LSE CPNSS

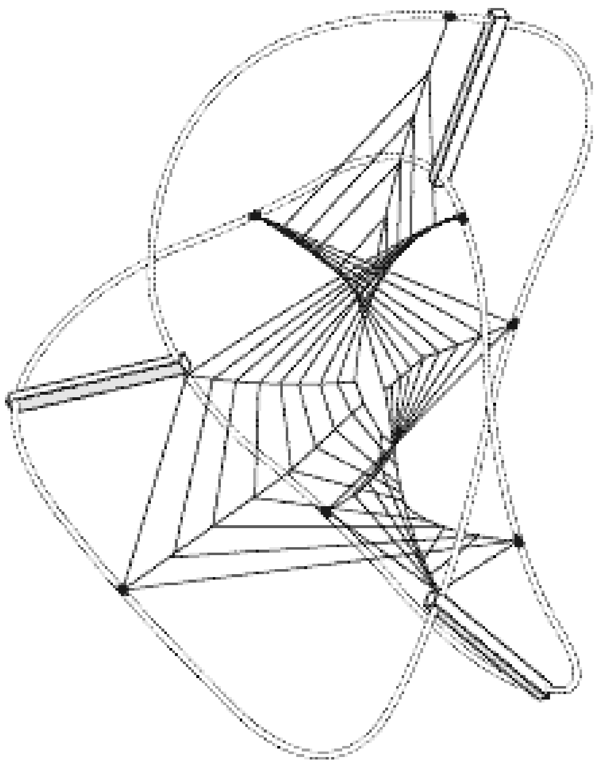
Centre for Philosophy of Natural and Social Science

Causality: Metaphysics and Methods

Technical Report 12/03

Practice Ahead of Theory: Instrumental Variables, Natural Experiments and Inductivism in Econometrics

Julian Reiss



Editor: Julian Reiss

Practice Ahead of Theory:

Instrumental Variables, Natural Experiments and Inductivism in Econometrics*

Julian Reiss
Centre for Philosophy of Natural and Social Science
London School of Economics
Houghton St
London WC2A 2AE
j.reiss@lse.ac.uk

October 2003

* Work on this paper was conducted under the AHRB project *Causality: Metaphysics and Methods*. I am very grateful to the Arts and Humanities Research Board for funding. Thanks to Nancy Cartwright for comments on an earlier draft.

Abstract

This paper argues that the field of econometrics divides into camps of sinners and preachers. Specifically, I look at two sins condemned by the preachers: the making of causal background assumptions when identifying parameters for estimation and the estimation of parameters that have no or little theoretical significance. I argue that practitioners in the natural experiments literature happily commit these sins and that this is for the good of the discipline.

We comfortably divide ourselves into a celibate priesthood
of statistical theorists, on the one hand,
and a legion of inveterate sinner-data analysts, on the other.
The priests are empowered to draw up lists of sins
and are revered for the special talents they display.
Sinners are not expected to avoid sins;
they need only confess their errors openly.

E. Leamer

1 Introduction

I purport to fulfil two aims with this article. The first, more general aim is to expose John Stuart Mill's scepticism of the inductive method in political economy as premature. He argued that, since there is no "crucial experiment" in the moral sciences, they could not be built bottom-up from experimental facts to more general truths. Against this, I maintain that, whether or not there are "crucial" experiments in political economy, econometricians have developed methods that mimic experiments, hence allow causal inference and hence make an inductive economic methodology possible.

The second aim relates to the details of the econometricians' stand-ins for experiments. As the quote from Edward Leamer, which makes the epigram for this article, illustrates, econometric methodology divides into a set of official dogmas that in their totality make up the saintly textbook scriptures and an implicit methodology found in sinful practice. That second aim, then, is to revere the sinners and damn the priests. Till not long ago, it was part of the official doctrine to identify parameters of interest on the basis of knowledge about correlations only. In practical applications, however, free use of institutional, historical and in particular causal knowledge is made. I want to argue that if the official methodology was followed, a particular technique in econometrics would yield misleading conclusions. In practical applications, by contrast, econometricians do not follow the official line but exploit the right kind of non-correlational knowledge to identify their parameters correctly.

The two seemingly disparate claims are in fact connected. This is due to a second sin theoretical econometricians tell us not to commit: to estimate parameters that have no theoretical significance. At least since the heyday of the econometrics of the Cowles Commission, it has been received wisdom that econometrics is the junior partner of economic theory and that its function is to test implications of this theory and to estimate theoretically identified economic parameters. In the case studies I am going to look at, little respect is paid to theory; right to the contrary, the parameters that are estimated are hardly related to theoretical claims at all. This is, then, a second way in which practice is ahead of theory: practice implicitly commits to an inductive methodology similar to the kind I am trying to defend in this paper.

Before going into the details of all this, I will provide some background by reviewing Mill's argument against inductivism in political economy and why it is still

relevant in section 2. Sections 3 and 4 provide the main discussion regarding the saints and sinners story. Section 5 wraps it up by projecting an inductive methodology based the mimicked econometric experiments of sections 3 and 4.

2 Background: Mill's Methodology

In the view of many an economist and methodologist, economics is an abstract and deductive science. It is abstract in a twofold sense. First, its basic principles do not describe actual states of affairs but "pure" states of affairs that would obtain if it were not for the operation of disturbing factors. Second, the basic principles are unspecific and stand in need of fitting in for concrete results. Take as a basic principle the desire to seek wealth or the rule to maximise utility. In order to derive anything from these principles, it must be specified what the object consists in in the concrete case and what the alternative means are towards the object. Further, economics is deductive in that concrete or, at any rate, more concrete results are derived from the basic principles in conjunction with assumptions about the structure of an economic system of interest deductively. Evidence at best plays the role of testing whether in one's deductions a disturbing factor has been left out.

This image of economics dates back to at least the writings of John Stuart Mill.[1] He constructed it because he thought economics could not be an inductive science, building up from particular items of evidence to general knowledge. His argument against induction in economics makes two claims. First, Mill says, the economic world is too complex to make inductive inference. This claim is based on his method of difference which presupposes there being two situations which differ only with respect to presence of the one causal mechanism of interest. To find these situations, he then claimed, in political economy would be excessively rare because of the complexity involved.

The second claim is that experimental control is hardly ever possible in the social world. In natural phenomena, even if their natural occurrence is complex, single causal mechanisms can be isolated experimentally and hypotheses tested on the basis of the experimental set up. Now given the social world is complex and it does not allow of experiments, the prospects for the inductive method are dire. However, Mill further claims that the basic principles of economics are known *a priori*. This has to do with the fact that it is the science of an aspect of human conduct, and we are all first-hand acquainted with that. The only difficulty that remains is to find all the disturbing factors that will get us from the basic principles to the concrete circumstance. To summarise, then, Mill's argument is this. Science proceeds either inductively or deductively. For the stated reasons, economics cannot be an inductive science. Thus, if it is a science at all, it has to proceed deductively. In economics we are lucky in so far as its basic principles are known *a priori*. Thus economics not only must be but it also can be a deductive science.

Even a cursory glance at modern economics shows that by and large economists

[1] See especially his 1948/1830.

still follow Mill's precepts.[2] To be fair, his "fundamental laws of behaviour"—the desires to seek wealth and avoid work—have been replaced by the principle of utility maximisation, and his demand to deduce more concrete results from the abstract principles has been made more precise by the demand to provide a mathematical model of the concrete situation, but the general outlook remains the same.

The purpose of this paper is to show that Mill's argument is wrong. While I do agree with the truism that the social world is complex and that conducting controlled experiments is hardly ever an option, inductive methods have improved since his day, and nowadays there are substitutes for controlled experiments, which enable inductive causal inference even in complex situations where no control is possible. The premiss in his argument that I am attacking is thus the implicit premiss that inductive inference presupposes either natural or experimental isolation of a single causal mechanism.

The case studies I use to make the point about inductive inference are drawn from the natural experiments movement in econometrics. This movement can be regarded exactly as an answer to the challenge Mill posed. Suppose we would like to know whether military service causes earnings later in life. Finding a correlation between the two variables would not suffice even if we can assume that earnings do not cause military service because of the temporal order. This is due to the reasons Mill gives. There is simply a myriad of factors that could be responsible for earnings, some of which may be related to earnings too. Suppose that, on average, fitter people both tend to join the army and tend to have higher wages. A positive correlation, then, would be indicative of this common cause rather than of a direct causal connection.

Ideally, hence, we would divide up (a representative sample of) the male population into two groups such that all other causes of earnings are identically distributed and send one group to serve and one group not to. But, as Mill noticed, such a social experiment is, for whatever reasons, impossible. Proponents of *natural* experiments now aim at finding situations that resemble a genuine controlled experiment without explicit control. For example,[3] the draft for the Vietnam war was organised by assigning numbers from 1 to 365 at random to every day of a given birth year of potential recruits. The men were potential draftees up to a certain threshold number, *e.g*. 195 for men born in 1950 and 125 in 1951. Assignment to draft status, thus, can be regarded as independent of other causes of the effect. Now if there remains a correlation between draft status and earnings, there is reason to interpret it causally.

Natural experiments are a species of a broader approach in econometrics, called instrumental variables estimation. This approach is used when a relationship between two variables, say, X and Y is potentially confounded by other variables. The instrument is defined as a third variable Z which has two characteristics: it is correlated with X but uncorrelated with the error term of Y. In the following section I want to demonstrate that the instrumental variables approach makes valid causal

[2] See for example the discussion in Hausman 1992.
[3] See Angrist 1996.

inference only under a set of very strong assumptions about the socio-economic system studied and that it is epistemically preferable to make explicitly causal assumptions.

3 Can We Get Causes from Statistics?

"Correlation is not causation" is a well-rehearsed slogan. The purpose if this section is to argue that prior knowledge about causal relations is not only necessary to identify causal parameters from statistics but that it is also often easier to have than knowledge about correlations.

The reason to re-rehearse the slogan is that econometrics textbooks tend to avoid explicitly causal language in their descriptions of estimation procedures. For example, in simultaneous equations, variables divide into endogenous and exogenous variables, and that means that they are either determined within the model or outside it. Importantly, without qualification, the "determined" can be read both functionally or causally. Or sometimes a model is called structural if its form is given by the underlying theory. But no word is lost on whether theory specifies functional or causal relations. To give a third example, the error term in a regression model is often said to represent "stochastic disturbances" or "shocks" as well as omitted factors, measurement error and other "influences". Again, it is unclear whether these terms carry a causal meaning or not.

The neglect of causality in econometrics has its roots deep in the origins of modern statistics itself. Thus Karl Pearson, famous for being one of the founders of modern statistics, is at least equally famous for his attacks against causal language and his substitution of the Humean concepts of association and correlation for causal concepts.[4] For him, causal language was the language of a bygone age; the language of the new era is that of association and correlation.

This sentiment carried over to early econometrics. For example, in one of his volumes on business cycles, Wesley Mitchell muses:[5]

> In the progress of knowledge, causal explanations are commonly an early stage in the advance toward analytic description. The more complete the theory of any subject becomes in content, the more mathematical inform, the less it invokes causation.

Surveying more recent contributions (a growing number of exceptions notwithstanding), one still gets the impression that, in order to be scientific, causal language has to be avoided, and the Humean language of correlation and association employed.[6] In particular, many textbooks contain "recipes" for econometric inference that give the impression that econometrics can proceed without causal background assumptions.

Suppose we are interested in the old monetarist question of whether money (X) causes income (Y). We can expect the estimator $\hat{\beta}$ of the causal effect of X on Y of

[4] See *e.g*. Pearson 1911.
[5] Mitchell 1927, p. 55, quoted from Hammond 1996, p. 10
[6] See also Pearl 1997 who makes a similar remark about statistics in general.

the standard regression model:[7]

$$Y = \alpha + \beta X + \varepsilon \tag{3.1}$$

to be biased for a number of different reasons. First, it is possible that X is measured with error, and that the measurement error is correlated with Y's error term ε. That X is measured with error is more than likely; it has been a matter of dispute what the "right" definition of money should be for about as long as economists thought about the quantity theory. The construction of the money variable has also been one of the major criticisms of Milton Friedman's epic studies of the role of money in the economy. Second, the relationship can be confounded by an unobserved variable. (3.1) is simplified in so far as the vector of known confounders (*e.g.* a time trend) has been omitted as a regressor on the right hand side. However, not all potential confounders can be measured, and thus there might always be residual correlation between X and ε. Third, there may be feedback effects from income to money. This is also a point that came up in the debate about Friedman's analysis of the role of money. Critics tended to understand the monetarist position to assert that money is the only cause of economic activity, but Friedman emphasised time and again that he thinks of money merely as one cause (the principal cause, however), and that it is well possible, especially in the short run, that causality runs from income to money.

For any of these three reasons, X might be correlated with ε, which biases the estimator for β. A standard technique to solve that problem is so-called instrumental variables estimation. An instrumental variable Z has two characteristics:

IV-1 Z is correlated with X

IV-2 Z is uncorrelated with ε.

If such a variable can be found, an estimator of the parameter of interest can be constructed as follows:

$$\hat{\beta} = \mathrm{corr}(Z, Y)/\mathrm{corr}(Z, X). \tag{3.2}$$

Now, depending on the additional assumptions made about the relations between the variables of interest, the definition of an instrumental variable may be too weak for a causal interpretation of the estimand. To see that most clearly, let us proceed in three stages with increasingly stronger assumptions made about the relations between the variables. The basic structure is the following:

[7] The hat signifies that this is the *estimator* of the true parameter rather than the true parameter itself.

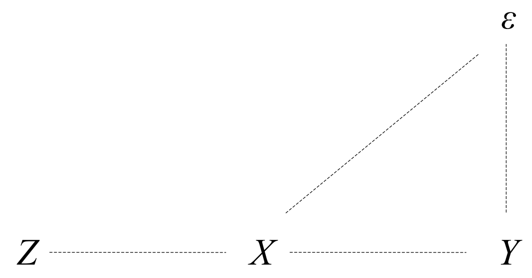


Figure 3.1 Instrumental Variables—The Basic Structure

X and Y are the variables of interest, Z is the purported instrument and ε is an error term. A dashed line marks correlation between two variables. Y's error term is correlated with X, which is why we need an instrument to estimate β in the first place.

The first assumption that must be made is a Reichenbach principle, which takes us from correlations to causal relations. This is never really discussed within econometrics but without it, no causal inference can be made in observational studies. If we allow for "brute correlations", *i.e.*, correlations that cannot be accounted for in terms of causal relations, then any attempt to infer causal relations from statistics—no matter how strong the additional assumptions about the system of interest—is futile. Any residual correlation could then always be "brute". An almost identical remark make Dan Hausman and Jim Woodward:[8]

> So when X_j [the putative effect variable] "wiggles" in response to an intervention with respect to X_i [the putative cause variable] (or in general without any change in its parents) one has a covariation between X_i and X_j that cannot be explained by X_j causing X_i or by a common cause of X_i and X_j [this follows from their definition of an intervention]. If there are no restrictions on the ways that a correlation between X_i and X_j might arise, one could go no further. But if one takes **CM1** [their version of the Reichenbach principle] for granted, then one can conclude that X_i causes X_j.

The particular version of the Reichenbach principle I am going to assume is the following:

RP Any two variables A and B are correlated if and only if either (a) A causes B, (b) B causes A, (c) a common cause C causes both A and B or (d) any combination of (a) – (c).

Let us call two variables that are related by any of the causal structures (a) – (d) *causally connected* (and *causally unconnected* in case they fall under none of the structures). I do not want to argue for the truth or plausibility of the particular version of the principle here. But let me stress again that we need some principle of the kind, and since investigations in economics typically face causally complex systems, we

[8] Hausman and Woodward forthcoming, p. 5

cannot, for example, exclude *a priori* the possibility of simultaneous causation or any other combination of the simpler structures (a) – (c).

The second general assumption we must make is the transitivity of causal relations. Again, without it, any attempt to infer causal relations from statistics would be futile. Thus:

> **T** For any three variables A, B and C, if A causes B and B causes C, then A causes C.

I will restrict the analysis to linear equations of the general form $X_j = \alpha_{j1}X_{j1} + \alpha_{j2}X_{j2} + \ldots + \alpha_{jn}X_{jn} + \varepsilon_j$. For these, I define the concept of "functional correctness of a structural equation" as follows:

> **FC** A structural equation $X_j = f(X_{j1}, X_{j2}, \ldots, X_{jn}, \varepsilon_j)$ is functionally correct if and only if it represents the true functional (but not necessarily causal) relations among its variables.

On the basis of these, let us examine what further assumptions we must make about the system of interest in order for the instrumental variables technique to yield causally correct conclusions. I will proceed in three stages, each with different causal assumptions added to the general assumptions.

Stage 1: No assumptions added.

If no causal assumptions are added to **RP**, **T** and **FC**, one can easily show that the instrumental variables technique can yield causally incorrect conclusions. The system consists of two functionally correct equations:

$$Y = \alpha X + \varepsilon \qquad (3.3)$$

$$X = \beta Z + \varepsilon. \qquad (3.4)$$

The claim is that Y is correlated with Z if and only if X causes Y. However, it is possible that the correlation between X and Y arises from other causal relations. Consider the structure in Figure 3.2.

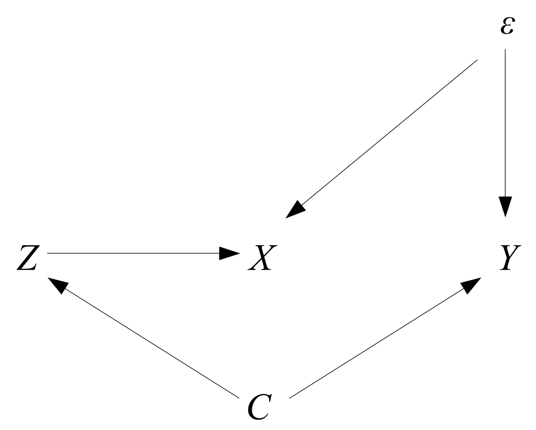


Figure 3.2 A counterexample to IV under RP, T and FC

Nothing in the assumptions prevents this situation. In addition to (3.3) and (3.4), we now have:

$$Z = \gamma C + \mu \text{ and} \tag{3.5}$$
$$Y = \delta C + v. \tag{3.6}$$

Since C is a common cause of Z and Y, these two variables will be correlated. Further, Z causes X, which also implies that they are correlated. Z and ε may or may not be correlated as can be seen from the following derivation:

$$Z = \gamma/\delta Y - \gamma v/\delta + \mu \text{ [substitute } C \text{ from (3.6) into (3.5)]}$$
$$= \gamma\alpha\beta/\delta Z + \gamma/\delta(1 + \alpha)\varepsilon - \gamma v/\delta + \mu \text{ [substitute for Y from (3.3) and (3.4)]}$$
$$= \zeta(1 + \alpha)\varepsilon - \zeta v + \zeta\delta/\gamma\mu \text{ [define } \zeta = 1/(\delta/\gamma - \alpha/\beta) \text{ and rearrange].}$$

$$\mathrm{E}(Z\varepsilon) = \zeta(1 + \alpha)\,\mathrm{E}(\varepsilon^2) - \zeta\,\mathrm{E}(\varepsilon v) + \zeta\delta/\gamma\,\mathrm{E}(\varepsilon\mu), \tag{3.7}$$
[multiply by ε and take expectations]

which may or may not be zero. It seems unlikely that such exact cancellations should occur but it is not impossible. One way to ensure that exact cancellations do not obtain that has been offered in the literature is to assume that correlations are stable under parameter changes.[9] In this example, we see that the lack of correlation between Z and ε obtains only under a very specific parameterisation (one that makes (3.7) true).

Alternatively, we can assume that ε represents the net effect of *all other causes* on X and Y. The above counterexample could not obtain because there could not be a cause of Y, C, which is not represented by ε. That the technique yields causally correct conclusions in general I have proved in stage 2.

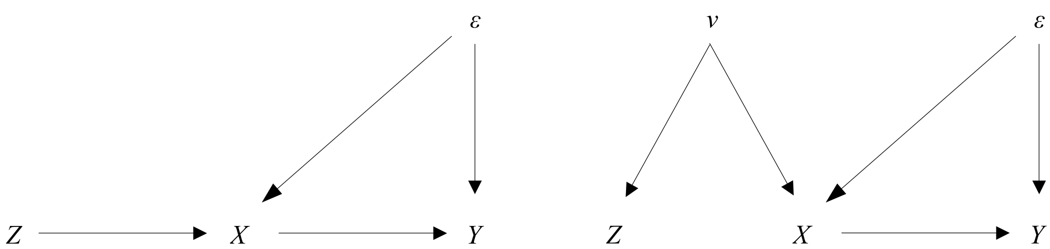
[9] Pearl 2000, p. 48

Stage 2: The error term represents the net effect of all other causes (except the ones modelled). If Z and Y are correlated, RP tells us that they must be causally connected. Since there is a third variable, X, I will first show that Z and Y cannot be causally connected on a route that excludes X.

Can Z and Y be causally connected on a route that excludes X? Under the given assumptions, Z cannot cause Y either directly or otherwise via a route that excludes X because its influence would be represented by ε, which we have excluded. Y cannot cause Z unmediated by X either, since by transitivity ε would cause Z and by **RP** they would be correlated. Nor can there be a common cause (that causes Z and Y on a route that excludes X) because its influence would be represented by ε. Hence, the causal connection between Z and Y must run through X.

Hence there remain three possibilities: Y causes Z through X, X is a common cause of Y (or a common cause causes Z and Y through X) and Z and Z causes Y through X. Y cannot cause Z through X because, as before, that would imply that ε is correlated with Z. A common cause now may either be between X and Y or between X and Z. If it is between X and Y (which means that its influence is represented by ε), then X must cause Z and thus ε causes Z by transitivity and would be correlated with it. However, there may be a common cause between Z and X. In this case Z and Y would be correlated if only if X causes Y, and thus Z would be a valid instrument.

The last remaining possibility is that Z causes Y via X. Again, in this case Z and Y would be correlated if and only if X causes Y, and Z is a valid instrument. The two causal structures in which Z is a valid instrument are depicted in Figure 3.3.



(a) Z causes Y via X.

(b) A common cause causes Y via X

Figure 3.3 Two Valid Instruments to Identify *β*

Thus, under the above assumptions, the instrumental variables technique makes causally correct inferences. The difficulty with the technique is, however, that it is, taken literally, unoperationable. By its very nature, the error term is ε unobservable. Hence it is not possible to test statistically whether Z is or is not uncorrelated with the error term. This a problem in particular because even slight correlations can severely bias the estimator.[10]

The irony is that one motivation to use the instrumental variables technique is

[10] See Pearl 1993, Bartels 1991.

exactly that there are unobservable common causes between *X* and *Y*. Now, if that is so, it seems hard to see how an instrument in the sense of a variable that satisfies **IV-1** and **IV-2** should be found (if a variable is not measurable, *a fortiori* its correlation with another variable is not measurable).

Stage 3: Assuming Z is an "experimental handle".

In practice, econometricians identify an instrument on the basis of causal background knowledge, which often derives from an institutional and/or historical analysis of the case at hand. Statisticians sometimes judge background assumption of this kind inadmissible, "subjective".[11] But apart from the fact that the whole approach would be ineffective unless these background assumptions (in addition to the general assumptions) are made, there are two further justifications for making the causal assumptions explicit. First, the structure of assumptions shows that the system *Z-X-Y* is equivalent to an experimental set-up in which *Z* is used as a "switch variable" to test the hypothesis whether *X* causes *Y*. Let us call a variable Z an "experimental handle" if it satisfies the following assumptions:[12]

EH-1 *Z* causes *X*

EH-2 *Z* causes *Y* if at all only through *X* (*i.e.*, not directly or via some other variable)

EH-3 *Z* and *Y* do not have causes in common (except those that might cause *Y* via *Z* and *X*).

Under **RP**, **T** and the assumption that error term represent the net effect of all other causes, these assumptions are almost equivalent with **IV-1** and **IV-2**. The only difference is that an instrument does not have to *cause* the putative cause; it merely cannot be its effect (**EH-1** rules out the structure given in Figure 3.3b).

But **EH-1 – 3** make the similarity with a controlled experiment more easily visible. A randomised clinical trial (RTC) is a paradigmatic example for a controlled experiment. In an ideal[13] RTC, the random allocation is the experimental handle *Z*. It (and only it) causes whether or not a subject will receive treatment (*X*). It does not have a causal influence on recovery (*Y*) which is not mediated by *X*. And the allocation itself is not caused by anything that also causes recovery. It is in fact **RP** which ensures that randomisation is successful (if the random allocation "happened" to produced biased groups, the resulting correlation of a cause of recovery with the allocation would have to have a causal explanation, which would violate either **EH-2** or **EH-3**).

In the literature about scientific experiments, "switches" or "handles" are similarly defined. For instance, James Woodward defines an intervention *I*, his equivalent, as follows:[14]

---

[11] Even Judea Pearl once seems to have thought this, see his 1993.

[12] *Cf.* Cartwright forthcoming.

[13] I qualify RTC here in order to avoid discussing problems with compliance *etc*.

[14] Woodward 1997, p. S30

1. I changes the value of X possessed by U_i [the "experimental unit"] from what it would have been in the absence of the intervention and this change in X is entirely due to I.[15]
2. I changes Y, if at all, only through X and not directly or through some other route. […]
3. I is not correlated with other causes of Y besides X (either via a common cause of I and Y or for some other reason) except [those that cause Y via Z and X].[16]

Thus, the instrumental variables approach receives additional credence by identifying it as a species of experimentation. The second justification is that **EH-1** to **EH-3** give us a nice algorithm to assess the design of studies using the instrumental variables technique. As I will discuss in detail below, James Heckman criticised Joshua Angrist's study of the causal effect of veteran status on civil earnings on the basis of what he calls a "behavioural model" that relates random draft lottery number (the putative instrument Z) with investment in training (a cause of the putative effect Y).[17] Such a relation would clearly invalidate the results of Angrist's study. Using the experimental handle interpretation, we have a simple algorithm to test for potential failures: list all other causes Y besides X and test for each whether Z is causally related to it. Importantly, there are more tests available than statistical tests. In particular, causal hypotheses can often be tested without the requirement that variables involved are measurable in the statisticians' sense. For example, there may simply be no data on training investments available. But it could still be known by other means whether or not employers typically invest more in people with high draft lottery numbers.[18]

The purpose of the next section is to show that indeed in practical applications, the assumptions made to identify the instrument are **EH-1** – **EH-3** (rather than the assumptions about correlations).

4 Natural Experiments

In this section I want to give the evidence that in their applications, many econometricians do not eschew causal language, but right to the contrary, make precisely those causal background assumptions that are needed to turn the "instrument" into an "experimental handle" in the sense of the previous section. These assumptions are thus sufficient to identify an instrument in the (theoretical) econometrician's sense but not necessary as many other causal structures would imply the same correlations; and they are both necessary and sufficient for the identification of an experimental handle.

I am going to report three brief case studies. The first is probably most naturally

[15] The counterfactual formulation of this condition, though central to Woodward's overall approach, is not relevant for the purpose of this paper.
[16] As is evident from Woodward's clause 3, he does not assume a Reichenbach principle. However, as argued above, this is necessary to move from statistics to causes at all.
[17] See *e.g.* Heckman 1996.
[18] Heckman's argument is theoretical: because employers are rational agents, they make use of all available information, and that includes the draft lottery number. But there are surely more empirical means to determine the existence of that link that, nonetheless, fall short of a full-blown statistical test.

regarded as a "natural" experiment as there are two populations, which (so the authors assume) can be regarded as identical with respect to the distributions of the causes of the putative effect but which occur naturally, and one of which receives the treatment while the other does not. The second case resembles a "randomised" experiment because treatment and control group are allocated using a random mechanism, in this case a lottery. From the point of view of the econometrician, this still constitutes a natural experiment as he does not influence the allocation mechanism but rather merely observes allocation and outcome. In the third case the experimental handle is constructed out of observed data. This case of an experiment is probably most remote from the usual reading of a natural experiment because a model is used in order to measure the instrument in the first place. However, due to the nature of the assumptions involved the case can be made that the procedure still constitutes a natural experiment.

Case 1: Do Minimum Wages Cause Unemployment?

The standard competitive model of neoclassical economics implies that a minimum imposed on the wage rate will, if higher than the competitive rate, reduce employment. David Card and Alan Krueger[19] challenge this prediction or, to be more precise, the universal character of the prediction. The natural experiment they observe and from which they extract information that in their view conflicts with the competitive model is an increase in New Jersey's minimum wage from \$4.25 to \$5.05 per hour on April 1, 1992. They attempt to measure the impact of the increase on employment by comparing employment growth at fast-food restaurants in New Jersey and neighbouring eastern Pennsylvania. They conclude that there is no indication in the data that would confirm the standard model.

Relevant for the context of this paper is Card and Krueger's use of background assumptions for the identification of an experimental handle. Recall that an experimental handle Z has the properties (i) Z causes X, the putative cause of the effect of interest Y, (ii) Z causes Y if at all only via X and (iii) Z and Y have no causes in common (except causes of Z that cause Y via Z and X). It must further be assumed that operating the handle does not alter the causal laws that have X, Y or Z as cause or effect. Since the causal relation of interest here is that between an enforced level of wages and employment, I suggest to regard the minimum wage legislation as the instrument or putative experimental handle, actually paid wages as the putative cause and employment as the putative effect.

(a) Does the minimum wage law cause wages? To make sure that the first assumption is fulfilled, Card and Krueger choose the experimental population such that the right causal relation can be expected. Their population consists of 410 fast-food restaurants in New Jersey and Pennsylvania because: "fast-food stores are a leading employer of low-wage workers" (thus an increase in the minimum will be relevant), and "fast-food restaurants comply with minimum-wage regulations and

[19] Card and Krueger 1994 and 1995

would be expected to raise wages in response to a rise in the minimum wage".[20]

(b) Does the legislation cause employment through other channels than the actual wage? The most important confounder in this category would probably be if the minimum wage bill had been passed along with other bills that were likely to affect employment at the same time (for instance, if it had been passed along with a second bill that would changed health and safety requirements for employees in low-wage occupations). This was not the case, however.

Another channel through which legislation could affect employment is expectations. Card and Krueger measure the difference in employment growth between the time just before the actual increase and about half a year after the increase. Since the bill was introduced more than two years before the actual increase, however, it is possible that all its employment effects have already occurred before they even measured because employers expected the wage to increase and adjusted payrolls accordingly.

Card and Krueger were helped in this matter by the fact that the minimum wage bill remained contested till almost the time of the actual increase:[21]

> The scheduled 1992 increase gave New Jersey the highest state minimum wage in the country and was strongly opposed by business leaders in the state…
>
> In the two years between the passage of the $5.05 minimum wage and its effective date, New Jersey's economy slipped into recession. Concerned with the potentially adverse impact of a higher minimum wage, the state legislature voted in March 1992 to phase in the 80-cent increase over two years. The vote fell just short of the margin required to override a gubernatorial veto, and the Governor allowed the $5.05 rate to go into effect on April 1 before vetoing the two-step legislation. Faced with the prospect of having to roll back wages for minimum wage earners, the legislature dropped the issue. Despite a strong last-minute challenge, the $5.05 minimum rate took effect as originally planned.

(c) Does employment have any of its causes in common with the wage legislation? The easiest way to test this is by checking whether any of the causes of the minimum wage legislation also affects employment through a different channel than the actual wage.

Historically, minimum wages have often been raised in times of favourable economic conditions in order to minimise the potentially adverse effects of the increase. Therefore, in these cases results would be confounded by a common cause between legislation and employment, *viz*. "economic conditions". Card and Krueger control for this potential cofounder by employing the so-called differences-in-differences method. That is, they do not merely compare employment before and after the increase but rather employment growth in New Jersey with employment growth in eastern Pennsylvania. Since, so they assume, economic conditions should affect both states in the same way (they judge this on the basis of observing the socio-economic structure in the two states), differencing should filter out all causes of employment that the two states share, as for instance economic conditions or the season.

Further, Card and Krueger's project was once more assisted by the course of

[20] Card and Krueger 1994, pp. 773f.
[21] *ibid*., p. 773

events. Whatever may have been the original rationale for increasing the minimum wage, at the time the final decision was made (just before the actual increase, in March 1992), the economy was in a recession and thus economic conditions, if a cause at all, would rather have prevented the introduction of the new wage law.

Under these assumptions Card and Krueger find that the increase of the minimum wage actually raised rather than lowered employment in New Jersey, both as compared to Pennsylvania as well as to high-wage stores in New Jersey. Far be it from me to pretend that their results are unexceptionable. It has, for example, been claimed that the results are reversed if instead of the employment data from Card and Krueger's telephone survey payroll data are used.[22]

But the point to make here is not whether or not minimum wages actually cause an increase or a decrease in employment but a methodological point about the nature of the assumptions econometricians make to identify a valid instrument. Very clearly, the assumptions Card and Krueger make are not only sufficient to identify an instrument (*i.e.* that wage legislation is correlated with the actual wage increase but uncorrelated with the causes of employment other than the wage increase) but just the right assumptions to identify an experimental handle. And therefore, the causal conclusions they draw are valid conditional on the correctness of their assumptions.

Case 2: Does Military Service Cause Later Civilian Earnings?

For reasons of social justice, governments aim at compensating veterans for losses they suffer due to their serving in the military. One of the potential losses due to military service lies in lower wages veterans might receive in later civilian life. One cannot measure the impact of serving in the army on civilian earnings by simple regression of earnings on a dummy indicating whether or not someone has served because there are factors that may affect both variables and thus bias the estimator. For example, men may be more inclined to serve in the military because they have fewer job market opportunities, which also affects their future earnings.

Joshua Angrist[23] attempts to solve that problem by exploiting a random lottery that was conducted in order to prioritise draft eligibility for the Vietnam war. In particular, each birth date of a cohort of a given year was assigned a random sequence number (RSN) from 1 – 365. Depending on the Defense Department manpower needs, a ceiling was determined, i.e. a number up to which men were drafted. In this example, the instrument is the RSN, the putative cause is serving in the military and the putative effect civilian earnings.

(a) Does the RSN cause military service? Angrist estimates that draft eligibility raises the probability of veteran status by 10 – 15 percentage points for whites born 1950 – 52.[24] Given it can be assumed that the lottery is truly random and thus veteran status does not cause the RSN nor is there a common cause and assuming the Reichenbach principle, information about the probabilities would already be sufficient

[22] Neumark and Wascher 2000
[23] Angrist 1990
[24] See *ibid*., Table 2, p. 321.

to establish that the RSN causes serving in the military.

(b) Does the RSN cause earnings through channels other than serving in the military? As I will discuss momentarily, this is an assumption that actually was contested later on. However, once more the point is not to justify any of these studies but to investigate the kind of assumptions authors make in order to identify the instrument they use. With respect to this assumption, Angrist says:[25]

> The justification for estimation of the effects of military service in this manner is clear: it is assumed that nothing other than difference in the probability of being a veteran is responsible for differences in earnings by draft-eligibility status.

(c) Does the RSN have any causes in common with earnings? Due to the random nature of the lottery, this assumption is also particularly easy to defend. If it is truly random, it is not caused either by a factor that also causes earnings nor by a cause that causes it through earnings.

On the basis of these Angrist concludes that[26]

> Estimates based on the draft lottery indicate that as much as ten years after their discharge from service, white veterans who served at the close of the Vietnam era earned substantially less than nonveterans. The annual earnings loss to white veterans is on the order of $3,500 current dollars, or roughly 15 percent of yearly wage and salary earnings in the early 1980s.

James Heckman challenged Angrist's investigation in a way that is also illuminating from our point of view. He argued that the RSN is not a proper instrument because among other things "persons with high numbers are likely to receive more job training, because their likelihood of being drafted is reduced and firms have less likelihood of losing them".[27] That is, Heckman believes that the RSN causes earnings not only via actual military service but also whether firms expect employers to have to serve in the military and subsequently base their training decisions on their expectations. In both the original article and the criticism it is thus causal background knowledge that identifies the instruments rather than mere information about correlations.

Case 3: Does Money Cause Interest Rates?

The third case differs from the others in two respects. First, confounding in this case is not due to one or more common causes but due to the fact that there is simultaneous causation between the variables of interest. Or more precisely, it is known that the putative effect is a cause of the putative cause, and hence if there is causation in the direction of interest, then there is simultaneous causation. Second, the instrument is not straightforwardly observable but constructed by the econometrician.

James Hamilton attempts to measure the "liquidity effect", *i.e.* the effect a change in the money stock has on interest rates.[28] More specifically, the variables of interest are non-borrowed reserves (the putative cause) and the federal funds rate (the putative

---

[25] *ibid*., p. 320
[26] *ibid*., p. 330
[27] Heckman 1996, p. 461
[28] Hamilton 1997

effect). An ordinary regression of the form

$$i_t = \lambda R_t + \boldsymbol{\gamma}'\mathbf{x}_t + v_t^i, \tag{4.1}$$

where i is the federal funds rate, R reserves and $\mathbf{x}$ a vector of "other causes", will result in a biased estimator λ because the Federal Reserve actively chooses the supply of reserves each day in response to the value of the federal funds rate using a reaction function of the form

$$R_t = \delta i_t + \boldsymbol{\psi}'\mathbf{x}_t + v_t^R. \tag{4.2}$$

Hamilton exploits an institutional detail to solve the simultaneous causation problem. The US Treasury maintains an account with the Federal Reserve system which fluctuates in response to fiscal receipts and expenditures. Checks drawn on private banks reduce their reserve requirements the moment they are deposited into the Treasury's account. The Federal Reserve, in turn, attempts to take this effect into account and predicts daily fluctuations before announcing reserve requirements. However, it cannot do so perfectly and therefore the prediction error induces variation in the reserves available to banks independently of monetary policy. Hamilton proposes to take these forecast errors as instrument for measuring λ.

(a) Do the Forecast Errors Cause Reserves? As just stated, institutional background knowledge justifies the assumption that the Fed's forecast errors cause the reserves available to private banks. The Fed, however, does not publish their forecasts. Hamilton therefore constructs a forecast model whose residuals provide estimates of the actual Fed's forecast errors. Essentially, he estimates a third-order autoregression for the Federal Reserve deposits held by the Treasury, which is allowed to have different coefficients when the previous day's Treasury balance exceeds a certain level. In a regression of non-borrowed reserves on changes in the federal funds rate and other explanatory variables that include the forecast errors, the errors are highly correlated with reserves. Hamilton thus concludes:[29]

> There is no question that the forecast errors of my model of the Treasury balance prove to be an excellent instrument for nonborrowed reserves, in the sense that $\hat{\varepsilon}_t^U$ [the forecast error] *exerts a highly significant effect on* R_t.

(b) Do the forecast errors cause the federal funds rate via channels other than the non-borrowed reserves? Hamilton clearly proceeds on this assumption:[30]

> The claim will be that the error the Fed makes in forecasting the Treasury balance matters for the federal funds rate only insofar as it affects the supply of reserves available to private banks.

(c) Are there any causes common to the forecast errors and the federal funds rate? There is a pretty simple argument that there cannot be such common causes. The argument is that if there were such a cause, the Fed would make systematic errors in

---

[29] *ibid.*, p. 88, emphasis added
[30] *ibid.*, p. 82

their predictions of the Treasury balance. That, in turn, would imply that Hamilton's model (if the model represents the Fed's forecasts accurately) is misspecified. In order to avoid this kind of error, Hamilton tests his specification:[31]

> I investigated how good this specification is at describing the dynamics of the Treasury balance by using a battery of Lagrange multiplier tests. […]
>
> Hamilton (1996b) reports 43 separate Lagrange multiplier tests of the specification in Table 2, of which three are significant at the 5-percent level and none at the 1-percent level. I conclude that the model of Table 2 provides a good parsimonious description of the series U_t.

And hence he can assume:[32]

> However, the Fed is not able to anticipate the end-of-day Treasury balance perfectly, with the result that this variable induces substantial daily fluctuations in the reserves available to private banks for reasons that are completely exogenous with respect to monetary policy.

5 Mill, Experimentalism and Beyond

One purpose of the past two sections was to demonstrate that Mill's rejection of the inductive method in economics was premature. I do, of course, not dispute that causal inference remains a hard task and that variables that constitute experimental handles may be very rare. But I think that the case studies of the previous section provide reason for at least an attempt to establish an inductive social science methodology. The purpose of this section is twofold. First, I want to sketch what the broader picture of such a methodology based on low-level causal regularities could look like. Second, I want to provide a brief argument that a kind of "bottom-up" methodology as proposed here is genuinely different from the methodology Mill defended.

To begin with, let us assume that we have at least one well-established causal law of the kind discussed above. Where do we go from here? The first thing to note is that, more likely than not, the law will be tied to a specific population and not be easily exportable to other situations. In other words, the law is a *ceteris paribus* law, where the *ceteris paribus* condition describes the test population, rather than a general tendency law.[33] Take the minimum wage example and assume that Card and Krueger established that the change in the minimum wage legislation of April, 1992, caused an increase in employment. In my view, this result can neither establish that increasing minimum wage will always *result in* an increase in employment nor that it will always *tend to* increase employment. Rather, what the increase in minimum wage will do will crucially depend on the setting in which it is embedded. For example, it will depend on how high the minimum wage is already relative to the average wage paid, on the industry, on the rate of unemployment that obtains, on compliance, on particular company policies and maybe on a host of other factors. This suggests that the second

[31] *ibid.*, p. 86
[32] *ibid.*, p. 81
[33] This was argued at length in Reiss forthcoming a.

step on the research agenda is to investigate the robustness of the law by testing it in various alternative situations.

Francis Bacon's philosophy of science provides a very simple starting point as to how to proceed here. His inductive method consists of a three-stage process of *observation*, *classification* and *causal inference*.[34] At this point, the second step, classification, is relevant. Bacon proposed that the phenomena that have been collected in natural histories be arranged in tables. He suggested three kinds of tables: the *Table of Presence*, the *Table of Absence*, and the *Table of Degrees*. The first table lists instances that are as varied as possible but where the phenomenon of interest is present. In our example, we might find cases of positive employment effects of an increased minimum wage in different countries, regions, industries, under different economic conditions (*e.g.* across the business cycle) and in different institutional settings (*e.g.* under different labour market rules). The second lists instances that are as similar as possible to the first set but where the phenomenon is absent. For example, we might find cases of a negative or zero employment effect in the same country, region, industry or under the same economic and institutional conditions where we previously found a positive effect. The third table lists cases of the phenomenon where it differs in degrees. Again, we might find a whole continuum of employment effects of a minimum wage change, and arrange them accordingly.

Observing what other factors co-vary with the employment effect might give us clues about the social structure or mechanism that is responsible for the causal law that we found in the first place. Hypotheses about such a mechanism can then be further tested, by means of laboratory experiments, social experiments or additional natural experiments.

There is no natural stopping point for this process of investigating deeper and deeper into the mechanisms of social causal laws. Assume for instance that the only relevant factor for differences in employment effects is industry—different industries have different wage elasticities of labour demand independent of other factors. But these industry-specific elasticities themselves will not be anything like "natural constants". Rather they will be subject to, say, secular trends or variation due to change in tastes, culture, laws and so on. How deeply the research is going to analyse a given phenomenon will then mainly be a pragmatic matter, dependent on interests as well as practical considerations.

If we allow for a sympathetic reading, this is roughly what Card and Krueger do in their monograph-length study of the minimum wage phenomenon. Among other things, they analyse another natural experiment in California, evaluate cross-state evidence from changes in the Federal minimum wage and differential impact on different groups, *e.g.* teenagers and employers in the retail industry and re-examine time-series, cross-section, panel-data studies as well as evidence from Puerto Rico, Canada and the United Kingdom.

On the basis of this evidence they go further and attempt mechanistic explanations for the observed employment effect differences. But at this point they

[34] For a discussion, see Reiss forthcoming b.

fall back into a Millian deductive methodology instead of further developing their empirical approach. In particular, they investigate what kinds of modifications the standard model requires in order to be consistent with the data. For example, they find that in a market with a single firm with monopsony power, a moderate increase in minimum wage can increase employment. The basic intuition behind models such as these is that when firms have monopsony power in their industry, a higher minimum wage might force them to fill vacations that otherwise would have been left empty. Another type of models they consider is where worker's levels of effort (and thus, productivity) are dependent on the wage paid. It can also be shown that in these so-called efficiency wage models a moderate minimum wage increase can lead to higher employment.

There is, however, little fine tuning between these models and their empirical findings. For example, they do not consider whether monopsony power of firms in industries they are looking at is really the factor that makes the crucial difference or whether workers are really motivated by a wage slightly higher than the current minimum. Nonetheless, I find their basic approach to be fairly in line with the inductivist stance taken here.

Maybe Card and Krueger's combination of inductive (natural experiments) and deductive (economic-theoretical models) suggests that the Millian approach is, after all, not so different from the approach defended in this paper. It is of course true that no scientific methodology is either purely inductivist or purely inductivist. All serious methodologies combine elements of both. But I think there is an appreciable difference between the research agendas implied in the two methodologies.

For Mill, the fundamental laws of political economy are known *a priori* in a particular sense. This sense deviates from the Kantian *a priori*, which means "prior to *all* experience" because it means only "prior to specific investigations in political economy". Thus prior to any specific investigation, we know the principal causes of any effect in political economy. These laws, the so-called "laws of human nature", were established by introspection and casual observation, *i.e.* inductively. Together with what one might call "general initial conditions" (conditions shared by many cases), one can deduce abstract truths from the laws. The abstract truths are applicable to concrete circumstances only with qualifications; the qualifications concern the operation of "disturbing causes". Had we knowledge about disturbing causes, or could we experimentally control for them, political economy would be a science as exact as the best branches of natural science. However, more often than not disturbing causes are not known, and almost never are they known before the event. Specific experience can still play a role, *viz.* that of verifying whether or not the major causes of a phenomenon have been taken into account.

The methodology proposed here shares with Mill's that background knowledge plays a significant role. But they differ both in the kind of background knowledge that is required as well as in the way it is exploited. For the methodology proposed here, one needs no substantial assumptions that transcend the inquiry at hand. All assumptions about the causal relations that are needed to identify an experimental

handle may be true of the particular situation under scrutiny only. By contrast, Mill's laws of human nature are (or so he assumes) true of any economy at any place and any time.

Further, an application of the methodology proposed here yields new knowledge about a particular causal parameters of interest. That is, on the basis of causal background assumptions measured correlations are given a causal interpretation. The result of Mill's methodology is a prediction about a phenomenon of interest. Evidence comes into play by either confirming or disconfirming the prediction. But importantly, neither in case of confirmation nor in case of disconfirmation success of lack thereof can be attributed to any particular assumption in the set. If the prediction is successful, the whole amalgam is regarded as confirmed while if it is false, we know only that one of the assumptions is false.

Mill argued:[35]

> [I have] now shown that the method *à priori* in Political Economy, and in all the other branches of moral science, is the only certain or scientific mode of investigation, and that the *à posteriori* method, or that of specific experience, as a means of arriving at truth, is in applicable to these subjects…

I hope to have convinced the reader that this conclusion of Mill's was a bit overhasty, and that it might be well worth investigating the alternative he rejects.

[35] Mill 1948/1830, p. 152

References

Angrist, Joshua (1990), "Lifetime Earnings and the Vietnam Era Draft Lottery: Evidence from Social Security Administrative Records", *American Economic Review* **80**:3, 313-36.

Blalock, Hubert (1979), *Social Statistics*, New York: McGraw-Hill.

Bartels, Larry (1991), "Instrumental and 'Quasi-Instrumental' Variables", *American Journal of Political Science* 35:3, 777-800.

Card, David and Alan Krueger 1994: "Minimum Wages and Employment: A Case Study of the Fast-Food Industry in New Jersey and Pennsylvania", *American Economic Review* **84**:4, 772-93.

Card, David and Alan Krueger (1995), *Myth and Measurement: The New Economics of the Minimum Wage*, Princeton: Princeton University Press.

Cartwright, Nancy (forthcoming), "Causal Inference à la Herbert Simon: A Primer", *Causality: Metaphysics and Methods Technical Reports*, CPNSS, LSE.

Hamilton, James (1997), "Measuring the Liquidity Effect", *American Economic Review* **87**:1, 80-97.

Hammond, J. Daniel (1996), *Theory and Measurement: Causality Issues in Milton Friedman's Economics*, Cambridge: CUP.

Hausman, Daniel (1992), *The Inexact and Separate Science of Economics*. Cambridge: Cambridge University Press.

Hausman, Daniel and James Woodward (forthcoming), "Modularity and the Causal Markov Condition: A Restatement".

Heckman, James (1996), "Comment", *Journal of the American Statistical Association* **91**:434, 459-62.

Mill, John Stuart (1948/1830), *Essays on Some Unsettled Questions of Political Economy*. London: Parker.

Mitchell, Wesley (1927), *Business Cycles: The Problem and its Setting*. New York: National Bureau of Economic Research.

Neumark, David and William Wascher (2000), "Minimum Wages and Employment:

A Case Study of the Fast-Food Industry in New Jersey and Pennsylvania: Comment", *American Economic Review* **90**:5, 1362-96.

Pearl, Judea (1993), "Mediating Instrumental Variables", *Statistical Science* **8**:3, 266-73.

Pearl, Judea (1997), "The New Challenge: From a Century of Statistics to an Age of Causation", manuscript, University of California, Los Angeles.

Pearl, Judea (2000), Causality: Models, Reasoning, and Inference, Cambridge: CUP.
Pearson, Karl (1911), *The Grammar of Science*, London: Walter Scott.

Reiss, Julian (forthcoming a), "Social Capacities", to appear in a volume on Nancy Cartwright's philosophy of science, ed. by Stephan Harmann and Luc Bovens.

Reiss, Julian (forthcoming b), "Empiricism Reconsidered From a Baconian Perspective", *Causality: Metaphysics and Methods Technical Reports*, CPNSS, LSE.

Woodward, James (1997), "Explanation, Invariance and Intervention", *Philosophy of Science*, Supplement to **64**:4, pp. S26-41.