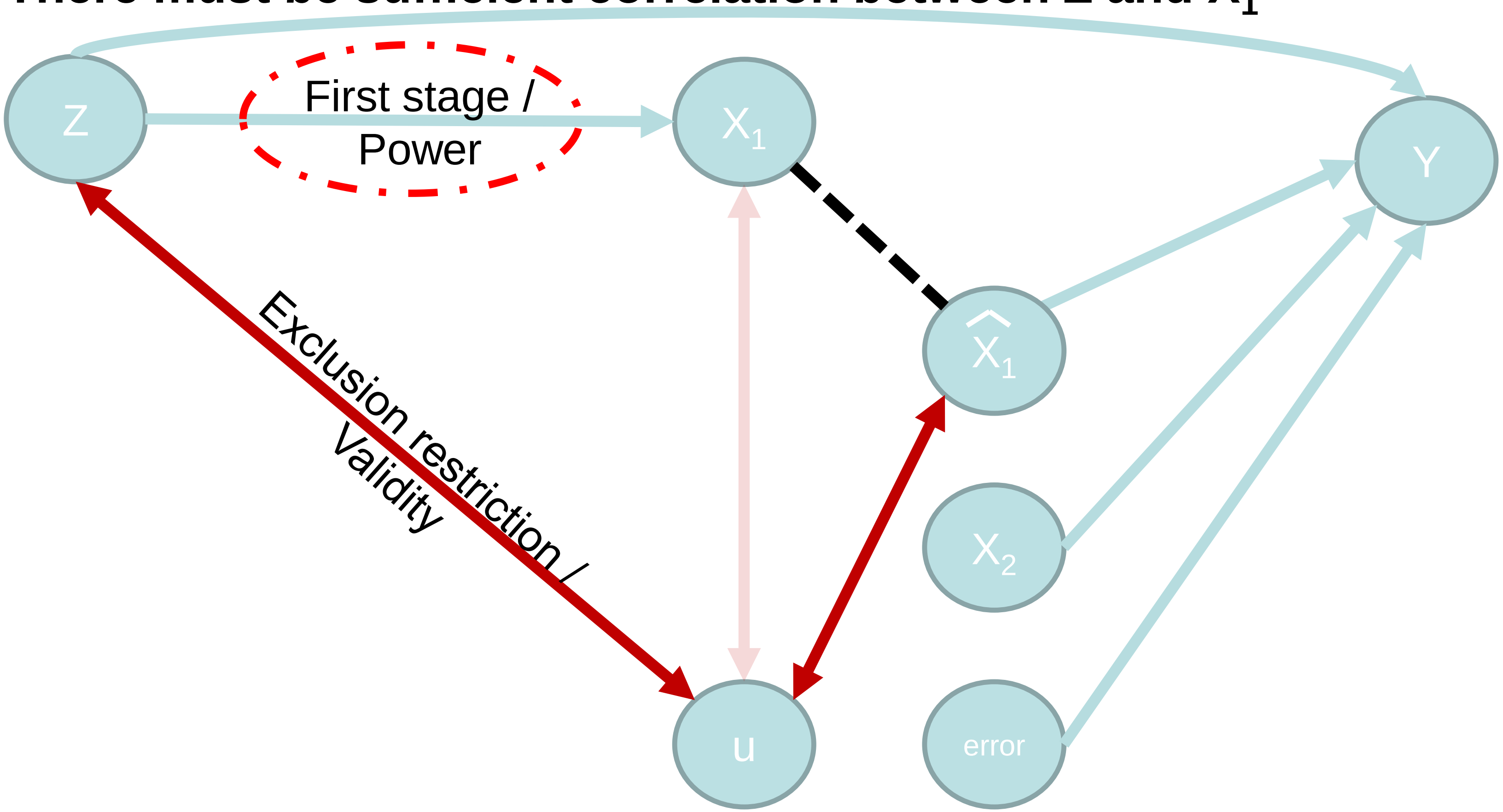
- **There must be no association between Z and u**



Two critical assumptions

UNIVERSITY OF LEEDS

- There must be no association between Z and u
- **There must be sufficient correlation between Z and X_1**



Two critical assumptions

UNIVERSITY OF LEEDS

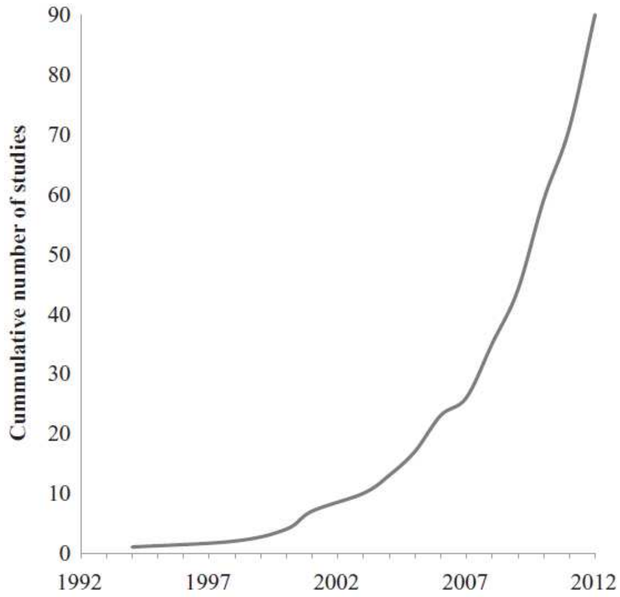
- There must be no association between Z and u (validity)
  - This cannot be demonstrated using the data
  - Researchers must appeal to existing theory and evidence

- There must be sufficient correlation between Z and X_1 (the first stage)
  - This can be proven using the data
  - However, in many empirical examples, if the first stage relationship is very strong, then the validity of the instrument is likely to become less tenable without further evidence
    - See Stock (2002)

A brief history

UNIVERSITY OF LEEDS

- Formal development of IV as a statistical method during the 1940s and 1950s
  - For example, H. Theil paper on two stage least squares published in 1953
- Use of IV in health economics started in the 1990s
  - Since then, interest in IV has grown



Cumulative number of studies listed in Medline and Embase that used (non-genetic) IV analysis for medical research over time (Davies, 2013)

The challenge is finding good instruments

UNIVERSITY OF LEEDS

Journal of Economic Perspectives—Volume 15, Number 4—Fall 2001—Pages 69–85

Instrumental Variables and the Search for Identification: From Supply and Demand to Natural Experiments

Joshua D. Angrist and Alan B. Krueger

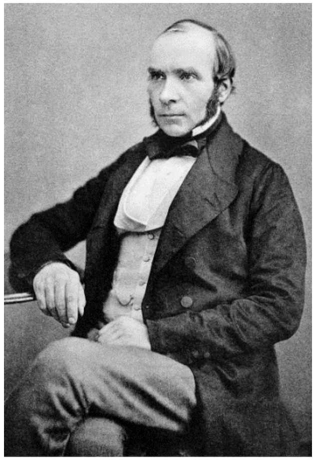
"Our view is that progress in the application of instrumental variables methods depends mostly on **the gritty work of finding** or creating **plausible experiments** that can be used to measure important economic relationships.

"Here the **challenges are not primarily technical** in the sense of requiring new theorems or estimators. Rather, progress comes from detailed institutional knowledge and the careful investigation and quantification of the forces at work in a particular setting.

"Of course, **such endeavours are not really new**. They have always been at the heart of good *empirical research."*

The visionaries

- With hindsight, it’s clear that some earlier researchers had fully understood the IV mechanisms
- Philip (or Sewall) Wright (1928)
  - Observed prices (x) and quantity demanded (y) are determined by the intersection of the supply and demand curves
  - Wanted to estimate the slope of the demand and supply curves for butter and flaxseed
  - An instrument for exogenous variation in supply was change in the weather (which affected crop growth but not product demand)
    - Hidden in Appendix B of a long report on agriculture and farming
- John Snow (1855)
  - For further discussion see Grootendorst (2007)

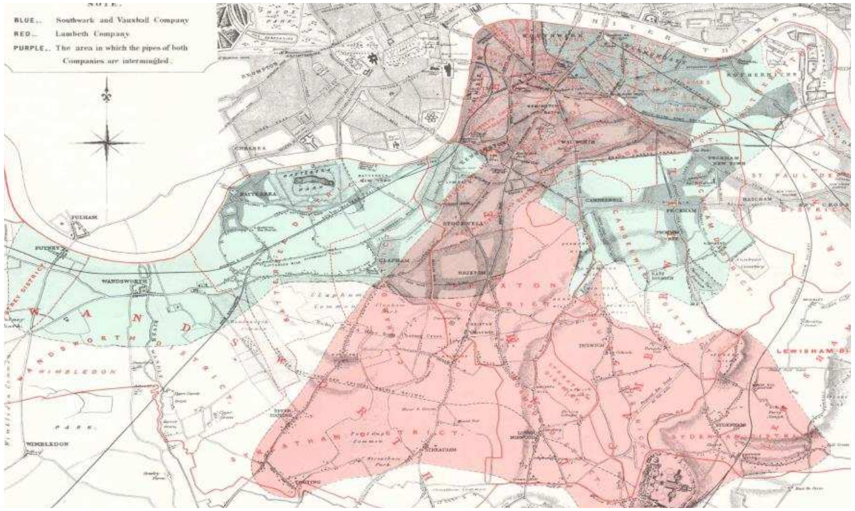


Snow (1855)

- John Snow (On the mode of the communication of cholera, 1855) used IV to investigate the hypothesis that cholera was waterborne, in contrast to other theories that it was airborne
- Observed correlation between water purity (X1) and cholera incidence (y)
- Problem: People drinking impure water were also more likely to be poor and living in unhealthy environments (X3s)
- Solution: Instrument (z) for pure water must be uncorrelated with other (observed and unobserved) determinants of cholera incidence

IV = The water supplier

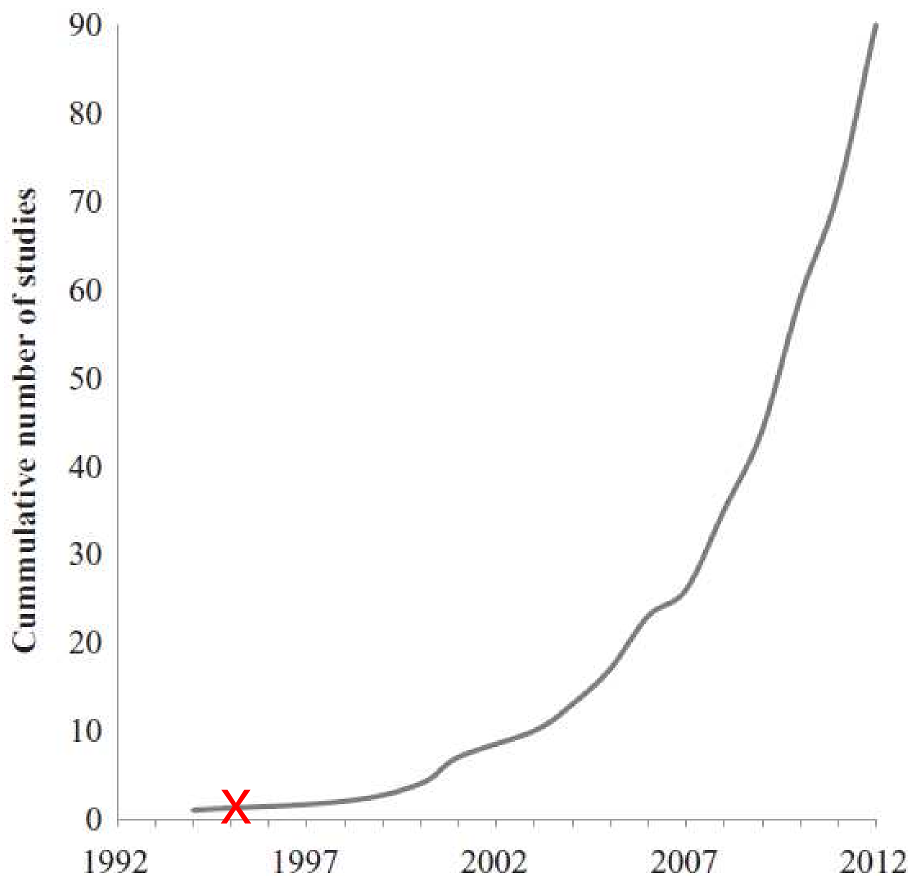
UNIVERSITY OF LEEDS



- People living in green area were served by the Southwark and Vauxhall Water Company
- The pink area by the Lambeth Water Company
- In 1849, both companies obtained their water from Central London
- In 1852, the Lambeth company moved its waterworks upriver to an area relatively free of sewage
- Death rates from cholera transmission during the London cholera epidemic 1853-54 in districts supplied by Lambeth were significantly different to those supplied by Southwark and Vauxhall

1994

UNIVERSITY OF LEEDS



Does More Intensive Treatment of Acute Myocardial Infarction in the Elderly Reduce Mortality?

Analysis Using Instrumental Variables

Mark McClellan, MD, PhD; Barbara J. McNeil, MD, PhD; Joseph P. Newhouse, PhD

Objective.—To determine the effect of more intensive treatments on mortality in elderly patients with acute myocardial infarction (AMI).

Design.—Analysis of incremental treatment effects using differential distances as instrumental variables to account for unobserved case-mix variation (selection bias) in observational Medicare claims data (1987 through 1991).

Main Outcome Measures.—Survival to 4 years after AMI.

Results.—Patients who receive different treatments differ in observable and unobservable health characteristics, biasing estimates of treatment effects based on standard methods of adjusting for observable differences. Patients' differential distances to alternative types of hospitals are strong independent predictors of how intensively an AMI patient will be treated and appear uncorrelated with health status. Thus, differential distances approximately randomize patients to different likelihoods of receiving intensive treatments. Comparisons of patient groups that differ

and comprehensive da able to evaluate medic providers,[4] outcomes r the fastest-growing fie search today. Its resul determine not only wh tive treatments have which practice patter ate,[5] but also which "best."[6,7] Yet, despite terest in and financial comes research, promi tion its validity.[8]

The critics' central c

McClellan et al (1994)

- McClellan et al (1994) used IV to estimate the effect on mortality of intensive treatment of heart attack (using e.g. catheterisation or CABG)
- Correlation between heart attack treatment (X1) and mortality (y), even after controlling for observables (X2s)

Table 2.—Estimated Cumulative Effect of Catheterization, Not Accounting for Selection Bias

Adjustment for Observable Differences Using ANOVA*	Percentage-Point Changes in Mortality Rates (SE)					
	1 d	7 d	30 d	1 y	2 y	4 y
None (unadjusted differences)	−9.4 (0.2)	−18.7 (0.2)	−19.2 (0.3)	−30.5 (0.3)	−34.0 (0.3)	−36.8 (0.3)
After adjustment for demographic differences (age, sex, race, and state)	−6.7 (0.2)	−13.7 (0.2)	−18.7 (0.3)	−26.0 (0.3)	−28.7 (0.3)	−30.4 (0.3)
After adjustment for demographic and comorbidity differences	−6.8 (0.2)	−13.5 (0.2)	−17.9 (0.3)	−24.1 (0.3)	−26.6 (0.3)	−28.1 (0.3)

- Problem: Whether or not patient receives more intensive treatment is, of course, correlated with unobserved factors that affect mortality (X3s)

IV = Distance to specialist hospital

UNIVERSITY OF LEEDS

- Solution: Instrument (z) for intensive treatment was differential distance to hospitals offering the intensive treatment
- Instrument relevance condition/first stage: Patients who live closer to hospitals that use more intensive treatments are more likely receive those treatments

Table 4.—Patient Characteristics by Differential Distance to a Catheterization or Revascularization Hospital*

Characteristic	Differential Distance ≤2.5 Miles (n=102 516)	Differential Distance >2.5 Miles (n=102 505)
Treatments		
Initial admit to catheterization hospital†	34.4	5.0
Initial admit to revascularization hospital†	41.7	10.7
Initial admit to high-volume hospital†	67.1	36.5
Catheterization within 7 d	20.7	11.0
Catheterization within 90 d	26.2	19.5
CABG‡ within 90 d	8.6	6.9
PTCA§ within 90 d	6.4	4.3

IV = Distance to specialist hospital

UNIVERSITY OF LEEDS

- Instrument exogeneity condition/exclusion restriction: Distance the patient lives from a hospital must be independent of the health status

Table 4.—Patient Characteristics by Differential Distance to a Catheterization or Revascularization Hospital*

Characteristic	Differential Distance ≤2.5 Miles (n=102 516)	Differential Distance >2.5 Miles (n=102 505)
Comorbid Disease Characteristics		
Cancer	1.9	1.9
Pulmonary disease, uncomplicated	10.4	10.9
Dementia	0.99	0.94
Diabetes	18.1	18.0
Renal disease, uncomplicated	2.0	1.9
Cerebrovascular disease	4.8	4.8

- This can only be demonstrated for observable characteristics; the authors must convince the reader that it also holds for unobservables!

Naïve results overestimated the impact

UNIVERSITY OF LEEDS

- IV estimates of the impact of catheterisation on mortality are considerably smaller than estimates that did not account for selection bias

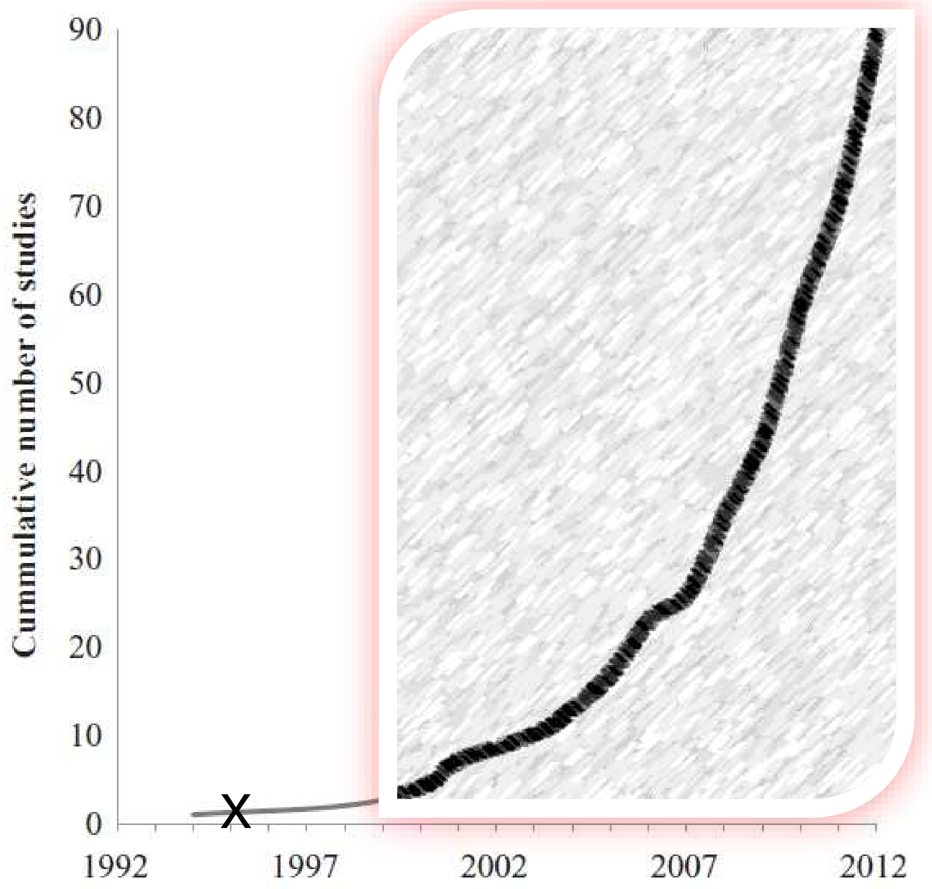
Table 7.—Instrumental Variable Estimates of the Effects of Patient Location, High-Volume Hospital, and Catheterization on Mortality at Indicated Time Intervals After Acute Myocardial Infarction

	Time After Acute Myocardial Infarction, Percentage-Point Change (SE)						
Average Effect	**1 d**	**7 d**	**30 d**	**1 y**	**2 y**	**3 y**	**4 y**
Catheterization within 90 d							
Cumulative	−8.8 (2.0)	−11.5 (2.5)	−7.4 (2.9)	−4.8 (3.2)	−5.4 (3.3)	−5.0 (3.2)	−5.1 (3.2)

- “A redirection of resources from marginal catheterisations and revascularisations to characterising and improving access to some [other] acute treatments could improve AMI mortality in the elderly and may reduce costs.”

Finding more instruments

UNIVERSITY OF LEEDS



2000 onwards

UNIVERSITY OF LEEDS

- **Distance**
  - To healthcare facility
    - McConnell (2005) and Pracht (2008) used differential distance between the nearest level I and level II trauma treatment centres
  - To other determinants of health
    - Zhao (2010) used distance from motorway junctions as instrument for location of fast food restaurants
    - Courtemanche (2011) used distance from Walmart headquarters as instrument for access to cheaper groceries
      - For similar examples, see review: Martin (2014)
- **Personal beliefs**
    - Auld (2005) used religiosity as an instrument for alcohol use
- **Calendar time**
    - Ho (2000) used day of the week of hospital admission for wait time for surgery (many physicians work only on week days)

2000 onwards

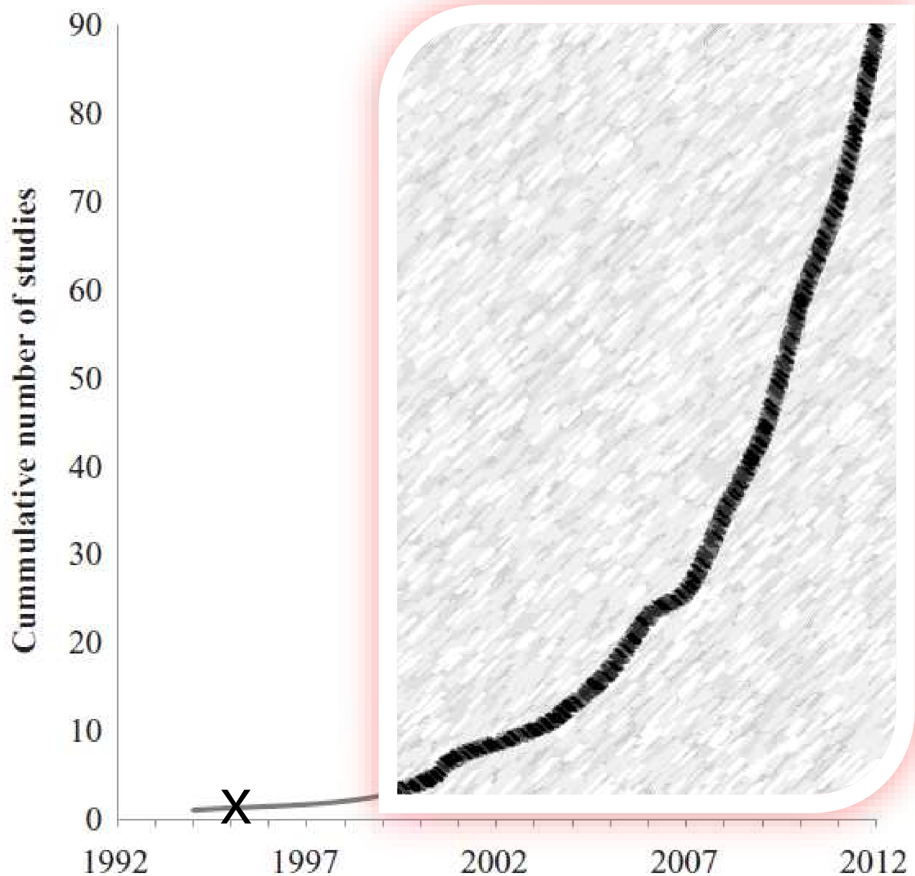
UNIVERSITY OF LEEDS

- **Exogenous shocks**
  - Sudden shifts in patient or physician behaviour
    - Shetty (2010) used publication of unexpected results in an RCT which led to sharp fall in hormone replacement therapy (HRT) use
    - Jenson (2015) used publication of unexpected results which led to sharply altered caesarean section rates for breech babies
  - Also known as ‘Fuzzy’ regression discontinuity design (RDD)
- **State laws, taxes, policies and prices**
    - Leigh (2004) used state-specific cigarette taxes as an instrument for smoking
    - Tonne (2010) used the London congestion charge zone as an instrument for exposure to traffic-related air pollution
    - Other studies have used differences in the minimum legal drinking age as an instrument for (illegal) alcohol consumption

More recent developments

X

X




Journal of Health Economics 31 (2012) 219–230
Contents lists available at SciVerse ScienceDirect
Journal of Health Economics
journal homepage: www.elsevier.com/locate/econbase

The medical care costs of obesity: An instrumental variables approach☆

John Cawley[a,b,c,*], Chad Meyerhoefer[d]

[a] Department of Policy Analysis and Management, Cornell University, United States
[b] Department of Economics, Cornell University, United States
[c] NBER, United States
[d] Department of Economics, Lehigh University, United States

ARTICLE INFO

Article history:
Received 7 October 2010
Received in revised form
29 September 2011
Accepted 11 October 2011
Available online 20 October 2011

JEL classification:
I1
I14
I18
D6

ABSTRACT

This paper is the first to use the method of instrumental variables (IV) to estimate the impact of obesity on medical costs in order to address the endogeneity of weight and to reduce the bias from reporting error in weight. Models are estimated using restricted-use data from the Medical Expenditure Panel Survey for 2000–2005. The IV model, which exploits genetic variation in weight as a natural experiment, yields estimates of the impact of obesity on medical costs that are considerably higher than the estimates reported in the previous literature. For example, obesity is associated with $656 higher annual medical care costs, but the IV results indicate that obesity raises annual medical costs by $2741 (in 2005 dollars). These results imply that the previous literature has underestimated the medical costs of obesity, resulting in underestimates of the economic rationale for government intervention to reduce obesity-related externalities.


IZA

DISCUSSION PAPER SERIES

IZA DP No. 9907

The Effect of Weight on Labor Market Outcomes: An Application of Genetic Instrumental Variables

Petri Böckerman
John Cawley
Jutta Viinikainen
Terho Lehtimäki
Suvi Rovio
Ilkka Seppälä
Jaakko Pehkonen
Olli Raitakari

Cawley (2012)

UNIVERSITY OF LEEDS

- Cawley and Meyerhoefer (2012) used IV to investigate impact of obesity on medical expenses in the U.S.
- Observed correlation between obesity (X_1) and medical expenses (y)

Marginal effects of BMI and obesity on annual medical expenditures (standard errors in parenthesis).

Population	Non-IV (total expenditures)	
	BMI	Obesity
Total (N = 23,689)	49 (9)	656 (113)
Men (N = 9852)	59 (11)	564 (128)
Women (N = 13,837)	47 (11)	749 (150)

Cawley (2012)

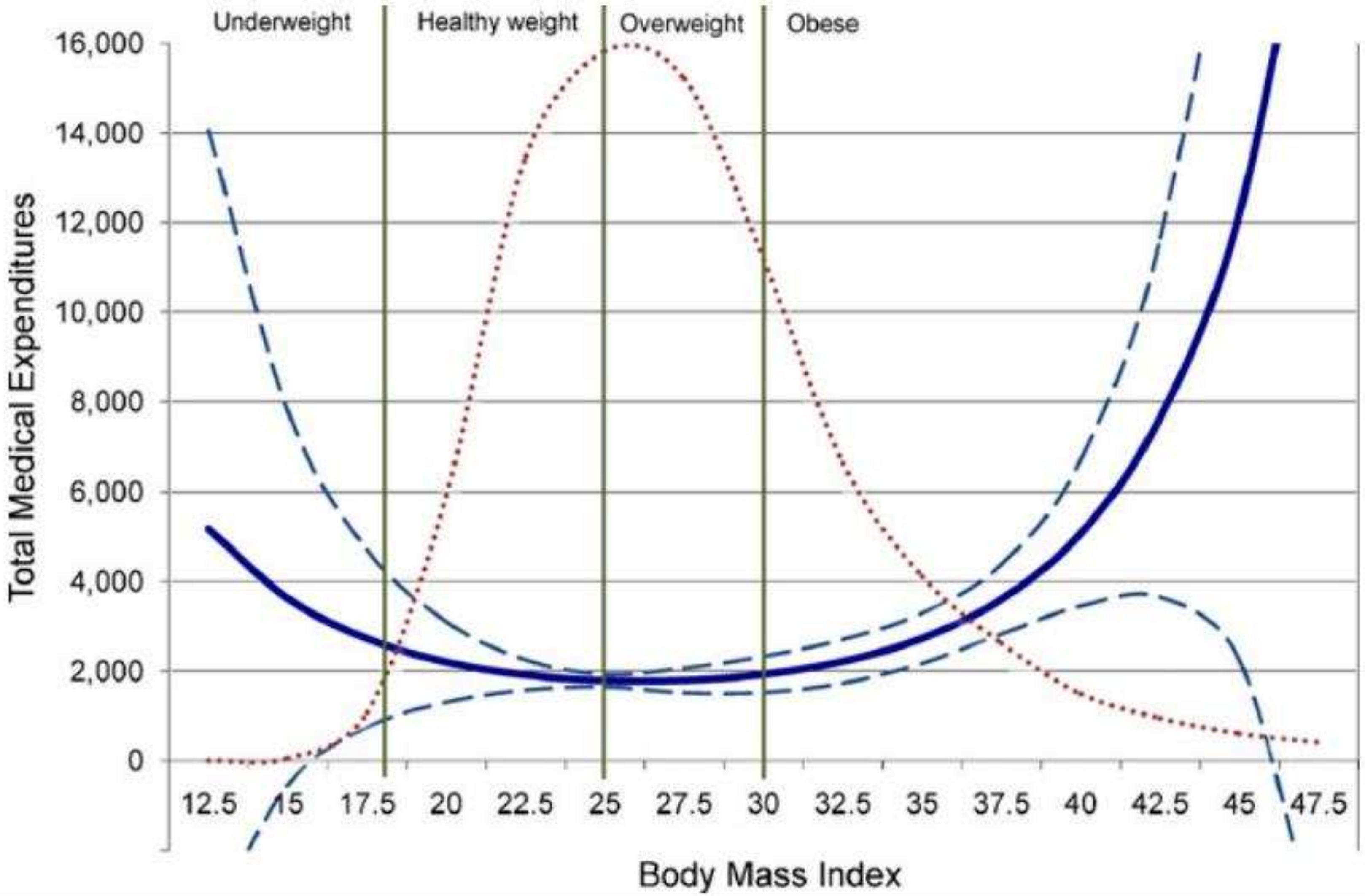
- Problem: Obesity is likely to be associated with other factors that affect medical costs (X3s)
  - People who have suffered a recent injury or chronic depression may have higher medical expenses and increased likelihood of obesity
    - Observed correlation between obesity and medical expenses would be an overestimate of the true causal effect
  - People with poorer access to healthcare, including the disadvantaged or poor, may be more likely to be obese
    - Observed correlation between obesity and medical expenses would be an underestimate of the true causal effect

IV = Weight of own child

- Solution: Instrument (z) for individual's body weight was BMI, BMI squared and BMI cubed of oldest biological child
  - Extensive evidence showing relationship between own BMI and BMI of biological relatives
- Instrument validity: extensive existing research showing no relationship between shared household environment on weight
  - For example, adoption studies consistently show that correlation between a child and its biological parents is comparable for children raised by their biological parents and children raised by adoptive parents

Non linear relationship

UNIVERSITY OF LEEDS



Naïve results underestimated the impact

UNIVERSITY OF LEEDS

- IV estimates of the impact of obesity on medical costs were considerably larger than estimates that did not account for endogeneity

Marginal effects of BMI and obesity on annual medical expenditures (standard errors in parenthesis).

Population	IV (total expenditures)	
	BMI	Obesity
Total ($N=23,689$)	$\boxed{149 (35)}$	$\boxed{2741 (745)}$

- "These results imply that the previous literature has underestimated the medical costs of obesity, resulting in underestimates of the economic rationale for government intervention to reduce obesity-related externalities."
- However, could it be that genes that affect weight may also directly affect medical costs, or lie next to genes that directly affect medical costs?

Bockerman et al (2016)

- Bockerman et al (2016) used IV to investigate the impact of obesity on labour market outcomes (earnings, employment and social income transfers)
- Problem: Obesity is likely to be associated with other factors, including rate of time preference, that affect medical costs (X3s)
  - Also, reverse causality if low income results in weight gain
- Solution: Instrument (z) for individual's body weight was an genetic risk factor score based on 32 SNPs shown to influence obesity
- Instrument validity: Authors appealed to existing research showing no relationship between the 32 BMI-related SNPs and intelligence
- The negative impact of weight on labour market outcomes appeared to be larger in the IV analysis

Warning!

UNIVERSITY OF LEEDS

CLINICAL OUTCOMES ASSESSMENT

Some Cautions on the Use of Instrumental Variables Estimators in Outcomes Research: How Bias in Instrumental Variables Estimators Is Affected by Instrument Strength, Instrument Contamination, and Sample Size

William H. Crown, PhD[1,], Henry J. Henk, PhD[2], David J. Vanness, PhD[3]*

[1]*OptumInsight Life Sciences, Waltham, MA, USA;* [2]*OptumInsight Life Sciences, Eden Prairie, MN, USA;* [3]*University of Wisconsin, Madison, WI, USA*

ABSTRACT

Objectives: To examine the performance of instrumental variables (IV) and ordinary least squares (OLS) regression under a range of conditions likely to be encountered in empirical research. **Methods:** A series of simulation analyses are carried out to compare estimation error between OLS and IV when the independent variable of interest is endogenous. The simulations account for a range of situations that may be encountered by researchers in actual practice—varying degrees of endogeneity, instrument strength, instrument contamination, and sample size. The intent of this article is to provide researchers with more intuition with respect to how important these factors are from an empirical standpoint. **Results:** Notably, the simulations indicate a greater potential for inferential error when using IV than OLS in all but the most ideal circumstances. **Conclusions:** Researchers should be cautious when using IV methods. These methods are valuable in testing for the presence of endogeneity but only under the most ideal circumstances are they likely to produce estimates with less estimation error than OLS.

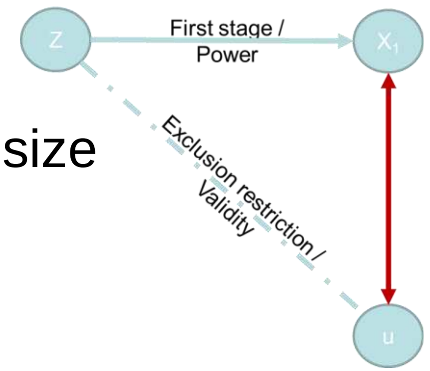
Keywords: bias, endogeneity, instrumental variables.

"Although the appeal of IV as a method for addressing endogeneity issues is undeniable, it is important to understand that the use of IV can do more harm than good.

"Rather common…are heroic attempts to 'find an instrument' whenever the slightest possibility of endogeneity arises."

Pitfalls

- Validity of the instrument
  - Just as the researcher never really knows how big the endogeneity problem is, nor can the validity of the instrument ever be proven
  - Vital that researcher **appeals to existing theory or empirical evidence** to make a convincing case
  - Falsification tests may be helpful, but still do not provide conclusive evidence
- In an effort to ensure the validity of the instrument, researchers may have a tendency to use weak instruments, i.e. weak correlation between Z and X_1
  - Large standard errors
  - Biased towards the OLS estimates if small sample size
  - Vital that researchers:
    - Report the first stage F statistics
    - Use a large sample size



Heterogeneous treatment effects

UNIVERSITY OF LEEDS

- IV parameters represent the Local Average Treatment Effect (LATE) – that is the average (causal) effect of X on Y for those whose treatment status has been changed by the instrument (i.e. **‘compliers’**)
  - This is because the estimates were identified through independent variation induced only by the chosen instrument(s)
    - e.g. in the McClellan (1994) study, the compliers are heart attack patients for whom the choice of hospital is affected by the differential distance to the specialist hospital
  - No information about **‘always-takers’** and **‘never-takers’** since they always/never take the treatment independently of Z
    - e.g. patients who will always (or never) attend the specialist hospital regardless of the differential distance
- LATE also assumes that there were no **‘defiers,**’ i.e. although the instrument need not have affected everyone, those who were affected were affected in the same way (monotonicity)
    - e.g. patients whose treatment is changed by the instrument would always be more (not less) likely to attend the specialist hospital as the differential distance decreases

Why is IV necessary in the first place?

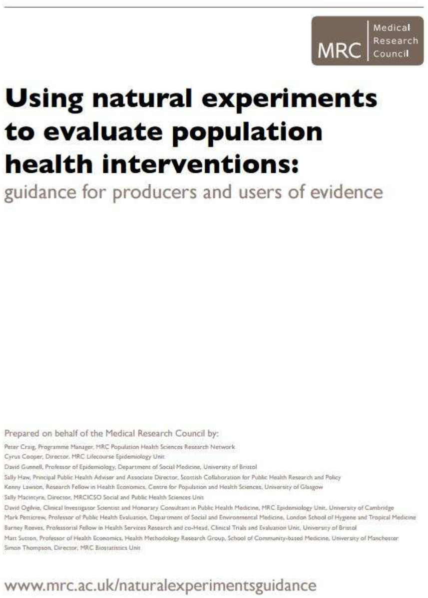
UNIVERSITY OF LEEDS

- Author must explain why IV is necessary
- If the potential bias of single equation method is of a small magnitude, then the loss of precision in IV estimates might not be justified
  - In some case better to report IV results as a sensitivity analysis?
  - Discussion of the expected direction of the bias in OLS estimates may be sufficient to inform policy making
- Unlike RCTs, IVs may take some explaining for policy makers?

The future of IV in health economics

UNIVERSITY OF LEEDS

- IV likely to remain popular
  - Increasing size and availability of panel datasets with variables across multiple domains, e.g. health, education, labour market, family history, etc.
  - Increased potential for data linkage across sectors, e.g. diverse administrative datasets and local geographic data related to the social determinants of health
  - Increased availability of genetic data
- Recent MRC guidance on natural experiments highlights IVs
- The main challenges are not technical, but in identifying good sources of natural experiments
- Good practice guidance for authors and reviewers of studies is vital to avoid erroneous interpretation


MRC Medical Research Council

Using natural experiments to evaluate population health interventions:
guidance for producers and users of evidence

Prepared on behalf of the Medical Research Council by:
Peter Craig, Programme Manager, MRC Population Health Sciences Research Network
Cyrus Cooper, Director, MRC Lifecourse Epidemiology Unit
David Gunnell, Professor of Epidemiology, Department of Social Medicine, University of Bristol
Sally Haw, Principal Public Health Adviser and Associate Director, Scottish Collaboration for Public Health Research and Policy
Kenny Lawson, Research Fellow in Health Economics, Centre for Population and Health Sciences, University of Glasgow
Sally Macintyre, Director, MRC/CSO Social and Public Health Sciences Unit
David Ogilvie, Clinical Investigator Scientist and Honorary Consultant in Public Health Medicine, MRC Epidemiology Unit, University of Cambridge
Mark Petticrew, Professor of Public Health Evaluation, Department of Social and Environmental Medicine, London School of Hygiene and Tropical Medicine
Barney Reeves, Professorial Fellow in Health Services Research and co-Head, Clinical Trials and Evaluation Unit, University of Bristol
Matt Sutton, Professor of Health Economics, Health Methodology Research Group, School of Community-based Medicine, University of Manchester
Simon Thompson, Director, MRC Biostatistics Unit

www.mrc.ac.uk/naturalexperimentsguidance

Volume 12 • Number 8 • 2009
VALUE IN HEALTH

Too Much Ado about Instrumental Variable Approach: Is the Cure Worse than the Disease?

Onur Baser, MS, PhD
STATinMED Research, LLC, and Department of Surgery, University of Michigan, Ann Arbor, MI, USA

"Do we have a method to control for both observed and unobserved bias?

"The answer is 'theoretically YES' but practical application is very limited because of the difficulty in finding the right instrument."

References: Examples of IV studies

UNIVERSITY OF LEEDS

Bockerman P, Cawley J, Viinikainen J, Lehtimäki T, Rovio S, Seppälä I, Pehkonen J, Raitakari O (2016) The Effect of Weight on Labor Market Outcomes: An Application of Genetic Instrumental Variables. IZA Discussion Paper No. 9907.

Cawley J, Meyerhoefer C (2012) The medical care costs of obesity: an instrumental variables approach. Journal of Health Economics.

Courtemanche C, Carden A (2011) Supersizing supercenters? The impact of Walmart Supercenters on body mass index and obesity. Journal of Urban Economics.

Ho V, Hamilton BH, Roos LL (2000) Multiple approaches to assessing the effects of delays for hip fracture patients in the United States and Canada. Health Services Research.

Howard D (2000) The impact of waiting time on liver transplant outcomes. Health Services Research.

Jensen VM, Wüst M (2015) Can Caesarean section improve child and maternal health? The case of breech babies. Journal of Health Economics.

McClellan M, McNeil BJ, Newhouse JP (1994) Does more intensive treatment of acute myocardial infarction in the elderly reduce mortality? Analysis using instrumental variables. Journal of American Medical Association.

Zhao Z and Kaestner R (2010) Effects of urban sprawl on obesity. Journal of Health Economics.

References: Reviews

UNIVERSITY OF LEEDS

Cawley J (2015) A selective review of the first 20 years of instrumental variables models in health-services research and medicine. Journal of Medical Economics.

Davies NM, Smith GD, Windmeijer F, Martin RM (2013) Issues in the reporting and conduct of instrumental variable studies: a systematic review. Epidemiology.

Martin A, Ogilvie D, Suhrcke M (2014) Evaluating causal relationships between urban built environment characteristics and obesity: a methodological review of observational studies. International Journal of Behavioral Nutrition and Physical Activity.

References: Commentary

UNIVERSITY OF LEEDS

Angrist JD, Krueger AB (2001) Instrumental Variables and the Search for Identification: From Supply and Demand to Natural Experiments. Journal of Economic Perspectives.

Angrist JD, Pischke JS (2009) Mostly Harmless Econometrics: An Empiricist's Companion. Princeton University Press.

Baser O (2009) Too much ado about instrumental variable approach: is the cure worse than the disease? Value in Health.

Bound J, Jaeger DA, Baker RM (1995) Problems with Instrumental Variables Estimation When the Correlation Between the Instruments and the Endogeneous Explanatory Variable is Weak. Journal of the American Statistical Association.

Crown WH, Henk HJ, Vanness DJ (2011) Some cautions on the use of instrumental variables estimators in outcomes research: how bias in instrumental variables estimators is affected by instrument strength, instrument contamination, and sample size. Value in Health.

Grootendorst P (2007) A review of instrumental variables estimation of treatment effects in the applied health sciences. Health Services and Outcomes Research Methodology.