An Extended Class of Instrumental Variables for the Estimation of Causal Effects

Karim Chalak*† Halbert White
Boston College UC San Diego

Current Draft: November 2009
First Draft: March 2005

Abstract We examine how structural systems can yield observed variables instrumental in identifying and estimating causal effects. We provide an exhaustive characterization of how structures determine exogeneity and exclusion restrictions that yield moment conditions supporting identification. This provides a comprehensive framework for constructing instruments and control variables. We introduce notions of conditioning and conditional extended instrumental variables (XIVs). These permit identification but need not be traditional instruments as they may be endogenous. We distinguish between observed XIVs and proxies for unobserved XIVs. We emphasize the importance of sufficiently specifying causal relations governing the unobservables.

JEL Classification Numbers: C10, C13, C21, C31, C51.
Keywords: *causality, conditional exogeneity, endogeneity, exogeneity, identification, instrumental variables, structural systems.*

*Karim Chalak is Assistant Professor of Economics, Department of Economics, Boston College, 140 Commonwealth Ave., Chestnut Hill, MA 02467 (email: chalak@bc.edu); and Halbert White is Chancellor's Associates Distinguished Professor of Economics, Department of Economics 0508, UCSD, La Jolla, CA 92093 (email: hwhite@ucsd.edu).

†The authors thank Kate Antonovics, Julian Betts, Gary Chamberlain, Graham Elliott, Marjorie Flavin, Clive Granger, Jinyong Hahn, James Heckman, Keisuke Hirano, Guido Imbens, Steffen Lauritzen, Arthur Lewbel, Mark Machina, Chuck Manski, Salvatore Modica, Judea Pearl, Dimitris Politis, Ross Starr, Elie Tamer, Ruth Williams, and the participants of the 4th annual Advances in Econometrics conference. All errors and omissions are the authors' responsibility.

1 Introduction

Suppose a researcher seeks to measure the effect of one variable on another from non-experimental data. When measurements on these variables alone do not suffice to identify effects of interest, what additional variables (if any) can one measure to identify and estimate effects? This paper addresses this question by studying the ways in which causal structures can yield observed variables that permit identifying and estimating causal effects.

Our focus here is not on how moment conditions permit identification of causal effects; rather, we study how the structure itself and the causal roles played by the variables of this structure give rise to such moment conditions. Specifically, we provide an *exhaustive characterization* of the possible ways in which various exclusion and exogeneity restrictions permit identification of effects. In some cases, the cause and its response (the dependent variable) suffice to identify effects. In other cases, additional variables play a key role. Because of their instrumental role in identifying effects and because they may not be traditional exogenous instruments, we call such variables *extended instrumental variables* (XIVs). Our characterization provides a taxonomy for the different types of XIVs and corresponding XIV estimators.

We emphasize the role that unobserved variables play in identifying effects. Classical structural systems do not specify structural equations for unobserved variables but only their joint distribution or some of its features (Hurwicz, 1950; Chesher, 2003). Correspondingly, classical identification results (cf. Hausman and Taylor, 1983) employ restrictions on the covariances or dependence of certain unobserved variables. But because such systems do not specify how dependence patterns among unobserved variables arise, they are not fully informative as to what additional measurements may aid in identification. Thus, researchers may make opaque or uninformed statistical assumptions in attempting to identify economically meaningful objects.

Typically, however, economic theory can specify structural relations without regard to the measurement abilities of outside observers. This knowledge provides insight into statistical dependence relations holding among all variables, both observed and unobserved. This then makes explicit the causal relations underlying the identification of effects and exposes both obstacles and opportunities for causal identification that could otherwise be missed.

For clarity and simplicity and because a self-contained analysis would not otherwise be possible, we formally treat only linear structures. Nevertheless, as we explain, the insights gained apply generally to nonlinear and nonseparable structures. Among our specific contributions are:

(1) We give new necessary and sufficient moment conditions for identification of effects in triangular structural systems using potentially non-classical (extended) instrumental variables (XIVs), without requiring the instruments to drive the cause of interest. (2) These results suggest feasible Hausman-Taylor-type XIV estimators. We give new consistency and asymptotic normality results for this class under conditions permitting arbitrary misspecification.

We then study how the identifying moment conditions can arise. (3) First, we specify all possible exogeneity relations that can identify effects of interest. These involve the unobserved causes of the response of interest together with *conditionally exogenous* and *conditioning* instruments. Either or both of these may be endogenous and therefore non-classical. (4) Second, we study exclusion restrictions involving the observed variables. This leads to a taxonomy of XIVs that (a) precede the cause of interest (*pre-cause*), (b) mediate the effect of the cause of interest (*intermediate-cause*), or (c) succeed the response (*post-response*).

(5) The cases we study exhaust the possibilities for identifying causal effects based on exogeneity and exclusion restrictions, providing a comprehensive framework for identifying effects that encompasses standard IV, control function, and other, less familiar approaches. Like Angrist, Imbens, and Rubin (1996, AIR), we provide a causal account of the role of each system component, but our account extends well beyond the standard IV framework considered by AIR.

(6) Our analysis reveals an important distinction between observed instruments and proxies for unobserved instruments. The latter fall outside the AIR framework. With linearity, proxy instruments can deliver consistent ILS-based estimates of effects, even though they are endogenous in the reduced form. In more general nonlinear or nonseparable systems, multiple proxies are needed to identify effects of interest (see Schennach, White, and Chalak, 2009).

(7) Our results explain the genesis of control variables (covariates) ensuring unconfoundedness (e.g., Rubin, 1974) or identification as in Altonji and Matzkin (2005) or Hoderlein and Mammen (2007). Here, the distinction between observed and unobserved confounders is important. When these are unobserved, we distinguish between *structural* and *predictive* proxies. Structural proxies mediate the effects of confounders on the cause or response, as in Pearl's (1995) "back-door" method. Predictive proxies complement and extend Pearl's back-door methods. In this context, we explain how the standard treatment of "omitted variables" bias is incomplete and why not all regression coefficients need have signs and magnitudes that make economic sense.

(8) We highlight the utility of Pearl's (1995) "front-door" method. In these structures, variables that mediate the effect of the endogenous treatment on the response (intermediate-cause instruments) can ensure identification. (9) We extend the Hausman-Taylor (1983) framework by showing that *conditional* covariance restrictions can also identify causal effects.

The plan of the paper is as follows. Section 2 specifies the data generating structural equation system. Section 3 gives necessary and sufficient moment conditions for identification of causal effects based on XIVs. Section 4 studies asymptotic properties of feasible XIV estimators with possible misspecification. Section 5 characterizes all possible exogeneity and exclusion restrictions ensuring causal identification. Sections 6 and 7 analyze these methods in detail, with particular attention to how causal relations among unobservables create opportunities and obstacles for identification. Section 6 treats single XIV methods; section 7 studies double XIV methods. Section 8 concludes with final remarks and a discussion of directions for future research. The Appendices contain formal proofs and a number of results supporting the discussion of the text.

2 Causal Data Generating Systems

We consider data generated by a triangular structural system $\mathcal{S} \equiv (V_0, r)$. For this, let $\mathbb{N} = \{0, 1, ...\}$ denote the integers, let $\mathbb{N}^+$ denote the positive integers, and let $\bar{\mathbb{N}}^+ \equiv \mathbb{N}^+ \cup \{\infty\}$.

Assumption 2.1 : ***Data Generation*** *The triangular structural system $\mathcal{S}$ generates a $k_0 \times 1$ random vector V_0, $k_0 \in \bar{\mathbb{N}}^+$, and $k_j \times 1$ random vectors V_j, $k_j \in \mathbb{N}^+$, $j = 1, ..., J$, $J \in \mathbb{N}^+$, as*

$$\begin{aligned} V_1 &\overset{c}{=} r_1(V_0) \\ V_2 &\overset{c}{=} r_2(V_1, V_0) \\ &\vdots \\ V_J &\overset{c}{=} r_J(V_{J-1}, ..., V_1, V_0), \end{aligned}$$

where structural response functions $r \equiv (r_1, ..., r_J)$ *are unknown vector-valued measurable functions.*

Triangular structures have long been of interest in economics and econometrics. Recent work of Imbens and Newey (2009) provides discussion and some instructive examples.

We use the $\overset{c}{=}$ notation to distinguish these relations from purely probabilistic relations, such as regression equations, and to emphasize that these structural equations represent directional "causal links" (Goldberger, 1972, p.979): the right-hand side *direct causes* $V_0, ..., V_{k-1}$ mechanistically determine the value of the corresponding left-hand side *response* V_k, but the converse is not true, i.e., V_k does not determine V_j for any $j \leq k$. Conceptually, the right-hand side variables of a structural relation can be manipulated without modifying any of the other relations.

This enables definition of causal effects as changes in V_k resulting from hypothetical interventions where V_j is set to some different value, V_j^*. Following the literature, we seek to measure features of the distribution of the effect of V_j on V_k. For example, when the derivative exists, we are interested in conditional averages of the marginal effect of continuously distributed V_j on V_k,

$$D_j r_k(V_{k-1}, ..., V_j, ..., V_0),$$

where $D_j \equiv (\partial/\partial v_j)$. White and Chalak (2009b) and Chalak and White (2009) provide a rigorous formalization; references here to notions of cause and effect are as formally defined there.

In triangular systems, there is an inherent ordering: "predecessor" variables may determine "successor" variables, but not vice versa. In particular, V_j *does not cause* V_k when $j \geq k$ (including $k = 0$). For $j < k$, we say that V_j does not *directly* cause V_k if $r_k(V_{k-1}, \ldots, V_0)$ defines a function constant in V_j. Otherwise, we say that V_j directly causes V_k. The variables V_0 are not determined by any other variable in the system[1]. Following White and Chalak (2009b), we call V_0 *fundamental variables.* Triangularity (also called "recursivity" or "acyclicality") also ensures that total or full effects, that is, the effect of one variable on another, channeled via all routes in the system, both direct and indirect, are well defined (see Chalak and White, 2009).

For now, we do not specify whether or not realizations of the random variables $V_0, V_1, \ldots, V_J$ are observed. Classically, structural systems do not specify causal relations among unobserved variables, apart from requiring that they do not depend on observed variables (e.g., Hurwicz, 1950). Thus, unobserved variables are fundamental in classical structural systems.

[1] The necessary existence of such variables for triangular systems is a consequence of Bang-Jensen and Gutin (2001, prop. 1.4.2).

In contrast, here we explicitly specify causal relations among *all* variables, observed or not. There are several reasons for this. First, structural relations emerge from the economic behavior of agents. In non-experimental settings, we view this behavior as logically prior to issues of observability by researchers. That is, we focus on situations where economic behavior is not affected by the econometrician's ability to observe some variables and not others; indeed, different researchers may have different observational abilities, due to differing access to data or differing budget or time constraints. Thus, though we may refer to "observables" or "unobservables," we understand these designations to be relative to an outside observer and not absolute.

Second, recursivity guarantees that causal notions are well defined. To ensure this for the entire system, one must specify structural relations for all variables, not just some. A related point is that, as Chalak and White (2009) prove, recursivity ensures the validity of Reichenbach's (1956) principle of common cause. This states that two random variables can exhibit correlation (or, more generally, stochastic dependence) only if one causes the other or if they share a common cause. Thus, stochastic dependence among unobserved variables *implies* the existence of causal relations. These relations may further entail dependence between observed and unobserved causes (endogeneity) for a given structural equation. Neglecting to understand this structure may cause one to overlook situations that present difficulties for identification. Further, as we show in detail below, the precise nature of the causal relations holding among the unobserved variables also creates opportunities for identification that are not evident otherwise.

Consequently, a main message from our analysis is that *sufficiently specifying the causal relations governing unobserved variables is of central importance to identifying causal effects.* Indeed, it is precisely the systematic examination of the causal relations that may hold among unobservables that makes it possible to characterize the ways in which effects of interest can be identified.

We thus do not favor researchers simply including "random disturbances" in their structural relations. Instead, we view such variables as unobserved causes, and we recommend that economic theory be used to the maximum extent to identify both the variables relevant to a given phenomenon and the structure relating them, e.g., exclusion restrictions and conditional independence.

Throughout, we focus on identifying and estimating suitable measures of the effect of one observable on another. In nonseparable systems of the sort specified in Assumption 2.1, fine details of the system structure (e.g., nonlinearity, separability, monotonicity) critically determine the precise form and meaning of the identifiable effects. The work of Heckman (1997), Abadie (2003), Matzkin (2003, 2004, 2008), Heckman and Vytlacil (2005), Hoderlein (2005), Heckman, Urzua, and Vytlacil (2006), Hoderlein and Mammen (2007), Imbens and Newey (2009), White and Chalak (2009a), Schennach, White, and Chalak (2009), and Song, Schennach, and White (2009), among others, demonstrates the rich possibilities and nuance that can arise.

Here, our main goal is to understand broadly how different exogeneity and exclusion restrictions can identify causal effects. As the work just cited shows, exploring the fine details of general structural systems requires more space than is available here. Thus, to achieve our main goal, we operate under assumptions that remove this sensitive dependence on structural details without sacrificing insight into how and why effects of interest are identifiable in some cases and not in others. For this, an effective strategy is to suppose that the system equations are linear. In this case, the identifiable effect measure is always interpretable as the marginal effect of V_j on V_k. Given linearity, this is also the average marginal effect, as well as the effect corresponding to the average marginal effect of treatment on the treated, etc. Nothing essential for understanding causal identification broadly is lost by the coincidence of these effect measures. Further, because the various exogeneity and exclusion restrictions ensuring identification do not critically depend on linearity, the insights gained apply to the general nonseparable case.

In some structures, linearity may facilitate identification of effects, masking complications that arise in more general cases. We carefully identify such situations. This provides insight into the robustness of identification methods to the form of the structural equations. When linearity merely simplifies exposition, the conditions ensuring identification of effects and the relation to other results in the literature, particularly classical results, become quite transparent.

A further advantage of imposing linearity is that the resulting estimators are parametric (rather than nonparametric, as they must necessarily be in the general case), so that the relation of these estimators to more familiar cases also becomes transparent. These advantages combine to make possible a clear and accessible exposition of the key principles affording identification generally that would otherwise be impossible in a self-contained article.

We now impose linear structure.

Assumption 2.2 : ***Linearity*** *For $j = 1, \ldots, J$, assume that r_j is linear, so that*

$$\begin{aligned} V_1 &\overset{c}{=} \alpha_1 V_0 \\ V_2 &\overset{c}{=} \beta_{2,1} V_1 + \alpha_2 V_0 \\ &\vdots \\ V_J &\overset{c}{=} \beta_{J,J-1} V_{J-1} + ... + \beta_{J,1} V_1 + \alpha_J V_0, \end{aligned}$$

where, for $j = 1, \ldots, J$, α_j is an unknown real $k_j \times k_0$ matrix, and for $j = 2, \ldots, J$, $i = 1, ..., j-1$, $\beta_{j,i}$ are unknown real $k_j \times k_i$ matrices. Let $E(V_0) = 0$, so $E(V_j) = 0$, $j = 1, ..., J$.

3 A Theorem for XIV Identification

We now focus attention on a single structural equation of interest:

$$V_j \overset{c}{=} \beta_{j,j-1} V_{j-1} + ... + \beta_{j,1} V_1 + \alpha_j V_0.$$

To proceed, we adapt our notation. Let $Y \equiv V_j$ denote the scalar response of interest. Let $X \equiv [X_1, \ldots, X_k]'$, $k \in \mathbb{N}^+$, denote observed causes of Y, let the vector U_y $(p \times 1, p \in \bar{\mathbb{N}} \equiv \mathbb{N} \cup \{\infty\})$ denote the remaining causes of Y, and write $\varepsilon \equiv U_y' \alpha_o$, for an unknown vector α_o, so that

$$Y \overset{c}{=} X' \beta_o + \varepsilon,$$

where the unknown vector β_o is the effect of interest. By convention, X omits observables known *not* to cause Y, ensuring that $Y \stackrel{c}{=} X'\beta_o + \varepsilon$ embodies known exclusion restrictions.

We let $Z \equiv [Z_1, \ldots, Z_\ell]'$ and $W \equiv [W_1, \ldots, W_m]'$ ($\ell, m \in \mathbb{N}$) denote observed variables potentially instrumental to identifying effects of interest, generated as in Assumptions 2.1 and 2.2. In Section 5, we formally distinguish Z and W based on their roles in identifying β_o.

We now state a result giving necessary and sufficient conditions for identifying β_o.

Theorem 3.1 *Suppose Assumptions 2.1 and 2.2 hold with* $Y \stackrel{c}{=} X'\beta_o + \varepsilon$.

Let $\tilde{Z}$ $(k \times 1)$ *and* $\tilde{W}$ $(\tilde{m} \times 1,\ \tilde{m} \in \mathbb{N})$ *be given by* $[\tilde{Z}', \tilde{W}']' = A\,[Y, X', Z', W']'$, *for given matrix* A, *where* $E[Z^*(Y, X')] < \infty$, *with* $Z^* \equiv \tilde{Z} - E(\tilde{Z} \mid \tilde{W})$ $(Z^* \equiv \tilde{Z}$ *if* $\tilde{m} = 0)$. *Then*

(*a*) $E\{Z^*\varepsilon\}$ *exists and is finite.*

(*b*) $E\{Z^*\varepsilon\} = 0$ *if and only if* β_o *satisfies*

$$E(Z^*X')\,\beta_o - E(Z^*Y) = 0. \tag{1}$$

(*c*) *If* $E\{Z^*\varepsilon\} = 0$, *then full identification holds* (*that is, there is a unique* β_o *satisfying eq.* (1)) *if and only if* $E(Z^*X')$ *is non-singular. Then*

$$\beta_o = E(Z^*X')^{-1}\,E(Z^*Y).$$

This result shares many common elements with classical treatments of IV methods, but it differs from classical formulations in two important ways. First, Theorem 3.1 explicitly uses the *extended* instrumental variables (XIVs) $[\tilde{Z}', \tilde{W}']' = A\,[Y, X', Z', W']'$ to construct "derived" or "residual" instruments, $Z^* \equiv \tilde{Z} - E(\tilde{Z} \mid \tilde{W})$, analogous to those of Hausman and Taylor (1983). As we demonstrate, the possible roles of Y, Z, and W in determining Z^* and the conditioning on $\tilde{W}$ can be non-classical. Second, we do not require Z^* to be a cause of X uncorrelated with ε. Instead, we focus on the crucial partial identification condition $E\{Z^*\varepsilon\} = 0$, and we leave open the causal relations underlying X and Z^*. By considering *which* causal structures admit a matrix A yielding $E\{Z^*\varepsilon\} = 0$ based on certain exclusion and exogeneity restrictions, we arrive at a characterization of causal structures identifying β_o that include a variety of interesting non-classical cases.

For this, we must also treat causal effects that are functions of other identified effects.

Corollary 3.2 *Suppose Assumptions 2.1 and 2.2 hold with* $Y \stackrel{c}{=} X'\beta_o + \varepsilon$, *as in Theorem 3.1. For* $H \in \mathbb{N}^+$, *let* $\theta_1, \ldots, \theta_H$ *be vectors of structural coefficients of* $\mathcal{S}$, *and let* $b(\cdot)$ *be a known measurable vector-valued function such that* $\beta_o = b(\theta_1, \ldots, \theta_H)$. *If* $\theta_1, \ldots, \theta_H$ *are each fully identified as in Theorem 3.1, then* β_o *is fully identified as* $b(\theta_1, \ldots, \theta_H)$.

4 Asymptotic Properties of XIV Estimators

The results of the previous section immediately suggest useful estimators. Thus, before embarking on a detailed examination of identification issues, we describe these briefly. This also helps to highlight the difference between structural and non-structural relations that may hold for given data and to further motivate the search for structures where causal effects can be identified.

The plug-in XIV estimator suggested by Theorem 3.1 is

$$\hat{\beta}_n^{XIV} \equiv [\mathbf{Z}^{*\prime}\mathbf{X}]^{-1}[\mathbf{Z}^{*\prime}\mathbf{Y}],$$

where $\mathbf{X}$, $\mathbf{Y}$, and $\mathbf{Z}^*$ denote $n \times k$, $n \times 1$, and $n \times k$ matrices containing n identically distributed observations on X, Y, and Z^*, respectively. Generally, Z^* may be unknown, so in practice we replace Z^* with an estimator, say $\hat{Z}^*$, giving a feasible XIV estimator

$$\hat{\beta}_n^{FXIV} \equiv [\hat{\mathbf{Z}}^{*\prime}\mathbf{X}]^{-1}[\hat{\mathbf{Z}}^{*\prime}\mathbf{Y}].$$

To construct $\hat{Z}^*$, we may need estimates of $\tilde{Z}$, $\tilde{W}$, and $E(\tilde{Z} \mid \tilde{W})$. To keep a tight focus here, however, we will abstract from all but the simplest statistical issues associated with this estimation, leaving the more involved aspects as interesting topics for further research.

We begin by noting that any useful feasible estimator should have the property that replacing Z^* with its estimator $\hat{Z}^*$ generally leaves the probability limit unchanged: $\hat{\beta}_n^{FXIV} - \hat{\beta}_n^{XIV} \xrightarrow{p} 0$. The next simple result, apparently not elsewhere available, shows that any such estimator is associated with a stochastic relation resembling a structural equation but with no structural content.

Proposition 4.1 *Suppose (X, Y, Z^*) are random vectors such that $S \equiv E(Z^*Y)$ is finite, and $Q \equiv E(Z^*X')$ is finite and non-singular. Then $\beta^\dagger \equiv Q^{-1}S$ exists and is finite, and there exists a random variable η such that $E(Z^*\eta) = 0$ and*

$$Y = X'\beta^\dagger + \eta.$$

If in addition $\{Z_i^* X_i'\}$ and $\{Z_i^* Y_i\}$ obey the law of large numbers (i.e., $\mathbf{Z}^{*\prime}\mathbf{X}/n \xrightarrow{p} Q$ and $\mathbf{Z}^{*\prime}\mathbf{Y}/n \xrightarrow{p} S$), then $\hat{\beta}_n^{XIV} \xrightarrow{p} \beta^\dagger$. If also $\hat{\beta}_n^{FXIV} - \hat{\beta}_n^{XIV} \xrightarrow{p} 0$ then $\hat{\beta}_n^{FXIV} \xrightarrow{p} \beta^\dagger$.

This result shows that *any* random vectors satisfying mild conditions can appear to obey a spurious "structural" equation

$$Y = X'\beta^\dagger + \eta$$

with proper instruments Z^* ($E(Z^*X')$ nonsingular and $E(Z^*\eta) = 0$). But this relation is simply a mathematical construct with no economic (i.e., structural) or causal content. Interventions to X need not impact Y, as Assumption 2.1 is not in force. Moreover, interventions to η are not even well defined, unless perhaps accomplished by manipulating X and/or Y.

The only symbolic difference between this relation and the structural relation of interest,

$$Y \stackrel{c}{=} X'\beta_o + \varepsilon,$$

is the $\stackrel{c}{=}$ appearing above, signaling the structure imposed by Assumption 2.1. Indeed, we use $\stackrel{c}{=}$ precisely because there is otherwise no explicit way to tell that one of these relations is structural and the other isn't. Our search for structures where causal effects can be identified is in fact the search for the economic relationships that force $\beta^\dagger$ to equal β_o.

Estimating Z^* generally has consequences for the asymptotic distribution of the feasible estimator. To illustrate, suppose that $\tilde{m} > 0$, that $\tilde{Z}$ and $\tilde{W}$ do not require estimation, and that the regression of $\tilde{Z}$ on $\tilde{W}$ is linear. That is, $E(\tilde{Z} \mid \tilde{W}) = \pi^* \tilde{W}$, where $\pi^* \equiv E(\tilde{Z}\tilde{W}')E(\tilde{W}\tilde{W}')^{-1}$ is unknown. The plug-in estimator for π^* is $\hat{\pi}_n \equiv \tilde{\mathbf{Z}}'\tilde{\mathbf{W}}(\tilde{\mathbf{W}}'\tilde{\mathbf{W}})^{-1}$, where $\tilde{\mathbf{Z}}$ and $\tilde{\mathbf{W}}$ are $n \times \ell$ and $n \times \tilde{m}$ matrices whose rows are identically distributed observations on $\tilde{Z}'$ and $\tilde{W}'$, corresponding to the observations on Y and X. Then

$$\begin{aligned} \tilde{\beta}_n^{FXIV} &\equiv [(\tilde{\mathbf{Z}} - \tilde{\mathbf{W}}\hat{\pi}_n')'\mathbf{X}]^{-1}(\tilde{\mathbf{Z}} - \tilde{\mathbf{W}}\hat{\pi}_n')'\mathbf{Y} \\ &= [\tilde{\mathbf{Z}}'(\mathbf{I} - \tilde{\mathbf{W}}(\tilde{\mathbf{W}}'\tilde{\mathbf{W}})^{-1}\tilde{\mathbf{W}}')\mathbf{X}]^{-1}[\tilde{\mathbf{Z}}'(\mathbf{I} - \tilde{\mathbf{W}}(\tilde{\mathbf{W}}'\tilde{\mathbf{W}})^{-1}\tilde{\mathbf{W}}')\mathbf{Y}]. \end{aligned}$$

Under plausible conditions $\tilde{\beta}_n^{FXIV} \xrightarrow{p} \beta^*$, say, and if in fact $E(\tilde{Z} \mid \tilde{W}) = \pi^*\tilde{W}$, then $\beta^* = \beta^\dagger$.

We now give asymptotic distribution results for $\tilde{\beta}_n^{FXIV}$, again without imposing any causal structure. For this, let $\eta_i^* \equiv Y_i - X_i'\beta^*$, $\zeta_i^* \equiv \tilde{Z}_i - \pi^*\tilde{W}_i$, and $\tilde{\eta}_i \equiv \eta_i^* - \tilde{W}_i'\ E(\tilde{W}\tilde{W}')^{-1}E(\tilde{W}\eta^*)$, $i = 1, ..., n$. By the definitions of β^* and π^*, $E(\zeta_i^*\eta_i^*) = 0$ and $E(\zeta_i^*\tilde{W}_i') = 0$, so $E(\zeta_i^*\ \tilde{\eta}_i) = 0$.

Theorem 4.2 *Suppose* $\{\tilde{W}, X, Y, \tilde{Z}\}$ *are random vectors such that* $E[\tilde{W}(\tilde{W}', X', Y)] < \infty$ *and* $E[\tilde{Z}(\tilde{W}', X', Y)] < \infty$, *and* $E[\tilde{W}\tilde{W}']$ *and* $Q^* = E[\zeta^* X']$ *are non-singular. Suppose further that*

(i) $\{\tilde{W}_i(\tilde{W}_i', X_i', Y_i)\}$ *and* $\{\tilde{Z}_i(\tilde{W}_i', X_i', Y_i)\}$ *obey the law of large numbers; and*

(ii) $n^{-1/2}\sum_{i=1}^n \zeta_i^*\eta_i^* = O_p(1)$ *and* $n^{-1/2}\sum_{i=1}^n \zeta_i^*\ \tilde{\eta}_i \xrightarrow{d} N(0, V^*)$, *where* V^* *is finite and positive definite. Then*

$$n^{1/2}(\tilde{\beta}_n^{FXIV} - \beta^*) \xrightarrow{d} N(0, Q^{*-1}V^*Q^{*\prime-1}).$$

This result is also apparently new, as it gives the asymptotic distribution of standard or Hausman-Taylor-type IV estimators with arbitrary misspecification. (For $\tilde{m} = 0$, put $\tilde{W}_i \equiv 1$.) Using $\hat{\pi}_n$ to construct $\tilde{\beta}_n^{FXIV}$ introduces additional terms into V^*, as $\tilde{\eta}_i = \eta_i^*$ when π^* is known.

Asymptotic distribution results also hold for feasible plug-in XIV estimators for effects identified by Corollary 3.2, say $\tilde{\beta}_n^{FXIV} = b(\tilde{\theta}_n^{FXIV})$, where $\tilde{\theta}_n^{FXIV}$ is a vector of FXIV estimators. These results follow easily from results for $\tilde{\theta}_n^{FXIV}$ by the delta method. The asymptotic covariance matrix can be complicated, so we give this result in Appendix B for easy reference.

To gain further insight, consider what happens when causal structure is present. Specifically, suppose the conditions of Proposition 4.1 hold and that $Y \stackrel{c}{=} X'\beta_o + \varepsilon$. Then

$$\beta^\dagger \equiv Q^{-1}E[Z^*(X'\beta_o + \varepsilon)] = \beta_o + Q^{-1}E(Z^*\varepsilon). \tag{2}$$

Thus, $\hat{\beta}_n^{XIV}$ converges generally to the true effect, β_o, plus an "effect bias," $\delta^* \equiv Q^{-1}E(Z^*\varepsilon)$. This vanishes if and only if $E(Z^*\varepsilon) = 0$.

As is easily seen, a sufficient condition for $E(Z^*\varepsilon) = 0$ is conditional non-correlation of $\tilde{Z}$ and ε, given $\tilde{W}$, $E(\tilde{Z}\,\varepsilon \mid \tilde{W}) = E(\tilde{Z} \mid \tilde{W})E(\varepsilon \mid \tilde{W})$. (Rewrite this as $E[(\tilde{Z} - E(\tilde{Z} \mid \tilde{W}))\varepsilon \mid \tilde{W}] = 0$, and take expectations.) It is also obvious that this must hold for *some* $\tilde{W}$ (including $\tilde{W} = 1$), since if this fails with $\tilde{W} = 1$, we have $E(Z^*\varepsilon) = E(\tilde{Z}\,\varepsilon) \neq E(\tilde{Z})E(\varepsilon) = 0$. Although conditional[2] non-correlation is thus necessary and sufficient to identify β_o when $E(Z^*X')$ is nonsingular, this is a special feature of the assumed linear structure.

In general nonlinear or nonseparable cases, the necessary and sufficient identifying condition is not conditional non-correlation, but conditional independence (cf. White and Chalak, 2009a, theorem 4.1). Even in linear structures, conditional independence is key for identifying other than average effects, such as quantile effects. It is even necessary for identifying average effects in linear structures, if the distribution of $(U_y, \tilde{W}, \tilde{Z})$ can be arbitrary, as we show in Appendix C. Thus, in what follows, we explicitly study which structures can generate $\tilde{Z}$ and $\tilde{W}$ ensuring conditional independence[3], $\tilde{Z} \perp U_y \mid \tilde{W}$. We call this *conditional* exogeneity; strict exogeneity, a classical condition, holds when $\tilde{m} = 0$, so that $\tilde{Z} \perp U_y$.

5 Exogeneity and Exclusion Restrictions in Structural Systems

Theorem 3.1 shows that effects of interest are identified if and only if $E(Z^*\varepsilon) = 0$ and $E(Z^*X')$ is non-singular. But it does not specify exactly how Z and W are generated. Nor does it specify how $\tilde{Z}$ and $\tilde{W}$ are to be chosen. We now study the possible ways a system can admit (conditional) exogeneity, $\tilde{Z} \perp U_y \mid \tilde{W}$, for specific choices of $\tilde{W}$ and $\tilde{Z}$, so that $E(Z^*\varepsilon) = 0$. We also study how Z and W are generated, with particular attention to both the causal relations among unobservables and the exclusion restrictions embodied in $\mathcal{S}$. The latter bears on the non-singularity of $E(Z^*X')$; the former has implications for conditional exogeneity.

We first consider structures where at least one relation $\tilde{Z} \perp U_y \mid \tilde{W}$ holds. Recall that $[\tilde{Z}', \tilde{W}']' = A[Y, X', Z', W']'$; without essential loss of generality, let A form $\tilde{Z}$ and $\tilde{W}$ as linear combinations of mutually exclusive groupings of Y, X, Z, and W. We distinguish Z and W by requiring that W not appear in $\tilde{Z}$ and that Z not appear in $\tilde{W}$. Accordingly, we call Z *(un)conditional* instruments and W *conditioning* instruments. There are fourteen possibilities involving W, represented in Table I. Omitting W yields fourteen more relations (not tabulated).

Table I: Possible Conditional Independence Relations

$(X,Y,Z) \perp U_y \mid W$	$(X,Y) \perp U_y \mid W$	$(X,Z) \perp U_y \mid W$ *	$(Y,Z) \perp U_y \mid W$
	$X \perp U_y \mid W$ *	$Y \perp U_y \mid W$	$Z \perp U_y \mid W$ *
$(Y,Z) \perp U_y \mid (W,X)$	$Y \perp U_y \mid (W,X)$	$Z \perp U_y \mid (W,X)$ *	
$(X,Z) \perp U_y \mid (W,Y)$	$X \perp U_y \mid (W,Y)$	$Z \perp U_y \mid (W,Y)$ *	
$Z \perp U_y \mid (W,X,Y)$ *			

[2] We take the unconditional case to be the special case of conditioning on a constant.

[3] Following Dawid (1979), we write $\tilde{Z} \perp U_y \mid \tilde{W}$ to denote that $\tilde{Z}$ is independent of U_y given $\tilde{W}$. If desired, one can interpret this as conditional non-correlation in strictly linear contexts.

Because $Y \overset{c}{=} X'\beta_o + \varepsilon$ where $\varepsilon \equiv U_y'\alpha_o$, we can rule out some of these possibilities *a priori*. In all but exceptional cases, the fact that U_y directly causes Y implies $Y \not\perp U_y$ and $Y \not\perp U_y \mid X$ (see Chalak and White, 2009). This then implies $(X, Y) \not\perp U_y$, $(Y, Z) \not\perp U_y$, $(X, Y, Z) \not\perp U_y$, and $(Y, Z) \not\perp U_y \mid X$. Similarly, $X \not\perp U_y \mid Y$, as conditioning on a common response generally renders X and U_y dependent (see Chalak and White, 2009). This then implies $(X, Z) \not\perp U_y \mid Y$. Except when W fully determines system variables, similar dependence holds when further conditioning on W. Six possibilities remain, denoted by * in Table I, along with their six analogs omitting W.

Table II provides a complete list of these twelve possible exogeneity relations, the basis of our characterization of identifiable structures.

Table II: Unconditional and Conditional Exogeneity Relations

Exogenous Causes (XC)	$X \perp U_y$
Exogenous Instruments (XI)	$Z \perp U_y$
Exogenous Causes and Instruments (XCI)	$(X, Z) \perp U_y$
Conditionally Exogenous Causes given Conditioning Instruments (XC|I)	$X \perp U_y \mid W$
Conditionally Exogenous Instruments given Causes (XI|C)	$Z \perp U_y \mid X$
Conditionally Exogenous Instruments given Response (XI|R)	$Z \perp U_y \mid Y$
Conditionally Exogenous Instruments given Causes and Response (XI|CR)	$Z \perp U_y \mid (X, Y)$
Conditionally Exogenous Instruments given Conditioning Instruments (XI|I)	$Z \perp U_y \mid W$
Conditionally Exogenous Causes and Instruments given Conditioning Instruments (XCI|I)	$(X, Z) \perp U_y \mid W$
Conditionally Exogenous Instruments given Causes and Conditioning Instruments (XI|CI)	$Z \perp U_y \mid (X, W)$
Conditionally Exogenous Instruments given Response and Conditioning Instruments (XI|RI)	$Z \perp U_y \mid (Y, W)$
Conditionally Exogenous Instruments given Causes, Response, and Conditioning Instruments (XI|CRI)	$Z \perp U_y \mid (X, Y, W)$

Observe that XC, XI, XCI, XC|I, XI|C, XI|R, and XI|CR involve a single type of instrumental variable: either (un)conditional instruments Z or conditioning instruments W but not both. On the other hand, XI|I, XCI|I, XI|CI, XI|RI, and XI|CRI involve both conditional and conditioning instruments. We refer to XIV methods that use a single type of instrumental variable as *single* XIV methods and to those that make use of both types as *double* XIV methods. In Section 6, we study single XIV methods. We study double XIV methods in Section 7.

To systematically analyze how system variables can be generated, we use the system's *direct causality matrix* $C_{\mathcal{S}} = [c_{gh}]$. This is an adjacency matrix in which every variable has a corresponding row and column. For simplicity, let $k_j = 1$, $j > 0$, for now. Then $(V_0', V_1, \ldots, V_J)'$ is $G \times 1$, $G \equiv k_0 + J$, so $C_{\mathcal{S}}$ is $G \times G$. An entry $c_{gh} = 0$ says that the row g variable does not directly cause the column h variable. An entry $c_{gh} = 1$ says that the row g variable is (or may be) a direct cause of the column h variable. By convention, a variable does not cause itself, so $c_{gg} = 0$, $g = 1, \ldots, G$. As we now explain, the direct causality matrix $C_{\mathcal{S}}$ for a system $\mathcal{S}$ generated under Assumption 2.1 has the form

$$C_{\mathcal{S}} = \begin{bmatrix} C_{\mathcal{S},1} & C_{\mathcal{S},2} \\ C_{\mathcal{S},3} & C_{\mathcal{S},4} \end{bmatrix} = \begin{array}{c|ccc|ccc} & V_{01} & \ldots & V_{0k_0} & V_1 & \ldots & V_J \\ \hline V_{01} & & & & * & \ldots & * \\ \vdots & & \mathbf{0} & & \vdots & \ddots & \vdots \\ V_{0k_0} & & & & * & \ldots & * \\ V_1 & & & & 0 & * & * \\ \vdots & & \mathbf{0} & & \vdots & \ddots & * \\ V_J & & & & 0 & \ldots & 0 \end{array}$$

Since V_0 is a vector of fundamental variables, these are not caused by any system variable. This implies that $C_{\mathcal{S},1}$ and $C_{\mathcal{S},3}$ are zero matrices. Unspecified entries ($*$) in $C_{\mathcal{S},2}$ and $C_{\mathcal{S},4}$ indicate elements taking either the values 0 or 1. In $C_{\mathcal{S},2}$, the $*$ entries reflect that elements of V_0 may or may not directly cause $V_1, ..., V_J$. Last, the assumed triangularity of $\mathcal{S}$ makes $C_{\mathcal{S},4}$ upper triangular, with zeros along the diagonal ($c_{gg} = 0$). Thus, $C_{\mathcal{S}}$ can explicitly specify all possible direct causality relationships, including those holding among unobserved variables, as specified in Assumption 2.1. It also refines Assumption 2.1 by explicitly specifying exclusion restrictions.

The recursivity of $\mathcal{S}$ can be concisely expressed in terms of $C_{\mathcal{S}}$:

Proposition 5.1 ***Acyclicality*** *Suppose that Assumption 2.1 holds for a structural system $\mathcal{S}$ with corresponding $G \times G$ direct causality matrix $C_{\mathcal{S}}$. Then $C_{\mathcal{S}}$ is such that for each $h \leq G$ and each set of h distinct elements, say $\{g_1, \ldots, g_h\}$, of $\{1, \ldots, G\}$, $c_{g_1 g_2} \times c_{g_2 g_3} \times \ldots \times c_{g_h g_1} = 0$.*

To illustrate how Assumptions 2.1 and 2.2 operate when, as is usual, only some variables are observed, suppose we observe X, Y, and Z, each driven by its own unobservable, U_x, U_y, and U_z, respectively. Although the unobservables need not be fundamental, we adopt the common convention that they are not caused by observables (if not, we can usually ensure this by suitable substitutions). We also specify that Y does not cause X. The direct causality matrix is then

$$
D_{\mathcal{S}} = \begin{bmatrix} D_{\mathcal{S},1} & D_{\mathcal{S},2} \\ D_{\mathcal{S},3} & D_{\mathcal{S},4} \end{bmatrix} = \begin{array}{c|cccccc} & U_x & U_y & U_z & X & Y & Z \\ U_x & 0 & * & * & 1 & 0 & 0 \\ U_y & * & 0 & * & 0 & 1 & 0 \\ U_z & * & * & 0 & 0 & 0 & 1 \\ X & 0 & 0 & 0 & 0 & * & * \\ Y & 0 & 0 & 0 & 0 & 0 & * \\ Z & 0 & 0 & 0 & * & * & 0 \end{array}
$$

As X, Y, and Z each have their own unobservable, $D_{\mathcal{S},2} = I$. As unobservables are not caused by observables, $D_{\mathcal{S},3} = \mathbf{0}$. As Y does not cause X, the $(2,1)$ element of $D_{\mathcal{S},4}$ is 0.

Now consider the possible causal relations among X, Y, and Z specified by $D_{\mathcal{S},4} = [d_{jk}]$. By Proposition 5.1, acyclicality imposes three constraints $d_{jk} \times d_{kj} = 0$ and two constraints $d_{jk} \times d_{k\ell} \times d_{\ell j} = 0$, $j, k, \ell \in \{1, 2, 3\}$. Assume X is a cause of Y, either direct or indirect; then $d_{12} = 1$ and/or $d_{13} \times d_{32} = 1$. Since $d_{21} = 0$, there are 9 possibilities that we depict in Table III, using arrows to denote direct causality in the obvious way. We label these cases in relation to instruments Z. These include the "pre-cause," "intermediate cause," and "post-response" instrument cases. Structures not excluding Z from the structural equation for Y appear in the second column of the first, second, and third rows. We refer to entry $(1, 1)$ of Table III as the non-causal case, entry $(1, 2)$ as the joint cause case, and entry $(1, 3)$ as the joint response case. We also sometimes refer to entry $(2, 2)$ as the common cause instrument case.

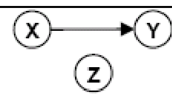
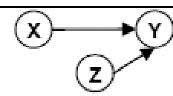
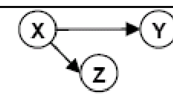
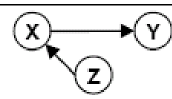
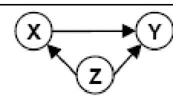
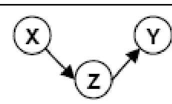
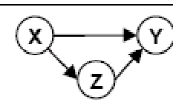
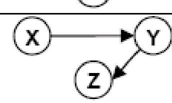
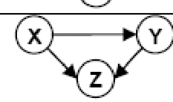
Non Causal, Joint Cause, and Joint Response			
Pre-Cause			
Intermediate Cause			
Post-Response			

Table III
Acyclic Causal Relationships among X, Y, and Z.

Classically, unobservables are treated as fundamental, either ruling out or ignoring causal relations among them. Considering just the presence or absence of statistical dependence among the unobservables, as is traditional, yields eight possibilities. Interacting these with the nine possibilities of Table III gives 72 possible systems.

But $D_{\mathcal{S},1}$ exposes a rich set of causal possibilities generating dependence patterns among unobservables. By Proposition 5.1, $D_{\mathcal{S},1} = [d'_{jk}]$ satisfies three constraints $d'_{jk} \times d'_{kj} = 0$ and two constraints $d'_{jk} \times d'_{k\ell} \times d'_{\ell j} = 0$, $j, k, \ell \in \{1, 2, 3\}$. This gives 25 possible structures among the unobservables. Interacting these with the possibilities of Table III gives 225 causal structures.

Each of these structures entails not only stochastic dependence but also conditional dependence among the unobservables, affording useful knowledge that would otherwise be unavailable.

In what follows, we study in detail leading examples of these structures, as well as more general structures with conditioning instruments, W. As a concise visual representation of the causality matrix, we use a form of causal graph suggested by White and Chalak (2009) in the spirit of Wright (1921, 1923), Spirtes, Glymour, and Scheines (1993), and Pearl (2000). To avoid misunderstanding, we emphasize that these graphs play no formal role in our analysis. They are solely a convenient means of depicting structural relations.

To describe these graphs, let $V \equiv (V_0', V_1, \ldots, V_J)$. The graph $G_{\mathcal{S}}$ for system $\mathcal{S}$ consists of a set of vertices (nodes), one for each element of V, and a set of arrows $\{a_{gh}\}$, corresponding to ordered pairs of distinct vertices. An arrow a_{gh} denotes that V_g directly causes V_h. We use solid nodes for observables and dashed nodes for unobservables. We link unobservables using a line with no arrow to indicate that these are stochastically dependent, without fully specifying their genesis. Here, this is due to either (i) one variable causing the other or (ii) both sharing a common cause that we do not depict. In the absence of structure to the contrary, the absence of lines between unobservables denotes independence.

For convenience, our graphs may represent vectors of variables as a single "vector" node. In this case, solid nodes denote observable vectors and dashed nodes denote vectors with some or all elements unobserved. An arrow from vector node Z to vector node X indicates that at least one component of Z is a direct cause of at least one element of X. A line with no arrow linking two dashed vector nodes denotes that these are dependent without specifying how this arises.

6 Single Extended Instrumental Variables

In this section, we study identification in the single XIV cases, XC, XI, XCI, XC|I, XI|C, XI|R, and XI|CR, paying particular attention to the causal role of the instruments.

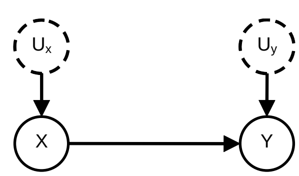
Let $U_x \equiv [U_{x_1}, \ldots, U_{x_k}]'$, $U_z \equiv [U_{z_1}, \ldots, U_{z_\ell}]'$, and $U_w \equiv [U_{w_1}, \ldots, U_{w_m}]'$ denote the unobserved causes of X, Z, and W, respectively. By assumption, each of these has mean zero.

6.1 Exogenous Causes

The simplest case is the classical exogenous cause (XC) case, where simple regression identifies the effect of X on Y. We represent this with the system $\mathcal{S}_1$ and its causal graph G_1:

(1) $X \overset{c}{=} \alpha_x U_x$
(2) $Y \overset{c}{=} X'\beta_o + U_y'\alpha_o$,

where $U_x \perp U_y$.



Graph 1 (G_1)
Exogenous Causes (XC)

Given $\mathcal{S}_1(1)$ and $U_x \perp U_y$, we have strict exogeneity by Dawid (1979, "D") lemma 4.2(i):

$$(XC) \quad \textit{Exogenous Causes}: \qquad X \perp U_y.$$

β_o is then fully identified if and only if $E(XX')$ is non-singular. The plug-in estimator for β_o is the usual OLS estimator, $\hat{\beta}_n^{XC} = (\mathbf{X}'\mathbf{X})^{-1}(\mathbf{X}'\mathbf{Y})$.

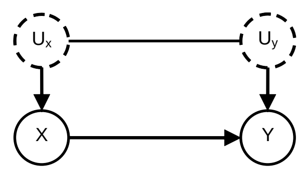
The analog of $E(XX')^{-1}E(XY)$ in nonseparable structures is $D_x E[Y \mid X = x]$. Under XC, this is also a structurally meaningful average marginal effect (e.g., Altonji and Matzkin, 2005).

Although $\mathcal{S}_1$ doesn't specify how $U_x \perp U_y$ arises, this independence does have causal content, due to Reichenbach's (1956) principle of common cause. Here, X and Y are correlated. Now, Y does not cause X. If $U_x \perp U_y$ because neither U_x nor U_y causes the other and U_x and U_y have no common cause, then neither can X and Y have a common cause (other than U_x, mediated solely by X). Thus, correlation between X and Y can only arise from the effect of X on Y. Alternatively, $U_x \perp U_y$ can hold even with a direct causal relation or common cause for U_x and U_y, as the converse of Reichenbach does not hold (see Chalak and White, 2009). Nevertheless, $U_x \perp U_y$ ensures that the effects of any common cause for X and Y, except one operating via X, are "neutralized." Again, the association between X and Y can only be due to the effect of X.

In experimental studies, randomization can ensure XC. In observational studies, XC is less plausible. To examine the failure of XC from a causal standpoint, consider $\mathcal{S}_2$ and its graph G_2:

(1) $X \stackrel{c}{=} \alpha_x U_x$
(2) $Y \stackrel{c}{=} X'\beta_o + U_y'\alpha_o,$

where $U_x \not\perp U_y$.



Graph 2 (G_2)
Endogenous Causes

For simplicity, let $U_x \not\perp U_y$ imply $E(U_x U_y') \neq 0$, so generally $E(XU_y'\alpha_o) \neq 0$. XC therefore fails, and β_o is not even partially identified, as moments involving unobservables appear in the moment equation $E(XY) = E(XX')\beta_o + E(XU_y'\alpha_o)$. Further, $\mathcal{S}_2$ does not specify how $U_x \not\perp U_y$ arises. By Reichenbach, either (*i*) U_x causes U_y, (*ii*) U_y causes U_x, or (*iii*) U_x and U_y share unobserved common causes, say U_0. U_x and U_y may also be dependent because they share common elements. Thus, the correlation between X and Y is due not just to X causing Y, but also or instead to the joint response of X and Y to U_y, to U_x, or to U_0.

In standard parlance, failure of XC is called *regressor endogeneity* and X is *endogenous*. In the treatment effect literature, this is *confoundedness of causes*. In $\mathcal{S}_2$, U_y, U_x, or U_0 are *confounding causes*. Thus, under Assumption 2.1, an endogenous cause shares an unobserved common cause with the response. Simultaneity is absent and is thus not responsible for endogeneity.

For XC, then, the presence of causal relations among unobservables creates obstacles for identification, whereas their absence creates opportunities.

6.2 Exogenous Instruments

When causes are endogenous, XC cannot identify β_o. But it is classical that "proper" instruments Z causing X can ensure identification. Z is *proper* if it is "valid," i.e. $E(Z\varepsilon) = 0$, and "relevant," i.e. $E(ZX') \neq 0$ (e.g., Hamilton, 1994 p.238; Hayashi, 2000 p.191; Wooldridge, 2002 p.83-84).

Wright (1928) first used instruments, which he called "curve shifters," to identify supply and demand elasticities (see Morgan, 1990; Angrist and Krueger, 2001; Stock and Trebbi, 2003). Wright states: "Such additional factors may be factors which (A) *affect* demand conditions without *affecting* cost conditions or (B) *affect* cost conditions without *affecting* demand conditions" (Wright, 1928, p.312; our italics). Clearly, Wright had in mind causality, not just correlation.

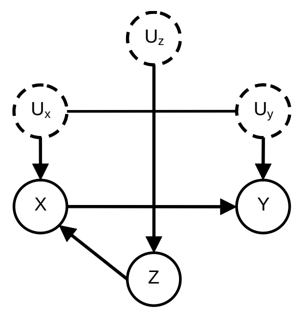
As we discuss next, standard IV methods fall into one of two causally meaningful subcategories, relating to instrument observability. As we show, the first case accords with the classical indirect least squares (ILS) interpretation for IV (Haavelmo, 1943, 1944). The second case extends this to situations where a suitable proxy for an unobserved instrument is available.

6.2.1 Observed Exogenous Instruments

Consider the classical triangular system $\mathcal{S}_3$ and its associated causal graph G_3:

$$
\begin{aligned}
&(1) \quad Z \overset{c}{=} \alpha_z U_z \\
&(2) \quad X \overset{c}{=} \gamma_x Z + \alpha_x U_x \\
&(3) \quad Y \overset{c}{=} X'\beta_o + U_y'\alpha_o,
\end{aligned}
$$

where $U_x \not\perp U_y$
and $(U_x, U_y) \perp U_z$



Graph 3 (G_3)
Observed Exogenous Instruments (OXI)

where γ_x is $k \times k$ (so $\ell = k$). As $U_x \not\perp U_y$, X is endogenous. Although $\mathcal{S}_3$ does not specify how this dependence arises, strict exogeneity holds, as D lemma 4.2(i) implies

$$\text{(XI)} \quad \textit{Exogenous Instruments}: \qquad Z \perp U_y.$$

Full identification holds if and only if $E(ZX')$ is non-singular, embodying the standard rank and order conditions. From $\mathcal{S}_3(1)$ and $\mathcal{S}_3(2)$, $E(ZX') = E(ZZ')\gamma_x + E(ZU_x')\alpha_x = E(ZZ')\gamma_x$. The non-singularity of γ_x and $E(ZZ')$ is thus necessary and sufficient for non-singularity of $E(ZX')$. The plug-in estimator is the familiar IV estimator, $\hat{\beta}_n^{XI} = (\mathbf{Z}'\mathbf{X})^{-1}(\mathbf{Z}'\mathbf{Y})$.

In $\mathcal{S}_3$, Z satisfies the following three causal properties essential to identifying β_o:

CP:OXI (Causal Properties for Observed Exogenous Instruments) *(i) Z causes X, and the effect of Z on X is identified via XC; (ii) Z indirectly causes Y, and the (total) effect of Z on Y is identified via XC; (iii) Z causes Y only via X.*

Condition (i) holds by $\mathcal{S}_3(1,2)$; (ii) holds by substituting $\mathcal{S}_3(2)$ into $\mathcal{S}_3(3)$, giving

$$(4) \qquad Y \overset{c}{=} Z'\pi_o + U_x'\alpha_x'\beta_o + U_y'\alpha_o,$$

where $\pi_o \equiv \gamma_x'\beta_o$. Condition (iii) enforces the exclusion restriction that Z does not enter $\mathcal{S}_3(3)$. Failure of any of these conditions results in XI failing to identify β_o in $\mathcal{S}_3$. Among other things, the presence of causal relations between U_z and (U_x, U_y) creates obstacles for identification; their absence creates opportunities.

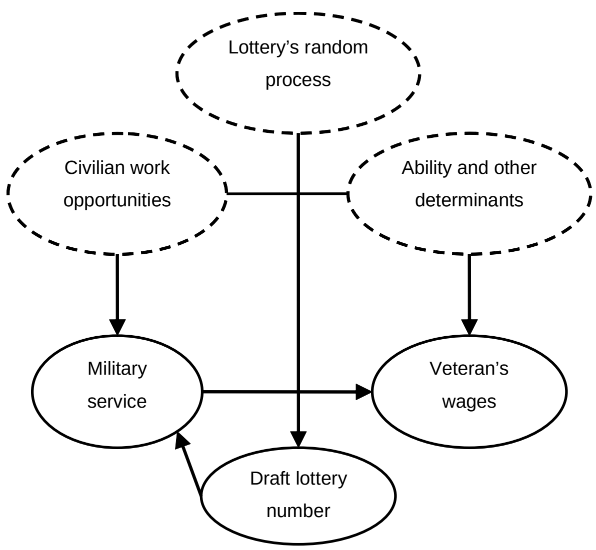
When it holds, CP:OXI justifies the classical ILS interpretation for IV (Haavelmo, 1943, 1944), as $\beta_o = \gamma_x'^{-1}\pi_o$, with γ_x and π_o identified by XC. Thus,

$$\beta_o = \gamma_x'^{-1}\pi_o = \{E(ZZ')^{-1}E(ZX')\}^{-1}\{E(ZZ')^{-1}E(ZY)\} = E(ZX')^{-1}E(ZY).$$

This motivates the construction of "derivative ratio" (DR) effects in more general settings. For example, Heckman (1997) and Heckman and Vytlacil (1999, 2001, 2005, 2007) have proposed local ILS (also called "local IV") estimators for such cases. Schennach, White, and Chalak (2009) give a detailed analysis of DR effects identified by XI in the general nonseparable case.

As Z is observable, we call Z *observed exogenous instruments* (OXI).

OXI Example Angrist (1990) provides an example of OXI where all causal elements are clear. Angrist is interested in the effect of Vietnam War military service on a veteran's post-war civilian wage.



Graph 4 (G_4)
OXI for the Effect of Military
Service on Veteran's Wage

Vietnam War service and civilian wage could be confounded by variables such as individual ability or education, as either might affect both joining the military and civilian wages. As military service is thus potentially endogenous, Angrist employs the Vietnam draft lottery number as an instrument. This number was randomly assigned by birth date; those whose lottery number fell below a certain threshold were "draft-eligible." Those with a lottery number above the threshold were not required to serve in the military.

Angrist (1990) assumes that the randomness of the lottery number makes it independent of unobserved factors affecting military service or civilian wages, and that the lottery number does not affect veteran's wages except via military service, as in G_4. If the data are indeed generated in this way, the lottery number satisfies CP:OXI and is a legitimate OXI.

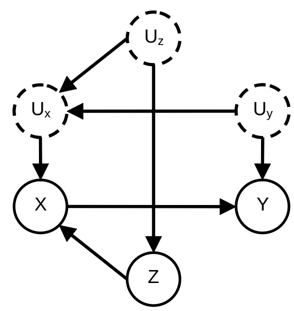
6.3 Proxies for Unobserved Exogenous Instruments

Satisfying CP:OXI in $\mathcal{S}_3$ requires that XC for Z identifies γ_x and π_o, as when Z is randomized. But randomized Z's can be just as hard to justify as randomized X's. On the other hand, Theorem 3.1, as is standard (e.g., Heckman, 1996), does not require $Z \perp U_x$: the effect of Z on X, usually estimated from a "first stage" regression, need not be identified, as we now demonstrate.

Consider $\mathcal{S}_5$ and its graph G_5:

(1) $U_x \stackrel{c}{=} \delta_y U_y + \delta_z U_z$
(2) $Z \stackrel{c}{=} \alpha_z U_z$
(3) $X \stackrel{c}{=} \gamma_x Z + \alpha_x U_x$
(4) $Y \stackrel{c}{=} X'\beta_o + U_y'\alpha_o$,

where $U_y \perp U_z$,



Graph 5 (G_5)
Proxies for Unobserved Exogenous Instruments (PXI)

with γ_x $k \times k$ (so $\ell = k$). Whereas in $\mathcal{S}_3$ we have $U_x \perp U_z$, here $U_x \not\perp U_z$. A key feature is that the dependence relations among U_x, U_y, and U_z arise because U_y and U_z are independent joint causes of U_x. Unlike $\mathcal{S}_1$ and $\mathcal{S}_3$, $\mathcal{S}_5$ specifies explicit causal relations among the unobservables.

It is easy to see that XI is satisfied, as $Z \perp U_y$. As we also have $Z \not\perp X$, Z is a proper instrument. Thus, Theorem 3.1 applies with $\tilde{Z} = Z$ and $\tilde{m} = 0$, identifying β_o.

The key difference between $\mathcal{S}_5$ and $\mathcal{S}_3$ is that the causal properties making Z instrumental for identifying β_o in $\mathcal{S}_5$ are satisfied *not* by Z but by U_z. If these were observable, they could act as proper instruments. In their absence, Z acts as their proxy. Accordingly, we call Z in $\mathcal{S}_5$ *proxies* for (unobserved) exogenous instruments (PXI). The necessary[4] causal properties are now:

CP:PXI (Causal Properties for Proxies for Exogenous Instruments) *(i) U_z indirectly causes X, and the total effect of U_z on X could be identified via XC had U_z been observed; (ii) U_z indirectly causes Y, and the total effect of U_z on Y could be identified via XC had U_z been observed; (iii) U_z causes Y only via X.*

[4] CP:OXI and CP:PXI account for both the successes and failures of standard IV. Appendix D contains a detailed causal description of how identification of β_o fails in the "irrelevant instrument," "invalid instrument," and "under-identified" cases. XI holds in the latter, but the exclusion restrictions are violated, so $E(Z^* X')$ is singular.

The CP:PXI conditions are analogs of CP:OXI with U_z replacing Z. In (i) and (ii), we refer to the total effect of U_z on X and Y. In (i), this includes not only the effect of U_z on X through Z, but also through U_x, and similarly for the effect on Y in (ii), ensured by the reduced form

$$(5) \qquad Y \overset{c}{=} Z'\pi_o + U_x'\alpha_x'\beta_o + U_y'\alpha_o \overset{c}{=} U_z'(\alpha_z'\pi_o + \delta_z'\alpha_x'\beta_o) + U_y'(\alpha_o + \delta_y').$$

CP:PXI(iii) enforces the exclusion restriction on Z, but Z need not (indirectly) cause Y.

The classical ILS account *fails* for Z in $\mathcal{S}_5$, as $U_x \not\perp U_z$ prevents XC from identifying either γ_x or π_o. Indeed, if we were to follow Pearl (2000, p.153-154), who advocates identification strategies that recover causal effects as functions only of identified effects, we would proceed no further.

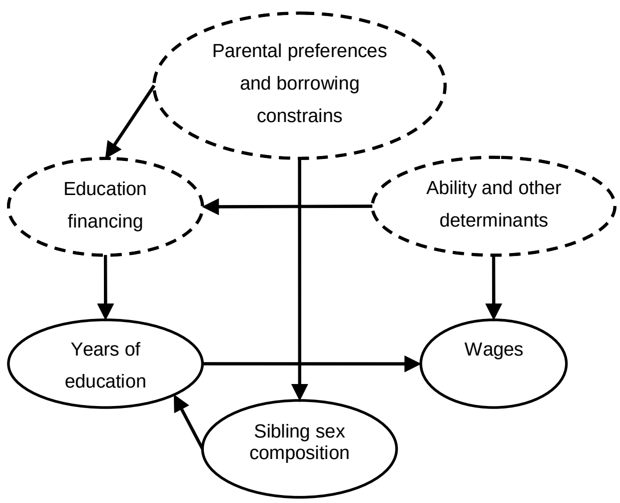
Interestingly, however, these identification failures exactly offset here. Since $\gamma_x = E(XZ')E(ZZ')^{-1} - \alpha_x E(U_x Z')E(ZZ')^{-1}$ and $\pi_o = E(ZZ')^{-1}E(ZY) - E(ZZ')^{-1}E(ZU_x')\alpha_x'\beta_o = \gamma_x'\beta_o$,

$$\begin{aligned} E(ZX')^{-1}E(ZY) &= \{E(ZZ')^{-1}E(ZX')\}^{-1}E(ZZ')^{-1}E(ZY) \\ &= \{\gamma_x' + E(ZZ')^{-1}E(ZU_x')\alpha_x'\}^{-1}\{\gamma_x' + E(ZZ')^{-1}E(ZU_x')\alpha_x'\}\beta_o \\ &= \beta_o. \end{aligned}$$

This even permits $\gamma_x = 0$, so Z need not cause X, whereas γ_x had to be invertible in $\mathcal{S}_3$.

Thus, for PXI, the *presence* of specific causal relations among unobservables identifies β_o.

PXI Example A number of applied papers that use standard IV implicitly employ PXI. Consider, for example, measuring the effect of education on future wages, as in Butcher and Case (1994, BC). As BC note, years of education and wages can be confounded by unobservables such as ability, making years of education potentially endogenous.



Graph 6 (G_6)
PXI for the effect of Education of wages

To address this, BC employ an individual's sibling sex composition as an instrument. BC discuss a variety of mechanisms by which household composition may be associated with children's educational attainment. These include parental preferences, investment in children's education under borrowing constraints, and developmental psychology. Further, BC argue that this association is unlikely to be related to future wages by means other than educational attainment.

In our terms, BC exploit correlation between a child's educational attainment and sibling sex composition without requiring this correlation to be solely due to the effect of one on the other. For example, parental preferences and their capacity to help finance their children's' education may generate a correlation between education level and sibling sex composition, as in G_6. If so, then sibling composition is a valid proxy for the unobserved exogenous instruments, parental preferences and borrowing constraints.

6.3.1 Nonseparability and Further Comments

The key differences between OXI and PXI become more stark when the response functions are nonparametric and nonseparable. For simplicity, let $\ell = k = 1$ and suppose $U_x \stackrel{c}{=} r_{u_x}(U_y, U_z), Z \stackrel{c}{=} r_z(U_z), X \stackrel{c}{=} r_x(Z, U_x), Y \stackrel{c}{=} r_y(X, U_y)$, and $U_y \perp U_z$ as in S_5. Thus, Z is a standard valid and relevant instrument, but it does *not* ensure identification: The PXI derivative ratio $D_z E(Y \mid Z = z)/D_z E(X \mid Z = z)$ is no longer a meaningful effect measure, as neither the numerator nor denominator have structural meaning, nor are they identically affected by the confounding. The cancellation that occurs for ILS in the linear PXI case does not hold generally. Instead, the ratio $D_{u_z} E(Y \mid U_z = u_z)/D_{u_z} E(X \mid U_z = u_z)$ has structural meaning parallel to the OXI case. The challenge is that this involves unobservables, U_z. Schennach, White, and Chalak (2009) show that in the nonseparable case, observing *two* proxies Z_1 and Z_2 for U_z suffices to fully identify this ratio. Here, the linear case masks deeper aspects of the general case.

Angrist, Imbens, and Rubin (1996) (AIR) provide an explicit causal account of IV in which they list sufficient conditions for the IV estimator to have a causal interpretation, namely that of "an average causal effect for a subgroup of units, the compliers." AIR do not impose linearity or separability and consider the case of binary treatment and observed exogenous binary assignment to treatment. Interestingly, PXI falls outside of AIR's framework, as the instrument proxy need not be random and need not determine the treatment. Chalak (2009b) provides a generalization of AIR to cases where only such proxies are available.

These substantive differences between OXI and PXI underscore the critical importance of the causal structure relating the unobservables: PXI differs from OXI *primarily with respect to the causal links among the unobservables*, and, consequently, with respect to the causal role of Z.

We call any Z satisfying OXI or PXI an *unconditional* instrumental variable, to distinguish it from the *conditional* instrumental variables discussed in Section 6.5 and Section 7.

For brevity, we omit treating the XCI case here (i.e., $(X, Z) \perp U_y$), taking this as a special case of XCI|I (i.e., $(X, Z) \perp U_y \mid W$), studied in Section 7.2.

6.4 Conditioning Instruments

In $\mathcal{S}_1$ and $\mathcal{S}_3$, treatment randomization neutralizes confounding (Fisher, 1935, ch.2); but randomization is rare in observational studies. In these studies, matching observations from treatment and control groups that share common causes or attributes provides a way forward. Conditioning on the information in these confounding variables permits interpreting the remaining conditional association between the potential cause and its response as the causal effect. Developments along these lines include ignorability and the propensity score (Rubin, 1974; Rosenbaum and Rubin, 1983), selection on observables (Barnow, Cain, and Goldberger, 1980; Heckman and Robb, 1985), control functions (Heckman and Robb, 1985; Imbens and Newey, 2009), the back-door method (Pearl, 1995), and predictive proxies (White, 2006; Chalak and White, 2007). Especially in labor economics, matching methods are well known and have been applied to study the distribution of earnings, policy evaluation, and the returns to education and training programs (Roy, 1951; Griliches, 1977; Heckman and Robb, 1985; Heckman, Ichimura, and Todd, 1998). Abadie and Imbens (2006) study large sample properties of matching estimators.

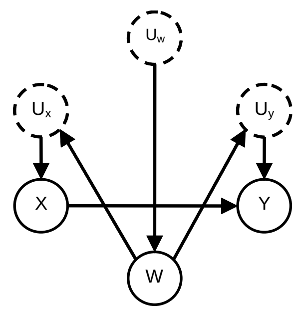
A crucial challenge for applications is to understand when control variables (also called "covariates") can be used to ensure unconfoundedness and what covariates to use. We address this issue by studying structures where *conditioning instruments* W proxy for unobserved confounding variables, ensuring conditional exogeneity and identification.

6.4.1 Observed Common Causes

Consider $\mathcal{S}_2$, where X is endogenous because $U_x \not\perp U_y$. Suppose that $U_x \not\perp U_y$ arises because U_x and U_y have a common cause. To gain insight, suppose we observe this confounding common cause, W. This violates our convention that observables do not cause unobservables, so this is just temporary. Thus, consider system $\mathcal{S}_{7a}$ and its causal graph G_{7a}:

(1) $W \stackrel{c}{=} \alpha_w U_w$
(2) $U_{x_1} \stackrel{c}{=} \gamma_{x_1} W$
(3) $U_{y_1} \stackrel{c}{=} \gamma_{y_1} W$
(4) $X \stackrel{c}{=} \alpha_{x_1} U_{x_1} + \alpha_{x_2} U_{x_2}$
(5) $Y \stackrel{c}{=} X'\beta_o + U'_{y_1}\alpha_{o1} + U'_{y_2}\alpha_{o2}$,

where $U_w \perp U_{x_2}, U_w \perp U_{y_2},$ and $U_{x_2} \perp U_{y_2}$,
with $U_x \equiv (U'_{x_1}, U'_{x_2})', U_y \equiv (U'_{y_1}, U'_{y_2})'$.



Graph $7a$ (G_{7a})
Conditioning Instruments

Regressor endogeneity ($X \not\perp U_y$) arises from correlation between U_{x_1} and U_{y_1}, resulting from the common cause W. The unobservables U_{x_2} and U_{y_2} ensure that X is not entirely determined by W and that Y is not entirely determined by X and W. (Below, we drop explicit reference to such independent sources of variation; their presence will be implicitly understood.)

In $\mathcal{S}_{7a}$, conditioning on W ensures that the remaining association between X and Y can only be interpreted as the causal effect of X on Y. $\mathcal{S}_{7a}$ ensures the key condition:

(XC|I) *Conditionally Exogenous Causes given Conditioning Instruments:* $X \perp U_y \mid W$.

We call W "conditioning instruments" to emphasize their instrumental role in ensuring identification. Imbens and Newey (2009) call such variables "control variables," also known as covariates. Although W has been called "exogenous" in the literature, Imbens (2004, p.5) notes that this is non-standard. Indeed, W is endogenous, as $W \not\perp U_y$. By instead referring to X as conditionally exogenous, to W as conditioning instruments, and to XC|I as a conditional form of exogeneity, we avoid confusion, clarify the causal roles of the variables involved, and facilitate the distinction between these instruments, standard instruments, and other extended instruments discussed here.

Putting $\tilde{Z} = X$ and $\tilde{W} = W$, the analysis of Section 4 shows that β_o is fully identified as

$$\beta_o = [E(\{X - E(X \mid W)\}X')]^{-1} E(\{X - E(X \mid W)\}Y)$$

if and only if $E(\{X - E(X \mid W)\}X')$ is non-singular. In the nonseparable case, XC|I identifies the derivative $D_x E[Y \mid W, X = x]$ as an average marginal effect analogous to β_o.

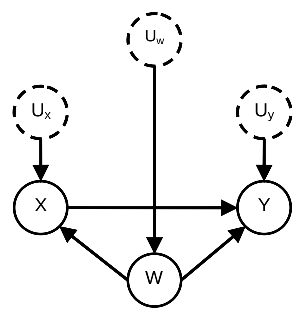
A feasible plug-in estimator for β_o based on linearity of $E(X \mid W)$ is

$$\hat{\beta}_n^{XC|I} \equiv \{\mathbf{X}'(\mathbf{I} - \mathbf{W}(\mathbf{W}'\mathbf{W})^{-1}\mathbf{W}')\mathbf{X}\}^{-1}\{\mathbf{X}'(\mathbf{I} - \mathbf{W}(\mathbf{W}'\mathbf{W})^{-1}\mathbf{W}')\mathbf{Y}\}.$$

This is the Frisch-Waugh (1933) partial regression estimator, obtained by regressing $\mathbf{Y}$ on the residuals from the regression of $\mathbf{X}$ on $\mathbf{W}$. This is also the OLS estimator for β_o in the regression of Y on X and W. This regression emerges naturally from $\mathcal{S}_{7a}$ by performing the substitutions enforcing our convention that observables do not cause unobservables, as represented in $\mathcal{S}_{7b}$:

(1) $W \overset{c}{=} \alpha_w U_w$
(2) $X \overset{c}{=} \gamma_x W + \alpha_x U_x$
(3) $Y \overset{c}{=} X'\beta_o + W'\gamma_o + U_y'\alpha_o$,

where $U_w \perp U_x$,
$U_w \perp U_y$, and $U_x \perp U_y$,



Graph 7b (G_{7b})
Conditioning Instruments

In writing $\mathcal{S}_{7b}$, we adjust the notation in the natural way. With the assumed independence, Theorem 3.1 applies with[5] $\tilde{Z} = (X, W)$ satisfying XC ($\tilde{m} = 0$). In $\mathcal{S}_{7b}$, both β_o, the direct (and full) effect of X on Y, and γ_o, the direct effect of W on Y, are identified.

Momentarily shifting attention to the effects of W, note that the full effect of W on Y is $\gamma_o + \gamma_x'\beta_o$, identified from a regression of Y on W only. Interestingly, in this regression, omitting the causally relevant X results in omitted variable bias *only if* one is interested in the direct rather than the full effect of W. When the full effect of W is of interest, including a variable caused by W results in *included* variable bias. The proper choice of variables thus depends on the specific effect of interest. The traditional account of the omitted variables "problem" is incomplete.

[5] In $\mathcal{S}_{7b}$ and $\mathcal{S}_{8a}$ below, W appears in $\tilde{Z}$, abusing our convention that W not appear in $\tilde{Z}$. But this is just for convenience, as the underlying identifying condition is indeed XC|I: $X \perp U_y \mid W$.

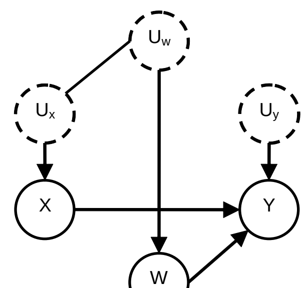
6.4.2 Proxies for Unobserved Common Causes

Structural Proxies $\mathcal{S}_{7a}$ ($\mathcal{S}_{7b}$) is not necessary for XC|I. Pearl's (1995; 2000, pp. 79-81) back-door structures, where an observable (W) mediates a link between X and Y, also ensure XC|I. Here, W acts either as a common cause (G_{7a}, G_{7b}), or as a response to the unobserved common cause and a cause of either Y or X (G_{8a}, G_{8b}). XC|I holds in G_{8a} and G_{8b}, as the unobserved confounding common cause causes Y via W (G_{8a}) or X via W (G_{8b}). As W is a structurally relevant proxy for the unobserved common cause, we may refer to W as a "structural proxy."

Specifically, let $\mathcal{S}_{8a}$ be given by

(1) $W \overset{c}{=} \alpha_w U_w$
(2) $X \overset{c}{=} \alpha_x U_x$
(3) $Y \overset{c}{=} X'\beta_o + W'\gamma_o + U_y'\alpha_o$,

where $U_w \not\perp U_x$,
$U_w \perp U_y$, and $U_x \perp U_y$.

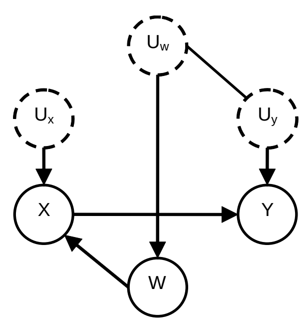


Graph 8a (G_{8a})
Conditioning Instruments

Similarly, let $\mathcal{S}_{8b}$ be given by

(1) $W \overset{c}{=} \alpha_w U_w$
(2) $X \overset{c}{=} \gamma_x W + \alpha_x U_x$
(3) $Y \overset{c}{=} X'\beta_o + U_y'\alpha_o$,

where $U_w \perp U_x$,
$U_w \not\perp U_y$, and $U_x \perp U_y$.



Graph 8b (G_{8b})
Conditioning Instruments

Note that in $\mathcal{S}_{8a}$, the causal relation between U_x and U_w is unspecified, so $\mathcal{S}_{8a}$ corresponds to three possible back-door structures each generating $U_w \not\perp U_x$. But in each case, X and W jointly satisfy XC in (3), so Theorem 3.1 holds with $\tilde{Z} = (X, W)$ and $\tilde{m} = 0$. We can identify β_o from the regression of Y on X and W. Here, W is a relevant exogenous variable correlated with X, so omitting W leads to omitted variable bias in measuring the direct (and full) effect β_o of X. Just as in $\mathcal{S}_{7b}$, however, the traditional account of omitted variables is not the whole story.

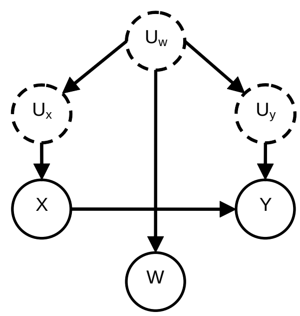
Similarly, $\mathcal{S}_{8b}$ corresponds to three possible back-door structures, each generating $U_w \not\perp U_y$. For concreteness, suppose U_y causes U_w. Now both X and W are endogenous, as $X \not\perp U_y$ and $W \not\perp U_y$. According to the textbooks, regressing Y on X and W should yield nonsense, due to this endogeneity. Nevertheless, this regression identifies β_o, because XC|I holds for $\tilde{Z} = X$, $\tilde{W} = W$, despite the failure of XC for $\tilde{Z} = (X, W)$.

What about the regression coefficients for W? In $\mathcal{S}_{8b}$, these have no causal interpretation, but only a predictive interpretation (see, e.g., White (2006) and White and Chalak (2009a)). Thus, *some regression coefficients have causal meaning* (those for the conditionally exogenous X), *but others do not* (those for the conditioning instruments W). That is, though all regression coefficients have predictive meaning, only some have structural meaning. This is an important example of what Heckman (2006) has termed "Marschak's maxim": we can identify some economically meaningful components of a structure (β_o) without having to identify the entire structure. Thus, *not all regression coefficients need have signs and magnitudes that make causal sense.* This has significant implications for how researchers can and should assess validity of regression results.

Predictive Proxies Pearl's back-door method does not exhaust the possibilities for XC|I. Using "predictive proxies" (White, 2006; Chalak and White, 2007; White and Chalak, 2009a) can also be viable. Structures such as $\mathcal{S}_9$, which violates Pearl's back-door criterion, can ensure XC|I:

(1) $W \overset{c}{=} \alpha_{w_1} U_{w_1} + \alpha_{w_2} U_{w_2}$
(2) $U_{x_1} \overset{c}{=} \delta_{x_1} U_{w_1}$
(3) $X \overset{c}{=} \alpha_{x_1} U_{x_1} + \alpha_{x_2} U_{x_2}$
(4) $U_{y_1} \overset{c}{=} \delta_{y_1} U_{w_1}$
(5) $Y \overset{c}{=} X'\beta_o + U'_{y_1}\alpha_{o1} + U'_{y_2}\alpha_{o2}$,

where $U_{w_1} \perp U_{w_2}, U_{w_2} \perp U_{x_2}$
$U_{w_2} \perp U_{y_2}$, and $U_{x_2} \perp U_{y_2}$,



Graph 9 (G_9)
Conditioning Instruments

with $U_w \equiv (U'_{w_1}, U'_{w_2})', U_x \equiv (U'_{x_1}, U'_{x_2})'$, and $U_y \equiv (U'_{y_1}, U'_{y_2})'$ so that $W \not\perp U_y$, and $X \not\perp U_y$.

In $\mathcal{S}_9$, U_{w_1} is an unobserved common cause for X and Y; the predictive proxy W may be a mismeasured version of U_{w_1}, for example. Again, both X and W are endogenous, but now W is structurally irrelevant. In $\mathcal{S}_9$, U_{w_1} is the confounding common cause of X and Y. Had U_{w_1} been observable, XC|I would ensure identification, as $X \perp U_y \mid U_{w_1}$. The structure of $\mathcal{S}_9$ need not always ensure $X \perp U_y \mid W$, so predictive proxies W may not always be viable conditioning instruments. Nevertheless, further conditions may ensure this. The key is W's ability to predict either X or U_y sufficiently well that neither contains additional information useful in predicting the other. Heuristically, W should contain the information in U_{w_1} whose knowledge makes X and Y conditionally independent and should not contain further information that may induce dependence between X and Y (see Chalak and White, 2007, theorem 2.8 and corollary 2.9). A simple case where this holds is when $\alpha_{w_2} = \mathbf{0}$ and W is a one-to-one function of U_{w_1}.

Just as in $\mathcal{S}_{8b}$, the regression coefficients associated with the conditionally exogenous regressors X have causal meaning, whereas the others do not: not all regression coefficients need have signs and magnitudes that make causal sense.

Although XC|I permits identification, the results of Hausman and Taylor (1983) do not apply. Their necessary order condition (prop. 4) *fails* here, as the number of unconstrained coefficients in $\mathcal{S}_9(5)$ $(k > 0)$ exceeds the number (zero) of "predetermined" (uncorrelated with U_y) variables. Similarly, as $\mathcal{S}_9$ imposes no error covariance restrictions ($U_x \not\perp U_y$, $U_x \not\perp U_z$, and $U_y \not\perp U_z$), Hausman and Taylor (1983, prop. 6) doesn't apply, and their sufficient condition for identification in the full information context (prop. 9) is not satisfied. Instead, a restriction on the *conditional* covariance of the unobserved causes, specifically $U_x \perp U_y \mid U_w$, ensures XC|I here, identifying β_o.

6.4.3 Further Comments

XC|I enables matching. Let Y_x be the value Y would take had X been set to x (the "potential outcome"). E.g., in $\mathcal{S}_9(5)$, $Y_x = x'\beta_o + U_y$. This and XC|I imply the key unconfoundedness (ignorability) condition $Y_x \perp X \mid W$ of Rosenbaum and Rubin (1983) (White, 2006, prop. 3.2). Note that here, as in Hirano and Imbens (2004), treatments need not be binary.

Given its role as an observable proxy for unobserved common causes of X and Y, we call W a vector of *common cause* instruments. Note that in contrast to XI, XC|I does not require $\ell = k$.

Because XC|I is so straightforward (in particular, there are no necessary exclusion restrictions), we do not provide causal properties for XC|I parallel to CP:OXI or CP:PXI. Nevertheless, we conjecture that in $\mathcal{S}_9/G_9$ XC|I implies (possibly with mild additional conditions) that X cannot cause W (see $\mathcal{S}_9(1)$). In particular, observe that if X causes W in $\mathcal{S}_9$, then conditioning on W, a common response of X and U_w, generally renders X and U_w dependent given W (see Chalak and White, 2009). As U_w causes U_y, this may lead XC|I to fail.

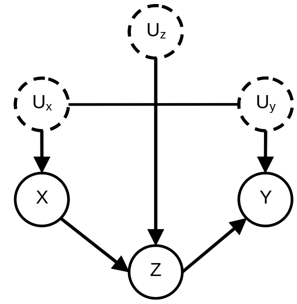
Systems $\mathcal{S}_{7b}$, $\mathcal{S}_{8a}$, $\mathcal{S}_{8b}$, and $\mathcal{S}_9$ are prototypical examples where identification via conditioning instruments obtains. In each, the presence of specific causal relations among the unobservables is key to ensuring XC|I. Chalak and White (2007) and White and Chalak (2009a) present further substantive analysis for the general nonlinear, nonseparable case.

6.5 Conditional Instruments

We now examine how *conditional* instruments Z can identify the effect of endogenous X on Y as the product of the effect of X on Z and that of Z on Y. To illustrate, consider system $\mathcal{S}_{10}$:

(1) $X \stackrel{c}{=} \alpha_x U_x$
(2) $Z \stackrel{c}{=} \gamma_z X + \alpha_z U_z$
(3) $Y \stackrel{c}{=} Z'\delta_o + U_y'\alpha_o$,

where $U_z \perp U_x$,
$U_z \perp U_y$, and $U_x \not\perp U_y$.



Graph 10 (G_{10})
Conditional Instruments

Substituting $\mathcal{S}_{10}(2)$ into $\mathcal{S}_{10}(3)$ with $\beta_o \equiv \gamma_z' \delta_o$ gives

$$(4) \qquad Y \overset{c}{=} X'\beta_o + U_z'\alpha_z'\delta_o + U_y'\alpha_o.$$

$\mathcal{S}_{10}$ doesn't specify how dependence between U_x and U_y arises. Clearly, X and Z are endogenous in (4), so neither XC nor XI can identify β_o. Also, XC|I fails, as $X \not\perp U_y \mid Z$. But $\mathcal{S}_{10}$ ensures

(XI|C) *Conditionally Exogenous Instruments given Causes:* $Z \perp U_y \mid X$

Theorem 3.1 applies with $\tilde{Z} = Z$ and $\tilde{W} = X$. This identifies δ_o from $\mathcal{S}_{10}(4)$ by XC|I with causes Z and conditioning instruments X. If γ_z can also be identified, then identification of $\beta_o \equiv \gamma_z'\delta_o$ follows from Corollary 3.2. But γ_z is identified from $\mathcal{S}_{10}(2)$ by XC, as $X \perp U_z$. In contrast to XI but like XC|I, XI|C does not require $\ell = k$.

The feasible plug-in estimator with linearity of $E(Z \mid X)$ is:

$$\hat{\beta}_n^{XI|C} \equiv \{(\mathbf{X}'\mathbf{X})^{-1}(\mathbf{X}'\mathbf{Z})\} \times \{[\mathbf{Z}'(\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}')\mathbf{Z}]^{-1}[\mathbf{Z}'(\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}')\mathbf{Y}]\}.$$

Although XI|C uses a single XIV Z to identify β_o, both Z and X play dual roles. Z is both a response for X and a cause for Y. The causes X are exogenous with respect to U_z in $\mathcal{S}_{10}(2)$ and are conditioning instruments with respect to U_y in $\mathcal{S}_{10}(3)$.

Just as for XI, a succinct set of causal properties characterizes structural identification:

CP:XI|C (Causal Properties of Conditionally Exogenous Instruments Given Causes) *(i) The effect of X on Z is identified via XC with exogenous causes Z; (ii) that of Z on Y is identified via XC|I with conditioning instruments X; (iii) If X causes Y, it does so only via Z.*

Note the exclusion restriction enforced by *(iii)*. As is easily seen, $\mathcal{S}_{10}$ satisfies CP:XI|C.

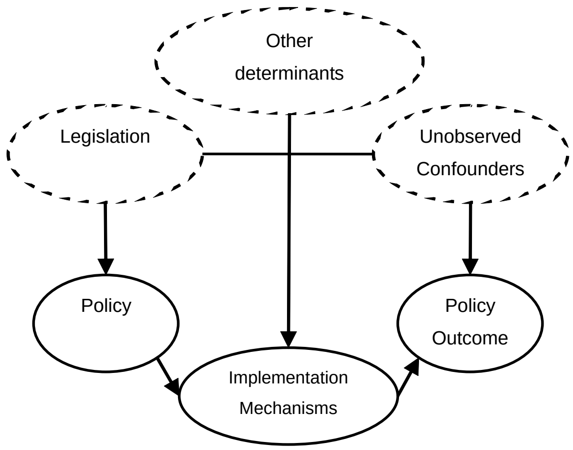
XI|C corresponds to Pearl's (1995, 2000) "front-door" method. Whereas the treatment effect literature applies XC|I (e.g., back door) to identify effects using covariates W unaffected by treatment to neutralize confounding, XI|C (front door) uses variables that *are* affected by (indeed, mediate) treatment. We thus call such Z's *intermediate-cause* instruments.

The structure of $\mathcal{S}_{10}$ can be analyzed via Hausman and Taylor (1983). As γ_z is identified by XC in $\mathcal{S}_{10}(2)$ and $U_z \perp U_y$, Hausman and Taylor (1983, prop. 6) ensures that residuals from a regression based on $\mathcal{S}_{10}(2)$ can play the role of derived instruments for Z in $\mathcal{S}_{10}(3)$, identifying δ_o.

With nonseparability, XI|C identifies $D_x E(Z \mid X = x) \times D_z E(Y \mid X = x, Z = z)$ as an analog of $\beta_o \equiv \gamma_z'\delta_o$. The structural content of this has yet to be explored. We also leave for future study structures where only proxies for mediating instruments are observed. There, the presence of causal relations among unobservables may help ensure identification, as for PXI.

6.5.1 XI|C Examples

Whereas standard IV methods exploit exogenous variation preceding a possibly endogenous policy to identify effects, XI|C exploits the way that the policy implementation channels the policy effects. Specifically, consider a policy that is endogenous, as it is affected by unobserved drivers of the policy outcome, as in G_{11}. E.g., consider the effect of per pupil funding on students' performance measured by standardized test scores (see e.g. Gordon and Vegas, 2005). Each school district collects taxes that determine per pupil funding. Because tax rates and family income correlate with household characteristics that can affect student performance, we may suspect that education funding is endogenous.



Graph 11 (G_{11})
Policy Evaluation by means of XI|C

One can recover the policy effect using policy-driven intermediate-cause instruments that in turn affect the policy outcome. These XIVs should not be driven by the unobserved common confounding causes of the policy and the response of interest, other than through the policy. In the education funding example, potential intermediate-cause instruments could be educational attainment and training of teachers, school investment in new technology, class size, etc.

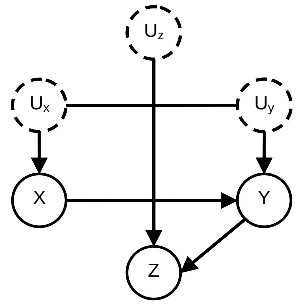
XI|C (and XIV methods generally) can also be used to test the validity of imposed assumptions. For instance, a controversial topic in experimental economics is whether to use cognition data in studying decisions or to only use classic choice data. Here, XI|C can be used to test whether information search fully reveals strategic thinking in experimental studies (e.g., Crawford, 2008). In these experiments, players must click on computer screen windows to access information (e.g., payoffs), yielding records of search patterns (occurrence, duration, adjacency, or repetition). If search patterns fully mediate strategic thinking, XI|C can identify effects of treatment on players' decisions. If treatments are also randomized across players, one has an over-identifying restriction useful for testing whether information search is fully revealing or if instead either information search and players' decisions are driven by other factors or treatment affects players' decisions through a channel other than information search.

6.6 Conditionally Exogenous Instruments given Response

We now study using $Z \perp U_y \mid Y$ (XI|R) to identify β_o. A structure generating this is one where Z (with $\ell = 1$) is a post-response instrument, as in $\mathcal{S}_{12}$:

(1) $X \overset{c}{=} \alpha_x U_x$
(2) $Y \overset{c}{=} X'\beta_o + U_y'\alpha_o,$
(3) $Z \overset{c}{=} Y\gamma_z + U_z'\alpha_z$

where $U_z \perp (U_x, U_y)$
and $U_x \not\perp U_y$.



Graph 12 (G_{12})

Substituting $\mathcal{S}_{12}(2)$ into $\mathcal{S}_{12}(3)$ with $\delta_o \equiv \beta_o\gamma_z$ gives

$$(4) \qquad Z \overset{c}{=} X'\delta_o + U_y'\alpha_o\gamma_z + U_z'\alpha_z.$$

This structure might look promising, as one may consider identifying $\beta_o \equiv \delta_o/\gamma_z$ by methods analogous to ILS. Nevertheless, $Z \perp U_y \mid Y$ cannot identify β_o here. Given the (implied) linearity of $E(Z \mid Y)$, the relevant moment equation for β_o is

$$\{E(ZY') - E(ZY')E(YY')^{-1}E(YY')\} - \{E(ZX') - E(ZY')E(YY')^{-1}E(YX')\}\beta_o = 0.$$

But $E(ZY') - E(ZY')E(YY')^{-1}E(YY') = 0$ and $E(ZX') - E(ZY')E(YY')^{-1}E(YX') = 0$, so condition ($c$) of Theorem 3.1 fails, leaving β_o undetermined. The problem is that δ_o is not identified, as X and Z are confounded by the unobserved common cause of X and Y. Prop. 7 of Hausman and Taylor (1983) also rules out identifying β_o using the estimated residuals from a regression based on $\mathcal{S}_{12}(3)$ as derived instruments, as $Y \perp U_z$.

We also have $Z \perp U_y \mid (X, Y)$ (XI|CR) in $\mathcal{S}_{12}$. Letting $\tilde{W} = [X', Y]'$, a similar analysis gives

$$\{E(ZY') - E(Z\tilde{W}')E(\tilde{W}\tilde{W}')^{-1}E(\tilde{W}Y')\} - \{E(ZX') - E(Z\tilde{W}')E(\tilde{W}\tilde{W}')^{-1}E(\tilde{W}X')\}\beta_o = 0.$$

But condition (c) of Theorem 3.1 again fails, as $E(ZY') - E(Z\tilde{W}')E(\tilde{W}\tilde{W}')^{-1}E(\tilde{W}Y') = 0$ and $E(ZX') - E(Z\tilde{W}')E(\tilde{W}\tilde{W}')^{-1}E(\tilde{W}X') = 0$.

In neither of these cases do post-response instruments identify effects. But the obstacle is essentially the lack of causal relations between U_z and (U_x, U_y). As Chalak (2009a) shows, the presence of specific causal structure relating these can either partially or fully identify β_o. A similar result holds for the case of joint response instruments of Table III. As the details are involved, we refer the interested reader to Chalak (2009a).

This exhausts the possibilities for structural identification for single XIVs.

7 Double Extended Instrumental Variables Methods

Economic theory can also suggest structures that permit identification of causal effects by jointly using conditional instruments, Z, and conditioning instruments, W. We now examine these.

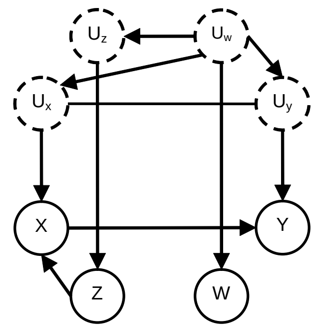
7.1 Conditionally Exogenous Instruments Given Conditioning Instruments

As for single XIVs, we distinguish between cases where we observe conditionally exogenous instruments and others where we only observe their proxies. For conditioning instruments, we focus on the use of predictive proxies, omitting explicit discussion of structural proxies for brevity.

7.1.1 Observed Conditionally Exogenous Instruments

Our first case is that of observed conditionally exogenous instruments given conditioning instruments (OXI|I). To illustrate, consider $\mathcal{S}_{13a}$ with associated graph G_{13a}:

(1) $U_z \overset{c}{=} \delta_z U_w$
(2) $[U_x', U_y']' \overset{c}{=} \delta_{x,y} U_w$
(3) $W \overset{c}{=} \alpha_w U_w$
(4) $Z \overset{c}{=} \alpha_z U_z,$
(5) $X \overset{c}{=} \gamma_x Z + \alpha_x U_x$
(6) $Y \overset{c}{=} X'\beta_o + U_y'\alpha_o$



Graph 13a (G_{13a})
Observed Conditionally Exogenous Instruments
Given Conditioning Instruments (OXI|I)

Substituting $\mathcal{S}_{13a}(5)$ into $\mathcal{S}_{13a}(6)$ and setting $\pi_o \equiv \gamma_x'\beta_o$ also gives

$$(7) \quad Y = Z'\pi_o + U_x'\alpha_x'\beta_o + U_y'\alpha_o.$$

$\mathcal{S}_{13a}$ specifies that U_x and U_y share a common cause, U_w, but leaves other causal or stochastic dependence between U_x and U_y unspecified. Nevertheless, $Z \perp U_y \mid U_w$. When W is a sufficiently good predictor for U_w, hence Z and U_y, this gives the key exogeneity relation:

(XI|I) *Conditionally Exogenous Instruments given Conditioning Instruments*: $Z \perp U_y \mid W$

Taking $\tilde{Z} = Z$ and $\tilde{W} = W$, it follows from Theorem 3.1 that β_o is identified as

$$\beta_o = [E(\{Z - E(Z \mid W)\}X')]^{-1} E(\{Z - E(Z \mid W)\}Y)$$

if and only if $E(\{Z - E(Z \mid W)\}X')$ is nonsingular. Note that, as for XI, $\ell = k$ is necessary.

No previous method can identify β_o in this case, as none of the other admissible conditional independence relationships hold in $\mathcal{S}_{13a}$.

A feasible plug-in for β_o based on linearity of $E(Z \mid W)$ is

$$\hat{\beta}_n^{XI|I} \equiv [\mathbf{Z}'(\mathbf{I} - \mathbf{W}(\mathbf{W}'\mathbf{W})^{-1}\mathbf{W}')\mathbf{X}]^{-1}[\mathbf{Z}'(\mathbf{I} - \mathbf{W}(\mathbf{W}'\mathbf{W})^{-1}\mathbf{W}')\mathbf{Y}].$$

This is a Hausman-Taylor estimator, but Hausman and Taylor's (1983) conditions need not hold.

In $\mathcal{S}_{13a}$, Z satisfies causal properties that parallel CP:OXI:

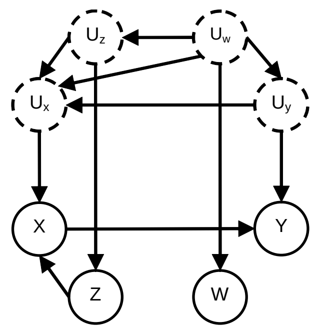
CP:OXI|I *(i) Z causes X, and the effect of Z on X is identified via XC|I with conditioning instruments W; (ii) Z indirectly causes Y, and the effect of Z on Y is identified via XC|I with conditioning instruments W; (iii) Z causes Y only via X.*

Similar to XI, the exclusion restriction (iii) plays a key role, and the effect of X on Y is identified as the ILS derivative ratio of the identified effects π_o and γ_x, $\beta_o = (\gamma_x')^{-1}\pi_o$. Note that the conditioning instruments need not be identical in (i) and (ii).

7.1.2 Proxies for Unobserved Conditionally Exogenous Instruments

As for all XIV methods, and as for XI and XC|I in particular, the true underlying instrument need not be observed; a suitable proxy suffices. For example, consider $\mathcal{S}_{13b}$ with graph G_{13b}:

(1) $U_z \overset{c}{=} \delta_z U_w$
(2) $U_y \overset{c}{=} \delta_y U_w$
(3) $U_x \overset{c}{=} \delta_{x_1} U_w + \delta_{x_2} U_z + \delta_{x_3} U_y$
(4) $W \overset{c}{=} \alpha_w U_w$
(5) $Z \overset{c}{=} \alpha_z U_z$
(6) $X \overset{c}{=} \gamma_x Z + \alpha_x U_x$
(7) $Y \overset{c}{=} X'\beta_o + U_y'\alpha_o$



Graph 13b (G_{13b})
Proxy for Unobserved Conditionally Exogenous Instruments Given Conditioning Instruments (PXI|I)

Substituting $\mathcal{S}_{13b}(6)$ into $\mathcal{S}_{13b}(7)$ and setting $\pi_o \equiv \gamma_x'\beta_o$ also gives

$$(8)\quad Y \overset{c}{=} Z'\pi_o + U_x'\alpha_x'\beta_o + U_y'\alpha_o \overset{c}{=} U_z'(\alpha_z'\pi_o + \delta_{x_2}'\alpha_x'\beta_o) + U_y'(\alpha_o + \delta_{x_3}'\alpha_x'\beta_o) + U_w'\delta_{x_1}'\alpha_x'\beta_o.$$

Here, $\mathcal{S}_{13b}$ specifies causal relations between U_x and U_y. XI|I holds when W is a suitable predictor for U_w, as $Z \perp U_y \mid U_w$. As for OXI|I, Theorem 3.1 applies with $\tilde{Z} = Z$ and $\tilde{W} = W$. Although CP:OXI|I fails for Z, U_z plays the key identifying role, and Z is its proxy. We thus call this the PXI|I case. The key causal properties are:

CP:PXI|I *(i) U_z indirectly causes X, and the full effect of U_z on X could be identified via XC|I with conditioning instruments W had U_z been observed; (ii) U_z indirectly causes Y, and the full effect of U_z on Y could be identified via XC|I with conditioning instruments W had U_z been observed; (iii) U_z causes Y only via X.*

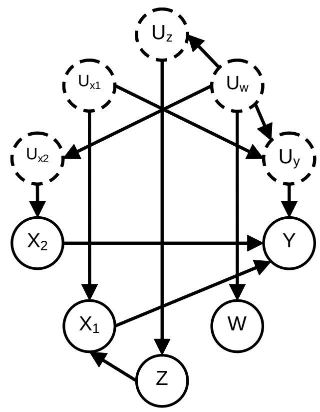
CP:PXI|I parallels CP:PXI, but identification in (i) and (ii) is via XC|I, not XC. As in CP:OXI|I, the conditioning instruments in (i) and (ii) may differ. (iii) imposes the key exclusion restriction. Our comments about PXI and ILS also apply here: the ILS derivative ratio of two inconsistent XC|I estimators remains informative for β_o in the linear case, but not in the nonseparable case. As with PXI, Z is not required to cause X in $\mathcal{S}_{13b}$.

7.2 Conditionally Exogenous Causes and Instruments Given Instruments

When conditioning instruments W render only a subvector X_2 of $X \equiv [X_1', X_2']'$ conditionally exogenous, no previous method identifies $\beta_o \equiv [\beta_1', \beta_2']'$. But identification obtains given conditional instruments Z for X_1 that are conditionally exogenous given W. Consider $\mathcal{S}_{14}$:

(1) $U_z \overset{c}{=} \delta_z U_w$
(2) $U_{x_2} \overset{c}{=} \delta_{x_2} U_w$
(3) $U_y \overset{c}{=} \delta_{y_1} U_w + \delta_{y_2} U_{x_1}$
(3) $W \overset{c}{=} \alpha_w U_w$
(2) $Z \overset{c}{=} \alpha_z U_z$
(3) $X_1 \overset{c}{=} \gamma_{x_1} Z + \alpha_{x_1} U_{x_1}$
(4) $X_2 \overset{c}{=} \alpha_{x_2} U_{x_2}$
(5) $Y \overset{c}{=} X_1'\beta_1 + X_2'\beta_2 + U_y'\alpha_o$,

where $U_{x_1} \perp U_w$.



Graph 14 (G_{14})
Conditionally Exogenous Instruments and Causes
Given Conditioning Instruments (XCI|I)

$\mathcal{S}_{14}$ ensures $(Z, X_2) \perp U_y \mid U_w$. Provided W is a suitable predictive proxy, this ensures

(XCI|I) *Conditionally Exogenous Causes and Instruments given Conditioning Instruments:* $(Z, X_2) \perp U_y \mid W$.

XCI|I is the special case of XI|I in which X_2 plays the role of a conditional instrument for itself. Letting $\tilde{Z} = [Z', X_2']'$ and $\tilde{W} = W$, Theorem 3.1 ensures that β_o is identified as

$$\beta_o = [E(\{\tilde{Z} - E(\tilde{Z} \mid W)\}X')]^{-1} E(\{\tilde{Z} - E(\tilde{Z} \mid W)\}Y)$$

if and only if $E(\{\tilde{Z} - E(\tilde{Z} \mid W)\}X')$ is nonsingular. Again, $\ell = k$ is necessary. The XCI|I plug-in estimator with linearity for $E(\tilde{Z} \mid W)$ is

$$\hat{\beta}_n^{XCI|I} \equiv [\tilde{\mathbf{Z}}'(\mathbf{I} - \mathbf{W}(\mathbf{W}'\mathbf{W})^{-1}\mathbf{W}')\mathbf{X}]^{-1}[\tilde{\mathbf{Z}}'(\mathbf{I} - \mathbf{W}(\mathbf{W}'\mathbf{W})^{-1}\mathbf{W}')\mathbf{Y}].$$

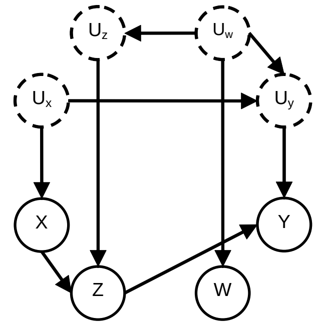
An analogous result holds for the XCI case ($\tilde{Z} \perp U_y$) with β_o identified as $\beta_o = E(\tilde{Z}X')^{-1}E(\tilde{Z}Y)$. The XCI plug-in estimator is $\hat{\beta}_n^{XCI} \equiv (\tilde{\mathbf{Z}}'\mathbf{X})^{-1}(\tilde{\mathbf{Z}}'\mathbf{Y})$.

7.3 Conditionally Exogenous Instruments Given Causes and Conditioning Instruments

A generalization of XI|C occurs when XI|C fails but conditioning instruments W and regressors X jointly render extended instruments Z conditionally exogenous, as in $\mathcal{S}_{15}$:

(1) $U_z \stackrel{c}{=} \delta_z U_w$
(2) $U_y \stackrel{c}{=} \delta_{y_1} U_w + \delta_{y_2} U_x$
(3) $W \stackrel{c}{=} \alpha_w U_w$
(4) $X \stackrel{c}{=} \alpha_x U_x$
(5) $Z \stackrel{c}{=} \gamma_z X + \alpha_z U_z$
(6) $Y \stackrel{c}{=} Z'\delta_o + U_y'\alpha_o$

where $U_x \perp U_w$.



Graph 15 (G_{15})
Conditionally Exogenous Instruments given Causes and Conditioning Instruments (XI|CI)

Substituting $\mathcal{S}_{15}(3)$ into $\mathcal{S}_{15}(4)$ with $\beta_o \equiv \gamma_z'\delta_o$ also gives

$$(7) \qquad Y \stackrel{c}{=} X'\beta_o + U_z'\alpha_z'\delta_o + U_y'\alpha_o.$$

The conditional exogeneity relationship ensured by $\mathcal{S}_{15}$ is:

(XI|CI) *Conditionally Exogenous Instruments given Causes and Conditioning Instruments:* $Z \perp U_y \mid (X, W)$.

Letting $\tilde{Z} = Z$ and $\tilde{W} = [X', W']'$, Theorem 3.1 identifies δ_o as

$$\delta_o = [E(\{Z - E(Z \mid \tilde{W})\}X')]^{-1} E(\{Z - E(Z \mid \tilde{W})\}Y).$$

Since $X \perp U_z$, XC identifies γ_z as $E(ZX')E(XX')^{-1}$. Clearly, $X \perp U_z$ can be replaced by a conditional relationship, such as XC|I: $X \perp U_z \mid W_1$, with suitable W_1.

Like XI|C, XI|CI embodies Pearl's (1995, 2000) front-door method. Corollary 3.2 gives

$$\beta_o = E(XX')^{-1}E(XZ') \times [E(\{Z - E(Z \mid \tilde{W})\}X')]^{-1} E(\{Z - E(Z \mid \tilde{W})\}Y).$$

The feasible plug-in estimator based on linearity of $E(Z \mid \tilde{W})$ is

$$\hat{\beta}_n^{XI|CI} \equiv (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Z} \times [\mathbf{Z}'(\mathbf{I} - \tilde{\mathbf{W}}(\tilde{\mathbf{W}}'\tilde{\mathbf{W}})^{-1}\tilde{\mathbf{W}}')\mathbf{X}]^{-1}[\mathbf{Z}'(\mathbf{I} - \tilde{\mathbf{W}}(\tilde{\mathbf{W}}'\tilde{\mathbf{W}})^{-1}\tilde{\mathbf{W}}')\mathbf{Y}].$$

The causal properties for XI|CI parallel those of CP:XI|C. In particular, the exclusion restriction that X causes Y only through Z remains necessary.

7.4 Conditionally Exogenous Instruments Given Causes, Response, and Conditioning Instruments

Analysis parallel to Section 6.5 shows that without specific structure among the unobservables, XI|RI ($Z \perp U_y \mid (Y, W)$) and XI|CRI ($Z \perp U_y \mid (X, Y, W)$) need not identify β_o. Nevertheless, suitable structure among the unobservables can partially or fully identify β_o; see Chalak (2009a).

7.5 Further Remarks and Summary

Reinforcing a main theme, we note that all of the double XIV methods rely on specific causal structure among the unobservables to identify β_o.

We also observe that XC, XI, XCI, XC|I, XI|I, and XCI|I form an exhaustive set of primitives, as other XIV methods, such as XI|C and XI|CI, identify causal effects as functions of effects identified by one or more of these primitives.

In the nonseparable case, double XIV methods can give structural meaning to derivatives of standard conditional expectations and ratios or products of these, analogous to our discussions in Section 6. The interpretation of these objects depends on the specific system structure; exploration of their detailed causal content is an area for further research.

As a helpful summary, Table IV displays the various XIV estimators of Sections 6 and 7.

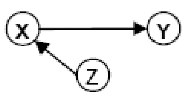
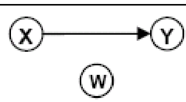
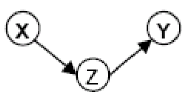
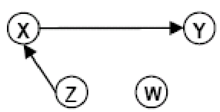
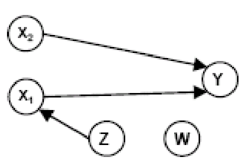
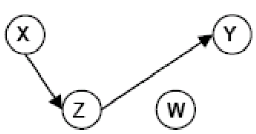
Causal Structure among Observables	**(Conditional) Exogeneity**	**Estimator**
	XC: $X \perp U_y$	$\hat{\beta}_n^{XC} \equiv (\boldsymbol{X'X})^{-1}(\boldsymbol{X'Y})$
	XI: $Z \perp U_y$	$\hat{\beta}_n^{XI} \equiv (\boldsymbol{Z'X})^{-1}(\boldsymbol{Z'Y})$
	XC|I: $X \perp U_y \mid W$	$\hat{\beta}_n^{XC\|I} \equiv [\boldsymbol{X'}(\boldsymbol{I} - \boldsymbol{W}(\boldsymbol{W'W})^{-1}\boldsymbol{W'})\boldsymbol{X}]^{-1}$ $[\boldsymbol{X'}(\boldsymbol{I} - \boldsymbol{W}(\boldsymbol{W'W})^{-1}\boldsymbol{W'})\boldsymbol{Y}]$
	XI|C: $Z \perp U_y \mid X$ and $X \perp U_z$	$\hat{\beta}_n^{XI\|C} \equiv (\boldsymbol{X'X})^{-1}(\boldsymbol{X'Z})$ $\times [\boldsymbol{Z'}(\boldsymbol{I} - \boldsymbol{X}(\boldsymbol{X'X})^{-1}\boldsymbol{X'})\boldsymbol{Z}]^{-1}$ $\times [\boldsymbol{Z'}(\boldsymbol{I} - \boldsymbol{X}(\boldsymbol{X'X})^{-1}\boldsymbol{X'})\boldsymbol{Y}]$
	XI|I: $Z \perp U_y \mid W$	$\hat{\beta}_n^{XI\|I} \equiv [\boldsymbol{Z'}(\boldsymbol{I} - \boldsymbol{W}(\boldsymbol{W'W})^{-1}\boldsymbol{W'})\boldsymbol{X}]^{-1}$ $\times [\boldsymbol{Z'}(\boldsymbol{I} - \boldsymbol{W}(\boldsymbol{W'W})^{-1}\boldsymbol{W'})\boldsymbol{Y}]$
	XCI|I: $(Z, X_2) \perp U_y \mid W$	$\hat{\beta}_n^{XCI\|I} \equiv [\tilde{\boldsymbol{Z}}' (\boldsymbol{I} - \boldsymbol{W}(\boldsymbol{W'W})^{-1}\boldsymbol{W'})\boldsymbol{X}]^{-1}$ $\times [\tilde{\boldsymbol{Z}}' (\boldsymbol{I} - \boldsymbol{W}(\boldsymbol{W'W})^{-1}\boldsymbol{W'})\boldsymbol{Y}]$ where $X \equiv [X_1', X_2']'$ and $\tilde{Z} \equiv [Z', X_2']'$
	XI|CI: $Z \perp U_y \mid (X, W)$ and $X \perp U_z$	$\hat{\beta}_n^{XI\|CI} \equiv [(\boldsymbol{X'X})]^{-1}(\boldsymbol{X'Z})$ $\times [\boldsymbol{Z}'(\boldsymbol{I} - \tilde{\boldsymbol{W}} (\tilde{\boldsymbol{W}}'\tilde{\boldsymbol{W}})^{-1} \tilde{\boldsymbol{W}}')\boldsymbol{Z}]^{-1}$ $\times [\boldsymbol{Z}'(\boldsymbol{I} - \tilde{\boldsymbol{W}} (\tilde{\boldsymbol{W}}'\tilde{\boldsymbol{W}})^{-1} \tilde{\boldsymbol{W}}')\boldsymbol{Y}]$ where $\tilde{W} \equiv [X', W']'$

Table IV
Extended Instrumental Variables Methods

8 Summary and Directions for Further Research

This paper examines the ways in which structural systems can yield observed variables that, along with the cause and response of interest, can play an instrumental role in identifying and estimating causal effects. We focus specifically on how structures can yield exogeneity and exclusion restrictions that ensure moment conditions supporting identification. This provides a comprehensive framework for constructing useful instruments and control variables. Because our results exhaust the possibilities, we ensure that there are no other opportunities for identification based on exogeneity and exclusion restrictions still to be discovered. We introduce notions of *conditional* and *conditioning* extended instrumental variables (XIVs). These need not be traditional instruments, as they may be endogenous. They nevertheless permit identification and estimation of causal effects. We analyze methods using these XIVs either singly or jointly.

As we show, it is important to distinguish between observed XIVs and proxies for unobserved XIVs. This distinction permits us to characterize the role of instruments in indirect least squares-like methods and to extend Pearl's (1995) back-door method. A main message emerging from our analysis is the central importance of sufficiently specifying the causal relations governing the *unobservables*, as these play a crucial role in creating obstacles or opportunities for identification.

Among the useful insights that emerge from our analysis is the understanding that in the presence of conditioning instruments, not all regression coefficients need have causal meaning. Thus, not all regression coefficients need have signs or magnitudes that make economic sense. We also see that the standard account of the "problem" of omitted variables is incomplete and possibly misleading. In order to properly understand what effects one is estimating with a given regression including or excluding causally relevant variables, it is necessary to have a clear understanding of the underlying causal relations. Another contribution is the demonstration that the Hausman-Taylor framework can be extended by considering conditional covariance restrictions.

In related work (White and Chalak, 2009a; Schennach, White, and Chalak, 2009; Song, Schennach, and White, 2009), we analyze nonparametric identification and estimation of effects for certain nonseparable versions of the cases discussed here. Investigating the remaining cases is an interesting focus for future work. White (2006), White and Chalak (2006), and White and Xu (2009) give conditional exogeneity tests. It is of interest to extend these for use with XIV methods. Here, we study identification of effects in causal structures specified a priori. By instead studying which causal matrices agree with a collection of given (observed) conditional independence relationships, one can develop methods for suggesting or ruling out potential causal structures. We leave this to future work.

Finally, XIV methods offer strategies for dealing with weak instruments: when standard instruments are weak, there may be extended instruments that are either less weak or not at all weak for identifying effects of interest. More generally, when one form of XIV is weak (e.g., $E(Z^* X')$ is near singularity) there may be another XIV that is either less or not at all weak. This is another interesting avenue for further research.

Appendix A: Mathematical Proofs

Proof of Theorem 3.1: (a) We have $E\{Z^*\varepsilon\} = E\{Z^*(Y - X'\beta_o)\} = E(Z^*Y) - E(Z^*X')\beta_o$. By Assumption 2.2, β_o is finite; since $E[Z^*(Y, X')] < \infty$, $E\{Z^*\varepsilon\} < \infty$. (b) Since $E\{Z^*\varepsilon\} = E(Z^*Y) - E(Z^*X')\beta_o$, the result follows immediately. (c) Given $E\{Z^*\varepsilon\} = 0$, $E(Z^*X')\beta_o - E(Z^*Y) = 0$. This system admits a unique solution β_o if and only if $E(Z^*X')$ is non-singular. ■

Proof of Corollary 3.2: Immediate. ■

Proof of Proposition 4.1: The existence and finiteness of $\beta^\dagger$ is trivial under the conditions given. Define $\eta \equiv Y - X'\beta^\dagger$. Then $Y = X'\beta^\dagger + \eta$. Also, $E(Z^*\eta) = E(Z^*(Y - X'\beta^\dagger)) = S - QQ^{-1}S = 0$. Given $\mathbf{Z}^{*\prime}\mathbf{X}/n \xrightarrow{p} Q$ and $\mathbf{Z}^{*\prime}\mathbf{Y}/n \xrightarrow{p} S$, $\hat{\beta}_n^{XIV} \xrightarrow{p} Q^{-1}S \equiv \beta^\dagger$ by Slutsky's theorem. $\hat{\beta}_n^{FXIV} - \hat{\beta}_n^{XIV} \xrightarrow{p} 0$ then immediately implies $\hat{\beta}_n^{FXIV} \xrightarrow{p} \beta^\dagger$. ■

Proof of Theorem 4.2: By definition of $\tilde{\beta}_n^{FXIV}$ and β^*, and with $\hat{Q}_n \equiv n^{-1}(\tilde{\mathbf{Z}} - \tilde{\mathbf{W}}\hat{\pi}_n')'\mathbf{X}$,

$$\tilde{\beta}_n^{FXIV} - \beta^* = [\hat{Q}_n^{-1}\ n^{-1}\sum_{i=1}^n (\tilde{Z}_i - \hat{\pi}_n\tilde{W}_i)Y_i] - \beta^*.$$

Letting $Q_n^* \equiv n^{-1}(\tilde{\mathbf{Z}} - \tilde{\mathbf{W}}\pi^{*\prime})'\mathbf{X}$ and adding and subtracting, we have

$$\begin{aligned}\tilde{\beta}_n^{FXIV} - \beta^* \quad &= \quad [Q_n^{*-1}\ n^{-1}\sum_{i=1}^n (\tilde{Z}_i - \pi^*\tilde{W}_i)Y_i] - \beta^* \\ &\quad + [\hat{Q}_n^{-1}\ n^{-1}\sum_{i=1}^n (\tilde{Z}_i - \hat{\pi}_n\tilde{W}_i)Y_i] - \ [Q_n^{*-1}\ n^{-1}\sum_{i=1}^n (\tilde{Z}_i - \pi^*\tilde{W}_i)Y_i].\end{aligned}$$

Replacing Y_i with $X_i'\beta^* + \eta_i^*$ gives

$$\begin{aligned}\tilde{\beta}_n^{FXIV} - \beta^* \quad &= \quad Q_n^{*-1}\ n^{-1}\sum_{i=1}^n (\tilde{Z}_i - \pi^*\tilde{W}_i)\eta_i^* \\ &\quad + [\hat{Q}_n^{-1}\ n^{-1}\sum_{i=1}^n (\tilde{Z}_i - \hat{\pi}_n\tilde{W}_i)\eta_i^*] - \ [Q_n^{*-1}\ n^{-1}\sum_{i=1}^n (\tilde{Z}_i - \pi^*\tilde{W}_i)\eta_i^*].\end{aligned}$$

Adding and subtracting again and multiplying by $n^{1/2}$ gives

$$\begin{aligned}n^{1/2}(\tilde{\beta}_n^{FXIV} - \beta^*) &= Q_n^{*-1}\ n^{-1/2}\sum_{i=1}^n (\tilde{Z}_i - \pi^*\tilde{W}_i)\eta_i^* \\ &\quad + \ (\hat{Q}_n^{-1} - Q_n^{*-1})\ n^{-1/2}\sum_{i=1}^n (\tilde{Z}_i - \pi^*\tilde{W}_i)\eta_i^* - \ \hat{Q}_n^{-1}\ n^{1/2}(\hat{\pi}_n - \pi^*)\ n^{-1}\sum_{i=1}^n \tilde{W}_i\eta_i^*.\end{aligned}$$

Now $n^{-1/2}\sum_{i=1}^n (\tilde{Z}_i - \pi^*\tilde{W}_i)\eta_i^* = O_p(1)$, $Q_n^* \xrightarrow{p} Q^*$, $\hat{Q}_n \xrightarrow{p} Q^*$, and $n^{-1}\sum_{i=1}^n \tilde{W}_i\eta_i^* \xrightarrow{p} E(\tilde{W}\eta^*)$, so

$$n^{1/2}(\tilde{\beta}_n^{FXIV} - \beta^*) = Q^{*-1}\ n^{-1/2}\sum_{i=1}^n (\tilde{Z}_i - \pi^*\tilde{W}_i)\eta_i^* - Q^{*-1}\ n^{1/2}(\hat{\pi}_n - \pi^*)\ E(\tilde{W}\eta^*) + o_p(1),$$

provided $n^{1/2}(\hat{\pi}_n - \pi^*) = O_p(1)$. We now verify this. With $\zeta_i^* \equiv \tilde{Z}_i - \pi^* \tilde{W}_i$, we have

$$n^{1/2}(\hat{\pi}_n - \pi^*) = n^{-1/2} \sum_{i=1}^{n} \zeta_i^* \tilde{W}_i' \; [n^{-1} \sum_{i=1}^{n} \eta_i \tilde{W}_i \tilde{W}_i']^{-1} = n^{-1/2} \sum_{i=1}^{n} \zeta_i^* \tilde{W}_i' \; E(\tilde{W}\tilde{W}')^{-1} + o_p(1),$$

given that $n^{-1/2} \sum_{i=1}^{n} \zeta_i^* \tilde{W}_i'$ is $O_p(1)$ and $n^{-1} \sum_{i=1}^{n} \tilde{W}_i \tilde{W}_i' \xrightarrow{p} E(\tilde{W}\tilde{W}')$, which hold under our assumptions. It follows that $n^{1/2}(\hat{\pi}_n - \pi^*) = O_p(1)$. This further implies

$$\begin{aligned} n^{1/2}(\tilde{\beta}_n^{FXIV} - \beta^*) \;\; &= \;\; Q^{*-1} \; n^{-1/2} \sum_{i=1}^{n} (\tilde{Z}_i - \pi^* \tilde{W}_i) \eta_i^* \\ &\quad - Q^{*-1} \; n^{-1/2} \sum_{i=1}^{n} \zeta_i^* \tilde{W}_i' \; E(\tilde{W}\tilde{W}')^{-1} \; E(\tilde{W}\eta^*) \; + \; o_p(1), \end{aligned}$$

so that with $\tilde{\eta}_i \equiv \eta_i^* - \tilde{W}_i' \; E(\tilde{W}\tilde{W}')^{-1} E(\tilde{W}\eta^*)$,

$$n^{1/2}(\tilde{\beta}_n^{FXIV} - \beta^*) = Q^{*-1} \; n^{-1/2} \sum_{i=1}^{n} \zeta_i^* \, \tilde{\eta}_i + o_p(1).$$

As $\{\zeta_i^* \, \tilde{\eta}_i\}$ obeys the central limit theorem by assumption, the result follows. ■

Proof of Proposition 5.1: Let $h \leq G$ and let $\{g_1, \ldots, g_h\}$ be a set of h distinct elements of $\{1, \ldots, G\}$ such that $c_{g_1 g_2} \times c_{g_2 g_3} \times ... \times c_{g_h g_1} = 1$. Then there exist $r_{g_1}, ..., r_{g_h}$ with

$$\begin{gathered} r_{g_1}(V_{g_1 - 1}, ..., V_{g_h}, ..., V_1, V_0) \\ r_{g_i}(V_{g_i - 1}, ..., V_{g_{i-1}}, ..., V_1, V_0), \quad i = 2, ..., h \end{gathered}$$

such that $r_{g_1}(V_{g_1 - 1}, ..., V_{g_h}, ..., V_1, V_0)$ is not constant in V_{g_h} and $r_{g_i}(V_{g_i - 1}, ..., V_{g_{i-1}}, ..., V_1, V_0)$ is not constant in $V_{g_{i-1}}$ for $i = 2, ..., h$. But this violates Assumption 2.1, a contradiction. ■

Appendix B: FXIV Estimators for Effects Identified by Corollary 3.2

Estimators for effects identified by Corollary 3.2 are $\tilde{\beta}_n^{FXIV} \equiv b(\tilde{\theta}_n^{FXIV})$, where $\tilde{\theta}_n^{FXIV} \equiv (\tilde{\theta}_{1,n}^{FXIV\prime}, ..., \tilde{\theta}_{H,n}^{FXIV\prime})'$ is a vector of estimators covered by Theorem 4.2. To give a result, let

$$\tilde{\theta}_{h,n}^{FXIV} \equiv [(\tilde{\mathbf{Z}}_h - \tilde{\mathbf{W}}_h \hat{\pi}_{h,n}')' \mathbf{X}_h]^{-1} [(\tilde{\mathbf{Z}}_h - \tilde{\mathbf{W}}_h \hat{\pi}_{h,n}')' \mathbf{Y}_h], \quad h = 1, ..., H,$$

where $\hat{\pi}_{h,n} \equiv \tilde{\mathbf{Z}}_h' \tilde{\mathbf{W}}_h (\tilde{\mathbf{W}}_h' \tilde{\mathbf{W}}_h)^{-1}$ and for $\tilde{\theta}_{h,n}^{FXIV}$, let Q_h^* and $\zeta_{hi}^* \, \tilde{\eta}_{hi}$ denote the analogs of Q^* and ζ_i^* $\tilde{\eta}_i$ defined for Theorem 4.2. Finally, let $Q^* \equiv diag(Q_1^*, ..., Q_H^*)$ and $\mathcal{Z}_i \equiv ((\zeta_{1i}^* \tilde{\eta}_{1i})', ..., (\zeta_{Hi}^* \tilde{\eta}_{Hi})')'$.

Proposition B.1 *Let the conditions of Theorem 4.2 hold for each element of* $\tilde{\theta}_n^{FXIV}$ *and let* $\theta^* \equiv plim \; \tilde{\theta}_n^{FXIV}$. *If* $n^{-1/2} \sum_{i=1}^{n} \mathcal{Z}_i \xrightarrow{d} N(0, V^*)$, *where* V^* *is finite and positive definite, then*

$$n^{1/2}(\tilde{\theta}_n^{FXIV} - \theta^*) \xrightarrow{d} N(0, Q^{*-1} V^* Q^{*\prime -1}).$$

Suppose further that b *is continuously differentiable at* θ^* *such that* $\nabla b(\theta^*)$ *(the gradient of* b *at* θ^**) has full column rank. Then with* $\tilde{\beta}_n^{FXIV} = b(\tilde{\theta}_n^{FXIV})$ *and* $\beta^* = b(\theta^*)$,

$$n^{1/2}(\tilde{\beta}_n^{FXIV} - \beta^*) \xrightarrow{d} N(0, \nabla b(\theta^*)' Q^{*-1} V^* Q^{*\prime -1} \nabla b(\theta^*)).$$

Appendix C: Necessity of Conditional Exogeneity for Identification in Linear Systems

We state the next result for generic random vectors (U, W, Z). For conciseness, we understand that "admissible" distributions are those satisfying the specified hypotheses. The absolute value of a matrix denotes the matrix of its absolute values.

Proposition C.1 *Suppose* (U, W, Z) *are random vectors such that* $E(U) = 0$, $E(Z'Z) < \infty$, *and* Z *is not a function solely of* W. *Let* $Z^* \equiv Z - E(Z \mid W)$ *and suppose that* $|E(Z^*U')| < \infty$. *Then* $E(Z^*U') = 0$ *for all admissible distributions of* (U, W, Z) *if and only if* $Z \perp U \mid W$.

Proof: To prove $E(Z^*U') = 0$ for all (U, W, Z) such that $Z \perp U \mid W$ is easy and is omitted. For the converse, let U, W, and Z be scalars without loss of generality, and write

$$E(Z^*U) = \int (z - \mu(w))\, u\, dF(u \mid w, z)\, dF(w, z) = \int (z - \mu(w)) \{\int u\, dF(u \mid w, z)\}\, dF(w, z),$$

where $\mu(w) \equiv E(Z \mid W = w)$ and $F(u \mid w, z)$ and $F(w, z)$ denote the distribution functions of $U \mid (W, Z)$ and (W, Z), respectively. Suppose that $Z \not\perp U \mid W$; we show that there exist admissible $F(u \mid w, z)$ and $F(w, z)$ such that $E(Z^*U) \neq 0$.

Specifically, choose $F(u \mid z, w)$ such that $\int u\, dF(u \mid z, w) = z - \mu(w)$. This choice is possible, as $Z \not\perp U \mid W$. It is admissible, as this implies $E(U) = E(E(U \mid W, Z)) = E(Z - \mu(W)) = E(E[Z - \mu(W) \mid W]) = 0$ and

$$E(Z^*U) = \int (z - \mu(w))^2\, dF(w, z),$$

which is easily shown to be finite by the conditional Jensen and Minkowski inequalities, given $E(Z'Z) < \infty$. As Z is not solely a function of W, there exists a distribution for (W, Z) such that $E(Z^*U) = \int (z - \mu(w))^2\, dF(w, z) > 0$, and the proof is complete. ■

Appendix D: Failures of Identification via XI from a Causal Perspective

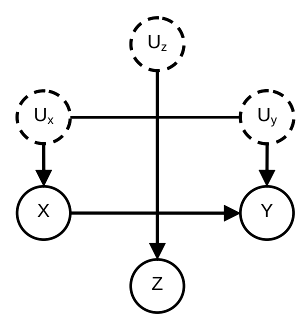
Here we examine how identification of β_o via the XI method fails in the standard "irrelevant instrument," "invalid instrument," and "under-identified" cases.

D.1 Irrelevant Exogenous Instruments

System $\mathcal{S}_{16}$ and its causal graph G_{16} depict the irrelevant XI case and demonstrate how an irrelevant XI satisfies neither CP:OXI nor CP:PXI. Let $\mathcal{S}_{16}$ be given by:

(1) $Z \overset{c}{=} \alpha_z U_z$
(2) $X \overset{c}{=} \alpha_x U_x$
(3) $Y \overset{c}{=} X'\beta_o + U_y'\alpha_o$,

where $U_x \not\perp U_y$,
$U_x \perp U_z$, and $U_y \perp U_z$,



Graph 16 (G_{16})
Irrelevant Exogenous Instruments

Although Z is valid and satisfies XI, it fails to identify β_o, because the effect of X on Y cannot be represented as the "ratio" of the effect of Z (resp. U_z) on Y and the effect of Z (resp. U_z) on X, as both of these are zero. In $\mathcal{S}_{16}$, neither CP:OXI(i) nor CP:PXI(i) hold, since neither Z nor U_z cause X, justifying the label "irrelevant exogenous variables." When $\ell = k$ (as assumed here), condition (c) fails in Theorem 3.1. From $\mathcal{S}_{16}(3)$ we have

$$E(ZY) - E(ZX')\beta_o = 0,$$

but $E(ZY) = 0$ and $E(ZX') = 0$, so identification via XI fails.

D.2 Endogenous Instruments

We next examine failure of XI, $Z \not\perp U_y$. Such endogenous Z's can occur in several ways.

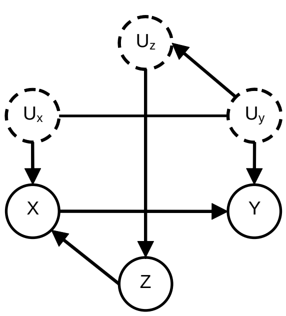
First, a potential instrument Z can be both irrelevant and endogenous, e.g., when Z doesn't cause X and $U_z \perp U_x$, but both U_x and U_z cause U_y. For relevant instruments, consider $\mathcal{S}_{17}$:

$$\begin{aligned}
&(1)\quad Z \overset{c}{=} \alpha_z U_z \\
&(2)\quad X \overset{c}{=} \gamma_x Z + \alpha_x U_x \\
&(3)\quad Y \overset{c}{=} X'\beta_o + U_y'\alpha_o,
\end{aligned}$$

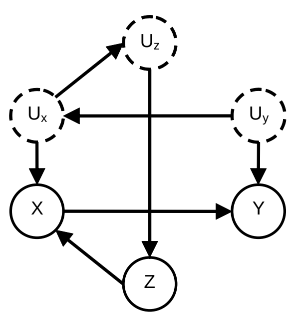
where $U_x \not\perp U_y$, $U_x \not\perp U_z$, and $U_y \not\perp U_z$. Substituting $\mathcal{S}_{17}(2)$ into $\mathcal{S}_{17}(3)$ with $\pi_o \equiv \gamma_x'\beta_o$ gives

$$(4)\quad Y \overset{c}{=} Z'\pi_o + U_x'\alpha_x'\beta_o + U_y'\alpha_o.$$

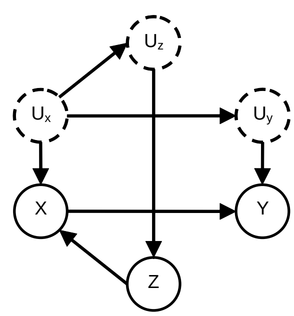
Because $U_y \not\perp U_z$, XI fails for Z. This can occur in several ways. For example, correlation between U_z and U_y can arise because either U_y causes U_z (G_{17a}, G_{17b}) or U_x causes both U_z and U_y (G_{17c}). CP:OXI(ii) fails, as Z and Y are confounded; and CP:PXI(ii) fails, as U_z and Y are confounded. CP:PXI(i) also fails, as U_z and X are also confounded.



Graph 17*a* (G_{17a})
Endogenous Instruments

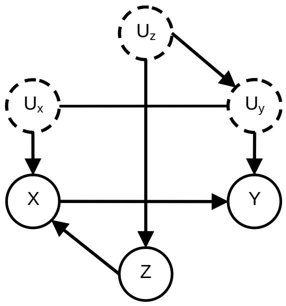


Graph 17*b* (G_{17b})
Endogenous Instruments

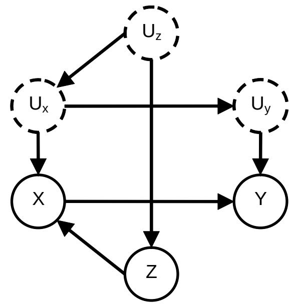


Graph 17*c* (G_{17c})
Endogenous Instruments

Alternatively, Z is endogenous when U_z affects U_y via a channel other than X. As we assume Z can't cause U_y, we need consider only the case where U_z causes Y via an intermediate cause other than X (G_{17d} and G_{17e}). Now CP:OXI(ii) fails as Z and Y are confounded. CP:PXI(iii) fails as U_z causes Y via an intermediate cause other than X; the effect of X on Y is thus no longer the ratio of the effect of U_z on Y and that of U_z on X.



Graph 17*d* (G_{17d})
Endogenous Instruments



Graph 17*e* (G_{17e})
Endogenous Instruments

From $\mathcal{S}_{17}(3)$ we have

$$E(ZY) = E(ZX')\beta_o + E(ZU_y'\alpha_o),$$

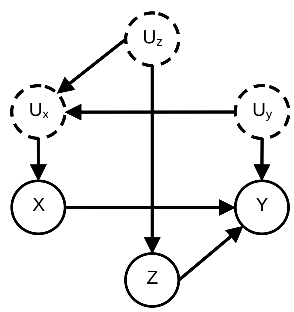
but $E(ZU_y'\alpha_o)$ does not vanish, so identification via XI fails.

D.3 Under-Identified Exogenous Instruments

Finally, consider what happens when instruments Z are valid and relevant, but the exclusion restriction (condition (i)) of Theorem 3.1 fails. Specifically, consider the system $\mathcal{S}_{18}$:

(1) $Z \overset{c}{=} \alpha_z U_z$
(2) $X \overset{c}{=} \alpha_x U_x$
(3) $Y \overset{c}{=} X'\beta_o + Z'\gamma_o + U_y'\alpha_o,$

where $U_x \not\perp U_y$,
$U_x \not\perp U_z$, and $U_y \perp U_z$,



Graph 18 (G_{18})
Under-identified XI

In this case, the regressors X are endogenous, as $U_x \not\perp U_y$, but we have that Z is relevant since $X \not\perp Z$, and valid since $Z \perp U_y$. It follows from $\mathcal{S}_{18}(3)$, however, that

$$\begin{aligned} E(ZY) &= E[Z(X'\beta_o + Z'\gamma_o + U_y'\alpha_o)] \\ &= E(ZX')\beta_o + E(ZZ')\gamma_o. \end{aligned}$$

Once again, stochastic identification of β_o via XI fails, this time due to the presence of the unknown (non-zero) γ_o. The problem is that Z (or U_z) causes Y but not solely via X. This violates CP:OXI(iii) and CP:PXI(iii).

When $E(ZX')$ is nonsingular, we can also write

$$E(ZX')^{-1}E(ZY) = \beta_o + E(ZX')^{-1}E(ZZ')\gamma_o.$$

Viewed in this way, the lack of stochastic identification appears as a form of "omitted variables bias," resulting from the failure to include Z in the instrumental variables regression. But one cannot resolve this problem by including Z, as then one is attempting to identify both β_o and γ_o, and there are not enough proper instruments for this. This is the classical "under-identified" case in which there are more right-hand side variables than valid instruments.

References

Abadie, A. (2003), "Semiparametric Instrumental Variable Estimation of Treatment Response Models," *Journal of Econometrics*, 113, 231–263.

Abadie, A. and G. Imbens (2006), "Large Sample Properties of Matching Estimators for Average Treatment Effects," *Econometrica*, 74, 235–267.

Altonji, J. and R. Matzkin (2005), "Cross Section and Panel Data Estimators for Nonseparable Models with Endogenous Regressors," *Econometrica*, 73, 1053-1102.

Angrist, J. (1990), "Lifetime Earnings and the Vietnam Era Draft Lottery: Evidence from the Social Security Administrative Records," *American Economic Review*, 80, 313- 336.

Angrist, J., G. Imbens, and D. Rubin (1996), "Identification of Causal Effects Using Instrumental Variables" (with Discussion), *Journal of the American Statistical Association*, 91, 444-455.

Angrist, J. and A. Krueger (2001), "Instrumental Variables and the Search for Identification: From Supply and Demand to Natural Experiments," *The Journal of Economic Perspectives*, 15, 69-85.

Bang-Jensen, J. and G. Gutin (2001). *Digraphs: Theory, Algorithms and Applications*. London: Springer Verlag.

Barnow, B., G. Cain, and A. Goldberger (1980), "Issues in the Analysis of Selectivity Bias," in E. Stromsdorfer and G. Farkas (eds.), *Evaluation Studies*, Vol. 5, San Francisco: Sage, pp.43-59.

Butcher, K. and A. Case (1994), "The Effects of Sibling Sex Composition on Women's Education and Earnings," *The Quarterly Journal of Economics*, 109, 531-563.

Chalak, K. (2009a), "Estimating Causal Effects in Structural Systems with Proportional Confounding," Boston College Department of Economics Discussion Paper.

Chalak, K. (2009b), "Estimating Local Average Treatment Effects When the Assignment to Treatment is Unobserved," Boston College Department of Economics Discussion Paper.

Chalak, K. and H. White (2007), "Identification with Conditioning Instruments in Causal Systems," UCSD Department of Economics Discussion Paper.

Chalak, K. and H. White (2009), "Independence and Conditional Independence in Structural Systems," Boston College Department of Economics Discussion Paper.

Chesher, A. (2003), "Identification in Nonseparable Models," *Econometrica*, 71, 1405-1441.

Crawford, V. (2008), "Look-ups as the Windows of the Strategic Soul: Studying Cognition via Information Search in Game Experiments," in A. Caplin and A. Schotter (eds.), *Perspectives on the Future of Economics: Positive and Normative Foundations*, Volume 1 in the series *Handbooks of Economic Methodologies*, Oxford University Press.

Dawid, A.P. (1979), "Conditional Independence in Statistical Theory" (with Discussion), *Journal of the Royal Statistical Society*, Series B, 41, 1-31.

Fisher, R.A. (1935). *The Design of Experiments*. Edinburgh: Oliver and Boyd.

Frisch, R. and F. Waugh (1933), "Partial Regressions as Compared with Individual Trends," *Econometrica*, 1, 939-953.

Goldberger, A. (1972), "Structural Equation Methods in the Social Sciences," *Econometrica*, 40, 979-1001.

Gordon, N. and E. Vegas (2005), "Educational Finance Equalization, Spending, Teacher Quality and Student Outcomes: The Case of Brazil's FUNDEF," in E. Vegas (ed.), *Incentives to Improve Teaching: Lessons from Latin America*. Washington, DC: The World Bank, 2005.

Griliches, Z. (1977), "Estimating the Returns to Schooling: Some Econometric Problems," *Econometrica*, 45, 1-22.

Haavelmo, T. (1943), "The Statistical Implications of a System of Simultaneous Equations," *Econometrica*, 11, 1-12.

Haavelmo, T. (1944), "The Probability Approach in Econometrics," *Econometrica*, 12 (Supplement), iii-vi and 1-115.

Hamilton, J.D. (1994). *Time Series Analysis*. Princeton: Princeton University Press.

Hausman, J. A. and W. E. Taylor (1983), "Identification in Linear Simultaneous Equations Models with Covariance Restrictions: An Instrumental Variables Interpretation," *Econometrica*, 51, 1527-1550.

Hayashi, F. (2000). *Econometrics*. Princeton: Princeton University Press.

Heckman, J. (1996), "Comment on 'Identification of Causal Effects Using Instrumental Variables' by Angrist, J., G. Imbens, and D. Rubin," *Journal of the American Statistical Association*, 91, 459-462.

Heckman, J. (1997), "Instrumental Variables: A Study of Implicit Behavioral Assumptions Used in Making Program Evaluations," *Journal of Human Resources*, 32, 441-462.

Heckman, J. (2006), "The Scientific Model of Causality," *Sociological Methodology* (forthcoming).

Heckman, J. and R. Robb (1985), "Alternative Methods for Evaluating the Impact of Interventions," in J. Heckman and B. Singer (eds.), *Longitudinal Analysis of Labor Market Data*. Cambridge: Cambridge University Press, pp.146-245.

Heckman J. and E. Vytlacil (1999), "Local Instrumental Variables and Latent Variable Models for Identifying and Bounding Treatment Effects," *Proceedings of the National Academy of Sciences*, 96, 4730-4734.

Heckman, J. and E. Vytlacil (2001), "Local Instrumental Variables," in C. Hsiao, K. Morimune, and J. Powell (eds.) in *Nonlinear Statistical Inference: Essays in Honor of Takeshi Amemiya*. Cambridge: Cambridge University Press, pp.1-46.

Heckman, J. and E. Vytlacil (2005), "Structural Equations, Treatment Effects, and Econometric Policy Evaluation," *Econometrica*, 73, 669-738.

Heckman, J. and E. Vytlacil (2007), "Evaluating Marginal Policy Changes and the Average Effect of Treatment for Individuals at the Margin," University of Chicago, Department of Economics Discussion Paper.

Heckman, J., H. Ichimura, and P. Todd (1998), "Matching as an Econometric Evaluation Estimator," *The Review of Economic Studies*, 65, 261-294.

Heckman, J., S. Urzua, and E. Vytlacil (2006), "Understanding Instrumental Variables in Models with Essential Heterogeneity," *Review of Economics and Statistics*, 88, 389–432.

Hirano, K. and G. Imbens (2004), "The Propensity Score with Continuous Treatments," in A. Gelman and X.-L. Meng (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 Discussion Paper.

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

Hurwicz, L. (1950), "Generalization of the Concept of Identification," in *Statistical Inference in Dynamic Economic Models*, Cowles Commission Monograph 10. New York: John Wiley.

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

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

Manski, C. (2003). *Partial Identification of Probability Distributions*. Berlin: Springer-Verlag.

Matzkin, R. (2003), "Nonparametric Estimation of Nonadditive Random Functions," *Econometrica*, 71, 1339-1375.

Matzkin, R. (2004), "Unobservable Instruments," Northwestern University Department of Economics Working Paper.

Matzkin, R. (2008), "Identification of Nonparametric Simultaneous Equations Models," *Econometrica*, 76, 945 - 978.

Morgan, M.S. (1990). *The History of Econometric Ideas*. Cambridge: Cambridge University Press.

Pearl, J. (1995), "Causal Diagrams for Empirical Research" (with Discussion), *Biometrika*, 82, 669-710.

Pearl, J. (2000). *Causality: Models, Reasoning, and Inference*. New York: Cambridge University Press.

Reichenbach, H. (1956). *The Direction of Time*. Berkeley: University of California Press.

Rosenbaum, P. R. and D. Rubin (1983), "The Central Role of the Propensity Score in Observational Studies for Causal Effects," *Biometrika*, 70, 41-55.

Roy, A. D. (1951), "Some Thoughts on the Distribution of Earnings," *Oxford Economic Papers* (New Series), 3, 135-146.

Rubin, D. (1974), "Estimating Causal Effects of Treatments in Randomized and Non-randomized

Studies," *Journal of Educational Psychology*, 66, 688-701.

Schennach, S., H. White, and K. Chalak (2009), "Estimating Average Marginal Effects in Nonseparable Structural Systems," UCSD Department of Economics Discussion Paper.

Song, S., S. Schennach, and H. White (2009), "Identification and Estimation of Nonseparable Models with Measurement Errors," UCSD Department of Economics Discussion Paper.

Spirtes, P., C. Glymour, and R. Scheines (1993). *Causation, Prediction and Search.* Berlin: Springer-Verlag.

Stock, J. and F. Trebbi (2003), "Who Invented Instrumental Variable Regression?" *Journal of Economic Perspectives*, 17, 177-194.

White, H. (2001). *Asymptotic Theory for Econometricians.* New York: Academic Press.

White, H. (2006), "Time Series Estimation of the Effects of Natural Experiments," *Journal of Econometrics*, 135, 527-566.

White, H. and K. Chalak (2006), "Parametric and Nonparametric Estimation of Covariate-Conditioned Average Effects," UCSD Department of Economics Discussion Paper.

White, H. and K. Chalak (2009a), "Identifying Structural Effects in Nonseparable Systems Using Covariates," UCSD Department of Economics Discussion Paper.

White, H. and K. Chalak (2009b), "Settable Systems: An Extension of Pearl's Causal Model with Optimization, Equilibrium, and Learning," *Journal of Machine learning Research*, 10 (Aug), 1759-1799.

White, H. and X. Lu (2009), "Granger Causality and Dynamic Structural Systems," UCSD Department of Economics Discussion Paper.

Wooldridge, J. M. (2002). *Econometric Analysis of Cross Section and Panel Data.* Cambridge: MIT Press.

Wright, P. G. (1928). *The Tariff on Animal and Vegetable Oil.* New York: Macmillan.

Wright, S. (1921), "Correlation and Causation," *Journal of Agricultural Research*, 20, 557-85.

Wright, S. (1923), "The Theory of Path Coefficients: A Reply to Niles' Criticism," *Genetics*, 8, 239-255.