Department of Economics, University College London

Information and admissible sets

Jeff Rowley

1^{st} July, 2014

Abstract

«Abstract here»

Acknowledgements. *«Acknowledgements here» I acknowledge financial support from the Economic and Social Research Council (ESRC).*

Notation

There is a probability space $(\Omega, \Sigma, \mathbb{P})$ on which are defined random variables (Y, D, X, Z, U). Here, (Y, D, X, Z) are observable with supports $(\mathcal{R}_Y, \mathcal{R}_D, \mathcal{R}_X, \mathcal{R}_Z)$, and U is unobservable with as yet unspecified support. I allow (X, Z, U) to be vectors, in which case the support is given by the Cartesian product of the supports of each element in the vector. I refer to Y as the dependent variable, to D as the endogenous variable, to X as the exogenous variable, to Z as the instrumental variable, and to U as unobservable heterogeneity. The logic of this naming convention will be made clear by the restrictions that are imposed upon these random variables in the main text. Lower case letters are used to represent specific values of these random variables.

I denote by $Y(d)$ the counterfactual value of Y when D is externally fixed, and by $D(z)$ the counterfactual value of D when Z is externally fixed. I denote by $\mathbb{E}$ the expectation operator, and by $\mathbb{1}$ the indicator function. Related to these concepts are the average causal effects $ACE(D \to Y)$ and $ACE(Z \to D)$ that are defined as $\mathbb{E}[Y(d_1) - Y(d_0)]$ and $\mathbb{E}[D(z_1) - D(z_0)]$ when (D, Z), respectively. To distinguish between population and sample quantities, I subscript sample quantities by n.

1 Introduction

I explore the effect of incorporating information on the identified set of values for a non-parametric model that permits non-random selection by agents into treatment groups, and that embeds an exclusion restriction and an independence restriction that characterise an instrumental variable.

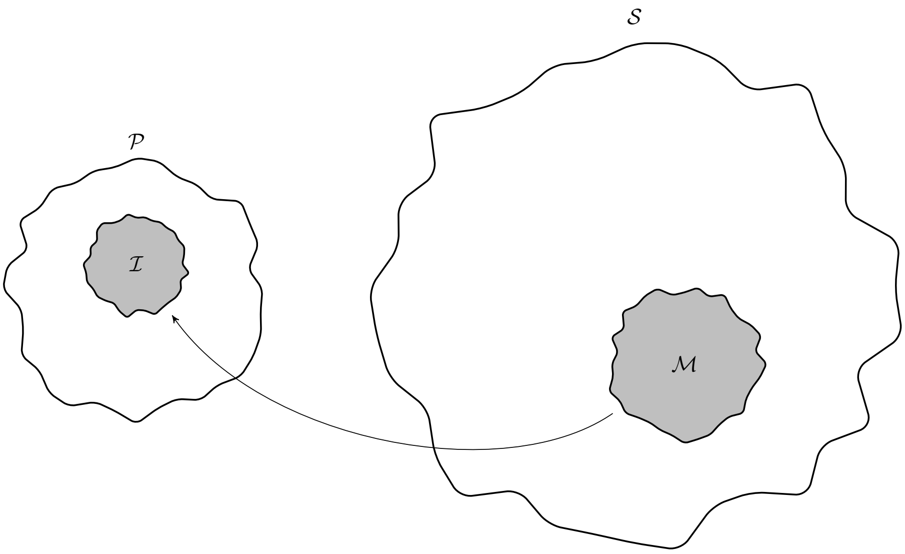
I consider the effect of combining many instrumental variables into a composite instrumental variable with many points of support on the identified set of values for counterfactual outcome distributions. Further, I consider the effect of enriching individual behaviour by allowing relevant exogenous variables to affect individual choice over treatment and outcomes. I establish the conditions under which the parameter of interest in the enriched model is equivalent to the parameter of interest in a model that does not explicitly account for the contribution of additional relevant exogenous variables beyond the endogenous variable of interest.

The model that I consider is partially identifying. That is, the restrictions on the set of admissible structures (or data generating processes) that are implied by the model are insufficient to exclude observationally equivalent structures; an identifying correspondence from a probability distribution to the set of structures that are admitted by the model is a one-to-many (or multivalued) function.

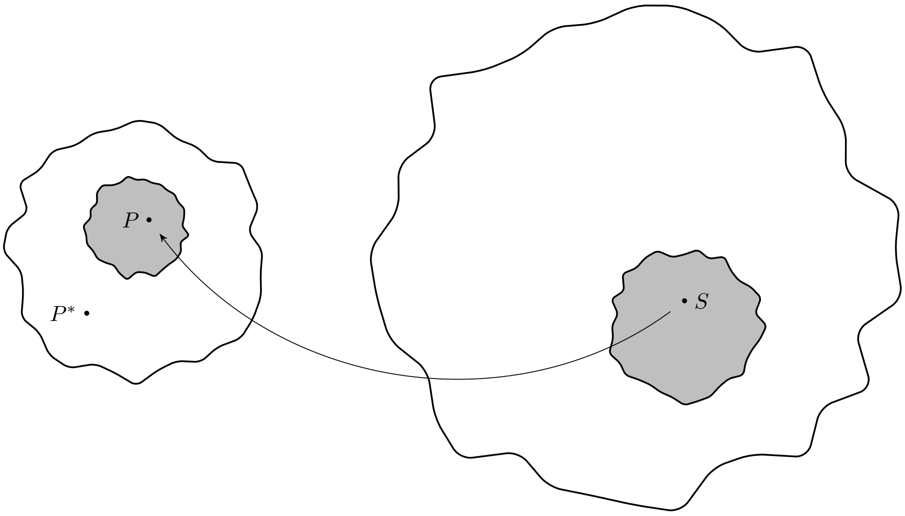
2 A threshold crossing model

I

Information and admissible sets

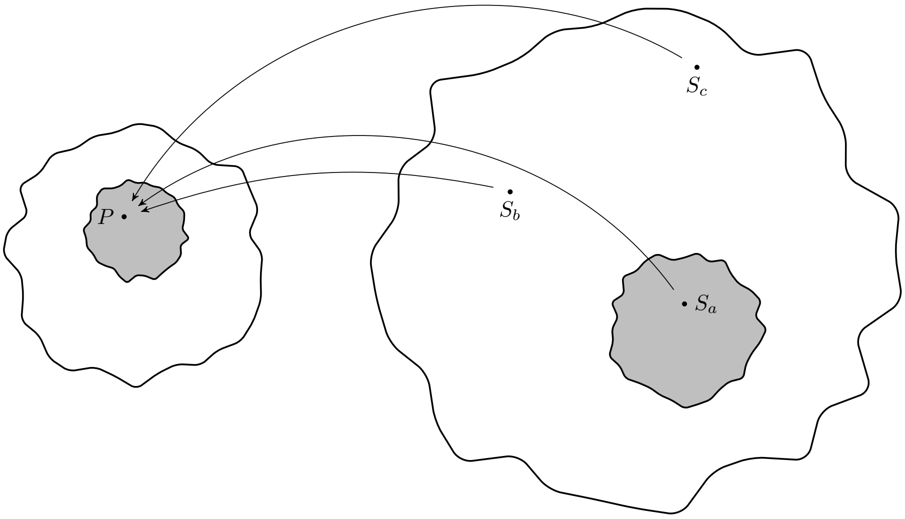


(a) A model $\mathcal{M}$ is a set of structures that forms a proper subset of the class of all structures $\mathcal{S}$. Each structure in $\mathcal{M}$ generates a probability distribution in the class of all probability distributions (of observable variables) $\mathcal{P}$. Then the image $\mathcal{I}$ is the set of all probability distributions that are generated by $\mathcal{M}$.

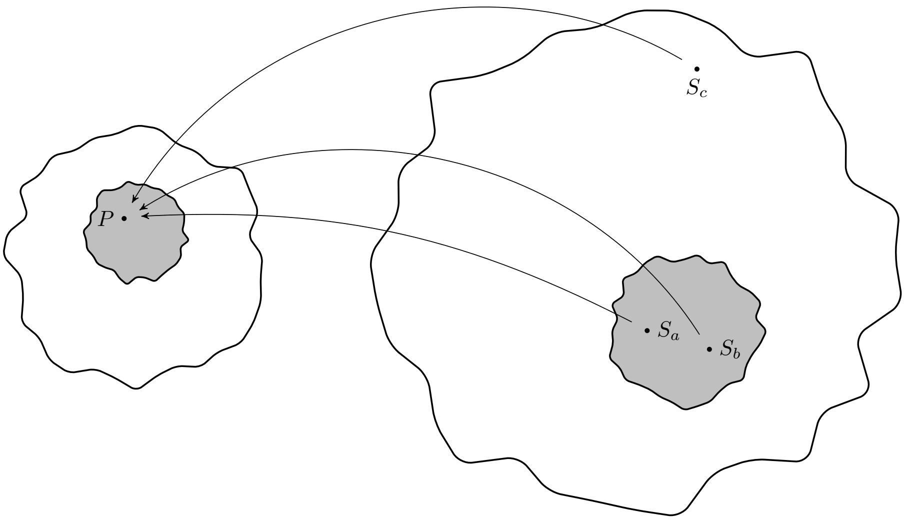


(b) A structure S is incompatible with data if it generates a probability distribution (of observable variables) P that is distinct from a realised probability distribution P^*. If all structures in $\mathcal{M}$ are incompatible with data then $\mathcal{M}$ is said to be observationally restrictive, and is falsified. This condition is equivalent to $P_a \in \mathcal{P} \setminus \mathcal{I}$.

Figure 1: Structures, models, probability distributions (of observable variables), and falsifiability.



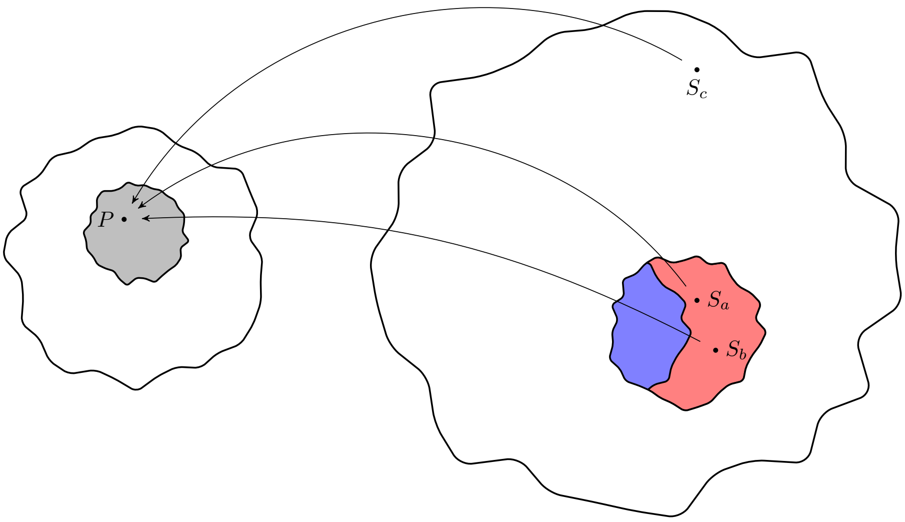
(a) A model $\mathcal{M}$ is said to identify a structure S if the probability distribution (of observable variables) P that is generated by S is distinct from those generated by other structures in $\mathcal{M}$. The structures S_a, S_b and S_c are said to be observationally equivalent as they all generate P but S_b and S_c are not admitted by $\mathcal{M}$. As S_a is the only structure that is admitted by $\mathcal{M}$ and that generates P, S_a is identified by $\mathcal{M}$. For completeness, $\mathcal{M}$ is said to be uniformly identifying if it identifies each structure that it admits.



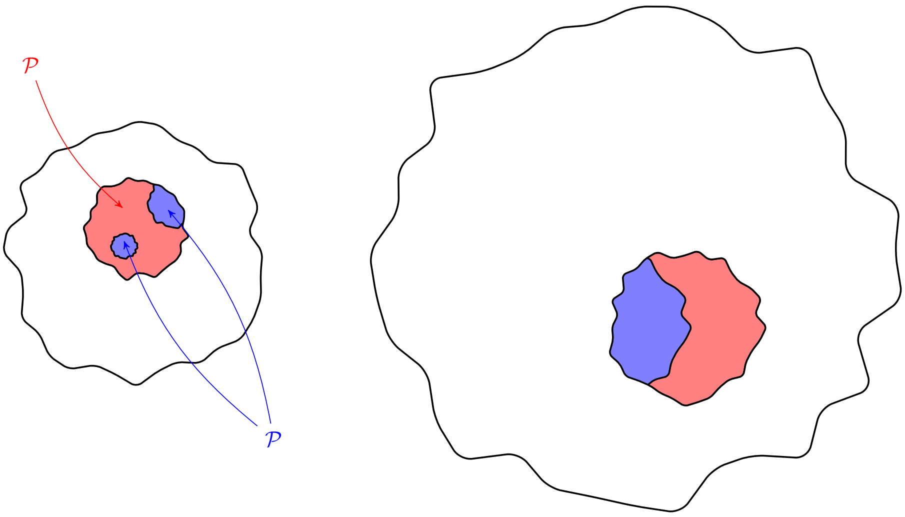
(b) As S_a and S_b are observationally equivalent and are both admitted by $\mathcal{M}$ then $\mathcal{M}$ does not identify either S_a or S_b. Nonetheless, as $\mathcal{M}$ restricts the set of observationally equivalent structures that generate P to S_a and S_b then $\mathcal{M}$ partially identifies S_a (and S_b to within $\{S_a, S_b\}$).

Figure 2: Identification and non-identification of a structure, and partial identification of a structure.

Information and admissible sets



(a) A structural characteristic Ψ is a function of a structure S. A model $\mathcal{M}$ can be partitioned such that structures in a partition deliver the same value for Ψ. Structures in the red partition $\mathcal{M}$ deliver the value a for Ψ, and structures in the blue partition $\mathcal{M}$ deliver the value b for Ψ. If Ψ is constant across all observationally equivalent structures that $\mathcal{M}$ admits then $\mathcal{M}$ is said to identify Ψ. As in Figure 2b S_a and S_b are not (separately) identified by $\mathcal{M}$ but Ψ is identified by $\mathcal{M}$ since $\Psi(S_a)$ is equal to $\Psi(S_b)$ (is equal to a).



(b) If $\mathcal{M}$ identifies Ψ for all structures in $\mathcal{M}$ then $\mathcal{M}$ is said to uniformly identify Ψ. Whether $\mathcal{M}$ uniformly identifies Ψ can be determined by the existence of an identifying correspondence G, a functional. The properties of G are given in the context of Figure 3b. Structures in $\mathcal{M}$ deliver the value a for Ψ, and structures in $\mathcal{M}$ deliver the value b for Ψ. The class of all probability distributions (of observable variables) is partitioned into the red partition $\mathcal{P}$ and into the blue partition $\mathcal{P}$. Probability distributions in $\mathcal{P}$ are generated by structures in $\mathcal{M}$, and probability distributions in $\mathcal{P}$ are generated by structures in $\mathcal{M}$. P is a probability distribution in $\mathcal{P}$, and P is a probability distribution in $\mathcal{P}$. Then $\mathcal{M}$ uniformly identifies Ψ if the value of $G(P)$ is a and if the the value of $G(P)$ is b, holding for any such P and P.

Figure 3: The identification of structural characteristics, and identifying correspondences.