Department of Economics, University College London

# Information and admissible sets

Jeff Rowley

27th June, 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 function.

## 2 A threshold crossing model

I

Information and admissible sets