*Statistical Science*
2014, Vol. 29, No. 3, 363–366
DOI: 10.1214/14-STS485
Main article DOI: 10.1214/14-STS480


# ACE Bounds; SEMs with Equilibrium Conditions

**Thomas S. Richardson and James M. Robins**

We congratulate the author on an enlightening account of the instrumental variable approach from the viewpoint of Econometrics. We first make some comments regarding the bounds on the ACE under the nonparametric IV model, and then discuss potential outcomes in the market equilibrium model.

## 1. ACE BOUNDS UNDER THE IV MODEL

We consider the model in which $X$ and $Y$ are binary, taking values in $\{0, 1\}$, while $Z$ takes $K$ states $\{1, \ldots, K\}$. We use the notation $X(z_i)$ to indicate $X(z = i)$, similarly $Y(x_j)$ for $Y(x = j)$. We consider four different sets of assumptions:

(i) $Z \perp\!\!\!\perp Y(x_0), Y(x_1), X(z_1), \ldots, X(z_K)$;
(ii) $Z \perp\!\!\!\perp Y(x_0), Y(x_1)$;
(iii) for $i \in \{1, \ldots, K\}$, $j \in \{0, 1\}$, $Z \perp\!\!\!\perp X(z_i), Y(x_j)$;
(iv) there exists a $U$ such that $U \perp\!\!\!\perp Z$ and for $j \in \{0, 1\}$, $Y(x_j) \perp\!\!\!\perp X, Z \mid U$.

Condition (i) is joint independence of $Z$ and all potential outcomes for $Y$ and $X$. (ii) does not assume independence (or existence) of counterfactuals for $X$. (iii) is a subset of the independences in (i), none of which involve potential outcomes from different worlds.[1] The counterfactual independencies (i), (ii), (iii) arise most naturally in the context where the instrument is randomized, as depicted by the DAG in Figure 1(a). Assumption (iii) may be read (via d-separation) from the Single-World Intervention Graph (SWIG)[2] $\mathcal{G}_1(z, x)$, depicted in Figure 1(b), which represents the factorization of $P(Z, X(z), Y(x), U)$, implied by the IV model.

Lastly (iv) consists of only three independence statements, but does assume the existence of an unobserved variable $U$ that is sufficient to control for confounding between $X$ and $Y$. No assumption is made concerning the existence of counterfactuals $X(z)$; confounding variables $(U^*)$ between $Z$ and $X$ are permitted (so long as $U^* \perp\!\!\!\perp U$). The DAG $\mathcal{G}_2$ and corresponding SWIG $\mathcal{G}_2(x)$ are shown in Figure 1(c), (d). In Richardson and Robins (2014), we prove the following.

THEOREM 1. *Under any of the assumptions* (i), (ii), (iii), (iv), *the set of possible joint distributions* $P(Y(x_0), Y(x_1))$ *are characterized by the* $8K$ *inequalities*:

$$P\bigl(Y(x_i) = y\bigr) \le P(Y = y, X = i \mid Z = z) + P(X = 1 - i \mid Z = z), \tag{1}$$

$$P\bigl(Y(x_0) = y, Y(x_1) = \tilde{y}\bigr) \le P(Y = y, X = 0 \mid Z = z) + P(Y = \tilde{y}, X = 1 \mid Z = z). \tag{2}$$

Thus a distribution $P(X, Y \mid Z)$ is compatible with the stated assumptions if and only if there exists a distribution $P(Y(x_0), Y(x_1))$ satisfying (1) and (2).

THEOREM 2. *Under any of the assumptions* (i), (ii), (iii), (iv) *for all* $i, j \in \{0, 1\}$, $P(Y(x_i) = j) \le g(i, j)$, *where*

$$\begin{aligned} g(i, j) \equiv \min\Bigl\{ &\min_z \bigl[P(X = i, Y = j \mid Z = z) + P(X = 1 - i \mid Z = z)\bigr], \\ &\min_{z, \tilde{z}: z \neq \tilde{z}} \bigl[P(X = i, Y = j \mid Z = z) + P(X = 1 - i, Y = 0 \mid Z = z) \\ &\quad + P(X = i, Y = j \mid Z = \tilde{z}) + P(X = 1 - i, Y = 1 \mid Z = \tilde{z})\bigr]\Bigr\}. \end{aligned}$$

*Thomas S. Richardson is Professor and Chair, Department of Statistics, University of Washington, Box 354322, Seattle, Washington 98195, USA (e-mail: thomasr@uw.edu). James M. Robins is Mitchell L. and Robin LaFoley Dong Professor, Department of Epidemiology, Harvard School of Public Health, 677 Huntington Avenue, Boston, Massachusetts 02115, USA (e-mail: robins@hsph.harvard.edu).*

[1] In other words, they do not involve both $Y(x_0)$ and $Y(x_1)$, nor $X(z_i)$ and $X(z_j)$ for $i \neq j$.

[2] See Richardson and Robins (2013) for details.

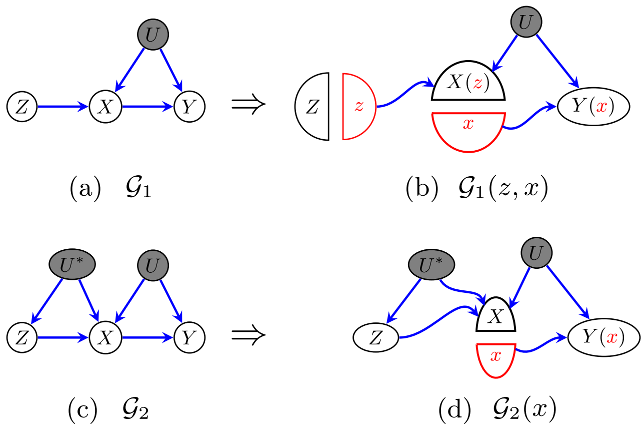


FIG. 1. (a) *IV model with no confounding between $Z$ and $X$*; (b) *SWIG representing $P(Z, X(z), Y(x), U)$*; (c) *IV model with confounding between $Z$ and $X$*; (d) *SWIG representing $P(Z, X, Y(x), U, U^*)$*.

*Furthermore, $P(Y(x_0))$ and $P(Y(x_1))$ are variation independent. Consequently,*

$$1 - g(1,0) - g(0,1) \le \mathrm{ACE}(X \to Y)$$
$$\le g(0,0) + g(1,1) - 1.$$

*These bounds are sharp.*

Note that to evaluate $g(i, j)$ requires finding a minimum over $K^2$ expressions. In the case where $K = 2$, these bounds reduce to those given by Balke and Pearl (1997), who assume (i).[3] Robins (1989) and Manski (1990) derived what are called the "natural bounds" on the ACE under the weaker assumption that $Z \perp\!\!\!\perp Y(x_0)$ and $Z \perp\!\!\!\perp Y(x_1)$. As noted by Imbens, without further assumptions these bounds are not sharp. However, the natural bounds are sharp under (i) or (iii), if, in addition, we assume there are no Defiers (an assumption that has testable implications). Cheng and Small (2006) considered bounds on the ACE when $K = 3$ under additional assumptions.

## 2. MARKET EQUILIBRIUM AND BICAUSAL MODELS

Imbens' clear description of the market equilibrium model is particularly informative. We also strongly endorse the author's contention that the RHS of systems of structural equations should be interpreted as describing potential outcomes for the LHS.[4]

However, we note that this position has important implications both for interpretation and inference. Furthermore, it does not seem to be universally accepted within Economics. LeRoy (2006) states that "economic models use the equality symbol with its usual mathematical meaning, not with the meaning of the assignment operator"; an approach that is clearly incompatible with an interpretation in terms of potential outcomes. For example, it becomes permissible to renormalize structural equations to change which variable is on the LHS.

It has also been argued that statistical analyses of such models should be invariant to the normalization; see Hillier (1990), Basmann (1963).[5] Contrary to Imbens' remark,[6] this alternative view does not appear to be motivated by considerations of measurement error. LeRoy (2006) makes clear that he does not believe that structural equations describe potential outcomes for endogenous variables and does not discuss issues relating to measurement.[7] Rather, this appears to be a fundamental difference in interpretation.

The market equilibrium model specifies potential outcomes for $Q_t^d(p)$, $Q_t^s(p)$:

$$Q_t^d(p) = \alpha^d + \beta^d p + \varepsilon_t^d, \tag{3}$$

$$Q_t^s(p) = \alpha^s + \beta^s p + \varepsilon_t^s, \tag{4}$$

and imposes the equilibrium condition:[8]

$$Q_t^d(p) = Q_t^s(p). \tag{5}$$

[3]Dawid (2003) working in a non-counterfactual framework also established the bounds for $K = 2$ under the DAG in Figure 1(a); however, his proof also applies to Figure 1(c). Robins and Greenland (1996) observed that the Balke–Pearl bounds were also sharp under (ii).

[4]Pearl (2000), Lauritzen (2001), Lauritzen and Richardson (2002) argue that these are not really "equations" but are better viewed as "assignments" in computer languages, for example, $y \leftarrow x + 1$; see also Strotz and Wold (1960), page 420.

[5]For example, Greene (2003), page 401, states (in the context of the IV model): "one significant virtue of [the Limited Information Maximum Likelihood Estimator] is its invariance to normalization of the equations."

[6]Footnote 8, page 331.

[7]For example, LeRoy (2006), page 23, states that "The assumption that it makes sense to delete one or more of the structural equations and replace the value of the internal variable so determined by a constant without altering the other equations [...] is virtually never satisfied in economic models since each external variable typically affects equilibrium values of more than one internal variable." He goes on to assert "In fact, it is difficult to think of nontrivial models in any area of research in which the [...] assumption is satisfied."

[8]To simplify notation, throughout we work directly in terms of log price and log quantity.

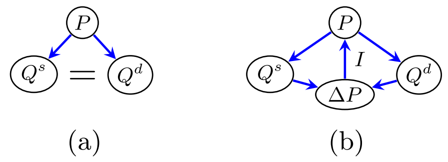


FIG. 2. (a) *Attempt to depict the bicausal model*; (b) *a schematic showing the deterministic system* (6)–(8); *the edge* $\xrightarrow{I}$ *denotes that* $P$ *is the integral of* $\Delta P$; *see Iwasaki and Simon* (1994).

Strotz and Wold (1960) described such systems as *bicausal*. It should be observed that the model does not specify potential outcomes for price ($P_t(q_s, q_d)$), nor does it view price as externally determined (i.e., exogenous). Instead price is determined implicitly as a consequence of the equilibrium condition. In this regard, the model might be regarded as incomplete: Indeed Haavelmo (1958) is quite critical of this model for failing to offer any *explanation* as to how the equilibrium price is determined. The model also falls outside the scope of non-parametric structural equation models (NPSEM) (see, e.g., Pearl, 2000), which require one equation for each endogenous variable;[9] likewise the model defies standard graphical representation, though see Figure 2(a).

A related question concerns whether there exist dynamic acyclic (i.e., recursive) systems of structural equations that lead to the equilibrium distribution corresponding either to a cyclic system of structural equations or a bicausal system.[10] Fisher (1970) provides just such a "correspondence principle" under which the distribution implied by a cyclic linear SEM is obtained as a time average of a deterministic set of first order difference equations reaching a static equilibrium subject to stochastic boundary conditions. The correspondence assumes that the equilibration time is very fast relative to the interval between observations so the time averaged variables are in deterministic equilibrium. Fisher also derived conditions on the coefficient matrices of a cyclic SEM that are required in order for the system to reach equilibrium; in fact he further required that each subset of structural equations also have this property.

However, Fisher's correspondence presumes a normalization under which each variable is associated with a single equation (as in an NPSEM), and hence would not apply to a bicausal system. Richardson (1996), Chapter 2, described a system of finite difference equations that gives rise to the bicausal system (3)–(5):

$$\text{Consumers:}\quad Q^d_{t+(k+1)\delta}(p_{t+k\delta}) = \alpha^d + \beta^d p_{t+k\delta} + \varepsilon^d_t, \tag{6}$$

$$\text{Suppliers:}\quad Q^s_{t+(k+1)\delta}(p_{t+k\delta}) = \alpha^s + \beta^s p_{t+k\delta} + \varepsilon^s_t, \tag{7}$$

$$\text{Merchants:}\quad P_{t+(k+1)\delta}(q^d_{t+k\delta}, q^s_{t+k\delta}, p_{t+k\delta}) = p_{t+k\delta} + \lambda(q^d_{t+k\delta} - q^s_{t+k\delta}), \tag{8}$$

for $k = \{0, \dots, \delta^{-1} - 1\}$. Note that the disturbances $(\varepsilon^d_t, \varepsilon^s_t)$ represent boundary conditions and hence remain fixed during the interval $[t, t+1)$. As in Fisher's correspondence, the observed variables correspond to limiting time-averages over a unit interval:

$$\overline{Q}^d_t = \lim_{\delta\to 0} \delta \sum_{k=0}^{\delta^{-1}-1} Q^d_{t+k\delta}, \qquad \overline{Q}^s_t = \lim_{\delta\to 0} \delta \sum_{k=0}^{\delta^{-1}-1} Q^s_{t+k\delta},$$

$$\overline{P}_t = \lim_{\delta\to 0} \delta \sum_{k=0}^{\delta^{-1}-1} P_{t+k\delta}.$$

Under suitable conditions on the coefficients, $(\overline{Q}^d_t, \overline{Q}^s_t, \overline{P}_t)$ obey equations (3)–(5). Note that Merchants' equation (8) which includes $P$, leads to the equilibrium condition (5) that does not.[11] It might be objected to the proposed model that there is no disturbance term in equation (8). The explanation for this is that the disturbance terms in the nonrecursive model correspond to constant factors in the deterministic evolution. The equation for price gives the change in price during a small interval (length $\delta$) to the discrepancy between supply and demand. Adding a disturbance term would say that throughout the observation period (length 1) the Merchants' reaction to change in price was off by a constant factor, so that even if quantities supplied and demanded were identical, the Merchants would change

[9] Indeed LeRoy (2006) argues against the interpretation of structural equations in terms of potential outcomes on the grounds that this interpretation, as advanced by Pearl, requires a one-to-one mapping between equations and endogenous variables that he argues, does not make sense for the market equilibrium model.

[10] Analysis of this question was stimulated by a heated debate that arose between Wold, who advocated a recursive, regression-based approach to demand analysis, and Haavelmo and the Cowles Commission who advocated simultaneous equations. See Haavelmo (1943), Bentzel and Wold (1946), Wold and Juréen (1953), Bentzel and Hansen (1954), Strotz and Wold (1960), Basmann (1963); historical overviews are given by Morgan (1991), Epstein (1987).

[11] In causal terms, this model is similar to one presented in Wold (1959). Wold viewed his model as a formalization of Cournot's theories.

the price. Thus, if we add an error $\varepsilon_t^p$ the model will not, in general, arrive at equilibrium within the unit interval.[12]

Iwasaki and Simon (1994) represent equilibrating mechanisms via "causal influence diagrams" in which the derivatives of variables are included. Under this scheme, model (6)–(8) is represented by the graph in Figure 2(b). This example serves to show that time averages of (deterministic) equilibrating systems need not have a structural equation for each variable. See also (Dash and Druzdzel, 2001) for related work.

## ACKNOWLEDGMENTS

This work was supported by the US National Institutes of Health Grant R01 AI032475; Richardson was also supported by the US National Science Foundation Grant CNS-0855230.

## REFERENCES

BALKE, A. and PEARL, J. (1997). Bounds on treatment effects from studies with imperfect compliance. *J. Amer. Statist. Assoc.* **92** 1171–1176.

BASMANN, R. L. (1963). The causal interpretation of non-triangular systems of economic relations (with discussion). *Econometrica* **31** 439–453.

BENTZEL, R. and HANSEN, B. (1954). On recursiveness and interdependency in economic models. *Rev. Econom. Stud.* **22** 153–168.

BENTZEL, R. and WOLD, H. (1946). On statistical demand analysis from the viewpoint of simultaneous equations. *Skand. Aktuarietidskr.* **29** 95–114. MR0017907

CHENG, J. and SMALL, D. S. (2006). Bounds on causal effects in three-arm trials with non-compliance. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **68** 815–836. MR2301296

DASH, D. and DRUZDZEL, M. J. (2001). Caveats for causal reasoning with equilibrium models. In *Proceedings of the Sixth European Conference on Symbolic and Quantitative Approaches to Reasoning with Uncertainty* (*ECSQARU*), *Toulouse*, *France* (S. Benferhat and P. Besnard, eds.). *Lecture Notes in Artificial Intelligence* **2143** 192–203. Springer, Berlin.

DAWID, A. P. (2003). Causal inference using influence diagrams: The problem of partial compliance. In *Highly Structured Stochastic Systems* (P. J. Green, N. L. Hjort and S. Richardson, eds.). *Oxford Statist. Sci. Ser.* **27** 45–81. Oxford Univ. Press, Oxford. MR2082406

EPSTEIN, R. J. (1987). *A History of Econometrics*. *Contributions to Economic Analysis* **165**. North-Holland, Amsterdam. MR0918969

FISHER, F. M. (1970). A correspondence principle for simultaneous equation models. *Econometrica* **38** 73–92.

[12]Having said this, the equations (6) and (7) still imply that producers and consumers make systematic errors in computing prices over a time-scale of length $\delta$.

GREENE, W. H. (2003). *Econometric Analysis*, 5th ed. Prentice Hall, Upper Saddle River, NJ.

HAAVELMO, T. (1943). The statistical implications of a system of simultaneous equations. *Econometrica* **11** 1–12. MR0007954

HAAVELMO, T. (1958). Hva kan statiske likevektsmodeller fortelle oss? *Nationaløkonomisk Tidsskrift* **96 (Suppl.)** 138–145 (in Norwegian). English translation published as: What can static equilibrium models tell us? *Economic Inquiry* **12** (1974) 27–34.

HILLIER, G. H. (1990). On the normalization of structural equations: Properties of direction estimators. *Econometrica* **58** 1181–1194. MR1079413

IWASAKI, Y. and SIMON, H. A. (1994). Causality and model abstraction. *Artificial Intelligence* **67** 143–194. MR1281657

LAURITZEN, S. L. (2001). Causal inference from graphical models. In *Complex Stochastic Systems* (*Eindhoven*, 1999). *Monogr. Statist. Appl. Probab.* **87** 63–107. Chapman & Hall/CRC, Boca Raton, FL. MR1893411

LAURITZEN, S. L. and RICHARDSON, T. S. (2002). Chain graph models and their causal interpretations. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **64** 321–361. MR1924296

LEROY, S. F. (2006). Causality in economics. Technical report, Univ. California, Santa Barbara.

MANSKI, C. F. (1990). Non-parametric bounds on treatment effects. *American Economic Review* **80** 351–374.

MORGAN, M. S. (1991). The stamping out of process analysis in econometrics. In *Appraising Economic Theories* (N. de Marchi and M. Blaug, eds.) 237–272. Edward Elgar, Cheltenham.

PEARL, J. (2000). *Causality*: *Models*, *Reasoning*, *and Inference*. Cambridge Univ. Press, Cambridge. MR1744773

RICHARDSON, T. S. (1996). Models of feedback: Interpretation and discovery. Ph.D. thesis, Carnegie-Mellon Univ.

RICHARDSON, T. S. and ROBINS, J. M. (2013). Single World Intervention Graphs (SWIGs): A unification of the counterfactual and graphical approaches to causality. Technical Report 128, Center for Statistics and the Social Sciences, Univ. Washington, Seattle, WA.

RICHARDSON, T. S. and ROBINS, J. M. (2014). Assumptions and bounds in the instrumental variable model. Preprint.

ROBINS, J. M. (1989). The analysis of randomized and non-randomized AIDS treatment trials using a new approach to causal inference in longitudinal studies. In *Health Service Research Methodology*: *A Focus on AIDS* (L. Sechrest, H. Freeman and A. Mulley, eds.). U.S. Public Health Service, Washington, DC.

ROBINS, J. M. and GREENLAND, S. (1996). Identification of causal effects using instrumental variables: Comment. *J. Amer. Statist. Assoc.* **91** 456–458.

STROTZ, R. H. and WOLD, H. O. A. (1960). Recursive vs. nonrecursive systems: An attempt at synthesis. *Econometrica* **28** 417–427. MR0120034

WOLD, H. O. A. (1959). Ends and means in econometric model building. In *Probability and Statistics*: *The Harald Cramér Volume* (U. Grenander, ed.) 355–434. Almqvist & Wiksell, Stockholm. MR0109088

WOLD, H. O. A. and JURÉEN, L. (1953). *Demand Analysis*. Wiley, New York.