Causal Bounds and Observable Constraints for Non-deterministic Models

Roland R. Ramsahai   RAMSAHAI@STATSLAB.CAM.AC.UK

Statistical Laboratory
Centre for Mathematical Sciences
University of Cambridge
Wilberforce Road
Cambridge CB3 0WB, UK

Abstract

Conditional independence relations involving latent variables do not necessarily imply observable independences. They may imply inequality constraints on observable parameters and causal bounds, which can be used for falsification and identification. The literature on computing such constraints often involve a deterministic underlying data generating process in a counterfactual framework. If an analyst is ignorant of the nature of the underlying mechanisms then they may wish to use a model which allows the underlying mechanisms to be probabilistic. A method of computation for a weaker model without any determinism is given here and demonstrated for the instrumental variable model, though applicable to other models. The approach is based on the analysis of mappings with convex polytopes in a decision theoretic framework and can be implemented in readily available polyhedral computation software. Well known constraints and bounds are replicated in a probabilistic model and novel ones are computed for instrumental variable models without non-deterministic versions of the randomization, exclusion restriction and monotonicity assumptions respectively.

Keywords: instrumental variables, instrumental inequality, causal bounds, convex polytope, latent variables, directed acyclic graph

1. Introduction

Conditional independence relations represent equality constraints on the parameters of a joint probability distribution. Such relations cannot be empirically validated if they involve latent variables. Collections of latent conditional independencies may imply inequality constraints on parameters of the observable distribution. The classical motivation is the *instrumental variable* (IV) model (Durbin, 1954; Angrist et al., 1996). It includes the IV, A, and inference is required about the effect of a variable, B, on another, C, in the presence of latent confounders, U. The IV model is defined by $A \not\!\perp\!\!\!\perp B$, $C \perp\!\!\!\perp A \,|\, (B, U)$ and $U \perp\!\!\!\perp A$. The latter two involve the latent variable U so it was traditionally thought that the model could not be empirically verified (Imbens and Angrist, 1994). However Pearl (1995) derived the 'instrumental inequality'

$$\max_B \sum_C \left\{ \max_A \mathrm{P}(C, B \,|\, A) \right\} \leq 1, \tag{1}$$

a set of constraints which are implied by and can be used to falsify the discrete IV model. To compute the constraints, Pearl (1995) defines the IV model as a deterministic counterfactual model (Rubin, 1974). Such models involve latent deterministic relations, which marginalise to produce observed probabilistic relationships, and are technically equivalent to structural equation models (Strotz and Wold, 1960) and other functional models (Heckerman and Shachter, 1995).

Without intervention data and further assumptions, the causal effect of B on C cannot be point identified (Durbin, 1954; Angrist et al., 1996), but, using the deterministic counterfactual model, it can be bounded with the joint distribution of A, B and C (Pearl, 1995; Robins, 1989; Manski, 1990). Thus making it possible to acquire non-trivial information about the effect of the intervention when intervention studies cannot be conducted; because of ethical, financial or other reasons. Using the deterministic counterfactual approach and linear programming software developed by Balke (1995), the constraints on the causal effect of B on C were improved by Balke and Pearl (1997) and extended to other models by Kaufman et al. (2009). This linear programming approach within a deterministic counterfactual model has become the standard tool for computing such constraints, with some exceptions (Geiger and Meek, 1998; Kang and Tian, 2006).

As a technical construct for computations, deterministic counterfactual models are widely accepted as valuable. Applications of deterministic counterfactual models assume there are underlying deterministic relations (Angrist et al., 1996) and pose no issues if the determinism can be practically justified. For certain applications though, for example, mutations that cause cancer (Aalen and Frigessi, 2007), subject matter knowledge suggests that assumptions about the existence of deterministic mechanisms are unrealistic and spawns controversy (Dawid, 2000). Even if an analyst is unaware of the type of mechanisms involved in their study, it would be desirable to avoid deterministic counterfactuals if alternative computations are no more difficult. The method in §2 provides such an alternative, which is agnostic to whether the underlying mechanisms are probabilistic or deterministic, to deriving falsifiable constraints and causal bounds of the type previously described. The method described does not use counterfactuals, which has certain advantages (Dawid, 2000), but more importantly demonstrates that the determinism in the models is unnecessary.

In this discussion, causal inference is formalized within standard decision theory (Spirtes et al., 1993; Pearl, 1993), with conditional independence assumptions (Lauritzen, 2001; Dawid, 2002). The model has been successfully applied in defining direct effects (Geneletti, 2007) and dynamic treatment strategies (Dawid and Didelez, 2010), to name a few. The approach in §2 and throughout uses this framework and it is compared to the counterfactual framework in §3. The method is based on the analysis of convex polytopes and can be implemented in standard polytope representation software such as Polymake (Gawrilow and Joswig, 2000) or PORTA (Christof and Loebel, 1998). Known constraints, which have been previously derived using deterministic counterfactuals, are derived in §5. Graphical models for representing causal assumptions are described in §4. Non-trivial modifications of the computation technique are considered in §6, §7 and §8 to derive novel constraints and causal bounds when various assumptions in the IV model are weakened.

Example 1 *Consider an IV model of partial compliance, where $A \in \{1,2\}$ is treatment assigned, $B \in \{0,1\}$ is treatment taken and C is an outcome of interest. The counterfactual IV model involves counterfactual variables (B_1, B_2) which represent a unit's deterministic compliance behaviour when A is set to 1 (no treatment) or 2 (treatment) respectively. Analyses of this model often make the monotonicity assumption $B_2 \geq B_1$, meaning that a unit which does not take treatment if assigned it, will never take it.*

The deterministic counterfactual framework only allows the compliance behaviour in Example 1 to be modelled as deterministic. Monotonicity assumptions, used to compute bounds in IV models, in the literature (Pearl, 1995; Balke and Pearl, 1997) are imposed in models which are stronger than necessary. It is shown in §8 that the known bounds and novel bounds can be derived for a weaker model. In the context of Example 1, monotonicity in the weaker model is equivalent to assuming that units are more likely to take treatment if assigned it than if not assigned it.

The counterfactual IV model in Example 1 uses the exclusion restriction assumption (Imbens and Angrist, 1994). This assumption restricts C to be a deterministic function of compliance behaviour and treatment taken only. Stochastic exclusion restrictions are considered within the deterministic counterfactual framework in Hirano et al. (2000). In a weaker fully probabilistic model, the exclusion restriction assumption $C \perp\!\!\!\perp A \,|\, (B,U)$ is used in §2 to replicate results which were derived under the stronger model (Balke and Pearl, 1993; Pearl, 1995; Balke and Pearl, 1997). Whilst varying the strength of the exclusion restriction, novel constraints are derived in §7 with the probabilistic approach. This allows a sensitivity analysis to the non-deterministic exclusion restriction, which is important when assumptions involve unobservable variables (Shepherd et al., 2006).

Another assumption in the IV model in Example 1 is that treatment assignment is independent of compliance behaviour. In the probabilistic framework, novel constraints are computed for a weaker IV model with $U \perp\!\!\!\perp A$, as described in §6. Applications to data are given in §9. The IV model provides motivation for this discussion but the approach extends to other models. The notation used throughout is listed in Appendix A.

2. Computation of Constraints in the Instrumental Variable Model

Consider a model involving the random variables A, B, C and U, where the state space of A is $\{1,2\}$, B is $\{0,1\}$ and C is $\{0,1\}$. U is unobservable by definition so no assumption is made about it. Let $\vec{v}^* = (\zeta_{00.1}^*, \zeta_{01.1}^*, \ldots, \zeta_{11.2}^*)$ be a random vector with components $\zeta_{cb.a}^*$, which are random variables that are functions of U, where $\zeta_{cb.a}^* = \mathrm{P}(C=c, B=b \,|\, A=a, U)$. Similarly, let $\vec{v} = (\zeta_{00.1}, \zeta_{01.1}, \ldots, \zeta_{11.2})$ be a fixed vector of probabilities that are not functions of U, where $\zeta_{cb.a} = \mathrm{P}(C=c, B=b \,|\, A=a)$. Let $\vec{\tau}^* = (\eta_0^*, \eta_1^*, \delta_1^*, \delta_2^*)$, where

$$\eta_b^* = \mathrm{P}(C=1 \,|\, B=b, U), \quad \delta_a^* = \mathrm{P}(B=1 \,|\, A=a, U). \tag{2}$$

Since $\vec{\tau}^*$ is a vector of probabilities then $\vec{\tau}^* \in \mathcal{T}$ since the components of $\vec{\tau}^*$ satisfy the axioms of probability, where $\mathcal{T} = [0,1]^4$. To derive falsifiable constraints on $\vec{v}$ for the IV model, it is necessary to determine the set of $\vec{v}$ which does not satisfy the assumptions in the IV model. Under the IV model, $C \perp\!\!\!\perp A \,|\, (B,U)$, which implies that

$$\mathrm{P}(C, B \,|\, A, U) = \mathrm{P}(C \,|\, B, U) \mathrm{P}(B \,|\, A, U), \tag{3}$$

and $\vec{v}^*$ can be parameterised by $\vec{\tau}^*$. The relation in Equation (3) together with the codes in (2) define a mapping $\Xi : \vec{\tau}^* \in \mathcal{T} \to \vec{v}^* \in \mathcal{V}$, where $\vec{\tau}^*$ is unrestricted by the IV model and $\mathcal{V} = \Xi(\mathcal{T})$ contains all $\vec{v}^*$ which obey the IV model. Since the components of each $\vec{v}^*$ obey the axioms of probability then $\mathcal{V} \subseteq \mathcal{Z}$, where $\mathcal{Z} \subset [0,1]^8$ is the intersection of the hyperplanes defined by $\sum_{c,b} \zeta_{cb.a}^* = 1$ for $a \in \{1,2\}$.

Under the IV model, $U \perp\!\!\!\perp A$, which implies that $\zeta_{cb.a} = \mathrm{E}_U(\zeta_{cb.a}^*)$ and thus all $\vec{v}$ that obey the IV model lie in $\mathcal{H}$, where $\mathcal{H}$ is the set of all possible convex combinations of all $\vec{v}^* \in \mathcal{V}$ or the convex

hull of $\mathcal{V}$. Let $\hat{\mathcal{V}} = \Xi(\hat{\mathcal{T}})$ and $\hat{\mathcal{H}}$ be the convex hull of $\hat{\mathcal{V}}$, where $\hat{\mathcal{T}}$ is the collection of extreme vertices of $\mathcal{T}$. The vertices $\hat{\mathcal{T}}$ and $\hat{\mathcal{V}}$ are partially listed in Figure 1 (top) and the transformation $\Xi(\cdot)$ is represented in Figure 1 (bottom).

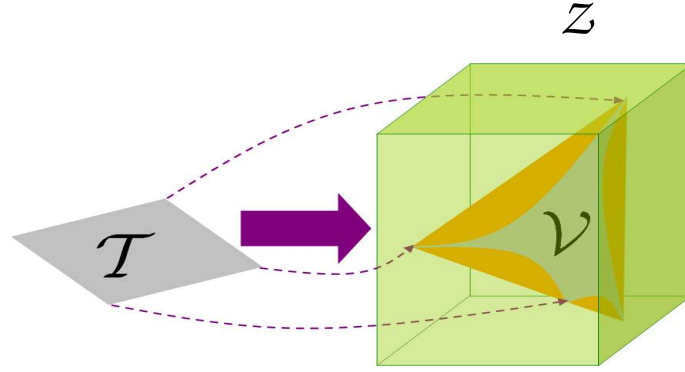| η_0^*	η_1^*	δ_1^*	δ_2^*		$\zeta_{00.1}^*$	$\zeta_{01.1}^*$	$\zeta_{10.1}^*$	$\zeta_{11.1}^*$	$\zeta_{00.2}^*$	$\zeta_{01.2}^*$	$\zeta_{10.2}^*$	$\zeta_{11.2}^*$
0	0	0	0		1	0	0	0	1	0	0	0
0	0	0	1	$\rightarrow$	1	0	0	0	0	1	0	0
$\vdots$	$\vdots$	$\vdots$	$\vdots$		$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
1	1	1	1		0	0	0	1	0	0	0	1

Figure 1: Transformation of extreme vertices (top) of polytope (bottom).

Since $\mathcal{H} = \hat{\mathcal{H}}$, from Theorem 1 in Appendix B, then all $\vec{v}$ that obey the IV model lie in $\hat{\mathcal{H}}$. The proof of Theorem 1 does not use the specific form of $\Xi(\cdot)$, only its monotonicity in each coordinate. A program such as Polymake (Gawrilow and Joswig, 2000) or PORTA (Christof and Loebel, 1998) can be used to transform the representation of $\hat{\mathcal{H}}$ in terms of its extreme vertices to a representation in terms of its facets or inequalities. The inequalities are constraints which are satisfied by $\vec{v}$ if $\vec{v}$ obeys the IV model. This specific computation produces the falsifiable 'instrumental inequality' constraints in (1) and is exactly the approach of Dawid (2003).

It is possible for the randomization or exclusion restriction assumption to fail without violation of any of the constraints in (1). This is because there are distributions $P(C, B, A, U)$ which either violate the assumption $U \perp\!\!\!\perp A$ or Equation (3) but give rise to margins $P(C, B, A)$ that obey the inequalities in (1). For example, if all $\vec{v}^*$ lie in $\mathcal{H} \backslash \mathcal{V}$ and randomization holds then the exclusion restriction in Equation (3) is not satisfied but all $\vec{v} \in \mathcal{H}$, which means that the inequalities in (1) are satisfied. I conjecture that the condition that U has a certain small state space is sufficient to imply that it is possible for the IV model to fail without violation of any of the constraints in (1).

3. Geometry of Counterfactuals and Latent Variables

The binary IV model can be re-parameterised by replacing U in $P(C, B | A, U)$ with $\vec{\tau}^*$ and considering $P(C, B | A, \vec{\tau}^*)$. The polytope $\mathcal{H}$ represents the model for $P(C, B | A)$ and, since $\mathcal{H} = \hat{\mathcal{H}}$, a computationally and empirically indistinguishable model is formed by restricting $\vec{\tau}^*$ to $\hat{\mathcal{T}}$. In this minimal representation of the model, where $\vec{\tau}^* \in \hat{\mathcal{T}}$, the parameters $\eta_b^* \in \{0, 1\}$ and $\delta_a^* \in \{0, 1\}$. Therefore η_b^* is a deterministic function of B and the latent variable U and can be interpreted as

a counterfactual variable which is the value of C when $B = b$ for a given $\vec{\tau}^*$, and similarly for δ_a^*. Similar comments are given in Lauritzen (2004).

If U is interpreted as the collection of variables which define a unit then $\vec{\tau}^*$ is the vector of potential responses for a unit and each vertex of the polytope corresponds to a certain type of unit. In the partial compliance model of Example 1, the vertex $\vec{\tau}^* = (0,0,0,0)$ corresponds to a unit which is classified as *never recover* (response is 0 regardless of treatment taken) and a *never taker* (treatment taken is 0 regardless of treatment assigned).

The probabilistic model is parameterised by $\vec{\tau}^*$ over the entire polytope $\mathcal{T}$ whereas the counterfactual model is parameterised by $\vec{\tau}^*$ only at the extreme vertices of the polytope $\mathcal{T}$. In special cases where latent determinism is realistic then such a parameterisation is meaningful and assumptions about the non-existence of certain vertices of the polytope or $\vec{\tau}^* \in \hat{\mathcal{T}}$ can potentially be justified. If latent determinism is known to be unrealistic (Aalen and Frigessi, 2007) and the reparameterisation is a technical construct then it may be wise to steer clear of any interpretation beyond simply saying that they are the vertices of the polytope defining the model.

The concepts are demonstrated in the reformulation of the monotonicity assumption in §8. The deterministic counterfactual approach assumes latent determinism and interprets the vertices as having real meaning. Under the deterministic interpretation, the monotonicity assumption implies that certain vertices are not valid for the model. The probabilistic approach defines monotonicity as a constraint on the latent conditional distributions to lie in a particular half-space, still allowing probabilistic behaviour.

4. Causal Graphical Models

The IV model considered so far, that is, without causal assumptions, is relatively simple. However extensions of it will be considered later and it will be useful, though not vital, to use graphical models to represent the assumptions involved. Graphs that are useful for representing conditional independence and causal assumptions are described in §4.1 and §4.2 respectively.

4.1 Directed Acyclic Graph

A purely probabilistic directed acyclic graph (DAG) (Lauritzen, 1996) consists of a set of *vertices* or *nodes*, $\mathcal{N}$, and a set of *directed edges*, $\mathcal{E}$. If $\lambda_1, \lambda_2 \in \mathcal{N}$ and $(\lambda_1, \lambda_2) \in \mathcal{E}$ then $(\lambda_2, \lambda_1) \notin \mathcal{E}$. It is said that there is a directed edge from λ_1 to λ_2, this is written as $\lambda_1 \rightarrow \lambda_2$ and λ_1 is called a *parent* of λ_2. In a DAG which represents the probability distribution of a set of random variables, $\mathcal{X}$, every $\lambda \in \mathcal{N}$ corresponds to a random variable $\mathcal{X}_\lambda \in \mathcal{X}$. The probability distribution function has the form

$$\mathrm{P}(\mathcal{X}) = \prod_{\lambda \in \mathcal{N}} \mathrm{P}\{\mathcal{X}_\lambda \,|\, \mathcal{X}_{\mathrm{pa}(\lambda)}\}, \tag{4}$$

where 'pa$(\cdot)$' is the set of 'parents' of a node. This factorisation property is equivalent to a collection of conditional independence relations, which can be derived from the DAG using the concepts of 'd-separation' (Verma and Pearl, 1988) and a 'moral graph' (Lauritzen et al., 1990). The observational assumptions of the IV model can be represented by the DAG in Figure 2 (left).

4.2 Augmented Directed Acyclic Graph

The notation '$||$' (Lauritzen, 2001) is used for intervention conditioning and is equivalent to the 'do$(\cdot)$' notation (Goldszmidt and Pearl, 1992) and the '$\mathrm{P}_{\mathrm{man}(\cdot)}$' notation (Spirtes et al., 1993). Using
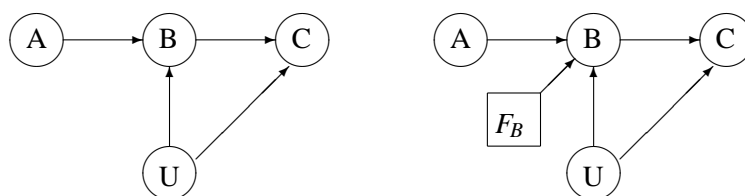
Figure 2: DAG which represents observational assumptions of the instrumental variable model (left) and augmented DAG for IV model, which includes causal assumptions (right).

the notation, $P(C||B=b)$ is the probability of C given that B is actively forced to take the value b, and not passively observed to take the value b, as in $P(C|B=b)$.

To derive intervention constraints, the assumptions represented by the augmented DAG (Spirtes et al., 1993; Pearl, 1993; Lauritzen, 2001; Dawid, 2002) in Figure 2 (right) are considered, where $ACE(B \to C) = \alpha = P(C=1||B=1) - P(C=1||B=0)$ is the causal effect of interest. The intervention node F_B is a *regime indicator* decision variable which represents the way in which the value of B arises. Conditional independence relations can be derived in the same way as for the purely probabilistic DAGs since the probability distribution, conditional on F_B, still factorises according to Equation (4). The node F_B takes the values '*idle*', 0 or 1. If $F_B = idle$ then B takes a random value given by $P\{B|pa(B)\}$, but if F_B is either 0 or 1 then $B = F_B$. Using previous notation $P(C||B=b) = P(C|F_B=b)$. The relation $C \perp\!\!\!\perp B | (F_B=b, U)$ holds from the definition of F_B but is not represented in Figure 2 (right). Square nodes are decision nodes which represent fixed strategies, whereas circle nodes are random nodes which represent random variables.

The augmented DAGs which represent the IV model without randomization and the exclusion restriction are given in Figure 3.
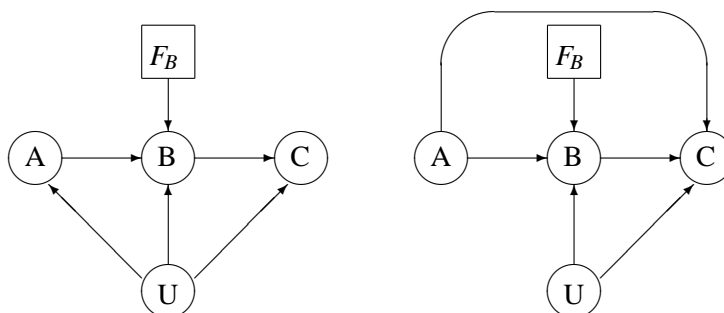

Figure 3: Augmented DAGs which represent the causal IV model without randomization (left) and without exclusion restriction (right).

The assumptions represented by the augmented DAGs in Figure 3 will be used in §6 and §7 respectively to derive constraints. The augmented DAG in Figure 2 (right) still applies under monotonicity since no extra conditional independences are assumed.

5. Applications of Computation

Many results in the literature are recovered by specific applications of the general method described in §2. It is based on parameterising with the factors of Equation (4) and transforming them according to a mapping. By defining the appropriate mapping, the various constraints are obtainable. Key requirements are the monotonicity of the mapping and that the space of valid parameters is the convex hull of the transformed polytope. Constraints on other quantities, such as $P(C|A)$, $P(B|A)$, $P(C|B)$ etc., can be derived but some interesting examples are given in §5.1 and §5.2.

5.1 Falsifiable Constraints

Some applications, such as studies with partial compliance, require constraints involving the distribution $P(C,B|A)$, whereas others can only identify the pairwise conditional distributions $P(C|A)$ and $P(B|A)$. For example, Mendelian randomization in genetic epidemiology involves the use of a genotype (A) as an instrument for the effect of a phenotype (B) on a disease (C). However only genotype-phenotype and genotype-disease data is usually available (Didelez and Sheehan, 2007) and thus constraints involving $P(C|A)$ and $P(B|A)$ are needed.

To derive the constraints, consider the monotone mapping $\vec{\tau}^* \longmapsto (\vec{\gamma}^*, \vec{\theta}^*)$ for the IV model of Figure 2 (left), where $\gamma_{ca}^* = P(C = c|A = a, U)$ and $\theta_{ba}^* = P(B = b|A = a, U)$. Since $U \perp\!\!\!\perp A$ then $(\vec{\gamma}, \vec{\theta})$ lies in the convex hull of the set of $(\vec{\gamma}^*, \vec{\theta}^*)$ which satisfy the IV model, where $\gamma_{ca} = P(C = c|A = a)$ and $\theta_{ba} = P(B = b|A = a)$. Similarly to the approach in §2, the constraints

$$\theta_{01} + \theta_{02} \geq \gamma_{01} - \gamma_{02},$$
$$\theta_{01} + \theta_{02} \geq \gamma_{02} - \gamma_{01},$$
$$\theta_{11} + \theta_{12} \geq \gamma_{01} - \gamma_{02},$$
$$\theta_{11} + \theta_{12} \geq \gamma_{02} - \gamma_{01},$$

are obtained, which are the same as in Ramsahai (2007).

5.2 Bounds on Fixed Interventions

From the motivating Mendelian randomization example in §5.1, it may be necessary to obtain causal bounds in terms of the pairwise conditional distributions $P(C|A)$ and $P(B|A)$. Consider the model in Figure 2 (right). Since $C \perp\!\!\!\perp F_B | (B, U)$ and $U \perp\!\!\!\perp F_B$ then $P(C||B) = \sum_U P(C|B, U)P(U)$. This implies that $(\vec{\gamma}, \vec{\theta}, \alpha)$ lies in the convex hull of $(\vec{\gamma}^*, \vec{\theta}^*, \alpha^*)$, where $\alpha^* = P(C = 1|B = 1, U) - P(C = 1|B = 0, U)$ and $\alpha = E_U(\alpha^*)$. Therefore the monotone mapping $\vec{\tau}^* \longmapsto (\vec{\gamma}^*, \vec{\theta}^*, \alpha^*)$ can be used to compute constraints on $(\vec{\gamma}, \vec{\theta}, \alpha)$. The results of the computation are given in Appendix C and are the same as those derived in Ramsahai (2007).

Similarly, constraints and causal bounds in terms of the identifiable $\zeta_{cb.a}$ parameters can be obtained by considering the mapping $\vec{\tau}^* \longmapsto (\vec{v}^*, \alpha^*)$. The constraints involving the identifiable $\zeta_{cb.a}$ parameters only are the same as those obtained in §2, which are given in (1), and the rest constrain α. The bounds on α are given in Appendix C and are the same as those of Dawid (2003), which are derived by Balke and Pearl (1997) in a deterministic model.

6. Relaxing the Randomization Assumption in the Instrumental Variable Model

It is possible for treatment assignment in a partial compliance study, which is suitable for an IV model, to have invalid randomization, for example, if the doctor involved is aware of the health

status of the patients. To analyze such a study, an analyst may opt for a model without randomization or at least assess the effect of the assumption on the inference. Both require constraints to be derived for a model without $U \perp\!\!\!\perp A$, as in Figure 3 (left). The decision framework is used here for computations without any assumptions of determinism. It is irrelevant to the computation whether U causes A, A causes U or both have a common cause. This is because the model in Figure 3 (left) only makes assumptions about distributions in the observational regime and the regime with intervention on B, since it includes the regime indicator F_B. No F_A or F_U regime indicators are included so no assumptions are made about interventions on A or U.

If there is data on $P(C|A)$ and $P(B|A)$ but not $P(C|B,A)$ then constraints and bounds involving $\vec{\gamma}$ and $\vec{\theta}$ are useful. Without the randomization assumption, $U \perp\!\!\!\perp A$, $(\vec{\gamma},\vec{\theta},\alpha)$ does not necessarily lie in the convex hull of $(\vec{\gamma}^*,\vec{\theta}^*,\alpha^*)$ and similarly for the other applications in §5. Assuming the exclusion restriction in Equation (3) still holds, $\vec{\tau}^*$ still fully parameterises $P(C,B|A,U)$. Consider the monotone mapping $\Xi_i(\cdot): \vec{\tau}^* \longmapsto \vec{v}_i^*$, where $\vec{v}_i^* = (\gamma_{0i}^*,\gamma_{1i}^*,\theta_{0i}^*,\theta_{1i}^*,\alpha^*)$ for $i = 1,2$, which can be expressed as

$$\alpha^* = \eta_1^* - \eta_0^*, \qquad \gamma_{0i}^* = (1-\eta_0^*)(1-\delta_i^*)+(1-\eta_1^*)\delta_i^*, \qquad \theta_{0i}^* = 1 - \delta_i^*$$
$$\gamma_{1i}^* = \eta_0^*(1-\delta_i^*)+\eta_1^*\delta_i^*, \qquad\qquad\qquad \theta_{1i}^* = \delta_i^*.$$

The transformation of $\hat{\mathcal{T}}$ by $\Xi_i(\cdot)$ is given in Figure 4. Since the relations $C \perp\!\!\!\perp F_B\,|\,(B,U)$, $U \perp\!\!\!\perp F_B$

η_0^*	η_1^*	δ_i^*		γ_{0i}^*	γ_{1i}^*	θ_{0i}^*	θ_{1i}^*	α^*
0	0	0		1	0	1	0	0
0	0	1		1	0	0	1	0
0	1	0		1	0	1	0	1
0	1	1	$\rightarrow$	0	1	0	1	1
1	0	0		0	1	1	0	-1
1	0	1		1	0	0	1	-1
1	1	0		0	1	1	0	0
1	1	1		0	1	0	1	0

Figure 4: Transformation of $\hat{\mathcal{T}}$ by $\Xi_i(\cdot)$ to the polytope which represents the IV model without randomization, in terms of the pairwise conditional distributions $P(C|A)$ and $P(B|A)$.

and $C \perp\!\!\!\perp A\,|\,(B,U)$ follow from Figure 3 and $C \perp\!\!\!\perp B\,|\,(U,F_B = B)$,

$$P(C||B) = \sum_A \sum_U P(C|B,U)P(U|A)P(A) = E_A(\alpha_A'), \tag{5}$$

where $\alpha_A' = \sum_U P(C|B,U)P(U|A)$. Since $P(C|A) = E_{U|A}\{P(C|A,U)\}$ and $P(B|A) = E_{U|A}\{P(B|A,U)\}$ then $\vec{w}_i$ lies in the convex hull of the set of $\vec{v}_i^*$, where $\vec{w}_i = (\gamma_{0i},\gamma_{1i},\theta_{0i},\theta_{1i},\alpha_i')$, and the method of §2 computes the tight constraints

$$0 \le \gamma_{0i} + 2\gamma_{1i} - \theta_{0i} + \alpha_i',$$
$$0 \le \gamma_{0i} + \theta_{0i} + \alpha_i',$$
$$0 \le \gamma_{1i} + \theta_{0i} - \alpha_i',$$
$$0 \le 2\gamma_{0i} + \gamma_{1i} - \theta_{0i} - \alpha_i',$$

or

$$\max\{\ \gamma_{0i} + \theta_{0i} - 2, -\gamma_{0i} - \theta_{0i}\ \} \le \alpha_i' \le \min\{\ -\gamma_{0i} + \theta_{0i} + 1, \gamma_{0i} - \theta_{0i} + 1\ \},$$

for all i. The constraints are tight since the vertices of the convex hull are a subset of the vertices of the transformed polytope and any vertex is achievable if the value of U, corresponding to the vertex, occurs with probability one. Since $\alpha = E_A(\alpha'_A)$ from Equation (5) then

$$\min_i \left[\max \left\{ \begin{array}{c} \gamma_{0i} + \theta_{0i} - 2 \\ -\gamma_{0i} - \theta_{0i} \end{array} \right\} \right] \leq \alpha \leq \max_i \left[\min \left\{ \begin{array}{c} -\gamma_{0i} + \theta_{0i} + 1 \\ \gamma_{0i} - \theta_{0i} + 1 \end{array} \right\} \right].$$

These bounds always span zero and are tight since the bounds on α'_i are achievable by α if $P(A = i) = 1$. If marginal A data are available, the bounds can be improved to

$$\text{ACE}(B \to C) \geq \Sigma_i \left[\max \left\{ \begin{array}{c} \gamma_{0i} + \theta_{0i} - 2 \\ -\gamma_{0i} - \theta_{0i} \end{array} \right\} P(A = i) \right],$$

$$\text{ACE}(B \to C) \leq \Sigma_i \left[\min \left\{ \begin{array}{c} -\gamma_{0i} + \theta_{0i} + 1 \\ \gamma_{0i} - \theta_{0i} + 1 \end{array} \right\} P(A = i) \right],$$

or

$$-1 + E_A(|\gamma_{1A} - \theta_{0A}|) \leq \text{ACE}(B \to C) \leq 1 - E_A(|\gamma_{0A} - \theta_{0A}|). \tag{6}$$

Although the expression in (6) bounds the unobservable causal effect, there are no falsifiable constraints to invalidate the model. The bounds in (6) always span zero.

If a sample from $P(C, B | A)$ is available, the mapping $\vec{\tau}^* \longmapsto \vec{v}_i^*$ can be used to compute observable constraints and causal bounds. The computation is possible since $P(C, B | A) = E_{U|A}\{P(C, B | A, U)\}$, which implies that $\vec{w}_i$ lies in the convex hull of the set of $\vec{v}_i^*$, where $\vec{v}_i^* = (\zeta_{00.i}^*, \zeta_{01.i}^*, \zeta_{10.i}^*, \zeta_{11.i}^*, \alpha^*)$ and $\vec{w}_i = (\zeta_{00.i}, \zeta_{01.i}, \zeta_{10.i}, \zeta_{11.i}, \alpha'_i)$. The bounds $-\zeta_{01} - \zeta_{10} \leq \text{ACE}(B \to C) \leq \zeta_{00} + \zeta_{11}$ are obtained, where $\zeta_{cb} = P(C = c, B = b)$. All of the results in this section still hold if the state space of A is extended to $\{1, 2, \ldots, l\}$ but the state space of (B, C) kept binary.

The bounds on $\text{ACE}(B \to C)$ by the ζ_{cb} parameters are derived by Manski (1990) in a model involving (B, C, C_0, C_1) under the assumptions that the potential outcomes (C_0, C_1) for a unit are the same regardless of how treatment is assigned, that is, whether by intervention or observation, and thus C is a deterministic function of (B, C_0, C_1) for a unit. The derivation, of the bounds on $\text{ACE}(B \to C)$ by the ζ_{cb} parameters, given here only requires the analogous assumptions $U \perp\!\!\!\perp F_B$ and $C \perp\!\!\!\perp F_B | (B, U)$. The additional variable A used here, which satisfies the condition $C \perp\!\!\!\perp A | (B, U)$, trivially exists by constructing a variable $A = B$. Also, the conditional independence assumption $A \perp\!\!\!\perp F_B$ represented in Figure 3 (left) is unnecessary since it is not used in the derivation.

7. Relaxing the Exclusion Restriction in the Instrumental Variable Model

The exclusion restriction assumption may often be inapplicable, for example, if patients in a study with partial compliance become aware of their treatment assignment and this affects their outcome. There could be a direct relation between treatment assignment A and the outcome C, for which the model in Figure 3 (right) would be appropriate. The probabilistic nature of the exclusion restriction within the decision framework allows the strength of the direct relation to be varied and the sensitivity of inference to this assumption to be assessed.

A weaker alternative to the exclusion restriction assumption, $C \perp\!\!\!\perp A | (B, U)$, in the binary IV model is $0 \leq |\eta_{b1}^* - \eta_{b2}^*| \leq \varepsilon$ for $b = 0, 1$, where $\eta_{ba}^* = P(C = 1 | B = b, A = a, U)$ and $0 \leq \varepsilon \leq 1$.

η_{01}^*	η_{02}^*	η_{11}^*	η_{12}^*	δ_1^*	δ_2^*
0	0	0	0	0	0
0	0	0	0.5	0	1
0	0	0.5	0	1	0
0	0	0.5	1	1	1
0	0	1	0.5	0	0
0	0	1	1	0	1
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$

$$\downarrow$$

$\zeta_{00.1}^*$	$\zeta_{01.1}^*$	$\zeta_{10.1}^*$	$\zeta_{11.1}^*$	$\zeta_{00.2}^*$	$\zeta_{01.2}^*$	$\zeta_{10.2}^*$	$\zeta_{11.2}^*$
1	0	0	0	1	0	0	0
1	0	0	0	0	0.5	0	0.5
0	0.5	0	0.5	1	0	0	0
0	0.5	0	0.5	0	0	0	1
1	0	0	0	1	0	0	0
1	0	0	0	0	0	0	1
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$

Figure 5: Transformation to the extreme vertices corresponding to the polytope which represents the IV model with the weaker exclusion restriction, for $\varepsilon = 0.5$, in terms of the distribution $P(C, B | A)$.

The condition $\varepsilon = 0$ is equivalent to the exclusion restriction. For $\varepsilon = 1$, there are no constraints on $(\eta_{b1}^*, \eta_{b2}^*)$ other than the axioms of probability and there are no falsifiable constraints or causal bounds for the IV model without the exclusion restriction. The augmented DAG in Figure 3 (right) does not represent any assumptions about ε but assumptions about ε are required to obtain non-trivial constraints and bounds. The application of the technique is considered for $\varepsilon = 0.5$. Consider the mapping of $\vec{\tau}^* = (\eta_{01}^*, \eta_{02}^*, \eta_{11}^*, \eta_{12}^*, \delta_1^*, \delta_2^*)$ to $\vec{\nu}^* = (\zeta_{00.1}^*, \zeta_{01.1}^*, \ldots, \zeta_{11.2}^*)$ for a model with $A \in \{1, 2\}$ and $B, C \in \{0, 1\}$. The transformation of some of the extreme vertices are given in Figure 5. Use of the technique produces the causal bounds in Appendix D and the constraints

$$\zeta_{00.1} + \zeta_{10.2} - \zeta_{10.1} - \zeta_{00.2} \leq 1,$$
$$\zeta_{10.1} + \zeta_{00.2} - \zeta_{00.1} - \zeta_{10.2} \leq 1,$$
$$\zeta_{11.1} + \zeta_{01.2} - \zeta_{01.1} - \zeta_{11.2} \leq 1,$$
$$\zeta_{01.1} + \zeta_{11.2} - \zeta_{11.1} - \zeta_{01.2} \leq 1,$$

which is a weaker version of the 'instrumental inequality' of Equation (1) and can be violated if the IV model with the weak exclusion restriction, $\varepsilon = 0.5$, is invalid. By adding the component $P(C | B, A, U)$ to $\vec{\nu}^*$, causal bounds on $P(C | A, F_B = B) = \sum_U P(C | B, A, U) P(U)$ can be derived for each A and used to compute bounds on $\text{ACE}(B \to C)$ since $P(C | F_B = B) = \sum_A P(C | A, F_B = B) P(A)$. Similarly to the bounds in §6, these bounds are tight. Although ε is currently defined as a constant, similar computations can be done if ε is allowed to be a function of b, that is, $|\eta_{b1}^* - \eta_{b2}^*|$ has a different range for each b.

7.1 Bounds on Direct Effects

Without assuming $C \perp\!\!\!\perp A \,|\, (B,U)$, if intervention on A is possible then the direct effect of A on C can be bounded with parameters of the distribution under no intervention. Consider extending the sample space of F_B to include the random regime d_A, which represents the regime in which $P(B|A,U,F_B = d_{a^*}) = P(B|A = a^*, U)$. Consider the controlled direct effect (CDE) (Didelez et al., 2006) and the random regime direct effect (RRDE)

$$
\begin{aligned}
\mathrm{CDE}(B) &= \mathrm{E}(C|F_B = B, F_A = 2) - \mathrm{E}(C|F_B = B, F_A = 1) \\
&= \mathrm{E}_U\{\mathrm{E}(C|B, A = 2, U) - \mathrm{E}(C|B, A = 1, U)\} \\
&= \mathrm{E}_U(\eta^*_{B2} - \eta^*_{B1}),
\end{aligned}
$$

$$
\begin{aligned}
\mathrm{RRDE}(a^*) &= \mathrm{E}(C|F_B = d_{a^*}, F_A = 2) - \mathrm{E}(C|F_B = d_{a^*}, F_A = 1) \\
&= \mathrm{E}_U[\mathrm{E}_B\{\mathrm{E}(C|B, A = 2, U) - \mathrm{E}(C|B, A = 1, U)|A = a^*, U\}] \\
&= \mathrm{E}_U\{(\eta^*_{12} - \eta^*_{11})\delta^*_{a^*} + (\eta^*_{12} - \eta^*_{11})(1 - \delta^*_{a^*})\}.
\end{aligned}
$$

The RRDE is called the NDE in Didelez et al. (2006) but Robins and Richardson (2010) argue that the parameter being referred to as NDE in Didelez et al. (2006) is not the same as the NDE in Pearl (2001). Thus a separate name is given here to RRDE. By considering the mapping of $\vec{\tau}^*$ to the vector with $\vec{v}^*$ and the extra component $\eta^*_{B2} - \eta^*_{B1}$, the bounds on $\mathrm{CDE}(B)$ of Cai et al. (2008) can be replicated. Similarly by mapping $\vec{\tau}^*$ to a vector with $\vec{v}^*$ and $\mathrm{E}_B(\eta^*_{B2} - \eta^*_{B1}|A = a^*, U)$, bounds on $\mathrm{RRDE}(a^*)$ are obtained, which are identical to the bounds on $\mathrm{NDE}(a^*)$ in Sjölander (2009). Unlike here, both references use counterfactuals and use the definition of CDE and NDE, sometimes called pure direct effect (Robins and Greenland, 1992), given in Pearl (2001).

8. Monotonicity Assumption in the Instrumental Variable Model

The monotonicity assumption in the literature (Imbens and Angrist, 1994; Angrist et al., 1996) is formulated in a deterministic model. In a partial compliance study, a patient may be more likely to take treatment under assignment but it may not be reasonable to assume that their behaviour is deterministically related to treatment assignment. A monotonicity assumption in a weaker probabilistic model is considered here and can be expressed mathematically for the binary IV model by $\delta^*_2 \geq \delta^*_1$, from Equation (2). It restricts the space of the vector of probabilities so the constraints are at least as strong as without it.

The IV model considered in this section includes the exclusion restriction and the randomization assumption, as in the augmented DAG in Figure 2 (right). As an illustrative example, consider the computation of falsifiable constraints and causal bounds on φ given $\vec{\gamma}$, without monotonicity, where $\varphi = \mathrm{ACE}(A \to B) = \theta_{01} - \theta_{02}$. This computation produces only trivial results. Under monotonicity, $\mathcal{T}$ must be redefined to omit all $\vec{\tau}^*$ which do not satisfy it. Therefore all of the vertices with $\delta^*_2 < \delta^*_1$ or $\varphi^* \geq 0$ should be removed to redefine $\hat{\mathcal{T}}$, where $\varphi^* = P(B = 1|A = 2, U) - P(B = 1|A = 1, U)$. The required mapping is $\vec{\tau}^* \longmapsto (\vec{\gamma}^*, \varphi^*)$ over the domain of the restricted $\mathcal{T}$. The transformation is given in Figure 6 and the non-trivial constraints obtained are

$$
\max\left\{\begin{array}{c} \gamma_{01} - \gamma_{02} \\ -\gamma_{01} + \gamma_{02} \end{array}\right\} \leq \varphi \leq 1 \quad \Longleftrightarrow \quad |\gamma_{01} - \gamma_{02}| \leq \varphi \leq 1, \tag{7}
$$

or $\mathrm{ACE}(A \to B) \geq |\gamma_{01} - \gamma_{02}| = |\mathrm{ACE}(A \to C)|$. This makes sense since it is assumed that $\varphi^* \geq 0$ and B lies on the causal pathway from A to C. The transformed polytope in Figure 6 is 3 dimensional

η_0^*	η_1^*	δ_1^*	δ_2^*		γ_{01}^*	γ_{11}^*	γ_{02}^*	γ_{12}^*	φ^*
0	0	0	0		1	0	1	0	0
0	0	0	1		1	0	1	0	1
0	0	1	0		–	–	–	–	–
0	0	1	1		1	0	1	0	0
0	1	0	0		1	0	1	0	0
0	1	0	1		1	0	0	1	1
0	1	1	0		–	–	–	–	–
0	1	1	1	$\rightarrow$	0	1	0	1	0
1	0	0	0		0	1	0	1	0
1	0	0	1		0	1	1	0	1
1	0	1	0		–	–	–	–	–
1	0	1	1		1	0	1	0	0
1	1	0	0		0	1	0	1	0
1	1	0	1		0	1	0	1	1
1	1	1	0		–	–	–	–	–
1	1	1	1		0	1	0	1	0

Figure 6: Transformation to the extreme vertices corresponding to the polytope which represents the IV model, with the exclusion restriction, randomization and monotonicity assumptions, in terms of the distribution $P(C\,|\,A)$. The dashes correspond to points ruled out by monotonicity, which are represented by ∘'s in Figure 7.

since $\gamma_{11}^* = 1 - \gamma_{01}^*$ and $\gamma_{12}^* = 1 - \gamma_{02}^*$; its projection in $(\gamma_{01}^*, \gamma_{02}^*, \varphi^*)$ space is given in Figure 7. The 6 •'s are the vertices which are not removed after assuming monotonicity and the 6 ∘'s, which correspond to dashes in Figure 6, are the vertices which are removed. Figure 7 clearly demonstrates that the constraints without monotonicity are trivial whereas those with it are not.

To determine the effect of the monotonicity assumption on the constraints and bounds with $(\vec{\gamma}, \vec{\theta})$ in §5, the same mapping is used as in the derivation of the bivariate bounds on α but applied to the restricted $\mathcal{T}$ and $\hat{\mathcal{T}}$ formed by removing the appropriate vertices. The falsifiable constraints obtained are $\theta_{01} - \theta_{02} \geq |\gamma_{01} - \gamma_{02}|$ (equivalent to Equation (7)) and the causal bounds

$$\text{ACE}(B \to C) \geq \max \left\{ \begin{array}{c} 2\gamma_{01} - \gamma_{02} + \theta_{01} - 2 \\ \gamma_{01} - 2\gamma_{02} - \theta_{02} \\ \gamma_{01} + \theta_{01} - 2 \\ -\gamma_{02} - \theta_{02} \\ \gamma_{01} - \gamma_{02} + \theta_{01} - \theta_{02} - 1 \end{array} \right\},$$

$$\text{ACE}(B \to C) \leq \min \left\{ \begin{array}{c} 2\gamma_{01} - \gamma_{02} - \theta_{01} + 1 \\ \gamma_{01} - 2\gamma_{02} + \theta_{02} + 1 \\ \gamma_{01} - \theta_{01} + 1 \\ -\gamma_{02} + \theta_{02} + 1 \\ \gamma_{01} - \gamma_{02} - \theta_{01} + \theta_{02} + 1 \end{array} \right\}.$$

By considering the analogous mapping for $\vec{v}^*$, as defined in §2, the falsifiable constraints of Balke and Pearl (1997) and the causal bounds of Balke and Pearl (1993) for the IV model with mono-
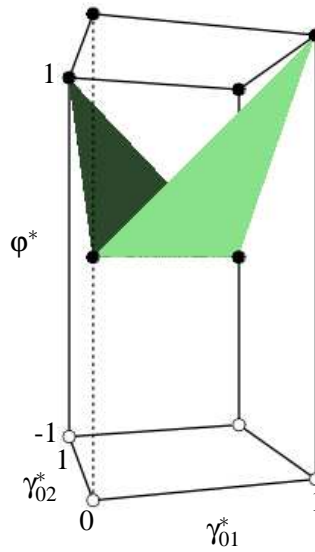
Figure 7: The extreme vertices which satisfy monotonicity are the 6 •'s, the 6 ∘'s are those which do not. The convex hull of the transformed polytope for the IV model with the monotonicity assumption is the region above the shaded surface and without the monotonicity assumption is the entire cuboid.

tonicity are recovered. The bounds correspond to results in Robins (1989) and Manski (1990). Thus the IV model with monotonicity, introduced in this section, is empirically and computationally indistinguishable from the IV model with 'no defiers' considered in Example 1 and Angrist et al. (1996).

9. Data Analysis

The relative frequencies for two data sets are given in Table 1 and described below.

9.1 Lipid Research Clinics Coronary Data

Consider the Lipid Research Coronary Primary Prevention Trial (Lipid Research Clinic Program, 1984), which was analysed by Efron and Feldman (1991) and Balke and Pearl (1997). Subjects were randomized into two groups, 172 men were given the placebo and 165 were given the treatment, and the subjects' cholesterol levels were measured. There was partial compliance with the treatment assigned.

9.2 Vitamin A Supplementation

Another example of partial compliance is the study of Vitamin A supplementation in northern Sumatra, described by Sommer and Zeger (1991). The study consisted of children in 450 villages, 11588 children (221 villages) were assigned to the control group and 12094 (229 villages) to the treatment group.

Data Set	a	$\hat{\zeta}_{00.a}$	$\hat{\zeta}_{01.a}$	$\hat{\zeta}_{10.a}$	$\hat{\zeta}_{11.a}$
Lipid Research	1	0.919	0	0.081	0
Clinic Program	2	0.315	0.139	0.073	0.473
Vitamin A	1	0.0064	0	0.9936	0
Supplementation	2	0.0028	0.0010	0.1972	0.7990

Table 1: Relative frequencies derived from the data sets in Lipid Research Clinic Program (1984) and Sommer and Zeger (1991).

In these trials, the relative frequencies are the maximum likelihood estimates of the parameters $\vec{v}$. Bounds on ACE$(B \to C)$ are computed under various assumptions from the data in Table 1 and are given in Table 2. Sampling uncertainty is ignored here but can be properly considered using techniques described in Ramsahai and Lauritzen (2011).

Study	Assumptions	Lower bound	Upper bound
Lipid	IV model	0.392	0.780
Research	IV, no randomization	-0.145	0.855
Clinic	IV, partial exclusion restriction ($\varepsilon = 0.5$)	0.050	0.855
Program	IV, monotonicity	0.392	0.780
	IV model	-0.1946	0.0054
Vitamin A	IV, no randomization	-0.587	0.413
Supplement.	IV, partial exclusion restriction ($\varepsilon = 0.5$)	-0.392	0.212
	IV, monotonicity	-0.1946	0.0054

Table 2: Causal bounds on ACE$(B \to C)$ computed from Lipid Research Clinic Program (1984) and Vitamin A Supplementation Study under various assumptions.

From Table 2, the imposition of the monotonicity assumption has no effect and is unnecessary for these data sets. However the randomized treatment assignment is important since the bounds computed without randomization are very wide and not much can be inferred about the causal effect. Even though the bounds are much wider for the Lipid Research Clinic Program (1984) data, under the partial exclusion restriction with $\varepsilon = 0.5$, it can still be deduced that there is a positive causal effect.

10. Discussion

The methods given here are applied while relaxing various assumptions that are often used in the deterministic counterfactual IV model. By removing the assumption that there are latent deterministic mechanisms, it is shown that the same bounds and constraints are obtained and that the models are empirically equivalent §3. The results for models which relax the randomization and exclusion restriction assumptions are valuable for sensitivity analyses. They are also useful for applications in which some of the assumptions in the IV model are known to be false.

In §7, the constraints and bounds were computed for the IV model with a partial exclusion restriction for $\varepsilon = 0.5$. It is not obvious how the methods described in this paper can be extended to

compute bounds and constraints as a function of ε but that would be worthy of future investigation. Such results would show how the bounds vary with ε and whether the data places any restrictions on the possible values of ε.

The ideas discussed can be extended to other models involving conditional independence since it is the factorization of the probability distribution which determines the algebraic structure of the polytope representing the model. The model must satisfy the condition that the observable distributions lie in the convex hull of the latent distribution. The vector of parameters $P(X)$ always lies in the convex hull of $P(X|U)$ but there is no guarantee that the factorisation of $P(X|U)$ produces any non-trivial constraints, where X and U are collections of observed and unobserved variables respectively. However there may be non-trivial constraints on conditional probabilities derived from $P(X)$.

Acknowledgments

Appendix A. Notation

The symbols used throughout are listed below.

$$
\begin{aligned}
\zeta^*_{cb.a} &= P(C=c, B=b \,|\, A=a, U),\\
\eta^*_b &= P(C=1 \,|\, B=b, U),\\
\delta^*_a &= P(B=1 \,|\, A=a, U),\\
\alpha &= P(C=1 \,||\, B=1) - P(C=1 \,||\, B=0),\\
\alpha^* &= P(C=1 \,|\, B=1, U) - P(C=1 \,|\, B=0, U),\\
\gamma^*_{ca} &= P(C=c \,|\, A=a, U),\\
\theta^*_{ba} &= P(B=b \,|\, A=a, U).
\end{aligned}
$$

The symbols with * are functions of U and the corresponding symbols without * are the marginals over U.

Appendix B. Equivalence of Convex Hulls

Theorem 1 $\mathcal{H} = \hat{\mathcal{H}}$.

Proof Following Dawid (2003), since $\hat{\mathcal{V}} \subseteq \mathcal{V}$ and $\hat{\mathcal{H}}$ is the minimal convex set containing $\hat{\mathcal{V}}$ then $\hat{\mathcal{H}} \subseteq \mathcal{H}$.

Let $m(\vec{v}^*)$ be an affine function, that is a linear function plus a constant, of $\vec{v}^*$, which returns a scalar, for $\vec{v}^* \in \mathcal{V}$. A closed half space in $[0,1]^8$ that contains $\vec{v}$ is defined by an affine function inequality $m(\vec{v}^*) \geq 0$ or $m\{\Xi(\vec{\tau}^*)\} \geq 0$ for $\vec{\tau}^* \in \mathcal{T}$. From Equations (3) and (2), $m\{\Xi(\vec{\tau}^*)\}$ is a monotonic function of any single component of $\vec{\tau}^*$ when the other three are fixed. Therefore the minimum of $m\{\Xi(\vec{\tau}^*)\}$ over $\mathcal{T}$ is attained for some $\vec{\tau}^* \in \hat{\mathcal{T}}$. Therefore

$$m\{\Xi(\vec{\tau}^*)\} \geq 0 \text{ for all } \vec{\tau}^* \in \hat{\mathcal{T}} \Rightarrow m\{\Xi(\vec{\tau}^*)\} \geq 0 \text{ for all } \vec{\tau}^* \in \mathcal{T}.$$

This means that any half space containing $\hat{\mathcal{V}}$ also contains $\mathcal{V}$. Since $\hat{\mathcal{H}}$ is the intersection of all half spaces containing $\hat{\mathcal{V}}$ then $\mathcal{V} \subseteq \hat{\mathcal{H}}$. Since $\hat{\mathcal{H}}$ is convex and $\mathcal{H}$ is the minimal convex set containing $\mathcal{V}$ then $\mathcal{H} \subseteq \hat{\mathcal{H}}$. ∎

Appendix C. Causal Bounds for Binary Instrumental Variable Model

For the binary IV model of Figure 2 (right), bounds on α in terms of $\zeta_{cb.a}$ are

$$
\alpha \geq \max \left\{
\begin{array}{c}
\zeta_{00.1} + \zeta_{11.2} - 1 \\
\zeta_{11.1} + \zeta_{00.2} - 1 \\
-\zeta_{01.1} - \zeta_{10.1} + \zeta_{11.1} - \zeta_{10.2} - \zeta_{11.2} \\
-\zeta_{10.1} - \zeta_{11.1} - \zeta_{01.2} - \zeta_{10.2} + \zeta_{11.2} \\
-\zeta_{01.1} - \zeta_{10.1} \\
-\zeta_{01.2} - \zeta_{10.2} \\
-\zeta_{00.1} - \zeta_{01.1} + \zeta_{00.2} - \zeta_{01.2} - \zeta_{10.2} \\
\zeta_{00.1} - \zeta_{01.1} - \zeta_{10.1} - \zeta_{00.2} - \zeta_{01.2}
\end{array}
\right\},
$$

and

$$
\alpha \leq \min \left\{
\begin{array}{c}
1 - \zeta_{10.1} - \zeta_{01.2} \\
1 - \zeta_{01.1} - \zeta_{10.2} \\
\zeta_{00.1} - \zeta_{01.1} + \zeta_{11.1} + \zeta_{00.2} + \zeta_{01.2} \\
\zeta_{00.1} + \zeta_{01.1} - \zeta_{01.2} + \zeta_{00.2} + \zeta_{11.2} \\
\zeta_{00.1} + \zeta_{11.1} \\
\zeta_{00.2} + \zeta_{11.2} \\
\zeta_{10.1} + \zeta_{11.1} + \zeta_{00.2} + \zeta_{11.2} - \zeta_{10.2} \\
\zeta_{00.1} - \zeta_{10.1} + \zeta_{11.1} + \zeta_{10.2} + \zeta_{11.2}
\end{array}
\right\}.
$$

For the binary IV model of Figure 2 (right), bounds on α in terms of γ_{ca} and θ_{ba} are

$$
\alpha \geq \max \left\{
\begin{array}{c}
2\gamma_{01} - \gamma_{02} + 2\theta_{01} - 3 \\
\gamma_{01} + \theta_{01} - 2 \\
\gamma_{02} + \theta_{02} - 2 \\
-\gamma_{01} + 2\gamma_{02} + 2\theta_{02} - 3 \\
-\gamma_{01} + \gamma_{02} - \theta_{01} + \theta_{02} - 1 \\
-\gamma_{01} - \theta_{01} \\
-\gamma_{02} - \theta_{02} \\
\gamma_{01} - 2\gamma_{02} - 2\theta_{02} \\
-2\gamma_{01} + \gamma_{02} - 2\theta_{01} \\
\gamma_{01} - \gamma_{02} + \theta_{01} - \theta_{02} - 1
\end{array}
\right\},
$$

and

$$\alpha \leq \min \left\{ \begin{array}{l} -2\gamma_{01} + \gamma_{02} + 2\theta_{01} + 1 \\ \gamma_{01} - 2\gamma_{02} + 2\theta_{02} + 1 \\ 2\gamma_{01} - \gamma_{02} - 2\theta_{01} + 2 \\ -\gamma_{01} + 2\gamma_{02} - 2\theta_{02} + 2 \\ \gamma_{01} - \gamma_{02} - \theta_{01} + \theta_{02} + 1 \\ -\gamma_{02} + \theta_{02} + 1 \\ \gamma_{01} - \theta_{01} + 1 \\ \gamma_{02} - \theta_{02} + 1 \\ -\gamma_{01} + \theta_{01} + 1 \\ -\gamma_{01} + \gamma_{02} + \theta_{01} - \theta_{02} + 1 \end{array} \right\}.$$

Appendix D. Causal Bounds for Instrumental Variable Model Without Exclusion Restriction

For the binary IV model without the exclusion restriction in §7, for $\varepsilon = 0.5$, the following bounds are obtained for $a = 1, 2$

$$2\{\mathrm{E}(C|A=a, F_B=1) - \mathrm{E}(C|A=a, F_B=0)\} \geq \max \left\{ \begin{array}{l} -2\zeta_{01.a} - 2\zeta_{10.a} \\ -\zeta_{01.a} - 2\zeta_{10.a} + \zeta_{11.a} + \zeta_{00.a'} - \zeta_{10.a'} - 1 \\ 2\zeta_{00.a} - 2 - \zeta_{01.a'} + \zeta_{11.a'} \\ -3\zeta_{01.a} - 2\zeta_{10.a} - \zeta_{11.a} - \zeta_{01.a'} + \zeta_{11.a'} \\ -3\zeta_{01.a} - 2\zeta_{10.a} + \zeta_{11.a} - 2\zeta_{10.a'} - 2\zeta_{11.a'} \\ -\zeta_{01.a} - 2\zeta_{10.a} - 3\zeta_{11.a} - 3\zeta_{01.a'} - 2\zeta_{10.a'} + \zeta_{11.a'} \\ 2\zeta_{11.a} + \zeta_{00.a'} - \zeta_{10.a'} - 2 \\ -2\zeta_{01.a'} - 2\zeta_{10.a'} - 1 \\ -3\zeta_{01.a} - 4\zeta_{10.a} - \zeta_{11.a} + 2\zeta_{10.a'} + 2\zeta_{11.a'} - 1 \\ \zeta_{01.a} + 2\zeta_{10.a} + 3\zeta_{11.a} - 3\zeta_{01.a'} - 4\zeta_{10.a'} - \zeta_{11.a'} - 2 \end{array} \right\},$$

and

$$2\{\mathrm{E}(C|A=a, F_B=1) - \mathrm{E}(C|A=a, F_B=0)\} \leq \min \left\{ \begin{array}{l} 2\zeta_{00.a} + 2\zeta_{11.a} \\ 2 - 2\zeta_{10.a} - \zeta_{01.a'} + \zeta_{11.a'} \\ 2 - \zeta_{01.a} - 2\zeta_{10.a} + \zeta_{11.a} - \zeta_{01.a'} + \zeta_{11.a'} \\ 1 + 2\zeta_{00.a'} + 2\zeta_{11.a'} \\ 2 - 2\zeta_{01.a} + \zeta_{00.a'} - \zeta_{10.a'} \\ 3\zeta_{00.a} - \zeta_{10.a} + 2\zeta_{11.a} + 2\zeta_{10.a'} + 2\zeta_{11.a'} \\ 3 - 3\zeta_{01.a} - 2\zeta_{10.a} - \zeta_{11.a} + \zeta_{00.a'} - \zeta_{10.a'} \\ 4 - 3\zeta_{01.a} - 2\zeta_{10.a} + \zeta_{11.a} - 2\zeta_{10.a'} - 2\zeta_{11.a'} \\ 4 + \zeta_{01.a} - 2\zeta_{10.a} - \zeta_{11.a} - 3\zeta_{01.a'} - 2\zeta_{10.a'} + \zeta_{11.a'} \\ 4 - \zeta_{01.a} + 2\zeta_{10.a} + \zeta_{11.a} - 3\zeta_{01.a'} - 4\zeta_{10.a'} - \zeta_{11.a'} \end{array} \right\}.$$

where $a' = 2$ if $a = 1$ and $a' = 1$ if $a = 2$.

References

O. O. Aalen and A. Frigessi. What can statistics contribute to a causal understanding? *Scandinavian Journal of Statistics*, 34(1):155–168, 2007.

J. D. Angrist, G. W. Imbens and D. B. Rubin. Identification of causal effects using instrumental variables. *J. Am. Statist. Assoc.*, 91(434):444–455, 1996.

A. Balke. Probabilistic counterfactuals: semantics, computation and applications. PhD Dissertation, University of California, Los Angeles, 1995.

A. Balke and J. Pearl. Non-parametric bounds on causal effects from partial compliance data. Technical Report R-199, Computer Science Department, University of California, Los Angeles, 1993.

A. Balke and J. Pearl. Bounds on treatment effects from studies with imperfect compliance. *Journal of the American Statistical Association*, 92(439):1171–1176, 1997.

Z. Cai, M. Kuroki, J. Pearl and J. Tian. Bounds on direct effects in the presence of confounded intermediate variables. *Biometrics*, 64(3):695–701, 2008.

T. Christof and A. Loebel. PORTA. Available online at URL: http://www.iwr.uni-heidelberg.de/groups/comopt/software/PORTA/, 1998.

A. P. Dawid. Conditional independence in statistical theory (with discussion). *Journal of the Royal Statisitical Society*, Ser. B, 41(1):1–31, 1979.

A. P. Dawid. Causal inference without counterfactuals (with discussion). *Journal of the American Statistical Association*, 95(450):407–448, 2000.

A. P. Dawid. Influence diagrams for causal modelling and inference. *International Statistical Review*, 70(2):161–189, 2002.

A. P. Dawid. Causal inference using influence diagrams: the problem of partial compliance. In P. J. Green, N. L. Hjort and S. Richardson, editors, *Highly Structured Stochastic Systems*, Oxford University Press, New York, 2003.

A. P. Dawid and V. Didelez. Identifying the consequences of dynamic treatment strategies: a decision theoretic overview. *Statistics Surveys*, 4:184–231, 2010.

V. Didelez, A. P. Dawid and S. Geneletti. Direct and indirect effects of sequential treatments. In *Proceedings of the Twenty Second Annual Conference on Uncertainty in Artificial Intelligence*, pages 138–146, AUAI Press, Arlington, Virginia, 2006.

V. Didelez and N. Sheehan. Mendelian randomisation as an instrumental variable approach to causal inference. *Statistical Methods in Medical Research*, 16(4):309–330, 2007.

J. Durbin. Errors in variables. *Review of the International Statistical Institute*, 22(1):23–32, 1954.

B. Efron and D. Feldman. Compliance as an explanatory variable in clinical trials. *Journal of the American Statistical Association*, 86(413):9–26, 1991.

E. Gawrilow and M. Joswig. Polymake: a framework for analyzing convex polytopes. In G. Kalai and G. M. Ziegler, editors, *Polytopes - Combinatorics and Computation*, Birkhäuser, Basel, 2000. Available online at URL: http://wwwopt.mathematik.tu-darmstadt.de/polymake/doku.php.

D. Geiger and C. Meek. Graphical models and exponential families. In *Proceedings of the Fourteenth Annual Conference on Uncertainty in Artificial Intelligence*, pages 156–165, Morgan Kaufmann, San Francisco, California, 1998.

S. Geneletti. Identifying direct and indirect effects in a non-counterfactual framework. *Journal of the Royal Statistical Society*, Ser. B, 69(2):199–215, 2007.

M. Goldszmidt and J. Pearl. Rank based systems: a simple approach to belief revision, belief update, and reasoning about evidence and actions. In *Proceedings of the Third International Conference on Knowledge Representation and Reasoning*, Morgan Kaufmann, San Mateo, California, 1992.

D. Heckerman and R. Shachter. Decision theoretic foundations for causal reasoning. *Journal of Artificial Intelligence Research*, 3:405–430, 1995.

K. Hirano, G. W. Imbens, D. B. Rubin and X.-H. Zhou. Assessing the effect of an influenza vaccine in an encouragement design. *Biostatistics*, 1(1):69–88, 2000.

G. W. Imbens and J. D. Angrist. Identification and estimation of local average treatment effects. *Econometrica*, 62(2):467–475, 1994.

C. Kang and J. Tian. Inequality constraints in causal models with hidden variables. In *Proceedings of the Twenty Second Annual Conference on Uncertainty in Artificial Intelligence*, pages 233–240, AUAI Press, Arlington, Virginia, 2006.

S. Kaufman, J. S. Kaufman and R. F. MacLehose. Analytic bounds on causal risk difference in directed acyclic graphs involving three observed binary variables. *Journal of Statistical Planning and Inference*, 139(10):3473–3487, 2009.

S. L. Lauritzen. *Graphical models*. Oxford University Press, Clarendon, Oxford, UK, 1996.

S. L. Lauritzen. *Causal inference from graphical models*. In O. E. Barndorff-Nielsen, D. R. Cox and C. Klüppelberg, editors, *Complex Stochastic Systems*, CRC Press, London, 2001.

S. L. Lauritzen. Discussion on causality. *Scandinavian Journal of Statistics*, 31(2):189–201, 2004.

S. L. Lauritzen, A. P. Dawid, B. N. Larsen and H. G. Leimer. Independence properties of directed Markov fields. *Networks*, 20(5):491–505, 1990.

Lipid Research Clinic Program. The lipid research clinics coronary primary prevention trial results, part I and II. *Journal of the American Medical Association*, 251(3):351–374, 1984.

C. F. Manski. Non-parametric bounds on treatment effects. *American Economic Review, Papers and Proceedings*, 80(2):319–323, 1990.

J. Pearl. Comment: graphical models, causality and interventions. *Statistical Science*, 8(3):266–269, 1993.

J. Pearl. Causal inference from indirect experiments. *Artificial Intelligence in Medicine*, 7(6):561–582, 1995.

J. Pearl. Direct and indirect effects. In *Proceedings of the Seventeenth Annual Conference on Uncertainty in Artificial Intelligence*, pages 411–420, Morgan Kaufmann, San Francisco, California, 2001.

R. R. Ramsahai. Causal bounds and instruments. In *Proceedings of the Twenty Third Annual Conference on Uncertainty in Artificial Intelligence*, pages 310–317, AUAI Press, Corvallis, Oregon, 2007.

R. R. Ramsahai and S. L. Lauritzen. Likelihood analysis of the binary instrumental variable model. *Biometrika*, 98(4):987–994, 2011.

T. S. Richardson and J. M. Robins. Analysis of the binary instrumental variable model. In R. Dechter, H. Geffner and J. Y. Halpern, editors, *Heuristics, probability and causality: a tribute to Judea Pearl*, College Publications, UK, 2010.

J. M. Robins. The analysis of randomized and non-randomized AIDS treatment trials using a new approach to causal inference in logitudinal studies. In L. Sechrest, H. Freeman and A. Mulley, editors, *Health Service Research Methodology: a Focus on AIDS*, U.S. Public Health Service, Washington D.C., 1989.

J.M. Robins and S. Greenland. Identifiability and exchangeability for direct and indirect effects. *Epidemiology*, 3(2):143–155, 1992.

J. M. Robins and T. S. Richardson. Alternative graphical causal models and the identification of direct effects. In P. Shrout, editor, *Causality and Psychopathology: Finding the Determinants of Disorders and Their Cures*, Oxford University Press, 2010.

D. B. Rubin. Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of Educational Psychology*, 66(5):688–701, 1974.

B. E. Shepherd, P. B. Gilbert, Y. Jemiai and A. Rotnitzky. Sensitivity analyses comparing outcomes only existing in a subset selected post randomization, conditional on covariates, with application to HIV vaccine trial. *Biometrics*, 62(2):332–342, 2006.

A. Sjölander. Bounds on natural direct effects in the presence of confounded intermediate variables. *Statistics in Medicine*, 28(4):558–571, 2009.

A. Sommer and S. L. Zeger. On estimating efficacy from clinical trials. *Statistics in Medicine*, 10(1):45–52, 1991.

P. Spirtes, C. Glymour and R. Scheines. *Causation, Prediction and Search*. Springer-Verlag, New York, 1993.

R. H. Strotz and H. O. A. Wold. Recursive vs non-recursive systems: an attempt at synthesis (part I of a triptych on causal chain systems). *Econometrica*, 28(2):417–427, 1960.

T. Verma and J. Pearl. Causal networks: semantics and expressiveness. In *Proceedings of the Fourth Workshop on Uncertainty in Artificial Intelligence*, pages 352-359, Elsevier Science, New York, 1988.