

Testing a Conditional Form of Exogeneity

Halbert White

Department of Economics, University of California, San Diego

Karim Chalak*

Department of Economics, Boston College

March 24, 2010

Abstract

We give two new approaches to testing a conditional form of exogeneity. This condition ensures unconfoundedness and identification of effects of interest in structural systems. As these approaches do not rely on the absence of causal effects of treatment under the null, they complement earlier methods of Rosenbaum (1987) and Heckman and Hotz (1989).

Keywords: Causality, Conditional Exogeneity, Nonparametric Test, Treatment Effect, Unconfoundedness.

JEL Classification Numbers: C12, C14, C21, C31.

1 Introduction

This note describes two new approaches for testing a conditional form of exogeneity ensuring unconfoundedness. Suppose data are generated by the structural equation

$$Y = r(D, Z, U), \tag{1}$$

where Y is a scalar response of interest, r is the unknown structural function, and D , Z , and U are drivers of Y . U is unobserved and may have any countable dimension. D and Z are observed and of finite dimension. We are interested in the effect of D on Y but not necessarily that of Z . D may contain binary, categorical, or continuously distributed elements.

As discussed by Altonji and Matzkin (2005), Hoderlein (2005, 2007), Hoderlein and Mammen (2007), White and Chalak (2008), and Imbens and Newey (2009), among others, a sufficient condition for identifying various effects of D on Y is that D is independent of U given X , where $X \equiv (W', Z')'$ is a vector of covariates (also known as control variables). When present, W

*Corresponding author: Karim Chalak, Department of Economics, Boston College, 140 Commonwealth Ave., Chestnut Hill, MA 02467, USA. Telephone: 1-617-552-6026, fax: 1-617-552-2308, email: chalak@bc.edu.

represents other observables that are not driven by Y or D , but that may be driven by U or that may drive D or U . As in Dawid (1979) ("D79"), we write

$$D \perp U \mid X. \tag{2}$$

We refer to (2) as *conditional exogeneity of D* , since it is a conditional form of the classical exogeneity condition, $D \perp U$.

It can also be shown that conditional exogeneity is, in a precise sense, necessary for identification of causal effects of interest (e.g., White and Chalak, 2008, theorem 4.1). Accordingly, it is important to have tests of this critical identifying condition. With $Y_d \equiv r(d, Z, U)$ representing potential responses, conditional exogeneity ensures the general unconfoundedness condition (see Hirano and Imbens, 2004)

$$\{Y_d, d \in \text{supp}(D)\} \perp D \mid X$$

ensuring identification of various treatment effects, where $\text{supp}(D)$ is the support of D .

Thus, to test conditional exogeneity, one can test unconfoundedness, as rejecting unconfoundedness implies rejecting conditional exogeneity. Rosenbaum (1987, section 3.2) and Heckman and Hotz (1989, section 4) propose tests for unconfoundedness for the case of a binary treatment D ($\text{supp}(D) = \{0, 1\}$). These tests rely on estimating a causal effect known to be zero; the finding of an apparent causal effect is then evidence against unconfoundedness. As suggested by Rosenbaum (1984) and Imbens (2004, section V.A), such tests can be divided into two categories. The first tests for an apparent causal effect across "unaffected units" of two control groups for which treatment is suppressed. The second tests for an apparent causal effect on "unaffected responses," variables known to be unaffected by the treatment.

Here, we propose two complementary new methods for testing conditional exogeneity that do not rely on the absence of causal effects of treatment. The first relies on an observable response R to D that is unrelated to the confounders U , conditional on X . If (2) holds, then R is independent of Y given D and X , a testable condition. The second relies on additional observable covariates S driven by the confounders U , covariates X , and other unobservables V that are independent of D given U and X , such as measurement errors. If (2) holds, then S is independent of D given X , a testable condition. An attractive feature of these methods is that they do not require D to be binary or r to be linear, separable between observables and unobservables, or monotonic in a scalar unobservable, as is often done in the literature.

2 Main results

Our first result permits using a variable R affected by treatment to test conditional exogeneity.

Proposition 1 *Suppose that Y is generated by (1) and that there exist observable finitely dimensioned random vectors W and R such that, with $X = (Z', W')'$,*

$$R = p(D, X, V),$$

for some measurable p , where V is a countably dimensioned unobservable vector such that

$$V \perp U \mid (D, X). \tag{3}$$

Then $D \perp U \mid X$ implies $Y \perp R \mid (D, X)$.

The conclusion that $Y \perp R \mid (D, X)$ generalizes the notion that, under the null, the linear regression of Y on R, D , and X should have zero coefficients for R .

Now consider the implications of a rejection, $Y \not\perp R \mid (D, X)$: either (i) $D \not\perp U \mid X$; (ii) $V \not\perp U \mid (D, X)$; (iii) (1) does not hold, i.e., R appears in r ; or (iv) any combination of these. Thus, to reject (2) given a finding that $Y \not\perp R \mid (D, X)$, we must maintain not only that R does not determine Y , but also that $V \perp U \mid (D, X)$.

This latter condition is often plausible. Specifically, when V is a measurement error, it is plausible that $V \perp (U, D, X)$; this implies (2) by D79 lemma 4.2(ii). For example, we may have $R = D + V$, in which case R and D denote "essentially equivalent treatments" (Rosenbaum, 1984). Thus, Proposition 1 supports general tests for unconfoundedness based on the availability of several equivalent treatment measures.

Our next result permits using covariates S driven by U to test conditional exogeneity.

Proposition 2 *Suppose that Y is generated by (1) and that there exist observable finitely dimensioned random vectors W and S such that, with $X = (Z', W')'$,*

$$S = q(X, U, V),$$

for some measurable q , where V is a countably dimensioned unobservable vector such that

$$D \perp V \mid (U, X).$$

Then $D \perp U \mid X$ implies $D \perp S \mid X$.

The conclusion that $D \perp S \mid X$ generalizes the notion that, under the null, the linear regression of D on S and X should have zero coefficients for S .

Now consider the implications of a rejection, $D \not\perp S \mid X$: either (i) $D \not\perp U \mid X$; (ii) $D \not\perp V \mid (U, X)$; (iii) (1) does not hold, i.e., S appears in r ; or (iv) any combination of these.

Thus, to reject (2) given a finding that $D \not\perp S \mid X$, we must maintain both that S does not determine Y and that $D \perp V \mid (U, X)$.

This latter condition is often plausible. A leading example occurs when S is an error-laden function of U , e.g., $S = U + V$ or $S = q_0(U) + V$, where V is a measurement error. As above, it is then plausible that $V \perp (U, D, X)$; this implies $D \perp V \mid (U, X)$ by D79 lemma 4.2(ii). For example, U can represent intellectual ability, and S can be a test score not already in W .

Comparing the assumptions of these results, we note that although neither R or S determine Y , their roles differ significantly: In Proposition 1, R is a response to D ; in Proposition 2, S is a vector of "supplemental" covariates. (We further discuss this below.) In particular, it is important that S depend on U , either directly or indirectly (e.g., through W), to ensure that the resulting test can have power. In contrast, R does not depend on U either directly, since U does not appear in p , or indirectly through V , as (3) must hold.

Tests based on either $Y \perp R \mid (D, X)$ or $D \perp S \mid X$ are straightforward, as only observables are involved. One can apply any of a growing number of parametric or nonparametric methods available in the literature. Su and White (2009) give a useful survey; White and Lu (2010a) discuss some straightforward regression-based procedures.

Other things equal, one may prefer the approach with fewer variables, as nonparametric procedures then place fewer demands on the data. Alternatively, to the extent that the underlying economics in a given application makes it more plausible to maintain the auxiliary assumptions of one of these approaches, then that approach may be preferred. Nevertheless, these approaches are complementary; nothing rules out their joint use.

Significantly, neither approach can deliver a test consistent against arbitrary failures of $D \perp U \mid X$. That is, neither result has a corresponding converse, even under some further plausible conditions. In fact, however, this is not a specific failure of these two approaches. Rosenbaum (1987) discusses similar properties of his tests based on absence of causal effects. As White and Lu (2010a) show, without further identifying information, there can be no test of any kind consistent against arbitrary alternatives, as easy examples prove the existence of cases where $D \perp U \mid X$ fails, but this can never be detected empirically (White and Lu, 2010a, example 7.A.(ii)). Thus, though the consequences of rejection are clear, one must be careful when failing to reject. It is always possible that conditional exogeneity fails, but in such a way as to evade detection. This underscores the usefulness of testing conditional exogeneity in multiple ways.

Finally, we justify referring to S as "supplementary" covariates by showing that if X is a sufficient set of covariates, then so is (X, S) .

Proposition 3 *Suppose the conditions of Proposition 2 hold. Then $D \perp U \mid X$ implies*

$D \perp U \mid (X, S)$.

This result raises the question of whether it is better to use a larger or smaller set of covariates. This issue has been addressed by Hahn (2004) and by White and Lu (2010b) for the case of binary D . The key idea is that if the goal is to increase the precision of a given measure of effect, then the answer depends on the nature of the underlying structure. Investigation of this issue for the case of general D is an interesting direction for further research.

3 Mathematical Appendix

Proof of Proposition 1: Given $V \perp U \mid (D, X)$, we have $(V, D, X) \perp U \mid (D, X)$ by D79 lemma 4.1 and lemma 4.2(i). Given $R = p(D, X, V)$, D79 lemma 4.2(i) gives that $R \perp U \mid (D, X)$. Now $D \perp U \mid X$ and $R \perp U \mid (D, X)$ imply $(D, R) \perp U \mid X$ by D79 lemma 4.3. Then $(D, R) \perp U \mid (D, X)$ by D79 lemma 4.2(ii). This implies $(D, R) \perp (U, X, D) \mid (D, X)$ by D79 lemma 4.1 and lemma 4.2(i). Given (1), D79 lemma 4.1 applied twice gives $Y \perp R \mid (D, X)$. ■

Proof of Proposition 2: $D \perp U \mid X$ and $D \perp V \mid (U, X)$ imply $D \perp (U, V) \mid X$ by D79 lemma 4.3. By D79 lemma 4.1 and Lemma 4.2(i), $D \perp (U, V, X) \mid X$. As $S = q(X, U, V)$, we have $D \perp S \mid X$ by D79 lemma 4.2(i). ■

Proof of Proposition 3: The conditions of Proposition 2 give $D \perp (U, V, X) \mid X$. As $S = q(X, U, V)$, this implies $D \perp (U, V, X) \mid (X, S)$ by D79 lemma 4.2(ii). It follows from D79 lemma 4.1(i) that $D \perp U \mid (X, S)$. ■

References

- Altonji, J., Matzkin, R., 2005. Cross Section and Panel Data Estimators for Nonseparable Models with Endogenous Regressors. *Econometrica* 73, 1053-1102.
- Dawid, P., 1979. Conditional Independence and Statistical Theory. *Journal of Royal Statistical Society, Series B* 41, 1-31.
- Hahn, J., 2004. Functional Restriction and Efficiency in Causal Inference. *Review of Economics and Statistics* 86, 73-76.
- Heckman, J., Hotz, V., 1989. Choosing Among Alternative Nonexperimental Methods for Estimating the Impact of Social Programs: The Case of Manpower Training. *Journal of American Statistical Society* 84, 862-874.
- Hirano, K., Imbens, G., 2004. The Propensity Score with Continuous Treatments, in Gelman, A., Meng, X. L. (Eds.), *Applied Bayesian Modeling and Causal Inference from Incomplete-Data Perspectives*, New York: Wiley, pp. 73-84.

Hoderlein, S., 2005. Nonparametric Demand Systems, Instrumental Variables and a Heterogeneous Population. Mannheim University Department of Economics Working Paper.

Hoderlein, S., 2007. How Many Consumers are Rational? Mannheim University Department of Economics Working Paper.

Hoderlein, S., Mammen, E., 2007. Identification of Marginal Effects in Nonseparable Models without Monotonicity. *Econometrica* 75, 1513-1518.

Imbens, G., 2004. Nonparametric Estimation of Average Treatment Effects under Exogeneity: A Review. *Review of Economics and Statistics* 86, 4-29.

Imbens, G., Newey, W., 2009. Identification and Estimation of Triangular Simultaneous Equations Models Without Additivity. *Econometrica* 77, 1481-1512.

Rosenbaum, P., 1984. From Association to Causation in Observational Studies: The Role of Tests of Strongly Ignorable Treatment Assignment. *Journal of American Statistical Association* 79, 41-48.

Rosenbaum, P., 1987. The Role of a Second Control Group in an Observational Study. *Statistical Science* 2, 292-306.

White, H., Chalak, K., 2008. Identifying Structural Effects in Nonseparable Systems Using Covariates. UCSD Department of Economics Discussion Paper.

White, H., Lu, X., 2010a. Granger Causality and Dynamic Structural Systems. *Journal of Financial Econometrics* 10, 193-243.

White, H., Lu, X., 2010b. Causal Diagrams for Treatment Effect Estimation with Application to Efficient Covariate Selection. UCSD Department of Economics Working Paper.