Causal Inference in Observational Settings

Peter Davis, University of Auckland

Seminar Series, LSHTM

Thursday, 12 April 2012

CSM Agenda Today

- **C**anvass
  - Sage Handbook/Reader
    - “Inference by Design”: outline, ideas, amendments
- **S**hare
  - Intellectual excitement
    - Philosophy, statistics, public health, econometrics
- **M**ethodological caveat

Presentation Outline

- Rationale, motivation

- Four “key” papers

- The Handbook
  – Volume I – Background
  – Volume II – Comparing like with like
  – Volume III – Panel data and instruments
  – Volume IV – Experimental analogues

- Concluding thoughts

What's at Issue

- **Fundamental issue of policy science**
  – how to draw "credible" (causal?) inferences from observational data
- **Causal identification via data analysis**
  – often a form of speculative post-mortem
- **Basic conundrum of causal reasoning**
  – impossible to observe unit response under alternative conditions

Rationale of Proposal

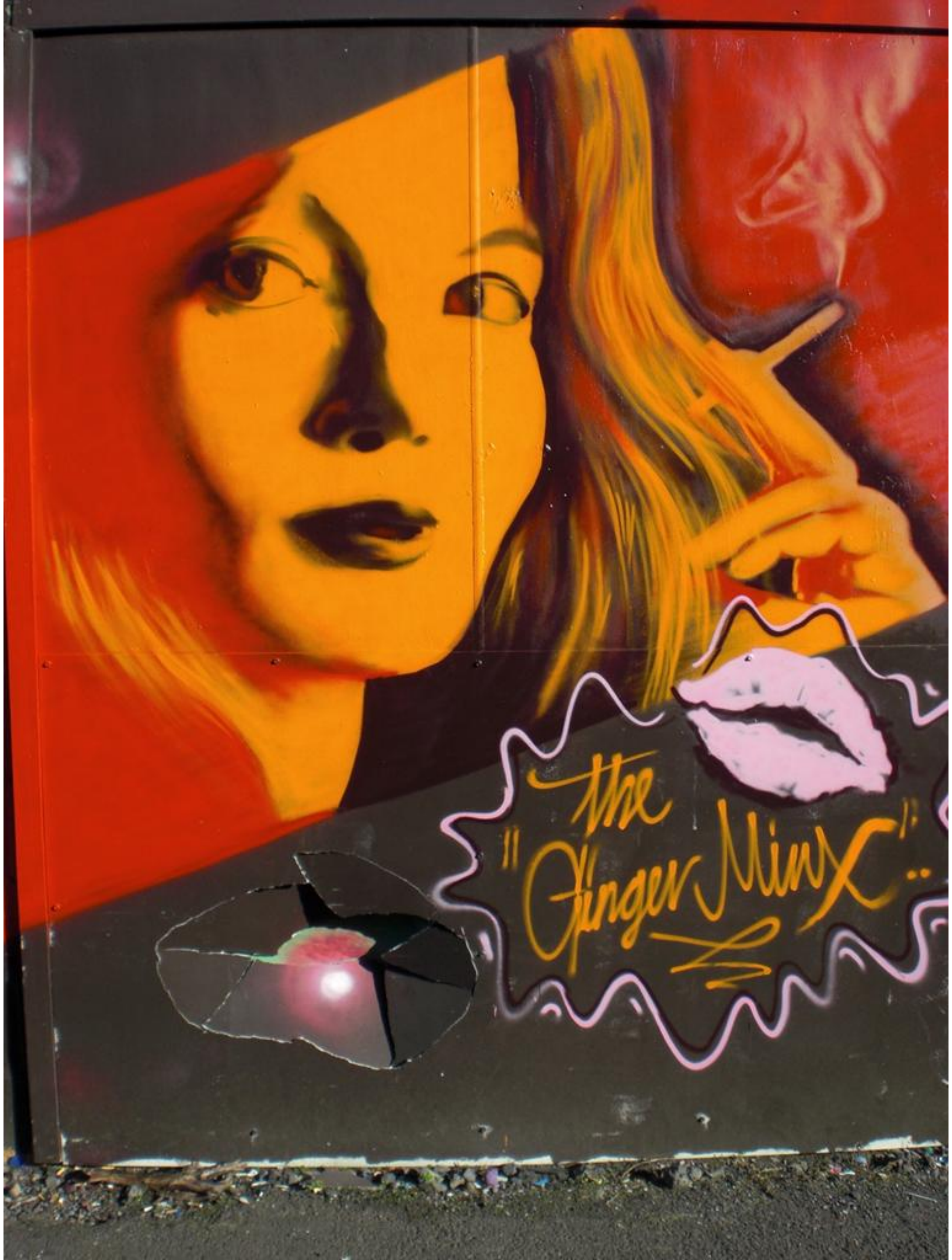
1. **Traditional statistical theory**
   mainly about representation not causation (i.e. sampling)

2. **Statistical inference=>causal inference**
   random assignment and manipulation of treatment conditions

3. **Counterfactual/potential outcomes**
   conceptually bridges experimental/observational settings

4. **Forward causation only**
   cause-to-effect (e.g. impact of policy intervention)

5. **Econometrics**
   a parallel community of policy practice


the "Ginger Minx"..

Four "Key" Papers

- Counterfactual thinking
  - *Estimating the effects of potential public health interventions*. Ahern et al. **AJE** 2009

- Using panel data
  - *Does marriage reduce crime?* Sampson et al. **Criminology** 2006

- Statistical reasoning
  - *Causal inference using potential outcomes*. Rubin, **JASA** 2005

- The econometric paradigm
  - *How better research design is taking the con out of econometrics*. Angrist/Pischke, **J Econ Persp** 2010

Ahern et al.

American Journal of Epidemiology

Vol. 169, No. 9
DOI: 10.1093/aje/kwp015
Advance Access publication March 6, 2009

Practice of Epidemiology

Estimating the Effects of Potential Public Health Interventions on Population Disease Burden: A Step-by-Step Illustration of Causal Inference Methods

Jennifer Ahern, Alan Hubbard, and Sandro Galea

Initially submitted March 26, 2008; accepted for publication January 13, 2009.

Causal inference methods allow estimation of the effects of potential public health interventions on the population burden of disease. Motivated by calls for epidemiologic research to be presented in ways that are more informative for intervention, the authors present a didactic discussion of the steps required to estimate the population effect of a potential intervention using an imputation-based causal inference method and discuss the assumptions of and limitations to its use. An analysis of neighborhood smoking norms and individual smoking behavior is used as an illustration. The implementation steps include the following: 1) modeling the adjusted exposure and outcome association, 2) imputing the outcome probability for each individual while manipulating the exposure by "setting" it to different values, 3) averaging these probabilities across the population, and 4) bootstrapping confidence intervals. Imputed probabilities represent counterfactual estimates of the population smoking prevalence if neighborhood smoking norms could be manipulated through intervention. The degree to which temporal ordering, randomization, stability, and experimental treatment assignment assumptions are met in the illustrative example is discussed, along with ways that future studies could be designed to better meet the assumptions. With this approach, the potential effects of an intervention targeting neighborhoods, individuals, or other units can be estimated.

causality; intervention studies; methods; population; residence characteristics; smoking; social environment

Abbreviations: GEE, generalized estimating equation; OR, odds ratio.

Most analyses of epidemiologic data apply a regression model such as linear or logistic regression. These models have in common that they estimate differences (relative or absolute) between outcomes (in terms of rates, risks, odds, or prevalences) associated with variations in exposure, while holding constant a set of covariates (1–3). These models estimate differences in outcomes that are stratum specific, because they are estimated within strata of the covariates specified in the model. Although such findings constitute the backbone of modern epidemiologic research (3), they represent only 1 approach to capturing the association between an exposure and an outcome. This approach tells us little about population disease burden or about how the disease burden might change if the exposure were modified.

One alternate approach, which could be more informative, would assess how a particular potential intervention on the exposure being studied might reduce disease burden across the population (2, 4). Several methods can estimate population parameters under hypothetical interventions. In simple situations, standardization can estimate a population-level causal effect (5, 6). Certain causal inference methods generalize standardization to situations with covariates that are continuous as well as categorical, covariates that are time dependent, models that include multiplicative interactions, and nonlinear model forms (5–10). Although many causal inference methods were developed to control time-dependent confounding, the machinery allows the estimation of population parameters under hypothetical interventions for cross-sectional studies. Causal inference analyses of epidemiologic data start with the specification of a causal effect that is of interest. The population average causal effect is specified as the difference in the outcome (e.g., the

Correspondence to Dr. Jennifer Ahern, School of Public Health, University of California, Berkeley, 101 Haviland Hall, Berkeley, CA 94720-7358 (e-mail: jahern@berkeley.edu).

Counterfactual – Neighbourhood Norms

- Population average causal effect
  - difference under one intervention vs. another (or none) by estimating counterfactual exposures->outcomes
- Epidemiological association smoking/norms
  - estimate counterfactual - impute new pattern of neighbourhood smoking norms and derive smoking levels
- Prevalence estimates if norms “manipulated”
  - 17% (versus 29%) if all neighbourhoods prohibitive

Ahern et al.

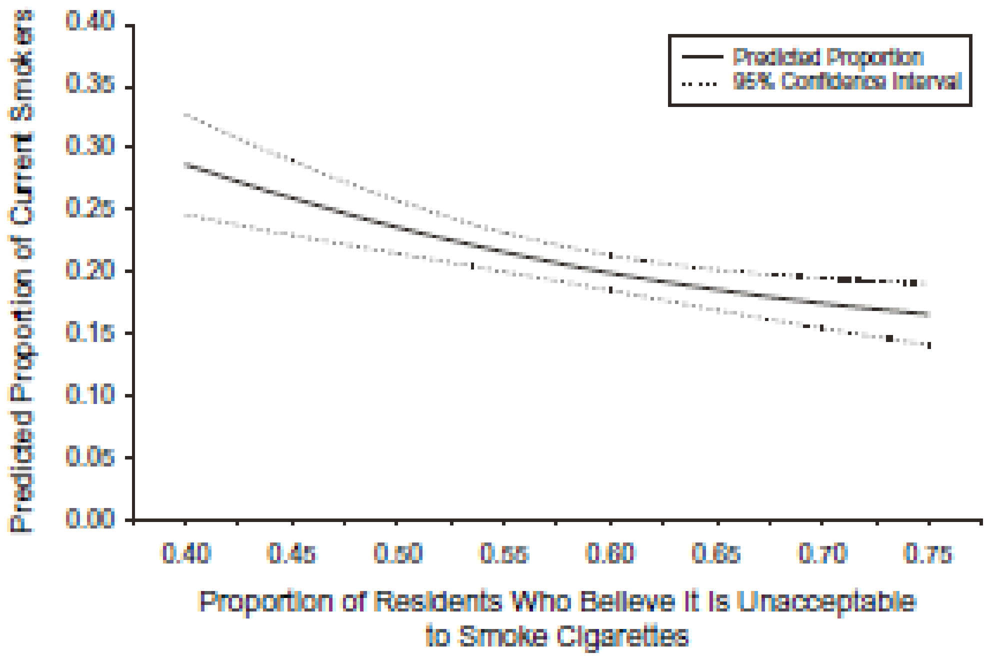


Figure 1. Predicted smoking prevalence corresponding to counterfactually "set" levels of neighborhood smoking norms, New York, New York, 2005.

Sampson et al.

DOES MARRIAGE REDUCE CRIME?
A COUNTERFACTUAL APPROACH TO
WITHIN-INDIVIDUAL CAUSAL EFFECTS*

ROBERT J. SAMPSON
Harvard University
JOHN H. LAUB
University of Maryland
CHRISTOPHER WIMER
Harvard University

KEYWORDS: marriage, crime, causality, counterfactual methods, life course

Although marriage is associated with a plethora of adult outcomes, its causal status remains controversial in the absence of experimental evidence. We address this problem by introducing a counterfactual life-course approach that applies inverse probability of treatment weighting (IPTW) to yearly longitudinal data on marriage, crime, and shared covariates in a sample of 500 high-risk boys followed prospectively from adolescence to age 32. The data consist of criminal histories and death records for all 500 men plus personal interviews, using a life-history calendar, with a stratified subsample of 52 men followed to age 70. These data are linked to an extensive battery of individual and family background measures gathered from childhood to age 17—before entry into marriage. Applying IPTW to multiple specifications that also incorporate extensive time-varying covariates in adulthood, being married is associated with an average reduction of approximately 35 percent in the odds of crime compared to nonmarried states for the same man. These results are robust, supporting the inference that states of marriage causally inhibit crime over the life course.

* We thank the Russell Sage Foundation (Grant # 85-01-23) for funding support and the following colleagues for advice: Chris Winship, Felix Elwert, David Harding, Steve Raudenbush, Guanglei Hong, Jamie Robins, and the reviewers of *Criminology*. Direct all correspondence to Robert J. Sampson, Department of Sociology, Harvard University, William James Hall, 33 Kirkland St., Cambridge, MA 02138 USA; e-mail: rsampson@wjh.harvard.edu.

CRIMINOLOGY VOLUME 44 NUMBER 3 2006 465

Using Panel Data - Marriage and Crime

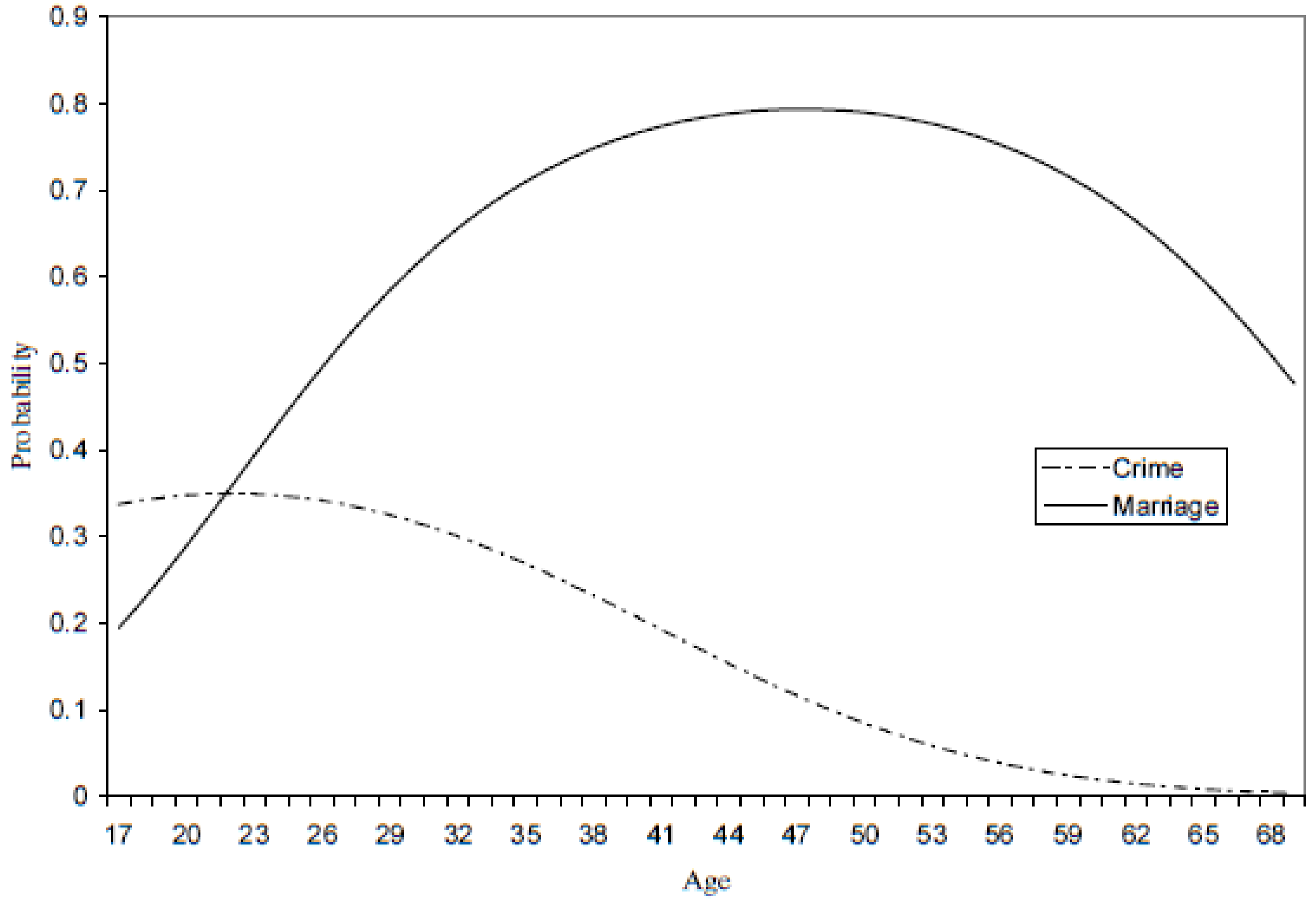
- Does marriage reduce crime?
  - issues of selection and confounding

- Longitudinal data available on "high-risk" men
  - within-individual analysis of role of marriage

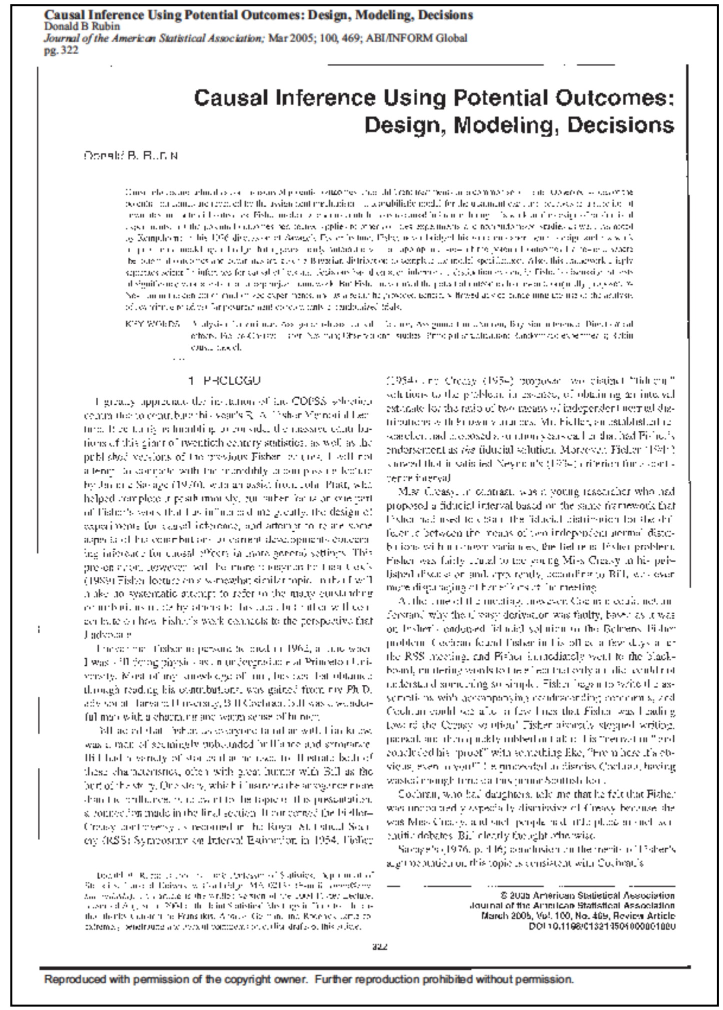
- Do states of marriage causally inhibit crime?
  - Yes – average 35% reduction compared to non-married

Sampson et al.

Figure 1. Predicted Crime and Marriage Probabilities by Age (Quadratic Model, N=2,585 Person-Years)



Rubin



Causal Inference Using Potential Outcomes: Design, Modeling, Decisions
Donald B Rubin
Journal of the American Statistical Association; Mar 2005; 100, 469; ABI/INFORM Global
pg. 322

Causal Inference Using Potential Outcomes: Design, Modeling, Decisions

Donald B. Rubin

1. PROLOGUE

© 2005 American Statistical Association
Journal of the American Statistical Association
March 2005, Vol. 100, No. 469, Review Article
DOI 10.1198/016214504000001880

322

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Statistical Reasoning - Design and Decisions

- Science and design vs. analysis and decisions
    - Fisher never related his work on likelihoods and models to his work on experimental design
- Neyman – potential outcomes of treatment
    - defines causal effects for both randomised and non-randomised studies (“Neyman-Rubin” model)
- Causal inference and assignment mechanism
    - assigns treatments to units (randomised in experiments), creating special type of missing data

Rubin

		Potential outcomes			
Units	*Covariates* X	*Treatment* $Y(1)$	*Control* $Y(0)$	*Unit-level Causal effects*	*Summary Causal effects*
1	X_1	$Y_1(1)$	$Y_1(0)$	$Y_1(1)$ v. $Y_1(0)$	Comparison of $Y_i(1)$ v. $Y_i(0)$ for a common set of units
⋮	⋮	⋮	⋮	⋮	
i	X_i	$Y_i(1)$	$Y_i(0)$	$Y_i(1)$ v. $Y_i(0)$	
⋮	⋮	⋮	⋮	⋮	
N	X_N	$Y_N(1)$	$Y_N(0)$	$Y_N(1)$ v. $Y_N(0)$	

Figure 1. "Science"—The Causal Estimand.

Angrist and Pischke

Journal of Economic Perspectives—Volume 24, Number 2—Spring 2010—Pages 3–30

The Credibility Revolution in Empirical Economics: How Better Research Design is Taking the Con out of Econometrics

Joshua D. Angrist and Jörn-Steffen Pischke

Just over a quarter century ago, Edward Leamer (1983) reflected on the state of empirical work in economics. He urged empirical researchers to "take the con out of econometrics" and memorably observed (p. 37): "Hardly anyone takes data analysis seriously. Or perhaps more accurately, hardly anyone takes anyone else's data analysis seriously." Leamer was not alone; Hendry (1980), Sims (1980), and others writing at about the same time were similarly disparaging of empirical practice. Reading these commentaries as late-1980s Ph.D. students, we wondered about the prospects for a satisfying career doing applied work. Perhaps credible empirical work in economics is a pipe dream. Here we address the questions of whether the quality and the credibility of empirical work have increased since Leamer's pessimistic assessment. Our views are necessarily colored by the areas of applied microeconomics in which we are active, but we look over the fence at other areas as well.

Leamer (1983) diagnosed his contemporaries' empirical work as suffering from a distressing lack of robustness to changes in key assumptions—assumptions he called "whimsical" because one seemed as good as another. The remedy he proposed was sensitivity analysis, in which researchers show how their results vary with changes in specification or functional form. Leamer's critique had a refreshing emperor's-new-clothes earthiness that we savored on first reading and still enjoy today. But we're happy to report that Leamer's complaint that "hardly anyone takes anyone else's data analysis seriously" no longer seems justified.

■ *Joshua D. Angrist is Ford Professor of Economics, Massachusetts Institute of Technology, Cambridge, Massachusetts. Jörn-Steffen Pischke is Professor of Economics, London School of Economics, London, United Kingdom. Their e-mail addresses are ⟨angrist@mit.edu⟩ and ⟨s.pischke@lse.ac.uk⟩.*

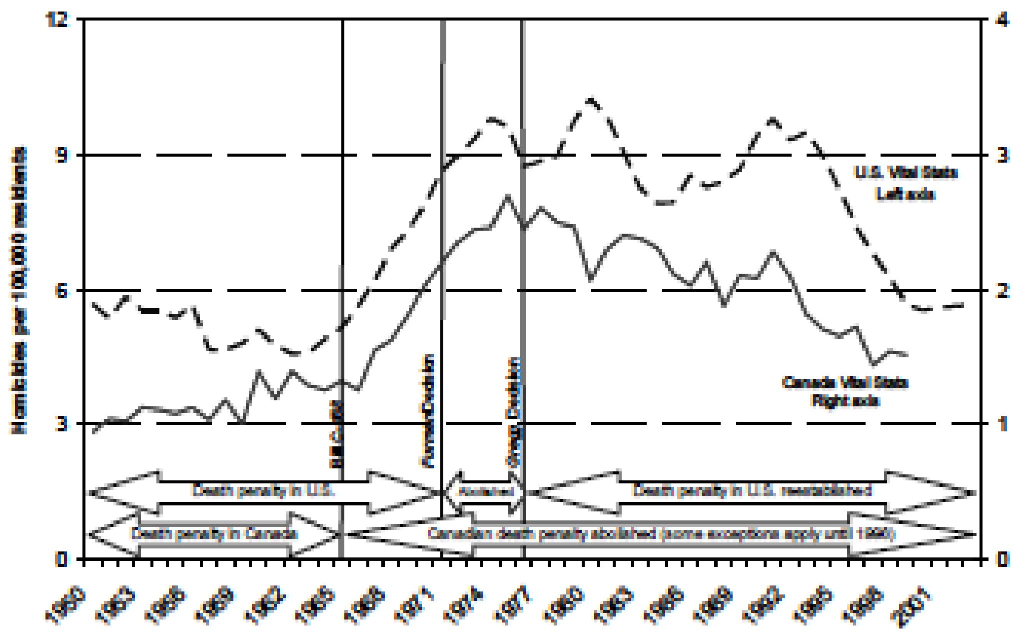
doi=10.1257/jep.24.2.3

Econometrics - “Better” Research Design

- “take the con out of econometrics” (1985)
  - Leamer “Hardly anyone takes data analysis seriously.”
- Better research design – quasi-experimental
  - Instrumental variables, regression discontinuity, differences-in-differences
- Has the design pendulum swung too far?
  - Lack of external validity; ignore the big questions?

Angrist and Pischke

Figure 1
Homicide Rates and the Death Penalty in the United States and Canada
(U.S. and Canada rates on the left and right y-axes, respectively)



Source: Donohue and Wolfers (2005).


CBAK
TMD
RTR

Sage Handbook Series

- Sage Benchmarks in Social Research Methods
- Four-volume readers
- 75 “readings”
- Previous examples
  - Social Statistics
  - Causality
  - Computational Social Science
- Working title: “Inference by Design”

Current Structure of Proposal

- Volume I – Background
  - Causal inference
  - Potential outcomes
  - “Evaluation research”
- Volume II – Comparing like with like
  - Matching methods
  - Propensity scoring
- Volume III – Panel data and instruments
  - Fixed effects
  - Difference-in-difference
  - Instrumental variables
- Volume IV – Experimental analogues
  - Regression discontinuity
  - Quasi-experiments, natural experiments
  - Field experiments

Volume I - Background

- A. Causal inference from observational data
- B. Potential outcomes and counterfactuals
- C. Programme and policy evaluation

Causal Inference from Observational Data

DATE	AUTHOR(S)	TITLE	SOURCE
1986	Holland	Statistics and causal inference	JASA
1999	Winship, Morgan	The estimation of causal effects from observational data	Ann Rev Sociol
2000	Little, Rubin	Causal effects in clinical and epidemiological studies	Ann Rev Public Health
2000	Sobel	Causal inferences in the social sciences	JASA
2005	Heckman	The scientific model of causality	Sociol Methodology
2007	Rubin	The design versus the analysis of observational studies for causal effects	Statistics in Medicine
2010	Gangl	Causal inferences in sociological research	Ann Rev Sociology

Causal Inference from Observational Data

- Holland
    - » The analysis of causation should begin with studying the effects of causes.
    - » No causation without manipulation.
- Sobel
    - » Only causal sequences are counterfactually regular.
- Rubin
    - » Observational studies can and should be designed to approximate randomized experiments as closely as possible.

Potential outcomes and counterfactuals

DATE	AUTHOR(S)	TITLE	SOURCE
1951	Roy	Some thoughts on the distribution of earnings	Oxford Economic Papers
1990	Holland	Rubin's model and its application to causal inference …	American Journal of Epidemiology
1991	Fearon	Counterfactuals and hypothesis testing …	World Politics
2001	Morgan	Counterfactuals, causal effect heterogeneity, and the Catholic School	Sociology of Education
2003	Harding	Counterfactual models of neighbourhood effects	American Journal of Sociology
2005	Rubin	Causal inference using potential outcomes	JASA
2006	Sampson et al.	Does marriage reduce crime? A counterfactual approach …	Criminology

Potential outcomes and counterfactuals

- Harding
  - » This study employs counterfactual models … to estimate the effects of neighborhood poverty …
- Sampson et al.
  - » Our approach is to extend “counterfactual” methods for time-varying covariates to a within-individual analysis of the role of marriage …

Programme and policy evaluation

DATE	AUTHOR(S)	TITLE	SOURCE
1969	Campbell	Reforms as experiments	Amer Psych
1975	Alwin, Sullivan	Issues of design and analysis in evaluation research	Sociological Methods & Res
1994	Imbens, Angrist	Identification and estimation of local average treatment effects	Econometrica
1999	Dehejia, Wahba	Causal effects in non experimental studies	JASA
2009	Ahern et al.	Estimating the effects of potential public health interventions …	American Journal of Epidemiology
2009	Imbens, Wooldridge	Recent developments in the econometrics of program evaluation	Journal of Economic Lit
2010	Angrist, Pischke	The credibility revolution in empirical economics: how better …	J Economic Perspectives

Programme and policy evaluation

- Ahern et al.
  - » Causal inference methods allow estimation of the effects of potential public health interventions …
- Alwyn, Sullivan
  - » The principal inferential device whereby the effects of various policies are made known involves the incorporation of valid comparison into research design …

Volume II – Comparing like with like

- D. Matching methods
- E. Propensity scoring

Matching methods

DATE	AUTHOR(S)	TITLE	SOURCE
1968	Cochran	The effectiveness of adjustment by sub-classification in removing bias …	Biometrics
1984	Rosenbaum, Rubin	Reducing bias in observational studies using sub-classification	JASA
1985	Rosenbaum, Rubin	Constructing a control group using multivariate matched sampling …	The American Statistician
1997	Smith	Matching with multiple controls to estimate treatment effects in observational studies	Sociological Methodology
1998	Heckman et al.	Matching as an econometric evaluation estimator	Review of Economic Stud
2003	Christakis, Iwashyna	The health impact of health care on families: a matched cohort study …	Social Science and Medicine
2004	DiPrete, Engelhardt	Estimating causal effects with matching methods …	Sociol Methods and Research
2005	Smith, Todd	Does matching overcome Lalonde's critique of nonexperimental estimators?	J Econometrics
2006	Morgan, Harding	Matching estimators of causal effects. Prospects, and pitfalls	Sociol Methods and Research
2008	Gilligan, Sergenti	Do UN interventions cause peace?	Q J Pol Sci
2010	Stuart	Matching methods for causal inference	Statistical Science

Matching methods

- Morgan, Harding
  - » … matching techniques can be used effectively to strengthen the prosecution of causal questions in sociology
- Stuart
  - » When estimating causal effects using observational data, it is desirable to replicate a randomized experiment as closely as possible by obtaining treated and control groups with similar covariate distributions.

Propensity scoring

DATE	AUTHOR(S)	TITLE	SOURCE
1983	Rosenbaum, Rubin	The central role of the propensity score in observational studies for causal effects	Biometrika
1997	Rubin	Estimating causal effects from large data sets using propensity scores	Ann Internal Medicine
2001	Hirano et al.	Efficient estimation of average treatment effects using the estimated propensity score	Econometrica
2002	Dehejia, Wahba	Propensity score-matching methods for nonexperimental causal studies	Rev Econom Statist
2002	Woodridge	Inverse probability weighted estimation for general missing data problems	J Econom
2004	Lunciford, Davidian	Stratification and weighting via propensity scores…	Statistics in Medicine
2006	Baser	Too much ado about propensity score models?	Value in Health
2007	Austin et al.	A comparison of the ability of different propensity score models to balance …	Statistics in Medicine

Volume III – Panel data and instruments

- F. Fixed effects
- G. Difference-in-difference
- H. Instrumental variables.

Fixed effects

DATE	AUTHOR(S)	TITLE	SOURCE
1998	Cherlin et al.	Effects of parental divorce on mental health throughout the life course	American Sociological Rev
1998	Duncan et al.	How much does childhood poverty affect the life chances of children?	American Sociological Rev
1999	Guo, van Wey	Sibship size and intellectual development	Am Soc Rev
2000	Conley, Bennett	Is biology destiny? Birth weight and life chances	Am Soc Rev
2004	Halaby	Panel models in sociological research: Theory and practice.	Annual Review of Sociology
2011	Gunasekara et al.	Change in income and change in self-rated health: Systematic review	SSM

Fixed effects

- Halaby
  - » The fundamental structure of panel data provides the analytical leverage for … the estimation of causal effects
- Duncan et al.
  - » We use whole-childhood data from the PSID to relate children's completed schooling and nonmarital fertility to parental income …
- Gunsekara et al.
  - » … the true causal short-term relationship between income and health … may be much smaller than that suggested by previous, mostly cross-sectional research.

Difference-in-difference

DATE	AUTHOR(S)	TITLE	SOURCE
2004	**Bertrand et al.**	**How much should we trust differences-in-differences estimates?**	**Quarterly Journal of Economics**

Instrumental variables

DATE	AUTHOR(S)	TITLE	SOURCE
1993	Manski	Identification of endogenous social effects: the reflection problem	Review of Economic Stud
1995	Bound et al.	Problems with instrumental variables estimation…	JASA
1996	Angrist et al.	Identification of causal effects using instrumental variables	JASA
1996	Heckman	Randomisation as an instrumental variable	Rev Econ Stat
1997	Staiger, Stock	Instrumental variables regression with weak instruments	Econometrica
2011	Denny	Instrumental variable estimation of the effect of prayer on depression	SSM
2011	Sovey, Green	Instrumental variables estimation in political science. A readers' guide	Am J Pol Sci

Instrumental variables

- Denny

  » Using Instrumental Variables estimation, which allows one to isolate exogenous variation in prayer, leads to the conclusion … there may be some benefit to prayer …

Volume IV – Experimental analogues

- I. Regression discontinuity
- J. Quasi-experiments and natural experiments
- K. Field experiments.

Regression discontinuity

DATE	AUTHOR(S)	TITLE	SOURCE
1960	Thistlethwaite, Campbell	Regression-discontinuity analysis: an alternative to ex post facto experiment	Journal of Educational Psych
1983	Berk, Rauma	Capitalising on non-random assignment to treatments: a regression discount ..	JASA
1995	Myer et al.	Workers' compensation and injury duration: evidence from a natural exp	American Economic Rev
1999	Berk, de Leeuw	An evaluation of California's inmate classification system using a generalised regression discount design	JASA
2001	Hahn et al.	Identification of treatment effects by regression discontinuity designs	Economtrica
2008	Imbens, Lemieux	Regression discontinuity designs: a guide to practice	J Econom

Quasi-experiments and natural experiments

DATE	AUTHOR(S)	TITLE	SOURCE
1985	Berk, Newton	Does arrest really deter wife battery?	Am Soc Rev
1994	Card and Krueger	Minimum wages and employment: a case study …	American Economic Review
1995	Myer et al.	Workers' compensation and injury duration: evidence from a natural exp	American Economic Rev
2002	Schneeweiss et al.	Quasi-experimental longitudinal designs to evaluate drug benefit policy..	Journal of Clinical Epidemiology
2009	Kirk	A natural experiment on residential change and recidivism …	American Sociological Rev
2010	Strully et al.	Effects of prenatal poverty on infant health	ASR

Field experiments

DATE	AUTHOR(S)	TITLE	SOURCE
2008	Clampet-Lundquist, Massey	Neighborhood effects on economic self-sufficiency: a reconsideration of the MTO experiment	American Journal of Sociology
2008	Ludwig et al.	What can we learn from neighbourhood effects from MTO?	American Journal of Sociology
2008	Sampson	Moving to inequality: neighborhoods and experiments meet social structure	American Journal of Sociology


TMD

Concluding Thoughts

- Can insistence on causal purity go too far?
  - Smoking and lung cancer; climate change
  - Status of predictive and descriptive work?
- Still plenty of “wriggle room” in clinical trials
  - Major investment in flu vaccine despite doubts
- But still a worthwhile criterion
  - Lack of clinical trial scrutiny for hip replacement
- The policy sciences need this credibility

Davis et al (Lancet article)

Davis et al. (Medical Care article)

ORIGINAL ARTICLE

Do Hospital Bed Reduction and Multiple System Reform Affect Patient Mortality?

A Trend and Multilevel Analysis in New Zealand Over the Period 1988–2001

Peter Davis, PhD, Roy Lay-Yee, MA,* Alastair Scott, PhD,† and Robin Gauld, PhD‡*

Background: The impact of hospital and system restructuring on the quality and pattern of care is an important issue of public policy concern.

Objective: To assess the effect on patterns of care and patient outcomes of a substantial reduction in public hospital bed availability and multiple reorganizations in New Zealand through the 1990s.

Research Design: Trend analysis using both tabular and multilevel techniques.

Subjects: Access to discharge data, amounting to 6,639,487 records, was secured for all 34 major public hospitals in New Zealand over the period 1988–2001.

Outcome Measures: Number of discharges, admission rate, access levels, mean length of stay, unplanned readmission rate, and 60-day postadmission mortality rate.

Results: Although the number of inpatient beds in use declined by one-third over the period and the national population grew by nearly one-fifth, discharge volumes increased significantly and rates of inpatient admission were maintained, as were access levels for vulnerable groups. These changes were accompanied by workload adjustments (a halving in length of stay and an increase by a quarter in readmission rates). Yet age-adjusted postadmission patient mortality decreased by a quarter over the period of study, a rate of decline that was slowed by the major workload adjustments but not by reform phase.

Conclusions: Other things being equal, a substantial reduction in inpatient bed availability can be effected in national public hospital systems, while largely maintaining access and quality of care. However, the workload adjustments that are required may slow improvements in patient outcomes.

Key Words: health system reform, patient outcomes, multilevel analysis

(*Med Care* 2007;45: 1186–1194)

From the Departments of *Sociology and †Statistics, University of Auckland, Auckland, New Zealand; and ‡Department of Preventive and Social Medicine, University of Otago, Dunedin, New Zealand.
Supported by the Health Research Council of New Zealand.
Reprints: Peter Davis, PhD, Department of Sociology, University of Auckland, Private Bag 92019, Auckland, New Zealand. E-mail: pb.davis@auckland.ac.nz.

ISSN: 0025-7079/07/4512-1186

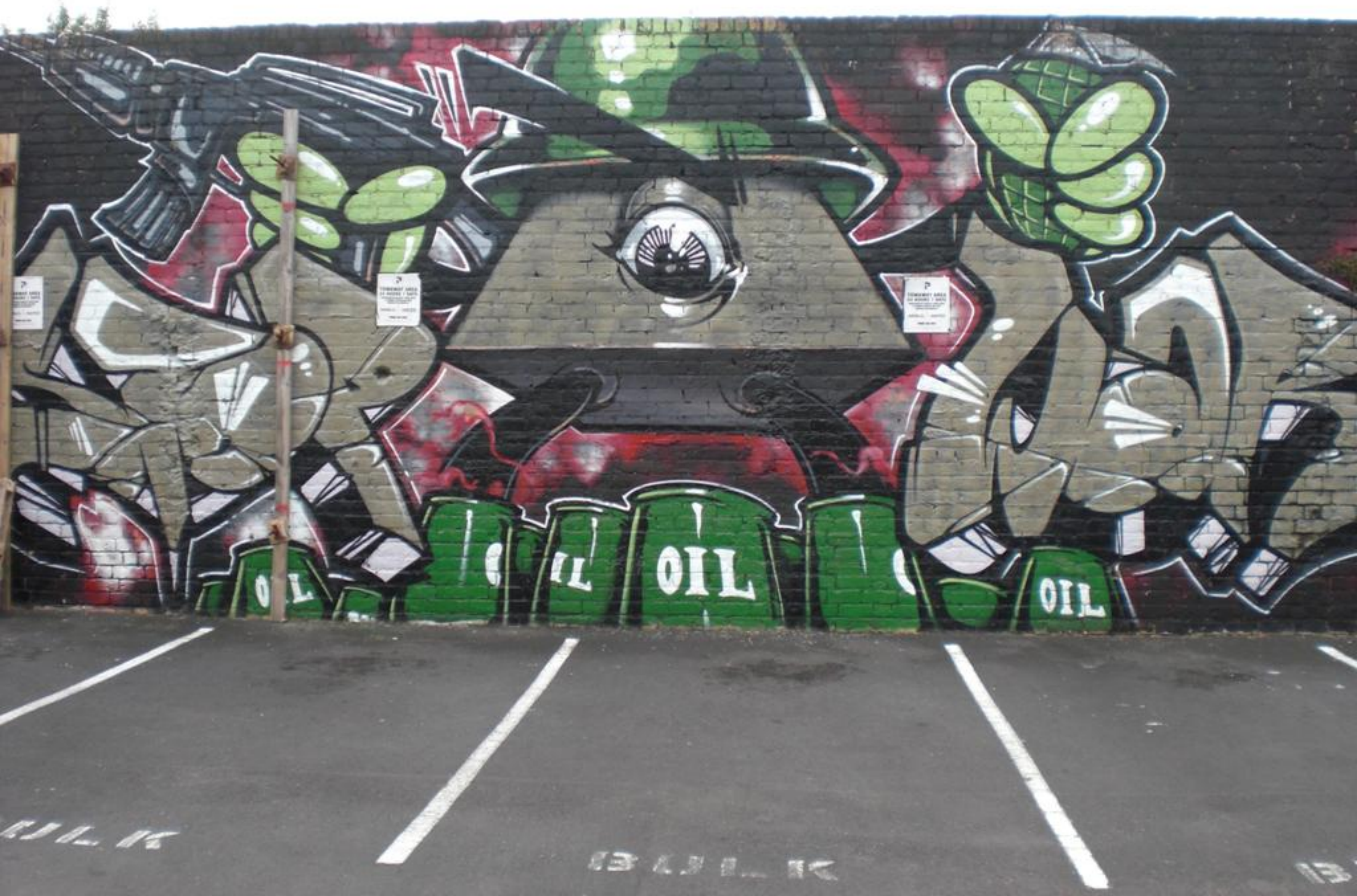
Internationally there has been a considerable change in the role of the hospital through the 1990s, with higher rates of admission, shorter periods of stay, and growing rates of outpatient and day care.[1] An important strand in this change in role was a conscious restructuring of hospital workforce and redesign of work in inpatient settings across the developed world.[2] Over this same period, many of these countries also underwent bouts of broader health reform.[3] New Zealand, where the government pays for 80% of health care and public institutions dominate the health system, was no exception. The country undertook 4 sets of changes to the publicly-funded health system up to 2001 (see Fig. 1), including a succession of public hospital sector reorganizations.[4] At the same time, in a related trend, the sector experienced a substantial reduction in the availability of inpatient beds.[5]

The substantive interest in the New Zealand case is 4-fold. First, it was one of a group of countries with national health service-type systems that implemented a suite of market-oriented reforms from the late-1980s to the mid-1990s (the others being Italy, Spain, Sweden, and the United Kingdom).[6] These reforms were typically intended to create a "market" for publicly-funded health services by instituting competitive tendering between government-purchasing agencies and service providers vying among one another to win contracts to provide public services, and also by transforming public hospitals into public corporations expected to function like private cost-conscious businesses. These were features of the second and third reform phases in New Zealand (see Fig. 1). Second, this suite of reforms probably went further and faster in New Zealand than anywhere else and were part of a broader reform thrust in economic and social policy. They also drew widespread popular and political opposition.[6] Third, New Zealand simultaneously experienced both a substantial reduction in availability of public hospital beds and 4 separate structural reorganizations (Fig. 1).[4] Fourth, even though many of these reform experiments were short lived, internationally, as Or has noted, "the lack of proper evaluation . . . is striking,"[7] particularly with concerns about possible effects on access and quality.[6]

Given the strength and coherence of the reform program, and its powerfully managerial and efficiency objectives,[4] 3 key questions arise. First, how did the performance

Concluding Thoughts

- Can insistence on causal purity go too far?
  - Smoking and lung cancer; climate change
  - Status of predictive and descriptive work?
- Still plenty of "wriggle room" in clinical trials
  - Major investment in flu vaccine despite doubts
- But still a worthwhile criterion
  - Lack of clinical trial scrutiny for hip replacement
- The policy sciences need this credibility


OIL
OIL
OIL
OIL
BULK