**Balancing, Conditioning, and Confounding in Causal Inference**

Clark Glymour

**Introduction**

A standard concern in scientific practice is to prevent erroneous estimates of causal effects produced by "confounding" of the association of an experimental treatment and an outcome variable by other variables, usually called "covariates." Recently, Fuller (2021) has argued that one strategy for avoiding confounding--"balancing" of covariates between treated and untreated groups--is unnecessary and insufficient for correctly estimating a causal relation between variables in randomized controlled trials (RCTs). His example is a fully deterministic system in which all causes of an outcome variable, except possibly the treatment variable, are known, a circumstance far different from that of most RCTs. The joint distribution of the measured variables is not available in his causal story, which is just made up without regard to any principle connecting causal relations and joint probabilities. Fuller remarks that epidemiologists (one at least) say they are not concerned about balancing but that some philosophers are. He does not explain why epidemiologists or others think balancing is not of concern for causal estimation. His discussion suggests that a philosophical exposition of well-understood aspects of confounding and its attempted remedies is needed in settings that are statistically and epistemologically more realistic.

**Fuller's Example**

I will consider only Fuller's first example, which imagines an RCT with the following data:



Fuller specifies "..two deterministic assumptions. The first assumption is

that causes act deterministically, that the set of causes present for an individual

fully determines whether or not that individual gets the outcome. The second assumption—

call it ‘reverse determinism’—is that whether or not an individual gets

the outcome fully determines whether or not there was a complete cause of the

outcome. Reverse determinism adheres to the slogan, ‘some effect, some (complete)

cause’. It discounts the possibility that any difference in the frequency of

Y between groups is due solely to Y’s spontaneously popping into existence,

uncaused. Together, forward determinism and reverse determinism imply that

any difference in the frequency of the outcome between groups is proof of

some relevant causal difference." (909)

" ...the investigators see from Table 1 that (i) X is positively correlated with Y: an increased frequency of X is accompanied by an increased frequency of Y. They also see that

(ii) all confounding causes of Y are balanced in the study; each confounding

cause is (perfectly) uncorrelated with exposure X. They suppose that (iii) if all

causes of Y are balanced except X and X is positively correlated with Y, then X

must have caused Y"

He makes his argument with a story about the causal mechanism in the case.

"In each participant, Y represents the presence of a clinically important protein biomarker.

The presence of Y is fully determined by the conjunction of C1 and C2.

The confounders C1 and C2 represent two other proteins, each coded by different

genes. C1 is a precursor for Y, while C2 is the enzyme that catalyses the

conversion of C1 to Y. The pathway is represented in Figure 1. X plays no part

in this mechanism, which is the only mechanism that produces Y. Thus, X does

not cause Y. How then can we explain the study results?

"The key is that to say C1 and C2 are balanced is not to say all that much of

use. Neither C1 nor C2 will cause Y without the other. Rather than the distribution

of C1 and the distribution of C2 provided by Table 1, we need to know

the distributions of C1&C2, C1&**~**C2, **~**C1&C2, and **~**C1&**~**C2. The frequencies

for C1 and for C2 in Table 1 are consistent with a range of possible

frequencies for these four conjunctions. It turns out that the actual frequencies

are those reported in Table 2.



He continues as follows:

"This table reveals that in the exposed group, 50% of individuals were positive for both C1 and C2 (C1&C2), which explains the 50% frequency of Y in that group because C1 and C2 are jointly sufficient for Y. However, in the unexposed group, the 50% of individuals who were positive for **C**1 (column ‘**C**1&**:C**2’) were not the same 50% of individuals who were positive for **C**2 ...Because neither **C**1 nor **C**2 will cause **Y** without the other, no one in the unexposed group was positive for **Y**.

"Exposure **X** plays no role in this causal story. Yet the researchers thought

that it must because all confounding causes were balanced and there was a

difference in outcome between the groups!... In a comparative group study with

positive results, the finding that all confounding causes are balanced is not

sufficient for inferring that X caused Y, as demonstrated by the folly of our

black box researchers."

Those stupid black box researchers! They thought they ought to explain why Y depends on X, an association as strong as between Y and any of the predictors in Table 2, when the true explanation, made up by Fuller, just leaves that association unexplained.

Philosophers can make up any causal story they wish to go with any probability distributions they imagine. (It is equally consistent with the marginal probabilities in Table 2 that X causes C1 & C2 or ~C1 & ~C2 and X alone causes Y.) Scientists are more constrained. They expect the effects in experiments to not have values uniquely determined by the experimentally manipulated variables and the measured covariates, and when they find a manipulated variable as strongly associated with the outcome as any of the covariates, they may take that as prima facie evidence of an effect of the manipulated variable, and, if not, require an explanation of that association. I say "may" because Fuller does not give us the full joint distribution, which might affect a reasonable scientific judgement of whether X in his data has any direct influence on Y. So let's turn from philosophy of science fiction to something else.

 **Confounding and the Causal Context**

In RCTs with a potential cause that has two values (treated, untreated) the (average) causal effect of a treatment variable, X, on an outcome of interest, Y, is the difference in a population mean of Y among those who are treated and, respectively, what the mean outcome would have been if those treated had not been treated. In models for which the treatment variable is continuous or approximately so, it is the expected change in Y were a random member of the population subjected to an intervention that would change X by 1 unit on whatever scale X and Y are measured--a change in Y that may in non-linear systems depend on the value of X before the intervention. These quantities are to be estimated from the observed respective frequencies of Y in the treated and untreated groups sampled from the population under assumptions of "strong ignorability" or "exchangeability"--essentially that the treatment (or absence of treatment) of any one unit has no influence on the treatment of any other or on its disposition to respond to the treatment it receives.

A variable Z might be said to potentially confound the relation between a potential cause, X, of Y if Z is associated in probability with both X and Y and Z is not both an effect of X and a cause of Y. Confounding may happen when the frequency of a "covariate" Z is different in the treated and the untreated groups--the two groups are *unbalanced* with respect to Z. In such cases variation in Z may contribute to the covariation of X and Y, and estimates of the average effect of X on Y will be "biased" by that covariation. Figure 1 illustrates the problem with a linear Gaussian parameterization of a graphical causal model for each member of the sample.

 Z Y = a Z + b X + Ey

 X = c Z + Ex

 Ex X Y Ey Ex independent of Z and Ey; Ey independent of X, Z, Ex

 Ex, Ey, Z ~ N(x, y, z; 0, 0 , 0)

Figure 1

In Figure 1, Ex is determined by the experimenter. The issue is simply illustrated with the linear model of Figure 1, for which the covariance of X and Y is:

COV(X,Y) = Exp[(XY) - ux, uy] = Exp[X(a Z + b X + Ey)] = b VAR(X) + a COV(X,Z).

The aim is to estimate b, which cannot be done from the observed sample covariance of X and Y unless the covariance of X and Z is zero. The independence of X and Z is sufficient for their zero covariance. For linear, Gaussian systems variables, as in the example, that independence can be estimated from a regression of X on Y finding in the sample that coefficient c is sufficiently close to zero by some statistical test, commonly a t-test. For binary variables, the independence of Z and X is sufficient for an estimate of the average effect of X on Y because in that case

z PR(Y | XZ) = z Pr(Y | X) Pr(X | Z) Pr(Z) = z Pr(Y | X) Pr(X) Pr(Z) = Pr(Y | X) P(X).

To avoid confounding when estimating the effect of X on Y, it is not *logically necessary* that covariates Z be independent of X. Z could be an effect of X but independent of Y conditional on X, or Z could be a cause of X but independent of Y conditional on X. In either case Z would not be a confounding variable.

It suffices to avoid confounding that covariate Z be independent of Y conditional on X, but that is not always necessary. Suppose X causes Y via a pathway that involves Z and Z has another source of variation besides X, as in Figure 2.

 Ez

 Ex X Z Y Ey

Figure 2

In Figure 2, Z is a mediator between X and Y. (The Ez variable insures that Z is not independent of Y conditional on X.) Conditioning on Z would statistically eliminate a causal pathway from X to Y, biasing the estimate of the effect of X on Y. Balancing treated and untreated groups for Z would require either intervening to break the influence of X on Z leaving the distribution of X unchanged but changing the influence of X on Y, or somehow manipulating the treatment assignments so that the conditional probability of Z given X is unchanged. A sensible investigator in this case would not try to balance the treated and untreated groups for Z, but on finding that X and Y are independent conditional on Z, would conclude that Z mediates the influence of X on Y.

The independence of Y and a covariate Z is also sufficient to avoid confounding by Z, but usually an experimenter has few if any methods available to try to guarantee that independence. A sample cannot be balanced against the outcome Y because before the experiment is done the distribution of Y is unknown.

Blocking strategies can sometimes be used to prevent unwanted influences of Y. In experiments to detect a theorized phenomenon, Y, predicted as an effect of X, physical or distance barriers may suffice to prevent the action of other potential causes of Y. Cognitive barriers, such as blinding and double blinding, aim to block some potential causal pathways from treatment to effect, but these are not confounding circumstances in the sense specified above, although unblocked the effect along the intended pathway could not be estimated.

Rather than deterministic relations among the measured variables, it is standardly assumed in modeling experiments that there are unrecorded variables other than the treatment, X, that influence Y and whose influences are not blocked by experimental designs--hence the inclusion of the Ey variables in Figure 1. What is essential to avoid confounding is that these "disturbance" variables are independent of X.

Randomization of assignment to treatment groups is a standard method to try to make covariates independent of treatment value. But randomization does not guarantee that in the sample obtained covariates are differently distributed in the treated and untreated groups. Various procedures have been proposed for reweighting units in the two groups, or resampling, to approximate balance. Another common strategy to avoid confounding is to estimate effects after conditioning the treatment variable on the joint distribution on the covariates, or some function of them. That is the strategy, for example, in multivariate regression. The intended result is that conditioning on covariate Z removes any component of the association of X and Y due to variation of Z. But conditioning on any or all covariates is risky without information about the causal relations among treatment X, outcome Y, and covariates Z. If Z is an effect of both X and Y as in Figure 3, then conditioning on Z will bias the estimate of the effect of X on Y.

 Ex X b Y Ey Y = b X + Ey

 a c `Z = a X + c Y + Ez

 Z Ez X = Ex

Figure 3

This is most easily seen if we assume a Gaussian linear model and without loss of generality let X, Y and Z, be standardized, each with mean 0 and variance 1, and a, b, c are the corresponding "path coefficients," i.e. coefficients in the corresponding standardized linear model. In the Gaussian case, the dependence of X, Y conditional on Z is given by the partial correlation of X, Y controlling for Z. The partial correlations of variables are then easily computed from the coefficients.

 XY.Z =

 (XY -  XZ yz) / ([1 -  XZ2) 1/2 (1 - YZ2 )1/2] = (b - (a+ bc)) / [(1 -(c +ab)2)1/2(1 - (a+ bc)2)1/2]

The partial correlation (and likewise the partial covariance) does not measure b. The experimenter would do better in such cases not to condition on the covariate. A similar result holds when X, Y and Z are binary variables. (This problem also arises when there is no deliberate conditioning but the selection of sample members depends on the values of X and Y, so-called sample selection bias.) In the case of Figure 3, balancing Z introduces no bias because it makes Z independent of X (as random variables) even though for each unit in the sample, the value of X for that unit influence the value it takes on for Z.

In general, it is sufficient for a covariate not to bias the estimate of the average effect of X on Y that X is independent of the covariate. Balancing the covariate--giving it the same distribution in the treated and untreated groups--is sufficient for its independence of treatment, X, providing there are no unmeasured confounding variables. In linear systems with that assumption it suffices to balance each covariate independently. In non-linear systems, including binary systems, conditioning or balancing on each covariate independently may be insufficient because the biasing effect of one covariate may vary with values of other covariates. In such cases, making *all* covariates jointly independent of treatment suffices to prevent confounding provided there are no unmeasured confounding variables.

Balancing measured covariates or conditioning on them is insufficient when there are unmeasured confounders. That is obvious in the example of Figure 1 if Z is unmeasured, because it's influence on X and Y cannot be removed by conditioning or by balancing other measured covariates. Further, consider Figure 4 where U is unobserved and the aim is to estimate the effect of X on Y.

 X Z1

 U

 Y Z2

Figure 4: Confounding by conditioning on a covariate. Disturbance variables are not depicted.

In Figure 4, X has no influence on Y but is associated with Y conditional on Z1 and Z2. Estimating the effect of X on Y is biased if Z1 is conditioned on (whether or not Z2 is also conditioned on.) By contrast, making the treated and untreated samples balanced with respect to Z1 and Z2 makes Z1 independent of X and removes the bias in estimating the average effect of X on Y.

**Balancing vs. Conditioning**

The previous examples illustrate that the value of simple conditioning and balancing in eliminating biases in the estimation of causal effects in RCTs depends on the actual causal relations among the measured variables and their joint distribution in the population sampled. Balancing covariates is never worse and sometimes better than conditioning on them, and has statistical advantages as well--conditioning monotonically reduces the effective sample size in test statistics for every variable conditioned on, and with discrete variables with a small but finite range of possible values that reduction can make statistical testing very difficult. But balancing is hard, especially when there are several covariates, and conditioning is easy. There are ways to use conditioning to avoid bias from conditioning on common effects and to minimize sample size reduction.

Consider Figure 4. An investigator could note that X and Y are independent, but are dependent conditional on Z1. That is a pattern of independencies and dependencies that occurs in many cases:

* when X is a cause of Z1 and Y is a cause of Z1,
* or when X and Z1 share a common cause and Y and Z1 share a common cause
* or when X is a cause of Z1 and Y and Z1 share a common cause.

 and X has no direct (with respect to Z1 and Z2) causal connection with Y.

Assuming it is known that the outcome, Y, is not a cause of the covariate Z1, that implies that Y and Z1 share a common cause which cannot be X or Z2 or both because X is independent of Y and of Z2. So the relation between Z1 and Y is confounded by some unobserved variable, U, and X is not a cause of Y[[1]](#footnote-1).

The logic of these inferences is systematized in the PC algorithm (Spirtes and Glymour, 1991) and some other search algorithms (Spirtes, et al, 1993; 2000). Given the prior information that X, Z1 and Z2 are not effects of Y, the PC algorithm would (from sufficient sample data for X, Z1, Z2 and Y) find the causal structure of Figure 4, except, of course, an algorithm cannot name the unmeasured common cause U or determine how many such unmeasured causes of Y and Z1 there are. The algorithm would in this case use only simple dependence and independence. If the case were modified so that Z2 causes Y and there is no unmeasured common cause of Z2 and Y, the algorithm would condition on Z2 to find that X is not a direct cause of Y. In this case, if X were experimentally manipulated, using that fact the algorithm would find that X is an indirect cause of Y mediated by Z1.

**Conclusion**

Confounding can be addressed by conditioning jointly on covariates or by arranging that covariates have the same joint distribution with respect to treated and untreated groups. Of course balancing can sometimes be unnecessary because conditioning will do, and in some cases balancing can actually impede correct causal inference. Balancing is sometimes better than conditioning, but can be difficult in experimental trials and impossible with non-experimental data. (In non-experimental data, the value of treatment for a unit is sometimes modified by multiplying it by the reciprocal of the conditional probability of the treatment value given the covariate values, called "inverse probability weighting." This procedure does not help if the covariate is an effect of the treatment and the outcome, or an effect of the treatment and related to the outcome by an unmeasured common cause.) But what to balance and what to condition on depend on features of the causal relations among the variables which the investigator may not know. There have been algorithms available for thirty years that help with that problem. Philosophers, if any, who are enlightened by Fuller's fanciful argument should catch up.

1. These inferences are correct on some extra minimality assumption, e.g., Faithfulness or Frugality. [↑](#footnote-ref-1)