Simultaneous equations models for discrete outcomes: coherence, completeness, and identification.*

Andrew Chesher and Adam M. Rosen
CeMMAP & UCL

August 21, 2012

Abstract

This paper studies simultaneous equations models for two or more discrete outcomes. These models may be incoherent, delivering no values of the outcomes at certain values of the latent variables and covariates, and they may be incomplete, delivering more than one value of the outcomes at certain values of the covariates and latent variates. We revisit previous approaches to the problems of incompleteness and incoherence in such models, and we propose a new approach for dealing with these. For each approach, we use random set theory to characterize sharp identification regions for the marginal distribution of latent variables and the structural function relating outcomes to covariates, illustrating the relative identifying power and tradeoffs of the different approaches. We show that these identified sets are characterized by systems of conditional moment equalities and inequalities, and we provide a generically applicable algorithm for constructing these. We demonstrate these results for the simultaneous equations model for binary outcomes studied in for example Heckman (1978) and Tamer (2003) and the triangular model with a discrete endogenous variable studied in Chesher (2005) and Jun, Pinkse, and Xu (2011) as illustrative examples.

Keywords: Discrete endogenous variables, Endogeneity, Incomplete models, Incoherent models Instrumental variables, Simultaneous equations models, Set Identification, Structural econometrics.

JEL Codes: C10, C14, C50, C51.

*This paper has benefitted from comments received at the March 2012 conference on recent developments in microeconometrics at Vanderbilt, an April 2012 seminar presented at Georgetown, the April 2012 UCL/CeMMAP conference on the use of random sets in economics, and the 2012 North American summer meetings of the Econometric Society at Northwestern. We gratefully acknowledge financial support from the UK Economic and Social Research Council through a grant (RES-589-28-0001) to the ESRC Centre for Microdata Methods and Practice (CeMMAP) and through the funding of the "Programme Evaluation for Policy Analysis" node of the UK National Centre for Research Methods, and from the European Research Council (ERC) grant ERC-2009-StG-240910-ROMETA.

1 Introduction

In this paper we consider the identifying power of simultaneous equations models in which endogenous variables, Y, are all discrete. The models involve covariates Z and latent variables U which are restricted to be stochastically independently distributed. The variables Y, Z and U, with support $\mathcal{Y}, \mathcal{Z}, \mathcal{U}$, respectively, may have any finite dimension.

The model imposes the structural relation

$$h(Y, Z, U) = 0, \tag{1.1}$$

which delivers a set of values of Y at each value (z, u) of (Z, U),

$$\mathcal{Y}(z, u; h) \equiv \{y : h(y, z, u) = 0\}.$$

The econometrician observes realizations (y, z) of the random vectors (Y, Z) such that $y \in \mathcal{Y}(z, u; h)$ when this set is nonempty When there are values of (z, u) for which $\mathcal{Y}(z, u; h)$ has more than one element, multiple outcomes are feasible, and the structural relation h is *incomplete*. When there are values of (z, u) for which $\mathcal{Y}(z, u; h)$ is empty, h is *incoherent*. For example, $h(y, z, u) = 0$ could be a set of conditions that characterize equilibrium behavior in a model of a game, and for certain (z, u) there could be multiple equilibria, a unique equilibrium, or no equilibrium.

Models that allow for $\mathcal{Y}(z, u; h)$ to be non-singleton are known to pose significant challenges for identification. The main challenge is the question of how the outcome Y is determined when the model is not guaranteed to deliver a unique y for every possible (z, u). We consider four distinct possibilities regarding the determination of Y in such cases, which in turn provide four distinct approaches for obtaining set identification of (h, G_U), where G_U is the marginal distribution of U. The first approach applies to settings where a null outcome is observed when the realization of (Z, U) gives no solution for Y. Although perhaps of limited practical applicability, this approach provides a useful backdrop for analysis of settings where null outcomes are never observed. The second approach, which is new to this paper, and the third approach both deal with multiple feasible outcomes (i.e. incompleteness) in the same way as the recent literature on set identification in models with multiple equilibria, but differ in the way they deal with the case where the set $\mathcal{Y}(Z, U; h)$ is empty.[1] For each approach, we use random set theory to obtain a sharp characterization of the identified set. Thus, in addition to introducing a new approach for dealing simultaneously with incoherence and incompleteness, a goal of our identification analysis is to illuminate the relationship between different approaches by analyzing each of them under the common framework of random set theory.

[1]Specifically, the second and third approaches deal with incompleteness by allowing any of the multiple solutions to obtain, as in e.g. Tamer (2003), Andrews, Berry, and Jia (2004), Aradillas-Lopez and Tamer (2008), Beresteanu, Molchanov, and Molinari (2011), Galichon and Henry (2011), and Kline and Tamer (2012)

Table 1: A summary of the determination of outcome Y for the approaches of section 3 when the model delivers no or multiple solutions.

	Restrictions for the determination of Y when:	
	$\mathcal{Y}(Z,U;h)=\emptyset$ (no solution)	$\#\mathcal{Y}(Z,U;h)\geq 2$ (multiple solutions)
Approach 1:	$Y=\phi$	$Y\in\mathcal{Y}(Z,U;h)$
Approach 2:	not observed	$Y\in\mathcal{Y}(Z,U;h)$
Approach 3:	$Y\in\mathcal{Y}$	$Y\in\mathcal{Y}(Z,U;h)$
Approach 4:	not observed	not observed

As in the recent literature, our general characterizations of identified sets across the different approaches take the form of conditional moment inequalities. We show that within the context of any particular model, some of these conditional moment inequalities can be combined to yield conditional moment *equalities*. Thus our characterizations are in fact collections of moment *equalities* and *inequalities*, and we provide an algorithm to collect the model's implied moment restrictions. The moment equalities correspond to well known probabilities of observable events in the simultaneous binary game studied by e.g. Bresnahan and Reiss (1991) and Tamer (2003), but our approach for detecting when inequalities may be strengthened to equalities applies more broadly. When applied to complete models, our characterization reduces to just a set of *moment equalities*, which coincide with the usual likelihood expressions if the model is fully parametric.

1.1 Approaches for Incompleteness and Incoherence

We now briefly discuss a variety of approaches to the problems of incompleteness and incoherence, beginning with the four that we formally investigate in Section 3, whose distinguishing features are summarized in Table 1 for convenience.

The first approach we consider assumes that if the set $\mathcal{Y}(z,u;h)$ is empty then *this event is recorded* as $Y=\phi$. So either a single value of Y is observed, or, *it is observed* that the process delivers no value of Y. This approach effectively treats these observations as data with censored outcomes. In this situation the probability distribution of Y given $Z=z$ has mass one on the *extended* support $\mathcal{Y}^*\equiv\mathcal{Y}\cup\{\phi\}$. While there are relatively few applications in which null outcomes are recorded, it is useful to study this case because it illuminates the issues posed by incoherence in the more commonly encountered situations in which null outcomes are not recorded.

The second approach is closely related, but like the rest of those we consider, does not require that null outcomes are recorded. Rather it takes the structural relation (1.1) as correct, and applies the logic that data where some outcome is always observed cannot have been generated by values of the unobservable that result in the feasible outcome set $\mathcal{Y}(Z,U;h)$ being empty, as then (1.1) would not hold. Following this logic, the distribution of observables is effectively treated as a truncated version of the distribution generated by G_U and the structural relation (1.1).

The third approach, previously considered by Beresteanu, Molchanov, and Molinari (2011), effectively takes an "anything goes" approach to emptiness of the feasible outcome set $\mathcal{Y}(Z, U; h)$.[2] Like the first two approaches, when $\mathcal{Y}(Z, U; h)$ is non-empty, $Y \in \mathcal{Y}(Z, U; h)$ is assumed, while if $\mathcal{Y}(Z, U; h)$ is empty all that is assumed is that $Y \in \mathcal{Y}$. The underlying reasoning is that when the model delivers an outcome set, or multiple equilibria, one of these occurs, but if the model delivers no feasible outcomes, i.e. no equilibria, then the model is completely silent as to how the outcome is determined and anything is possible.

The fourth approach was previously considered by Dagenais (1997) and Hajivassiliou (2008). It stipulates that the distribution of observables corresponds only to that subset of the sample space where the outcome is uniquely determined, i.e. where the set $\mathcal{Y}(Z, U; h)$ is singleton. Here the distribution of observables is given by the truncated distribution corresponding to this subset of the sample space. This restriction effectively renders the model both coherent and complete, and allows for construction of a likelihood function if the model is fully parametric.

In some earlier studies incompleteness and incoherence have been treated by restricting the space of admissible structural relations in a way that guarantees that the outcome set $\mathcal{Y}(Z, U; h)$ is unique. Like the approach of Dagenais (1997) and Hajivassiliou (2008), this effectively transforms the model into one that is both complete and coherent, and thus enables the use of maximum likelihood in fully parametric models, but instead achieves this aim through restrictions on h. Related conditions are imposed by Amemiya (1974) in a class of simultaneous Tobit models. Similarly, Gourieroux, Laffont, and Monfort (1980) and Blundell and Smith (1994) impose coherency conditions in their analysis of simultaneous equations regime-switching models and simultaneous equations models of qualitative or censored outcomes, respectively. For a more thorough review of this approach in a variety of models see e.g. Schmidt (1981) and Maddala (1983).

Other approaches to completing incomplete models include redefining the outcome variable or specifying a selection mechanism among multiple potential outcomes. In simultaneous binary response models of firm entry, Bresnahan and Reiss (1990, 1991) circumvent problems posed by incompleteness by exploiting the observation that their model uniquely determines the number of entrants. Bjorn and Vuong (1984) and Kooreman (1994) augment models which are incomplete by specifying that whenever $\mathcal{Y}(Z, U; h)$ contains more than one outcome, one of these is selected with fixed probability, which is then taken as a parameter to be estimated. More recently Bajari, Hong, and Ryan (2010) illustrate how to incorporate more complex equilibrium selection mechanisms into general discrete games of complete information. These approaches all yield specifications of the structural relation that result in a model that is complete and coherent. We show in Section 5 how our characterizations of the identified set apply in such contexts, reducing to a set of *moment equalities*, which coincide with the usual likelihood expressions if the model is fully parametric. An alternative approach to incoherence, but which allows for incompleteness, is to choose a sufficiently

[2]See Section D of the on-line Supplementary Appendix to Beresteanu, Molchanov, and Molinari (2011).

flexible solution concept, for example by assuming that strategies are rationalizable as in Aradillas-Lopez and Tamer (2008), or by allowing mixed strategies as in Beresteanu, Molchanov, and Molinari (2012). We do not explicitly consider the implications of mixed or rationalizable strategies, though these could be allowed in our setup through modification of the structural relation h.

We characterize the identified set delivered by the different approaches through a collection of conditional containment functional inequalities of the form

$$\tilde{G}_U(\mathcal{S}; h, z) \geq C_{\mathcal{T}(Y,Z;h)}(\mathcal{S}|z), \text{ a.e. } z \in \mathcal{Z}, \tag{**}$$

for some collection of sets $\mathcal{S} \subseteq \mathcal{U}$. Here $C_{\mathcal{T}(Y,Z;h)}(\mathcal{S}|z)$ is the conditional containment functional of a random set $\mathcal{T}(Y, Z; h)$, a set of values for the unobservables U. This set, and hence its containment functional, varies with the structural relation h under consideration, and with what is assumed about the realization of Y when $\mathcal{Y}(Z, U; h)$ is non-singleton. Thus it differs across the different approaches outlined above and considered formally in Section 3, but for each of them, and for every (h, z), is identified from knowledge of the joint distribution of (Y, Z). The quantity on the left hand side, $\tilde{G}_U(\cdot; h, z)$, as a function of $\mathcal{S} \subseteq \mathcal{U}$ is for each (h, z) a probability measure on $(\mathcal{U}, \mathcal{B}(\mathcal{U}))$, where $\mathcal{B}(\mathcal{U})$ denotes the Borel sets on $\mathcal{U}$. It can also vary across the different approaches in Section 3, and with (h, z), but not with the distribution of observables. It is a known functional of G_U. In fully parametric models it is either known by assumption, or known up to a finite dimensional parameter vector, for example when U is assumed multivariate normal.

1.2 Definitions of Complete and Coherent Functions and Models

To enable further discussion of these issues, we classify *structural relations* h as follows:

- h is **complete** if for almost every $(z, u) \in \mathcal{Z} \times \mathcal{U}$ the set $\mathcal{Y}(z, u; h)$ has cardinality no greater than one.
- h is **incomplete** if for some positive measure set of values of $(z, u) \in \mathcal{Z} \times \mathcal{U}$ the set $\mathcal{Y}(z, u; h)$ has cardinality exceeding one.
- h is **coherent** if for almost every $(z, u) \in \mathcal{Z} \times \mathcal{U}$ the set $\mathcal{Y}(z, u; h)$ is non-empty.
- h is **incoherent** if for some positive measure set of values of $(z, u) \in \mathcal{Z} \times \mathcal{U}$ the set $\mathcal{Y}(z, u; h)$ is empty.

Any structural relation h is necessarily either coherent or incoherent, and either complete or incomplete, and can be any of the four combinations of these.

A model comprises a collection of admissible structural relations $\mathcal{H}$ and conditional probability measures for U given Z, denoted $\left\{G_{U|Z}(\cdot|z), z \in \mathcal{Z}\right\}$. We use the following terminology regarding the classification of *models* according to the set of structural relations admitted:

- A model is **complete** if every $h \in \mathcal{H}$ is complete.
- A model is **incomplete** if there is at least one $h \in \mathcal{H}$ that is incomplete.
- A model is **coherent** if every $h \in \mathcal{H}$ is coherent.
- A model is **incoherent** if there is at least one $h \in \mathcal{H}$ that is incoherent.
- A model is **proper** if it both coherent and complete.

A model can therefore include structural relations h which fall into any or all of the four possible combinations of coherence and completeness. Particular structures admitted by an incoherent model may be coherent and particular structures permitted by an incomplete model may be complete. Our identification results deliver an identified set $\mathcal{D}^0(\mathcal{Z})$ of structural relations h, paired with distributions of unobservables.

Our terminology and classification of models follows that of Tamer (2003) and Lewbel (2007) in distinguishing between incoherent and incomplete models. Some earlier papers in the literature, e.g. Gourieroux, Laffont, and Monfort (1980) and Blundell and Smith (1994) use incoherence to mean either incoherence or incompleteness as we have defined them. This alternative definition renders coherence equivalent to the existence of a unique reduced form, as lucidly described by Lewbel (2007), and is also equivalent to our definition of *proper*. Many models encountered in econometrics are proper, and these allow direct derivation of the conditional distribution of outcomes given covariates from the conditional distribution of latent variables given covariates. Thus, in fully parametric models, these can allow for estimation by maximum likelihood.

1.3 Outline of the Paper

In Section 2 we formally lay out the restrictions of our model, absent assumptions regarding the determination of outcomes when there is the possibility of null outcome sets due to incoherence. We also provide two examples of simultaneous equations binary outcome models, which we refer back to extensively in the paper to illustrate key points. In Section 3 we describe the different approaches to incompleteness and incoherence studied, including what is assumed for the determination of Y when $h(y, Z, U) = 0$ holds for no $y \in \mathcal{Y}$. For each approach we provide a formal characterization of the identified set.[3] These characterizations apply quite generally to simultaneous discrete outcome models, but each entail a collection of conditional moment inequalities of the same cardinality as the set of closed subsets of $\mathcal{U}$. In Section 4 we show that in the context of any particular model, this characterization can be reduced to a much smaller number of conditional moment inequalities. This is done through the notion of core-determining sets, defined by Galichon and Henry (2011).

[3] By "identified set" we mean the set of pairs of structural functions and distributions of unobserved heterogeneity $(h, G_U) \in \mathcal{H} \times \mathcal{G}_U$ that can possibly generate the conditional distributions of Y given $Z = z$ for almost every z on the support of Z. That is, these sets constitute sharp bounds for (h, G_U).

We further show that some of these inequalities may in fact be strengthened to *equalities*, and we provide an algorithm for obtaining the resulting system of conditional moment equalities and inequalities that characterize the identified set. The analysis of this section applies to all four of the approaches laid out in Section 3. In Section 5 we show how our characterizations for the identified set simplify when the model is complete, producing a set of only conditional moment equalities, which coincide with the usual likelihood expressions if the model is fully parametric. All proofs are collected in the Appendix.

2 An Incoherent and Incomplete Discrete Outcome Model

We study models that incorporate the following restrictions:

Restriction A1: (Y, Z, U) are defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, endowed with the Borel sets on Ω. The support of Y, denoted $\mathcal{Y} \subseteq \mathbb{R}^{N_Y}$, is countable, and the support of (U, Z) is $\mathcal{U} \times \mathcal{Z} \subseteq \mathbb{R}^{N_U} \times \mathbb{R}^{N_Z}$.

Restriction A2: For each value $z \in \mathcal{Z}$ there is a proper conditional mass function for Y given $Z = z$, denoted $f^0_{Y|Z}(\cdot|Z = z)$, the collection of which is denoted $\{f^0_{Y|Z}, \mathcal{Z}\}$ These conditional mass functions are identified by the sampling process.

Restriction A3: If $\mathcal{Y}(Z, U; h) \neq \emptyset$, then (Y, Z, U) satisfies $h(Y, Z, U) = 0$. The function $h(\cdot, \cdot, \cdot) : \mathcal{Y} \times \mathcal{Z} \times \mathcal{U} \to \mathbb{R}$ belongs to a known class of functions $\mathcal{H}$.

Restriction A4: U is marginally distributed according to probability measure G_U on $(\mathcal{U}, \mathcal{B}(\mathcal{U}))$, where $\mathcal{B}(\mathcal{U})$ denotes the Borel sets on $\mathcal{U}$. G_U is known to belong to some family of probability measures $\mathcal{G}_U$.

Restriction A5: U and Z are stochastically independent.

Restriction A1 is standard. Restriction A2 stipulates that the conditional distribution of Y given $Z = z$ is identified for all $z \in \mathcal{Z}$. This holds under random sampling, but can also hold in other contexts. Restriction A3 imposes the model specification $h(Y, Z, U) = 0$, and additionally stipulates that the structural relation h belongs to some family of functions $\mathcal{H}$. In principal this class of functions could be parametrically, semiparametrically, or nonparametrically specified. Restriction A4 specifies that the measure G_U belongs to some class of probability measures $\mathcal{G}_U$ Thus $G_U(\mathcal{S})$ denotes the probability of the event $\{U \in \mathcal{S}\}$. Likewise, we use $G_{U|Z}(\mathcal{S}|z)$ to denote the probability of this event conditional on $Z = z$ for any $z \in \mathcal{Z}$. Like the family of structural relations $\mathcal{H}$, the class of probability measures $\mathcal{G}_U$ could be parametrically, semiparametrically, or nonparametrically specified. Restriction A5 formally states the instrumental variable restriction we impose throughout. Under this restriction $G_U(\mathcal{S}) = G_{U|Z}(\mathcal{S}|z)$ for all $\mathcal{S} \subseteq \mathcal{U}$ and $z \in \mathcal{Z}$, but in some places it will be useful to use $G_{U|Z}(\cdot|z)$ explicitly in the derivation of our results.

It is important to note that Restriction A3 is silent as to the realization of Y when the set $\mathcal{Y}(Z, U; h) = \emptyset$, i.e. when the structural relation $h(Y, Z, U) = 0$ holds for no $y \in \mathcal{Y}$. This is relevant when the model is incoherent. The approaches in Section 3 differ in what is assumed about the realization of Y in this case. For each approach we explicitly provide an additional restriction on the determination of Y when $\mathcal{Y}(Z,U;h)=\emptyset$, which in some cases requires extending the definition of the structural relation $h(y,z,u)$. In our various treatments of incoherent models below we will be precise about the implied relationship between the population probabilities $\{\mathbb{P}[\cdot|Z=z], z\in\mathcal{Z}\}$ and the identified quantities $\{f^0_{Y|Z},\mathcal{Z}\}$.

Before proceeding to identification analysis we provide two benchmark examples of models in the literature to which our analysis applies. Both models comprise multiple equation specifications for a bivariate binary outcome variable. While two equation systems of binary outcomes constitute the simplest incarnation of the class of models we study, our analysis applies more generally to models with any finite number of binary or non-binary discrete outcomes, such as for example those considered by Andrews, Berry, and Jia (2004), Ciliberto and Tamer (2009), Engers and Stern (2002), and Jovanovic (1989). In subsequent sections of the paper we refer back to the examples below to help illustrate the various approaches we consider.

Example 1: A Simultaneous Equations Model for Binary Outcomes:

Our first example is the simultaneous binary response model previously considered by Heckman (1978), Bresnahan and Reiss (1990,1991), and Tamer (2003), among others. This model has received much attention in the literature and as such provides a familiar setting in which to illustrate key concepts. The model is given by

$$Y_1 = 1\left[Z_1\beta_1+Y_2\delta_1+U_1>0\right], \tag{2.1}$$

$$Y_2 = 1\left[Z_2\beta_2+Y_1\delta_2+U_2>0\right], \tag{2.2}$$

where $Y\equiv(Y_1,Y_2)$, $(Y,Z)\in\mathcal{Y}\times\mathcal{Z}$ are observable, and $U=(U_1,U_2)$ is an unobservable 2-vector in $\mathbb{R}^2$. Typically in the prior literature U has been assumed to be normally distributed with mean zero and unknown covariance matrix Σ, with Σ_{11} diagonal elements normalized to one.

To write the model in the form of (1.1), one can use any structural relation $h(y,z,u)$ that takes value zero if and only if (2.1) and (2.2) are satisfied, for example,

$$h(Y,Z,U)=\left\{\begin{array}{c} Y_1\cdot|Z_1\beta_1+Y_2\delta_1+U_1|_- +(1-Y_1)\cdot|Z_1\beta_1+Y_2\delta_1+U_1|_+ \\ +Y_2\cdot|Z_2\beta_2+Y_1\delta_2+U_2|_- +(1-Y_2)\cdot|Z_2\beta_2+Y_1\delta_2+U_2|_+ \end{array}\right\},$$

where $|x|_-$ and $|x|_+$ denote the negative and positive part of x, respectively. Here the structural relation h is known up to the finite dimensional parameter vector $\theta\equiv\left(\beta_1',\delta_1,\beta_2',\delta_2\right)'$.

Whether the function h is incomplete or incoherent depends on the sign of the product of interaction terms $\delta_1\cdot\delta_2$. The top panels of Figure 1 illustrates the case for a particular z and θ where $\delta_1\cdot\delta_2>0$, with dark blue regions indicating values of u that entail multiple solutions. In the first panel, where $\delta_1,\delta_2<0$, the solution set is $\mathcal{Y}(z,u;h)=\{(0,1),(1,0)\}$ whenever u belongs to the rectangle $(-z_1\beta_1, -z_1\beta_1 - \delta_1] \times (-z_2\beta_2, -z_2\beta_2 - \delta_2]$. For all u outside this rectangle, there is a unique solution. The second panel depicts the case where $\delta_1, \delta_2 > 0$. In this case the solution set is $\mathcal{Y}(z, u; h) = \{(0,0), (1,1)\}$ whenever $u \in (-z_1\beta_1 - \delta_1, -z_1\beta_1] \times (-z_2\beta_2 - \delta_2, -z_2\beta_2]$. For all other u there is a unique solution for y. In these cases h is thus coherent but incomplete.

The two bottom panels in the figure illustrate the cases in which $\delta_1 \cdot \delta_2 < 0$. In each of these cases there is no solution for (y_1, y_2) whenever u belongs to the middle unshaded region. In the case where $\delta_1 > 0$ and $\delta_2 < 0$, this region is given by the set of u satisfying $-z_1\beta_1 - \delta_1 < u_1 \leq -z_1\beta_1$ and $-z_2\beta_2 \leq u_2 \leq -z_2\beta_2 - \delta_2$. When u lies in this rectangle, $y_1 = 1$ solves (2.1) only if $y_2 = 0$, but $y_2 = 0$ solves (2.2) only if $y_1 = 0$. When $\delta_1 < 0$ and $\delta_2 > 0$, the situation is symmetrical, with no solution for (y_1, y_2) when $-z_1\beta_1 < u_1 \leq -z_1\beta_1 - \delta_1$ and $-z_2\beta_2 - \delta_2 < u_2 \leq -z_2\beta_2$. In both cases with $\delta_1 \cdot \delta_2 < 0$ there are no regions where multiple solutions to (2.1) and (2.2) are admitted, so such structural relations are incoherent but not incomplete. The model however may be both incoherent and incomplete, as long as the sign of $\delta_1 \cdot \delta_2$ is not restricted a priori. Other simultaneous equations models for discrete outcomes may admit functions h that are both incomplete and incoherent, for example in simultaneous equations models with more than two binary endogenous outcome variables, such as that of Ciliberto and Tamer (2009). ■

Example 2: A Triangular Model for Binary Outcomes: Consider the *triangular* simultaneous equations model

$$
\begin{aligned}
Y_1 &= g_1(Y_2, Z_1, U_1), \\
Y_2 &= g_2(Z_2, U_2),
\end{aligned}
$$

with Y_1 again binary and Y_2 a discrete, ordered outcome, where g_1 and g_2 are weakly increasing in scalar unobservables U_1 and U_2, respectively. With (U_1, U_2) and Z independent and with suitable normalizations, but without further parametric structure on functional forms, g_2 is point identified while g_1 may only be set identified, see Chesher (2005) and Jun, Pinkse, and Xu (2011). The model may be written in the form of (1.1), for example as

$$h(Y, Z, U) = (Y_1 - g_1(Y_2, U_1))^2 + (Y_2 - g_2(Z, U_2))^2.$$

A special case is that where Y_1 and Y_2 are both binary, and where the functions g_1 and g_2 are restricted to have the linear threshold-crossing representation:

$$
\begin{aligned}
g_1(Y_2, Z_1, U_1) &= 1[Z_1\beta_1 + Y_2\delta_1 + U_1 > 0], && (2.3) \\
g_2(Z_2, U_2) &= 1[Z_2\beta_2 + U_2 > 0]. && (2.4)
\end{aligned}
$$

Like our first example, this system is an instance of Heckman's (1978) Multivariate Probit Model with Structural Shift. The added restriction, relative to example 1, that $\delta_2 = 0$ guarantees that

Heckman's *principal assumption* holds, and hence that the model is both complete and coherent, and therefore *proper*. Heckman (1978) provides conditions under which there is point identification when (U_1, U_2) are bivariate normally distributed, and provides a likelihood-based estimation procedure. Note however that without the assumption of bivariate normal errors, the model remains proper, though this alone does not guarantee point identification. In Section 5 we show how in such cases the approaches of this paper reduce to the classical treatment of this model, producing moment equalities that correspond precisely to likelihood equations when the distribution is parametrically specified. ■

3 Identification

In this section we consider four different approaches to the analysis of incoherent and incomplete models. The first three approaches deal with incompleteness in the same way as the recent literature on set identification in incomplete models. That is, the observed outcome is permitted to be any one of the set of outcomes compatible with an *incomplete* structural relation. These approaches differ however in how they handle the event $\mathcal{Y}(Z, U; h) = \emptyset$ when the structural relation is *incoherent*. The fourth approach, proposed by Dagenais (1997) and Hajivassiliou (2008), treats incoherence and incompleteness identically in a way that enables construction of a unique likelihood function. The properties of all four approach are summarized in Table 1.

In our first approach we assume that null outcomes are directly observed, much as censored outcomes in models of censoring. Data often include no observations with "null" outcomes, for example in the analysis of market structure, one never observes markets in which the market configuration is null or empty.[4] Nonetheless, this setting provides an instructive precursor to the rest of the approaches we consider, all of which propose alternative methods to deal with incoherence in settings where null outcomes are never observed.

Our goal throughout is (set) identification of the pair (h, G_U). To this end we define the zero level sets of the structural relation h as follows:

$$\forall y \in \mathcal{Y}, \qquad \mathcal{L}(y, z; h) \equiv \{u : y \in \mathcal{Y}(z, u; h)\} = \{u : h(y, z, u) = 0\}\,, \tag{3.1}$$

and sets on which $h(y, z, u) = 0$ has no solution as

$$\mathcal{L}(\emptyset, z; h) \equiv \{u : \mathcal{Y}(z, u; h) = \emptyset\} = \{u : \forall y \in \mathcal{Y},\ h(y, z, u) \neq 0\}\,, \tag{3.2}$$

where $\emptyset$ denotes the null or empty set.

In the following analysis we work with random sets $\mathcal{T}(Y, Z; h)$, which are constructed through

[4] A null outcome should not be confused with one in which no firm enters, in which case the outcome is the zero vector.

combinations of set theoretic operations applied to the level sets $\mathcal{L}(Y,Z;h)$. The definition and indeed support of the random sets $\mathcal{T}(Y,Z;h)$ varies with the different approaches below, as well as the structural relation h under consideration. In each case, $\mathcal{T}(\cdot,\cdot;h) : \mathcal{Y} \times \mathcal{Z} \to \mathcal{U}$ comprises the set of values of the unobservable U that are feasible given arguments (y,z). Applying the correspondence $\mathcal{T}(\cdot,\cdot;h)$ to random arguments Y and Z, we have the *random set* $\mathcal{T}(Y,Z;h)$. For any $z \in \mathcal{Z}$, the distribution of this random set conditional on $Z = z$ is completely characterized by its conditional containment functional $C_{\mathcal{T}(Y,Z;h)}(\cdot|z) : 2^{\mathcal{U}} \to [0,1]$, defined as

$$C_{\mathcal{T}(Y,Z;h)}(\mathcal{S}|z) \equiv \mathbb{P}\{\mathcal{T}(Y,Z;h) \subseteq \mathcal{S}|Z = z\}, \tag{3.3}$$

for any $\mathcal{S} \subseteq \mathcal{U}$.

The conditional moment inequalities used in our characterizations are all of the form

$$\tilde{G}_U(\mathcal{S};h,z) \geq C_{\mathcal{T}(Y,Z;h)}(\mathcal{S}|z), \text{ a.e. } z \in \mathcal{Z}, \tag{3.4}$$

for all closed $\mathcal{S} \subseteq \mathcal{U}$, where $\tilde{G}_U(\cdot;h,z)$ is a probability measure on $(\mathcal{U}, \mathcal{B}(\mathcal{U}))$. Both $\tilde{G}_U(\mathcal{S};h,z)$ and $C_{\mathcal{T}(Y,Z;h)}(\mathcal{S}|z)$ vary across the different approaches. The theorems of this section apply rather generally to any simultaneous discrete choice model that satisfies our restrictions. In the next section we show that in the context of any particular model we can reduce this characterization to a subset of these inequalities by showing it suffices to consider a collection of regions $\mathcal{S}$ in (3.4), that is much smaller than the collection of all closed subsets of $\mathcal{U}$. These are "core-determining" sets as in Galichon and Henry (2011). We further provide conditions whereby some of these inequalities reduce to equalities.

In studying these characterizations it is important to understand that the containment functional in (3.4) may vary with z because (i) as z varies the conditional distribution $f^0_{Y|Z}(\cdot|z)$ may vary and (ii) as z varies the conditional support of the random set $\mathcal{T}(Y,Z;h)$ may vary. As a result the identified set $\mathcal{D}^0(\mathcal{Z})$ is critically dependent on the support of Z.

3.1 Approach 1: Observed Null Outcomes

This approach requires that when the realizations of (Z,U) are such that no outcome y satisfies $h(y,Z,U) = 0$, this event is observed. We use ϕ as a place-holder for the "null" realization that Y takes when $\mathcal{Y}(Z,U;h) = \emptyset$. Thus we have

$$\{Y = \phi\} \Leftrightarrow \{\mathcal{Y}(Z,U;h) = \emptyset\} \Leftrightarrow \{U \in \mathcal{L}(\emptyset,z;h)\}. \tag{3.5}$$

and for all $z \in \mathcal{Z}$, $\mathbb{P}[\mathcal{Y}(Z,U;h) = \emptyset|Z = z] = \mathbb{P}[Y = \phi|Z = z]$. The probability allocated by $\mathbb{P}\{\cdot|Z = z\}$ over the extended support $\mathcal{Y}^* \equiv \mathcal{Y} \cup \phi$ is one. With the event $\{Y = \phi\}$ observed, $\mathbb{P}\{Y = y|Z = z\} = f^0_{Y|Z}(y|z)$ is identified for all $Y \in \mathcal{Y}^*$ and almost every $z \in \mathcal{Z}$.

To capture the knowledge that observation of $Y = \phi$ is logically equivalent to $U \in \mathcal{L}(\emptyset, z; h)$, we expand the definition of the structural relation $h(\cdot,\cdot,\cdot)$ to cover the extended support $\mathcal{Y}^*$ with the following restriction.

Restriction B1 (Censored Null Outcomes): The outcome variable Y has support $\mathcal{Y}^*$, and there is the equivalence of events $\{U \in \mathcal{L}(\emptyset, Z; h)\} \Leftrightarrow \{Y = \phi\}$. Furthermore, for any $(z, u) \in \mathcal{Z} \times \mathcal{U}$, $h(\phi, z, u) = 0$ if for all $y \in \mathcal{Y}$, $h(y, z, u) \neq 0$, and $h(\phi, z, u) \neq 0$ if $h(y, z, u) = 0$ for some $y \in \mathcal{Y}$.

Under Restriction B1, $h(\phi, z, u) = 0$ precisely when $h(y, z, u) = 0$ is solved by no $y \in \mathcal{Y}$. Therefore $\{Y = \phi\}$ implies that $h(\phi, Z, U) = 0$, and taken together with Restriction A3 this implies that $U \in \mathcal{T}_1(Y, Z; h)$ for all (Y, Z, U), where the random set $\mathcal{T}_1(Y, Z; h)$ is defined simply as

$$\mathcal{T}_1(Y, Z; h) \equiv \mathcal{L}(Y, Z; h), \tag{3.6}$$

for all $(Y, Z, h) \in \mathcal{Y}^* \times \mathcal{Z} \times \mathcal{H}$.

Because the event $\{Y = y \wedge Z = z\}$ implies that $U \in \mathcal{T}_1(y, z; h)$, any structural relation h and collection of conditional measures $\left\{G_{U|Z}, z \in \mathcal{Z}\right\}$ that can generate $\{f^0_{Y|Z}, \mathcal{Z}\}$ must satisfy for all $y \in \mathcal{Y}^*$,

$$G_{U|Z}(\mathcal{T}_1(y, z; h)\,|z) \geq f^0_{Y|Z}(y|z)\,, \text{ a.e. } z \in \mathcal{Z}. \tag{3.7}$$

It follows that for any set $\mathcal{S} \subseteq \mathcal{U}$,

$$G_{U|Z}(\mathcal{S}|z) \geq \sum_{y \in \mathcal{Y}^*} 1[\mathcal{T}_1(y, z; h) \subseteq \mathcal{S}] \times f^0_{Y|Z}(y|z)\,, \text{ a.e. } z \in \mathcal{Z}. \tag{3.8}$$

Under the *independence restriction* $U \perp\!\!\!\perp Z$ this simplifies to

$$G_U(\mathcal{S}) \geq \sum_{y \in \mathcal{Y}^*} 1[\mathcal{T}_1(y, z; h) \subseteq \mathcal{S}] \times f^0_{Y|Z}(y|z)\,, \text{ a.e. } z \in \mathcal{Z}, \tag{3.9}$$

for all $\mathcal{S} \subseteq \mathcal{U}$. The left hand side does not depend on z so that for each $h \in \mathcal{H}$ only the largest value of the summation on the right hand side is relevant. The quantity on the right of (3.9) is the conditional (on $Z = z$) containment functional of the random set $\mathcal{T}_1(Y, Z; h)$,

$$C_{\mathcal{T}_1(Y,Z;h)}(\mathcal{S}|z) \equiv \sum_{y \in \mathcal{Y}^*} 1[\mathcal{T}_1(y, z; h) \subseteq \mathcal{S}] \times f^0_{Y|Z}(y|z)\,, \tag{3.10}$$

which for any $h \in \mathcal{H}$ and $z \in \mathcal{Z}$ maps sets $\mathcal{S} \subseteq \mathcal{U}$ to the unit interval, and is identified from knowledge of $\{f^0_{Y|Z}, \mathcal{Z}\}$.

This delivers the following system of inequalities.

$$\forall \mathcal{S} \subseteq \mathcal{U}, \quad G_U(\mathcal{S}) \geq \sup_{z \in \mathcal{Z}} C_{\mathcal{T}_1(Y,Z;h)}(\mathcal{S}|z)\,, \tag{3.11}$$

which is equivalent to (**) with $\tilde{G}_U(\mathcal{S}; h, z) = G_U(\mathcal{S})$. As we establish in the following theorem, the pairs of (h, G_U) that satisfy this system of inequalities can generate the family of distributions $\{f^0_{Y|Z}, \mathcal{Z}\}$, so these pairs comprise the identified set.

Theorem 1 *Let Restrictions A1-A5 and B1 hold. Then the identified set of admissible pairs (h, G_U) associated with $\left\{f^0_{Y|Z}, \mathcal{Z}\right\}$ is*

$$\mathcal{D}^0(\mathcal{Z}) \equiv \left\{(h, G_U) \in \mathcal{H} \times G_U : \forall \mathcal{S} \in \mathsf{F}(\mathcal{U}),\ G_U(\mathcal{S}) \geq C_{\mathcal{T}_1(Y,Z;h)}(\mathcal{S}|z),\ a.e.\ z \in \mathcal{Z}\right\},$$

where $\mathsf{F}(\mathcal{U})$ denotes the set of all closed subsets of $\mathcal{U}$.

The theorem establishes that $\mathcal{D}^0(\mathcal{Z})$ is the identified set for (h, G_U) from knowledge of $\{f^0_{Y|Z}, \mathcal{Z}\}$. The proof of this and the analogous theorems for subsequent approaches in this section are based on a key result from random set theory, namely Artstein's inequality (Artstein (1983)) that characterizes the selections of a random set through inequalities based on its associated capacity or containment functional. This and related results have been previously used to establish sharp bounds for the parameters of econometric models by Beresteanu, Molinari, and Molchanov (2011a, 2011b), Galichon and Henry (2011), and Chesher, Rosen, and Smolinski (2011). In our model this characterization guarantees that, up to an ordered coupling, we can find for any $(h, G_U) \in \mathcal{D}^0(\mathcal{Z})$ a random vector $\tilde{U}$ that is (i) compatible with our restrictions and that (ii) delivers the observed distribution of the data conditional on almost every $z \in \mathcal{Z}$. Thus, for any $(h, G_U) \in \mathcal{D}^0(\mathcal{Z})$, Restrictions A1-A5 and B1 are satisfied and the implied conditional distribution of $Y|Z = z$ corresponds to that given by the conditional mass function $f^0_{Y|Z}(\cdot|Z = z)$, a.e. $z \in \mathcal{Z}$.

Example 1 (continued): In the context of the simultaneous binary model, observability of the event $\{Y = \phi\}$ trivially allows one to detect whether the structural relation h is incoherent according to whether $f^0_Y(\phi) > 0$, and hence to infer the sign of $\delta_1 \times \delta_2$. If $f^0_Y(\phi) = 0$ then the situation is that depicted in one of the top two panels, so that the model is coherent but incomplete. If instead $f^0_Y(\phi) > 0$ then h is incoherent but complete. In both cases the support of the random set $\mathcal{T}_1(Y, Z; h)$ conditional on $Z = z$ is the collection of regions:

$$\begin{aligned}
\mathcal{T}_1((0,0), z; h) &= \{u \in \mathcal{U} : u_1 \leq -z_1\beta_1 \wedge u_2 \leq -z_2\beta_2\}, \qquad (3.12)\\
\mathcal{T}_1((0,1), z; h) &= \{u \in \mathcal{U} : u_1 \leq -z_1\beta_1 - \delta_1 \wedge u_2 > -z_2\beta_2\},\\
\mathcal{T}_1((1,0), z; h) &= \{u \in \mathcal{U} : u_1 > -z_1\beta_1 \wedge u_2 \leq -z_2\beta_2 - \delta_2\},\\
\mathcal{T}_1((1,1), z; h) &= \{u \in \mathcal{U} : u_1 > -z_1\beta_1 - \delta_1 \wedge u_2 > -z_2\beta_2 - \delta_2\},\\
\mathcal{T}_1(\emptyset, z; h) &= \left\{u \in \mathcal{U} : \left\{\begin{array}{c} \{-z_1\beta_1 < u_1 < -z_1\beta_1 - \delta_1 \wedge -z_2\beta_2 - \delta_2 < u_2 < -z_2\beta_2\} \\ \vee \quad \{-z_1\beta_1 - \delta_1 < u_1 < -z_1\beta_1 \wedge -z_2\beta_2 < u_2 < -z_2\beta_2 - \delta_2\} \end{array}\right\}\right\}.
\end{aligned}$$

These sets correspond precisely to the regions $\mathcal{L}(Y, z; h)$ shown in Figure 1. Application of Theorem

1 then gives that for all closed $\mathcal{S} \subseteq \mathcal{U}$

$$G_U(\mathcal{S}) \geq \sup_{z \in \mathcal{Z}} C_{\mathcal{T}_1(Y,Z;h)}(\mathcal{S}|z),$$

if and only if (h, G_U) is in the identified set. ■

3.2 Approach 2: Truncation of Null Outcomes

Suppose now and for all remaining approaches that the model is incoherent but null outcomes are never observed. One may reason that observations of (Y, Z) must therefore correspond to realizations of (Z, U) such that $\mathcal{Y}(Z, U; h)$ is non-empty, even if the model allows $\mathcal{Y}(z, u; h) = \emptyset$. Taking the structural relation $h(Y, Z, U) = 0$ at face value, a realization of (Z, U) such that $\mathcal{Y}(z, u; h) = \emptyset$ then could not have resulted in the observed Y. For example, in a study of firm entry into markets one will never observe a market in which the "the empty set" obtains. In any observed market a non-null outcome will be found, e.g. $(0, 0)$ in a binary entry model in which neither firm enters. This is formalized with the following restriction.

Restriction B2 (Truncated Null Outcomes): Only realizations of $(Y, Z) = (y, z)$ that satisfy $y \in \mathcal{Y}(z, U; h)$ are observed and thus for each $z \in Z$, the distribution of Y conditional on $Z = z$ identified by the distribution of the data is that of $Y | \{Z = z \wedge \mathcal{Y}(Z, U; h) \neq \emptyset\}$.

Under Restriction B2, data are drawn from the subset of the sample space in which $\mathcal{Y}(Z, U; h)$ is non-empty, since null outcomes are never observed. Considering incoherence in this way is akin to assuming that the distribution of observable data is a truncated version of the underlying population distribution. Under this restriction the probability mass functions $\{f^0_{Y|Z}, \mathcal{Z}\}$ correspond to the distribution of U over the support $\mathcal{U} \setminus \mathcal{L}(\emptyset, z; h)$ so that for all $y \in \mathcal{Y}$,

$$f^0_{Y|Z}(y|z) = \mathbb{P}[Y = y|\mathcal{L}(Y, z; h) \neq \emptyset, Z = z], \text{ a.e. } z \in \mathcal{Z}. \tag{3.13}$$

Observation of (Y, Z) implies that $U \in \mathcal{T}_2(Y, Z; h)$ where

$$\mathcal{T}_2(Y, Z; h) \equiv \mathcal{L}(Y, Z; h), \tag{3.14}$$

which is identical to the definition of $\mathcal{T}_1(Y, Z; h)$ in Section 3.1, but now $Y = \phi$ is never observed and $f^0_{Y|Z}(\cdot|z)$ assigns unit mass to $\mathcal{Y}$.

Although the truncation probability is not observed in the data, it is a known function of (h, G_U), namely

$$G_{U|Z}(\mathcal{L}(\emptyset, z; h)|z) = \mathbb{P}[\mathcal{Y}(Z, U; h) = \emptyset|Z = z]. \tag{3.15}$$

Starting with (3.13) we then have for all $y \in \mathcal{Y}$

$$\begin{aligned} f^0_{Y|Z}(y|z) &= \mathbb{P}[Y = y|\mathcal{Y}(Z,U;h) \neq \emptyset, Z = z] \\ &= \frac{\mathbb{P}[Y = y \wedge \mathcal{Y}(Z,U;h) \neq \emptyset | Z = z]}{\mathbb{P}[\mathcal{Y}(Z,U;h) \neq \emptyset | Z = z]} \\ &= \frac{\mathbb{P}[Y = y|Z = z]}{1 - G_{U|Z}(\mathcal{L}(\emptyset, z; h)|z)}, \end{aligned} \tag{3.16}$$

where the third line follows because $Y = y \Rightarrow \mathcal{Y}(Z,U;h) \neq \emptyset$, and by (3.15).

Combining $U \in \mathcal{T}_2(y,z;h)$ with (3.15) we have that for all $y \in \mathcal{Y}$ and almost every $z \in \mathcal{Z}$,

$$G_{U|Z}(\mathcal{T}_2(y,z;h)|z) \geq \mathbb{P}[Y = y|Z = z] = \left[1 - G_{U|Z}(\mathcal{L}(\emptyset, z; h)|z)\right] \cdot f^0_{Y|Z}(y|z).$$

Then, using the independence condition $U \perp\!\!\!\perp Z$ it follows that

$$G_U(\mathcal{T}_2(y,z;h)) \geq [1 - G_U(\mathcal{L}(\emptyset, z; h))] \cdot f^0_{Y|Z}(y|z),$$

and dividing through by $\left[1 - G_{U|Z}(\mathcal{L}(\emptyset, z; h))\right]$,

$$G_U(\mathcal{T}_2(y,z;h)|\mathcal{Y}(z,U;h) \neq \emptyset) \geq f^0_{Y|Z}(y|z). \tag{3.17}$$

Consequently, for any $\mathcal{S} \subseteq \mathcal{U}$,

$$G_U(\mathcal{S}|\mathcal{Y}(z,U;h) \neq \emptyset) \geq \left(\sum_{y \in \mathcal{Y}} 1[\mathcal{T}_2(y,z;h) \subseteq \mathcal{S}] f^0_{Y|Z}(y|z)\right) = C_{\mathcal{T}_2(Y,Z;h)}(\mathcal{S}|z),$$

where $G_U(\mathcal{S}|\mathcal{Y}(z,U;h) \neq \emptyset)$ is the conditional probability that $U \in \mathcal{S}$ given $\mathcal{Y}(z,U;h) \neq \emptyset$. The quantity on the right, $C_{\mathcal{T}_2(Y,Z;h)}(\mathcal{S}|z)$, is the conditional containment functional of $\mathcal{T}_2(Y,Z;h)$. Thus, this inequality is precisely (**) with $\tilde{G}_U(\mathcal{S};h,z) = G_U(\mathcal{S}|\mathcal{Y}(z,U;h) \neq \emptyset)$. Indeed, the inequalities (3.2) taken over all closed sets and almost every $z \in \mathcal{Z}$ deliver the identified set.

Theorem 2 *Let Restrictions A1-A5 and B2 hold. Then the identified set of admissible pairs* (h, G_U) *associated with* $\left\{f^0_{Y|Z}, \mathcal{Z}\right\}$ *is*

$$\mathcal{D}^0(\mathcal{Z}) \equiv \left\{(h, G_U) \in \mathcal{H} \times G_U : \forall \mathcal{S} \in \mathsf{F}(\mathcal{U}), \ G_U(\mathcal{S}|\mathcal{Y}(z,U;h) \neq \emptyset) \geq C_{\mathcal{T}_2(Y,Z;h)}(\mathcal{S}|z), \ a.e. \ z \in \mathcal{Z}\right\},$$

where $\mathsf{F}(\mathcal{U})$ *denotes the set of all closed subsets of* $\mathcal{U}$*.*

Example 1 (continued): Using this approach can only be justified if the event $\{Y = \phi\}$ is never recorded. The support of the random set $\mathcal{T}_2(Y,Z;h)$ conditional on $Z = z$ is identical to that for $\mathcal{T}_2(Y,Z;h)$ of the first approach given in (3.12), but for the exclusion of $\mathcal{T}_1(\emptyset, z; \theta)$. ■

3.3 Approach 3: Indeterminate Allocation of Null Outcomes

The restriction considered in this section regarding the determination of outcomes when $\mathcal{Y}(Z,U;h) = \emptyset$ was previously considered by Beresteanu, Molchanov, and Molinari (2011, on-line supplement p. 53) as their Assumption D.1(ii). The restriction is the following.

Restriction B3 (Anything Goes): If (Z,U) are such that $U \in \mathcal{L}(\emptyset, Z; h)$, equivalently if $\mathcal{Y}(Z,U;h) = \emptyset$, then any $Y \in \mathcal{Y}$ is feasible.

Restriction B3 allows for Y to take any realization when (Z,U) have realizations (z,u) such that there is no $y \in \mathcal{Y}$ satisfying $h(y,z,u) = 0$. Combined with Restriction A3, this implies that when $\mathcal{Y}(Z,U;h)$ is non-empty, then $Y \in \mathcal{Y}(Z,U;h)$, while if $\mathcal{Y}(Z,U;h) = \emptyset$, then Y can take any realization on $\mathcal{Y}$. The idea behind the restriction is that $\mathcal{Y}(Z,U;h)$ are the solutions to the underlying structural model $h(Y,Z,U) = 0$, for example the set of equilibrium outcomes of a game. Thus it should be that $Y \in \mathcal{Y}(Z,U;h)$. However, when $\mathcal{Y}(Z,U;h) = \emptyset$, the model has no solutions, and is therefore silent as to what outcome will occur, so that any is taken to be feasible.

Here we have that observation of $\{Y = y \wedge Z = z\}$ implies that $U \in \mathcal{T}_3(y,z;h)$, where

$$\mathcal{T}_3(y,z;h) \equiv \{u : h(y,z,u) = 0\} \cup \{u : \mathcal{Y}(z,u;h) = \emptyset\}. \tag{3.18}$$

This is equivalently captured by imposing that (Y,Z,U) solve

$$\tilde{h}(y,z,U) = 0, \tag{3.19}$$

for the modified structural relation

$$\tilde{h}(y,z,u) \equiv 1[\mathcal{Y}(z,u;h) = \emptyset] \times h(y,z,u), \tag{3.20}$$

which is zero if $\mathcal{Y}(z,u;h)$ is the null set, and $h(y,z,u)$ otherwise. The sets $\mathcal{T}_3(Y,Z;h)$ then correspond to the zero level sets of $\tilde{h}$, $\mathcal{L}\left(Y,Z;\tilde{h}\right)$.

Using (3.18) we have that for almost every $z \in \mathcal{Z}$,

$$G_{U|Z}(\mathcal{T}(Y,Z;h)|z) \geq f^0_{Y|Z}(y|z).$$

Then applying the same logic as in Sections 3.1 and 3.2 and again using the independence Restriction A5, it follows that any (h, G_U) in the identified set must satisfy

$$G_U(\mathcal{S}) \geq C_{\mathcal{T}_3(Y,Z;h)}(\mathcal{S}|z), \text{ a.e. } z \in \mathcal{Z},$$

for any $\mathcal{S} \in \mathcal{U}$. This is again of the form (**), with $\tilde{G}_U(\mathcal{S};h,z) = G_U(\mathcal{S})$, but with a different containment functional than in the first two approaches, namely that of $\mathcal{T}_3(y,z;h)$. The following theorem establishes that this inequality indeed characterizes the identified set.

Theorem 3 *Let Restrictions A1-A5 and B3 hold. Then the identified set of admissible pairs* (h, G_U) *associated with* $\left\{f^0_{Y|Z}, \mathcal{Z}\right\}$ *is*

$$\mathcal{D}^0(\mathcal{Z}) \equiv \left\{(h, G_U) \in \mathcal{H} \times G_U : \forall \mathcal{S} \in \mathsf{F}(\mathcal{U}), \ G_U(\mathcal{S}) \geq C_{\mathcal{T}(Y,Z;h)}(\mathcal{S}|z), \ a.e. \ z \in \mathcal{Z}\right\},$$

where $\mathsf{F}(\mathcal{U})$ *denotes the set of all closed subsets of* $\mathcal{U}$*, with* $\tilde{h}$ *as defined in (3.20).*

Example 1 (continued): Like the previous approach this one can only be justified if the event $\{Y = \phi\}$ is never recorded. The support of the random set $\mathcal{T}_3(Y, z; h)$ for any $z \in \mathcal{Z}$ is given by

$$\begin{aligned}
\mathcal{T}_3((0,0), z; h) &= \{u \in \mathcal{U} : u_1 \leq -z_1\beta_1 \wedge u_2 \leq -z_2\beta_2\} \cup \mathcal{L}(\emptyset, z; h), \\
\mathcal{T}_3((0,1), z; h) &= \{u \in \mathcal{U} : u_1 \leq -z_1\beta_1 - \delta_1 \wedge u_2 > -z_2\beta_2\} \cup \mathcal{L}(\emptyset, z; h), \\
\mathcal{T}_3((1,0), z; h) &= \{u \in \mathcal{U} : u_1 > -z_1\beta_1 \wedge u_2 \leq -z_2\beta_2 - \delta_2\} \cup \mathcal{L}(\emptyset, z; h), \\
\mathcal{T}_3((1,1), z; h) &= \{u \in \mathcal{U} : u_1 > -z_1\beta_1 - \delta_1 \wedge u_2 > -z_2\beta_2 - \delta_2\} \cup \mathcal{L}(\emptyset, z; h).
\end{aligned} \tag{3.21}$$

If $\delta_1\delta_2 > 0$, as illustrated in the top panels of Figure 1, then h is incomplete but coherent so that $\mathcal{L}(\emptyset, z; h) = \emptyset$, and these sets correspond precisely to those of the second approach, equivalently the first four sets listed in (3.12). If instead $\delta_1\delta_2 < 0$, as illustrated in the bottom panels of Figure 1, then h is complete but incoherent. In these panels $\mathcal{L}(\emptyset, z; h)$ is the unshaded region in the center, and each region $\mathcal{T}_3(y, z; h)$ is the union of the cyan-shaded region $\mathcal{L}(y, z; h)$ with this region. ■

3.4 Approach 4: Truncation of Null and Non-Unique Outcomes

In this section we characterize the identified set that results from imposing a restriction previously considered by Dagenais (1997) and Hajivassiliou (2008). The restriction, formalized below, effectively removes incoherence and incompleteness from the model by assuming that observations correspond to the subset of the population distribution on which (Z, U) map to a unique outcome.

Restriction B4 (Truncated Non-Unique Outcomes): Only realizations of $(Y, Z) = (y, z)$ for which $\mathcal{Y}(z, U; h) = \{y\}$ are observed. Thus the distribution of Y conditional on $Z = z$ identified by the distribution of the data is that of $Y|\{Z = z \wedge \#\mathcal{Y}(z, U; h) = 1\}$, where $\#\mathcal{A}$ denotes the cardinality of $\mathcal{A}$.

This restriction implies that the family of probability mass functions $\{f^0_{Y|Z}, \mathcal{Z}\}$ identified from the data corresponds to

$$f^0_{Y|Z}(y|z) = \mathbb{P}[Y = y|Z = z|\,\#\mathcal{Y}(z, U; h) = 1],$$

We have that observation of $\{Y = y \wedge Z = z\}$ implies that $U \in \mathcal{T}_4(Y, Z; h)$, where

$$\mathcal{T}_4(y, z; h) \equiv \{u : \mathcal{Y}(z, u; h) = \{y\}\} = \mathcal{L}(y, z; h) \backslash \left\{\cup_{y' \neq y} \mathcal{L}(y', z; h)\right\} \tag{3.22}$$

where "\" denotes the set difference operator.

Proceeding in the same manner as for Approach 2, Section 3.2 starting with (3.13), but with the conditioning event $\{\#\mathcal{Y}(z, U; h) = 1\}$ replacing $\{\mathcal{L}(Y, Z; h) \neq \emptyset\}$, we have the inequalities

$$G_U\left(\mathcal{S}|\#\mathcal{Y}(z, U; h) = 1\right) \geq C_{\mathcal{T}_4(Y,Z;h)}\left(\mathcal{S}|z\right), \text{ a.e. } z \in \mathcal{Z}, \tag{3.23}$$

for all $S \in \mathcal{U}$, equivalently (**), with $\tilde{G}_U\left(\mathcal{S}; h, z\right) = G_U\left(\mathcal{S}|\#\mathcal{Y}(z, U; h) = 1\right)$ and the sets $\mathcal{T}_4(y, z; h)$ defined above. However, under Restriction B4 the model is effectively rendered complete, as the realization of unobservable U that coincides with observation of $(Y, Z) = (y, z)$ must be such that y solves $h(y, z, U) = 0$ *uniquely*. This implies that for each $z \in \mathcal{Z}$, the sets $\{\mathcal{T}(y, z; h) : y \in \mathcal{Y}\}$ form a partition of $\{u : \#\mathcal{Y}(z, U; h) = 1\}$. Thus, by applying identical reasoning as that for identification in *complete* models of Section 5 below, we can reduce this characterization to a simpler representation based solely on equality restrictions, as stated in the following theorem.

Theorem 4 *Let Restrictions A1-A5 and B4 hold. Then the identified set of admissible pairs* (h, G_U) *associated with* $\left\{f^0_{Y|Z}, \mathcal{Z}\right\}$ *is*

$$\mathcal{D}^0(\mathcal{Z}) \equiv \left\{\{h, G_U\} \in \mathcal{H} \times \mathcal{G}_U : \forall y \in \mathcal{Y},\ G_U\left(\mathcal{T}_4(y, z; h)|\#\mathcal{Y}(z, U; h) = 1\right) = f^0_{Y|Z}\left(y|z\right)\ \textit{a.e. } z \in \mathcal{Z}\right\}$$

Thus the identified set is characterized by equalities of the form

$$G_U\left(\mathcal{T}_4(y, z; h)|\#\mathcal{Y}(z, U; h) = 1\right) = f^0_{Y|Z}\left(y|z\right).$$

The quantity on the right hand side, $f^0_{Y|Z}\left(y|z\right)$, is point-identified. The quantity on the left varies with z, h, and G_U, and is a known function of these quantities. In a model that is fully parametric G_U and h are specified up to a finite-dimensional parameter, say θ. Then $G_U\left(\mathcal{T}_4(y, z; h)|\,|\mathcal{Y}(z, U; h)| = 1\right)$ is a parametric single-observation likelihood,

$$\ell\left(\theta; y, z\right) = G_U\left(\mathcal{T}_4(y, z; h)|\#\mathcal{Y}(z, U; h) = 1\right),$$

and the log-likelihood based on a random sample of observations $\{(y_i, z_i) : i = 1, ..., n\}$ is

$$\log L_n\left(\theta\right) = \sum_{i=1}^{n} \log \ell\left(\theta; y_i, z_i\right).$$

Then, under adequate rank or support conditions for the particular parametric specification, there may be point identification so that $\mathcal{D}^0(\mathcal{Z})$ is singleton, and θ may be consistently estimated by maximum likelihood. If these conditions do not hold, $\mathcal{D}^0(\mathcal{Z})$ may then still be a proper set, despite the characterization via moment *equalities* or likelihood functions.

Example 1 (continued): The sets on the support of $\mathcal{T}_4(Y,z;h)$ are the subsets of the level sets $\{\mathcal{L}(y,z;h) : y \in \mathcal{Y}\}$ on which $\mathcal{Y}(z,U;h)$ is singleton. In the top two panels of Figure 1 each set $\mathcal{T}_4(y,z;h)$ is given by the light blue portion of the set $\mathcal{L}(y,z;h)$; the dark blue region in the center belong to no $\mathcal{T}_4(y,z;h)$. This differs from the first three approaches, which allow observations to be generated from realizations of U in this region, and hence assign the dark blue region to sets $\mathcal{T}(y,z;h)$ for which $Y = y$ is one of the possible outcomes. When the structural relation h is complete, as in the bottom panels, the sets $\{\mathcal{T}_4(y,z;h) : y \in \mathcal{Y}\}$ are the same in this case as in the second approach, considered in Section 3.2. Thus the second and fourth approaches handle incompleteness identically, but differ in the way that they deal with incoherence. ■

4 Core-Determining Sets

Section 3 above characterizes the identified set for (h, G_U) under various restrictions on the observability of outcomes when the model is incoherent or incomplete. In each case the identified set can be written as the set of (h, G_U) satisfying the conditional moment inequalities

$$\tilde{G}_U(\mathcal{S}; h, z) \geq C_{\mathcal{T}(Y,Z;h)}(\mathcal{S}|z), \tag{4.1}$$

for all closed $\mathcal{S} \subseteq \mathcal{U}$ and almost every $z \in \mathcal{Z}$, where $\tilde{G}_U(\cdot; h, z)$ is a probability measure on $\mathcal{U}$ and $C_{\mathcal{T}(Y,Z;h)}(\mathcal{S}|z)$ is the conditional containment functional of a random set $\mathcal{T}(Y,Z;h)$. The measure $\tilde{G}_U(\cdot; h, z)$ and the random set $\mathcal{T}(Y,Z;h)$ differ with the imposed restrictions.

We now show that in the context of any particular model it can be sufficient to work with the inequalities (4.1) over a smaller collection of test sets $\mathcal{S}$. Specifically, we provide for each (h,z) a *core-determining* collection of sets, $\mathsf{Q}(h,z)$, such that if (4.1) holds for all $\mathcal{S} \in \mathsf{Q}(h,z)$, then it also holds for all closed $\mathcal{S} \subseteq \mathcal{U}$. Therefore (h, G_U) belongs to the identified set if and only if (4.1) holds for all $\mathcal{S} \in \mathsf{Q}(h,z)$, a.e. $z \in \mathcal{Z}$. The use of core-determining sets for identification analysis was initially put forward by Galichon and Henry (2011), who characterize core-determining classes of sets for incomplete models that satisfy a monotonicity requirement, which we do not require here. We further extend their definition of such collections by allowing them to be specific to the structural relation h and covariate value z.

The logic of our approach is similar to that of Chesher, Rosen, and Smolinski (2011) (CRS),although the analysis there features instruments that are independent of unobserved heterogeneity, and some of which are excluded from the structural relation. Here the covariates Z and unobserved heterogeneity U are also independently distributed, but Z is an argument of the structural relation. As a result, when conditioning on $Z = z$, the support of the random set $\mathcal{T}(Y, Z; h)$, denoted

$$\mathsf{S}(h, z) \equiv \{\mathcal{T} \subseteq \mathcal{U} : \mathbb{P}[\mathcal{T}(Y, Z; h) = \mathcal{T} | z] > 0\},$$

varies with z. This raises the question of how the support $\mathsf{S}(h, z)$ conditional on different values of z can be combined to produce core-determining sets, and in particular whether one needs to consider as test sets unions of sets in $\mathsf{S}(h, z)$ across different values of z. Lemma 1 below addresses this question.

Our construction first requires some further notation. We use $\tilde{G}_U$ to denote a probability measure on U, which may depend on both h and z, and $\mathcal{T}(Y, Z; h)$ a random set. The results in this section can be applied to any of the approaches laid out in Section 3 by taking $\tilde{G}_U$ and $\mathcal{T}(Y, Z; h)$ as the probability measures and random sets appearing in the characterizations of identified sets for each of the various approaches. For any set $\mathcal{S} \subseteq \mathcal{U}$ and any $(h, z) \in \mathcal{H} \times \mathcal{Z}$, define

$$\mathsf{T}^{\mathcal{S}}(h, z) \equiv \{\mathcal{T} \in \mathsf{S}(h, z) : \mathcal{T} \subseteq \mathcal{S}\}, \qquad \mathsf{T}_{\mathcal{S}}(h, z) \equiv \left\{\mathcal{T} \in \mathsf{S}(h, z) : \tilde{G}_U(\mathcal{T} \cap \mathcal{S}) = 0\right\},$$

which are the sets $\mathcal{T} \in \mathsf{S}(h, z)$ that are contained in $\mathcal{S}$ and that, up to zero measure $\tilde{G}(\cdot)$, do not hit $\mathcal{S}$, respectively. Define

$$\overline{\mathsf{T}}^{\mathcal{S}}(h, z) \equiv \mathsf{S}(h, z) / \left(\mathsf{T}^{\mathcal{S}}(h, z) \cup \mathsf{T}_{\mathcal{S}}(h, z)\right)$$

as those sets $\mathcal{T} \in \mathsf{S}(h, z)$ that belong to neither $\mathsf{T}^{\mathcal{S}}(h, z)$ nor $\mathsf{T}_{\mathcal{S}}(h, z)$. For any sets $\mathcal{A}, \mathcal{B}$, we use "$\mathcal{A} \subset \mathcal{B}$" to mean $\mathcal{A} \subseteq \mathcal{B}$ and $\mathcal{A} \neq \mathcal{B}$. Finally, for convenience we employ the following slight abuse of notation:

$$\forall \tilde{\mathcal{Y}} \subseteq \mathcal{Y}, \qquad \mathcal{T}\left(\tilde{\mathcal{Y}}, z; h\right) \equiv \bigcup_{y \in \tilde{\mathcal{Y}}} \mathcal{T}(y, z; h).$$

Lemma 1 *Let $z \in \mathcal{Z}$, $h \in \mathcal{H}$, $\mathcal{S} \subseteq \mathcal{U}$, and $\tilde{G}_U(\cdot) : \mathcal{B}(\mathcal{U}) \to [0, 1]$ be a probability measure on $(\mathcal{U}, \mathcal{B}(\mathcal{U}))$, where $\mathcal{B}(\mathcal{U})$ denotes the Borel sigma algebra on $\mathcal{U}$. Let $\mathcal{M}_{\mathcal{S}}(h, z)$ denote the union of all sets in $\mathsf{T}^{\mathcal{S}}(h, z)$,*

$$\mathcal{M}_{\mathcal{S}}(h, z) \equiv \bigcup_{\mathcal{T} \in \mathsf{T}^{\mathcal{S}}(h, z)} \mathcal{T}. \tag{4.2}$$

Suppose that

$$\tilde{G}_U(\mathcal{M}_{\mathcal{S}}(h, z)) \geq C_{\mathcal{T}(Y, Z; h)}(\mathcal{M}_{\mathcal{S}}(h, z) | z).$$

Then

$$\tilde{G}_U(\mathcal{S}) \geq C_{\mathcal{T}(Y, Z; h)}(\mathcal{S} | z).$$

Lemma 1 establishes that for any $(h, z) \in \mathcal{H} \times \mathcal{Z}$, in order for the inequalities (4.1) to hold for all closed $\mathcal{S} \subseteq \mathcal{U}$, it suffices to show that they hold only for those sets $\mathcal{S}$ that are unions of sets in $\mathsf{S}(h,z)$. However, the collection of core-determining sets may be further refined. Consider a collection of sets $\mathcal{S}_1,...,\mathcal{S}_J$ and a generic random set $\mathcal{T}$. The probability that $\mathcal{T}$ is contained in $\mathcal{S} \equiv \mathcal{S}_1 \cup \cdots \cup \mathcal{S}_J$:

$$\mathbb{P}\{\mathcal{T} \subseteq \mathcal{S}\} = \begin{pmatrix} \sum_{j=1}^{J} \mathbb{P}\{\mathcal{T} \subseteq \mathcal{S}_j\} + \mathbb{P}\left\{\mathcal{T} \subseteq \mathcal{S} \bigwedge_{j=1,...,J} \mathcal{T} \nsubseteq \mathcal{S}_j\right\} \\ -\sum_{k=2}^{J} (k-1) \cdot \mathbb{P}\{\mathcal{T} \subseteq \mathcal{S}_j \text{ for exactly } k \text{ sets } \mathcal{S}_j\} \end{pmatrix}. \tag{4.3}$$

The following lemma uses (4.3) to provide conditions whereby for any probability measure $\tilde{G}_U$

$$\tilde{G}_U(\mathcal{S}_j) \geq \mathbb{P}\{\mathcal{T} \subseteq \mathcal{S}_j\} \text{ for all } j = 1,...,J \ \Rightarrow \tilde{G}_U(\mathcal{S}) \geq \mathbb{P}\{\mathcal{T} \subseteq \mathcal{S}\}.$$

Lemma 2 *Let $z \in \mathcal{Z}$, $h \in \mathcal{H}$, and $\tilde{G}_U(\cdot) : \mathcal{B}(\mathcal{U}) \to [0,1]$ be a probability measure on $(\mathcal{U}, \mathcal{B}(\mathcal{U}))$, where $\mathcal{B}(\mathcal{U})$ denotes the Borel sigma algebra on $\mathcal{U}$. Suppose that $\mathcal{S} \subseteq \mathcal{U}$ admits the representation $\mathcal{S} = \mathcal{S}_1 \cup \cdots \cup \mathcal{S}_J$ such that*

(i) *$\tilde{G}_U(\mathcal{S}_j) \geq C_{\mathcal{T}(Y,Z;h)}(\mathcal{S}_j|z)$, $j = 1,...,J$;*

(ii) *$\sum_{k=2}^{J} (k-1) \cdot \mathbb{P}\{\mathcal{T} \subseteq \mathcal{S}_j \text{ for exactly } k \text{ sets } \mathcal{S}_j|z\} \geq \mathbb{P}\left\{\mathcal{T} \subseteq \mathcal{S} \bigwedge_{j=1,...,J} \mathcal{T} \nsubseteq \mathcal{S}_j|z\right\}$; and*

(iii) *$\tilde{G}_U(\mathcal{S}_j \cap \mathcal{S}_k) = 0$ for all $j \neq k$.*

Then $\tilde{G}_U(\mathcal{S}) \geq C_{\mathcal{T}(Y,Z;h)}(\mathcal{S}|z)$.

The following theorem combines the two prior lemmata to show that in general only a subset of unions of sets in $\mathsf{S}(h,z)$ are core-determining.

Theorem 5 *Let $\tilde{G}_U(\cdot) : \mathcal{B}(\mathcal{U}) \to [0,1]$ be a probability measure on $(\mathcal{U}, \mathcal{B}(\mathcal{U}))$. Let $\mathcal{Y}$ denote the support of Y, and for any $(h,z) \in \mathcal{H} \times \mathcal{Z}$ define*

$$\mathsf{Q}(h,z) \equiv \left\{\mathcal{T}(\mathcal{Y}', z; h) : \begin{pmatrix} \mathcal{Y}' \subseteq \mathcal{Y} \text{ such that for all nonempty } \tilde{\mathcal{Y}}' \subset \mathcal{Y}', \\ \tilde{G}_U\left(\mathcal{T}\left(\tilde{\mathcal{Y}}', z; h\right) \cap \mathcal{T}\left(\mathcal{Y}'/\tilde{\mathcal{Y}}', z; h\right)\right) > 0 \end{pmatrix}\right\},$$

Then $\tilde{G}_U(\mathcal{S}) \geq C_{\mathcal{T}(Y,Z;h)}(\mathcal{S}|z)$ for all $\mathcal{S} \in \mathsf{Q}(h,z)$ implies that the same inequality holds for all $\mathcal{S} \subseteq \mathcal{U}$, so that the collection of sets $\mathsf{Q}(h,z)$ is core-determining.

Note that from the set $\mathcal{Y}' = \{y\}$ it follows that $\mathsf{S}(h,z) \subseteq \mathsf{Q}(h,z)$. Theorem 5 can be applied by taking $\tilde{G}_U(\cdot)$ corresponding to any of the probability measures on the left hand side of the inequalities defining the identified sets of Theorems 1-4. The theorem then implies that these identified sets are characterized by the set of (h, G_U) that satisfy these inequalities for all core-determining test sets $\mathcal{S} \in \mathsf{Q}(h,z)$, and almost every $z \in \mathcal{Z}$. Moreover, the following corollary shows that some of these inequalities can be replaced by equalities, so that the identified set can be written as a collection of conditional moment inequalities and equalities.

Corollary 1 *Define*

$$\mathsf{Q}^E(h,z) \equiv \{\mathcal{S} \in \mathsf{Q}(h,z) : \forall y \in \mathcal{Y} \text{ either } \mathcal{T}(y,z;h) \subseteq \mathcal{S} \text{ or } \mathcal{T}(y,z;h) \cap \mathcal{S} = \emptyset\}.$$

Then, under the conditions of Theorem 5, the collection of equalities and inequalities

$$\begin{aligned} \tilde{G}_U(\mathcal{S}) &= C_{\mathcal{T}(Y,Z;h)}(\mathcal{S}|z), \text{ all } \mathcal{S} \in \mathsf{Q}^E(h,z), \\ \tilde{G}_U(\mathcal{S}) &\geq C_{\mathcal{T}(Y,Z;h)}(\mathcal{S}|z), \text{ all } \mathcal{S} \in \mathsf{Q}^I(h,z) \equiv \mathsf{Q}(h,z) \backslash \mathsf{Q}^E(h,z). \end{aligned}$$

holds if and only if $\tilde{G}_U(\mathcal{S}) \geq C_{\mathcal{T}(Y,Z;h)}(\mathcal{S}|z)$ *for all* $\mathcal{S} \in \mathsf{Q}(h,z)$.

Below we provide an algorithm to compute $\mathsf{Q}^I(h,z)$ and $\mathsf{Q}^E(h,z)$ for any h, z and $\tilde{G}(\cdot) = \tilde{G}_U(\cdot; h, z)$. We make use of the operator $\circledast$, for any collections of sets C_1 and C_2 defined as:

$$\mathsf{C}_1 \circledast \mathsf{C}_2 \equiv \left\{ \begin{array}{c} \mathcal{S}_1 \cup \mathcal{S}_2 : \mathcal{S}_1 \in \mathsf{C}_1,\ \mathcal{S}_2 \in \mathsf{C}_2,\ \mathcal{S}_1 \nsubseteq \mathcal{S}_2, \\ \mathcal{S}_2 \nsubseteq \mathcal{S}_1,\ \tilde{G}\{\mathcal{S}_1 \cap \mathcal{S}_2\} > 0 \end{array} \right\}.$$

Algorithm for the construction of core-determining sets $\mathsf{Q}(h,z)$.

1. Initialization. Set $\mathsf{Q}^*(h,z) = \mathsf{S}(h,z)$, $\mathsf{Q}^I(h,z) = \emptyset$, and $\mathsf{Q}^E(h,z) = \emptyset$.

2. For each $\mathcal{S} \in \mathsf{Q}^*(h,z)$ compute $\overline{\mathsf{T}}^{\mathcal{S}}(h,z)$ and set

$$\begin{aligned} \overline{\mathsf{Q}}^*(h,z) &\equiv \left\{\mathcal{S} \in \mathsf{Q}^*(h,z) : \overline{\mathsf{T}}^{\mathcal{S}}(h,z) = \emptyset\right\}, \\ \underline{\mathsf{Q}}^*(h,z) &\equiv \left\{\mathcal{S} \in \mathsf{Q}^*(h,z) : \overline{\mathsf{T}}^{\mathcal{S}}(h,z) \neq \emptyset\right\}. \end{aligned}$$

   (a) Set $\mathsf{Q}^E(h,z) = \mathsf{Q}^E(h,z) \cup \overline{\mathsf{Q}}^*(h,z)$.

   (b) Set $\mathsf{Q}^I(h,z) = \mathsf{Q}^I(h,z) \cup \underline{\mathsf{Q}}^*(h,z)$.

3. Set $\mathsf{Q}^*(h,z) = \underline{\mathsf{Q}}^*(h,z) \otimes \mathsf{S}(h,z)$.

   (a) If $\mathsf{Q}^*(h,z) = \emptyset$, stop. Set $\mathsf{Q}(h,z) = \mathsf{Q}^I(h,z) \cup \mathsf{Q}^E(h,z)$.

   (b) If $\mathsf{Q}^*(h,z) \neq \emptyset$, return to step 2.

In the Appendix we prove that for any (h, z) the above algorithm produces precisely the sets $\mathsf{Q}(h, z)$ characterized in Theorem 5. Initialization and the first iteration guarantees that all sets $\mathsf{S}(h, z)$ belong to $\mathsf{Q}(h, z)$ produced by the algorithm. All other core-determining sets characterized by the theorem are picked up in successive iterations. Moreover, in the process of constructing the core-determining sets, those which produce moment equalities as characterized in Corollary 1 are detected in step 2a. At termination the algorithm provides the collection of test sets $\mathcal{S} \in \mathsf{Q}^I(h, z)$ for which the containment inequality $\tilde{G}_U(\mathcal{S}) \geq C_{\mathcal{T}(Y,Z;h)}(\mathcal{S}|z)$ must hold, as well as test sets $\mathcal{S} \in \mathsf{Q}^E(h, z)$ for which the containment inequality must in fact hold with equality.

The algorithm extends that of CRS for single agent discrete choice models in a number of significant ways. First, as previously discussed, CRS focus primarily on cases where there is an instrument Z that is excluded from the structural relation relating outcomes (choices) to observed covariates X. Here the variable Z is an argument of the structural relation, so that the *support* of the random sets $\mathcal{T}(Y, Z; h)$ given $Z = z$ varies with z. Thus the algorithm in this paper must be applied separately for each $z \in \mathcal{Z}$. Second, the core-determining sets can vary with the probability measure $\tilde{G}(\cdot; h, z)$, because different measures may assign different probabilities to areas of overlap between sets $\mathcal{T}(y, z; h)$ and $\mathcal{T}(y', z; h)$, $y \neq y'$. If however $\tilde{G}(\cdot; h, z)$ is absolutely continuous with respect to Lebesgue measure, then whether the intersection of such regions has zero measure will not vary with $\tilde{G}(\cdot; h, z)$. The restrictions of CRS, and indeed most econometric models impose this, but it is not required here. Third, we do not require that the sets in $\mathsf{S}(h, z)$ are connected, as was required in CRS. Finally, as already described, the algorithm produces a collection of conditional moment inequalities *and equalities*. Conditions that deliver equalities are not provided in CRS.[5]

4.1 Illustration of Core-Determining Sets and Moment Conditions in the Simultaneous Binary Model

We now consider core-determining sets and collections of moment equalities and inequalities produced using each of the four approaches in section 3 in the context of Example 1. All conditional moment restrictions are to be understood as holding for almost every $z \in \mathcal{Z}$.

Approach 1: Observed Null Outcomes

In this case, if the structural relation h is incoherent, then a null outcome is observed with positive probability. Therefore, based on whether $\mathbb{P}[Y = \phi|Z = z] > 0$, the researcher can determine the sign of the product of the interaction parameters, i.e. the sign of $\delta_1\delta_2$. If $\mathbb{P}[Y = \phi|Z = z] > 0$ the sets $\mathcal{T}_1(y, z; h)$ coincide with the level sets $\mathcal{L}(y, z; h)$ for all $y \in \mathcal{Y}^*$. The level sets are disjoint, so that by Theorem 5 and Corollary 1, or equivalently, the algorithm, the identified set is characterized by a collection of moment *equalities*, one for each regions $\mathcal{L}(y, z; h)$, $y \in \mathcal{Y}^*$.

[5]Indeed, in the model of CRS the condition used to strengthen inequalities to equalities appears difficult to motivate, outside the special case where covariates are in fact restricted to be independent of unobserved heterogeneity. On the other hand simultaneous discrete models are well-known to produce some equalities as well as inequalities, at least in special cases such as the simultaneous binary model.

If on the other hand $\mathbb{P}[Y=\phi|Z=z]=0$ then the researcher knows that $\delta_1\delta_2 \geq 0$, so that the level sets correspond to those depicted in one of the top panels of Figure 1. In the first case, where $\delta_1, \delta_2 < 0$, the regions $\mathcal{T}_1((0,1), z; h)$ and $\mathcal{T}_1((0,1), z; h)$ overlap. Application of the algorithm thus produces

$$\mathsf{Q}^I(h,z) = \{\mathcal{T}_1((0,1), z; h), \mathcal{T}_1((1,0), z; h)\},$$

which yield moment inequalities, as well as the sets

$$\mathsf{Q}^E(h,z) = \{\mathcal{T}_1((0,0), z; h), \mathcal{T}_1((1,1), z; h), \mathcal{T}_1(\{(0,1),(1,0)\}, z; h)\},$$

which correspond to moment equalities. Applying the definition of the sets $\mathcal{T}_1(y,z;h)$ in (3.6), the inequality $G_U(\mathcal{S}) \geq C_{\mathcal{T}_1(Y,Z;h)}(\mathcal{S}|z)$ in Theorem 1, and the definition of the containment functional from (3.3), this gives:

$$\begin{aligned}
G_U(\mathcal{L}((0,0), z; h)) &= f^0_{Y|Z}((0,0)|z), \\
G_U(\mathcal{L}((0,1), z; h)) &\geq f^0_{Y|Z}((0,1)|z), \\
G_U(\mathcal{L}((1,0), z; h)) &\geq f^0_{Y|Z}((1,0)|z), \\
G_U(\mathcal{L}((1,1), z; h)) &= f^0_{Y|Z}((1,1)|z), \\
G_U(\mathcal{L}((0,1), z; h) \cup \mathcal{L}((1,0), z; h)) &= f^0_{Y|Z}((0,1)|z) + f^0_{Y|Z}((1,0)|z).
\end{aligned}$$

The three moment equalities generated in the simultaneous binary model were used to construct a likelihood based procedure by Bresnahan and Reiss (1991), and Tamer (2003) first illustrated how the additional moment inequalities could be used to provide additional information for estimation and inference.

Analogous reasoning for the case depicted in the upper right panel of Figure 1, where $\delta_1, \delta_2 > 0$ produces the system of conditional moment restrictions:

$$\begin{aligned}
G_U(\mathcal{L}((0,0), z; h)) &\geq f^0_{Y|Z}((0,0)|z), \\
G_U(\mathcal{L}((0,1), z; h)) &= f^0_{Y|Z}((0,1)|z), \\
G_U(\mathcal{L}((1,0), z; h)) &= f^0_{Y|Z}((1,0)|z), \\
G_U(\mathcal{L}((1,1), z; h)) &\geq f^0_{Y|Z}((1,1)|z), \\
G_U(\mathcal{L}((0,0), z; h) \cup \mathcal{L}((1,1), z; h)) &= f^0_{Y|Z}((0,0)|z) + f^0_{Y|Z}((1,1)|z),
\end{aligned}$$

since in this case the regions $\mathcal{L}((1,1), z; h)$ and $\mathcal{L}((0,0), z; h)$ overlap, while $\mathcal{L}((0,1), z; h)$ and $\mathcal{L}((0,1), z; h)$ overlap no other region. Note that in both cases, if G_U is parametrically specified the probabilities $G_U(y,z;h)$ can be calculated as a function of model parameters.

Approach 2: Truncation of Null Outcomes

This and each of the remaining approaches apply when $Y = \phi$ is never realized in the data. Thus the researcher cannot immediately distinguish whether the structural relation is incoherent as in Approach 1 from $\mathbb{P}[Y = \phi | Z = z]$, since it is always zero.

First consider the case of a structural relation satisfying $\delta_1 \delta_2 \geq 0$, as in the top panels of Figure 1. For any such h this approach produces precisely the same sets $\{\mathcal{T}_2(y, z; h) : y \in \mathcal{Y}\}$ as in Approach 1, and hence the same core-determining sets $\mathsf{Q}(h, z)$. Thus any pair (h, G_U) with incomplete h (i.e. $\delta_1 \delta_2 > 0$) is in the identified set if and only if the same moment restrictions laid out in Approach 1 are satisfied.

The treatment differs from Approach 1 for incoherent h. The support of $\mathcal{T}_2(Y, z; h)$ is now $\{\mathcal{L}(y, z; h) : y \in \mathcal{Y}\}$ i.e. all sets $\mathcal{L}(y, z; h)$ taken over $\mathcal{Y}$ rather than over the extended support $\mathcal{Y}^*$ which included ϕ. None of these sets overlap, and so the algorithm produces core-determining sets that are simply the level sets

$$\mathsf{Q}^E(h, z) = \{\mathcal{L}((0,0), z; h), \mathcal{L}((0,1), z; h), \mathcal{L}((1,0), z; h), \mathcal{L}((1,1), z; h)\}, \quad \mathsf{Q}^I(h, z) = \emptyset.$$

Application of Theorem 2 then produces the moment conditions

$$\forall y \in \mathcal{Y},\ G_U(\mathcal{L}(y, z; h) | U \notin \mathcal{L}(\emptyset, z; h)) = f^0_{Y|Z}(y|z).$$

Once again, in a model where G_U is parametrically specified, the probabilities on the left can be computed as a function of model parameters.

Approach 3: Anything Goes

The case of functions h satisfying $\delta_1 \delta_2 > 0$ is again treated identically as in approaches 1 and 2, yet functions h such that $\delta_1 \delta_2 < 0$, and which are hence incoherent, are treated differently. The support of the random set $\mathcal{T}_3(Y, z; h)$ given in (3.21) produces the core-determining sets

$$\mathsf{Q}^E(h, z) = \emptyset,$$

and

$$\mathsf{Q}^I(h, z) = \left\{ \begin{array}{c} \mathcal{T}_3((0,0), z; h), \mathcal{T}_3((0,1), z; h), \mathcal{T}_3((1,0), z; h), \mathcal{T}_3((1,1), z; h), \\ \mathcal{T}_3((0,0), z; h) \cup \mathcal{T}_3((0,1), z; h), \mathcal{T}_3((0,0), z; h) \cup \mathcal{T}_3((1,0), z; h), \\ \mathcal{T}_3((0,0), z; h) \cup \mathcal{T}_3((1,1), z; h), \mathcal{T}_3((0,1), z; h) \cup \mathcal{T}_3((0,1), z; h), \\ \mathcal{T}_3((0,1), z; h) \cup \mathcal{T}_3((1,1), z; h), \\ \mathcal{T}_3((0,0), z; h) \cup \mathcal{T}_3((0,1), z; h) \cup \mathcal{T}_3((1,0), z; h), \\ \mathcal{T}_3((0,0), z; h) \cup \mathcal{T}_3((0,1), z; h) \cup \mathcal{T}_3((1,1), z; h), \\ \mathcal{T}_3((0,0), z; h) \cup \mathcal{T}_3((1,0), z; h) \cup \mathcal{T}_3((1,1), z; h), \\ \mathcal{T}_3((0,1), z; h) \cup \mathcal{T}_3((1,0), z; h) \cup \mathcal{T}_3((1,1), z; h) \end{array} \right\}.$$

The resulting moment inequalities are given by

$$G_U\left(\mathcal{T}_3\left(\mathcal{Y}', z; h\right), z; h\right) \geq \sum_{y\in\mathcal{Y}'} f^0_{Y|Z}\left(y|z\right),$$

for all subsets $\mathcal{Y}' \subseteq \mathcal{Y}$. Because for each $y \in \mathcal{Y}$ the region $\mathcal{T}_3(y, z; h)$ contains $\mathcal{L}(\emptyset, z; h)$, all regions $\mathcal{T}_3(y, z; h)$ overlap with one another, so that all unions of subsets of the support of $\mathcal{T}_3(Y, z; h)$ belong to $\mathsf{Q}^I(h, z)$, resulting in 13 moment inequalities and no moment equalities.[6]

Approach 4: Truncated Non-Unique Outcomes

In this case the distribution of the data is assumed to be a truncated version of the population distribution, where conditional on any $Z = z$ regions of the unobservable that result in non-unique outcomes are truncated. By application of Theorem 4 we have the moment equalities

$$\forall y \in \mathcal{Y},\ G_U\left(\mathcal{T}_4(y, z; h)\,|\#\mathcal{Y}(z, U, h) = 1\right) = f^0_{Y|Z}(y|z),$$

where $\mathcal{T}_4(y, z; h)$ is as defined in (3.22). The left hand side probabilities are those given by the distribution G_U conditional on U not belonging to the center box in the panels of Figure 1 in all cases, incomplete or incoherent.

5 Identification in Complete Models

In this section we specialize our results to *complete* models, though we continue to allow the possibility that the model is incoherent. The resulting characterization follows by application of Theorem 5 and Corollary 1. We provide the formal result as a theorem to highlight the application of our methodology to such cases, and in order to help illustrate the relation between our characterization of identified sets in incomplete models to more familiar characterizations via moment equalities used to deliver point identification in complete models.

Theorem 6 *Suppose that, as in Theorems 1-4, the identified set for (h, G_U) is given by*

$$\mathcal{D}^0(\mathcal{Z}) \equiv \left\{(h, G_U) \in \mathcal{H}\times G_U : \forall \mathcal{S} \in \mathsf{F}(\mathcal{U}),\ \tilde{G}_U(\mathcal{S}; h, z) \geq C_{\mathcal{T}(Y,Z;h)}(\mathcal{S}|z),\ a.e.\ z \in \mathcal{Z}\right\}, \quad (5.1)$$

where for each $(h, G_U) \in \mathcal{H}\times G_U$, $\tilde{G}_U(\cdot; h, z)$ is a known functional of probability measure G_U, and is itself a probability measure on $(\mathcal{U}, \mathcal{B}(\mathcal{U}))$. Suppose in addition that the available restrictions require the model to be complete. Then the identified set is equivalently characterized by the following system of equalities:

$$\mathcal{D}^0(\mathcal{Z}) = \left\{(h, G_U) \in \mathcal{H}\times G_U : \forall y \in \mathcal{Y},\ \tilde{G}_U(\mathcal{T}(y, z; h); h, z) = \mathbb{P}[Y = y|Z = z],\ a.e.\ z \in \mathcal{Z}\right\}.$$

[6] There are $2^4 - 1$ subsets of the support of $\mathcal{T}_3(Y, z; h)$, since $\#\mathcal{Y} = 4$. These subsets include the null set and the set itself, which both result in uninformative inequalities given that G_U is a probability measure.

This result is straightforward. When the structural relation is complete the event$\{Y = y\}$, conditional on $Z = z$, occurs if and only if $U \in \mathcal{T}(y, z; h)$. This results in the conditional moment equalities in the characterization delivered by the theorem.

The term $\tilde{G}_U(\mathcal{T}(y, z; h); h, z)$ appearing on the left hand side here gives the probability distribution of the structural relation delivered by the probability distribution $\tilde{G}_U(\cdot; h, z)$, as

$$\tilde{G}_U(\mathcal{T}(y, z; h); h, z)) = \mathbb{P}\{h(y, z, U) = 0\} = \mathbb{P}[Y = y|Z = z]$$

where the uniqueness of the solution for y is guaranteed by completeness. The structural relation $h(y, z, U) = 0$ now delivers a structural *function* for Y as a function of (U, Z), and $\mathbb{P}[Y = y|Z = z]$ gives the distribution of that structural function conditional on $Z = z$.

We now return to our examples to illustrate the implications of completeness.

Example 1 (continued): Consider again Example 1, where Approach 1 is used, with $\mathbb{P}[Y = \phi|Z = z] > 0$, that is where null outcomes are explicitly observed in the data. As discussed in Section 4.1, the model is complete over the extended support $\mathcal{Y}^*$. From Theorem 6 the identified set is given by those (h, G_U) pairs that satisfy

$$\forall y \in \mathcal{Y}^*,\ G_U(\mathcal{L}(y, z; h)) = f^0_{Y|Z}(y|z),\ \text{a.e. } z \in \mathcal{Z}.$$

The probabilities on the left, $G_U(\mathcal{L}(y, z; h))$, are the probabilities that U falls within the five regions depicted in the bottom panels of Figure 1. ■

Example 2 (continued): Consider again the discrete outcome triangular model of Example 2. For any realization of Z_2, the function g_2 maps from the realization of U_2 to a unique realization of Y_2. For any realizations of Y_2, Z_1, and U_2, g_1 maps to a unique realization of Y_1. Hence the model is complete, with corresponding equality restrictions

$$\forall y \in \mathcal{Y},\ G_U(\mathcal{L}(y, z; h)) = f^0_{Y|Z}(y|z),\ \text{a.e. } z \in \mathcal{Z}. \tag{5.2}$$

Figure 2 depicts the regions $\mathcal{L}(y, z; h)$ in the case where both Y_1 and Y_2 are binary, when the g functions are given by the linear threshold crossing functions in (2.3) and (2.4). The figure shows these regions for a single value of z when $\delta_1 < 0$. Heckman (1978) provides conditions for point identification in the case where G_U is a bivariate normal distribution function with mean zero and unknown but suitably normalized variance. For that model (5.2) corresponds precisely to the resulting likelihood equations.

It is worth noting again that completeness alone does not guarantee identification. Even in the binary triangular model, if there are no parametric restrictions placed on the functions g_1 and g_2 or the distribution of U, there in general will not be point identification, although the equalities (5.2) still provide a characterization of the identified set. Chesher (2005) and Jun, Pinkse, and Xu

(2011) provide set identification results for such models. ■

6 Conclusion

Incoherent and incomplete models are well-known to cause problems for identification of structural parameters. In this paper we have used random set theory to characterize identified sets under a variety of different approaches to dealing with these issues in simultaneous discrete outcome models. One of these approaches presents a new way for dealing with incoherence and incompleteness, using recently developed methods for dealing with incompleteness, and a truncation approach for incoherence. We showed how the conditional containment functional of a particular random set can be used to characterize identified sets under each of the different approaches.

Our most general results delivered characterizations consisting of large collections of conditional moment inequalities. We then used the notion of core-determining sets to refine these characterizations. This was achieved by establishing that in the context of any particular model one can rely on a much smaller collection of conditional moment inequalities for characterization of the identified set. In addition, we showed that some of these conditional moment inequalities must in fact hold with equality. We provided an algorithm for constructing core-determining collections of sets in these models, and the algorithm distinguishes which of these correspond to moment inequalities and equalities in the characterization of the identified set.

Although it has been previously recognized that some incomplete models may produce some moment equality restrictions, general characterizations from the recent literature involve only moment inequalities. Our conditions for strengthening such inequalities to equalities are generally applicable to any simultaneous discrete outcome model. These conditions and our algorithm for producing the system of moment equalities and inequalities in such models is a novel contribution. Recognition of the moment equalities can potentially be of use in deriving conditions for point identification and for estimation. Indeed, we showed that in the special case where the model is complete, our characterization produces precisely the collection of conditional moment equalities obtained from classical analysis, for instance likelihood contributions in fully parametric models, and in these settings conditions for point identification and consistent estimation are well understood.

While our focus has been on the effect of incoherence and incompleteness on identification, our characterizations of identified sets as collections of moment equalities and inequalities lend themselves to recently developed inference methods such as those of Andrews and Shi (2009) and Chernozhukov, Lee, and Rosen (2009). There are however interesting and challenging complications for estimation, inference, and computation that may arise from the sheer number of restrictions when the support of the outcome space is large, as well as from conditioning variables with high dimension. These are important considerations, which, while beyond the scope of this paper, warrant further research.

Appendix A: Auxiliary Lemma

In this section we provide an auxiliary lemma used to prove theorems appearing in the main text.

Lemma 3 *Let $(Y, Z, U) \in \mathcal{Y} \times \mathcal{Z} \times \mathcal{U}$ be random vectors defined on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$, where $\mathcal{Y}$ is a finite set, such that for some function $h : \mathcal{Y} \times \mathcal{Z} \times \mathcal{U} \to \mathbb{R}$, $\mathbb{P}[h(Y, Z, U) = 0] = 1$. For each $z \in \mathcal{Z}$, let $\tilde{G}_U(\mathcal{S}; h, z)$ denote the probability of the event $\{U \in \mathcal{S}\}$ conditional on $Z = z$ for any $\mathcal{S} \subseteq \mathcal{U}$. Let $C_{\mathcal{T}(Y,Z;h)}(S; h, z) \equiv \mathbb{P}[\mathcal{T}(Y, Z; h) \subseteq \mathcal{S}|Z = z]$ denote the containment functional of the set $\mathcal{T}(Y, Z; h) \equiv \{u : h(Y, Z, u) = 0\}$ conditional on $Z = z$, and let $\mathsf{F}(\mathcal{U})$ denote the collection of all closed subsets of $\mathcal{U}$. Then if*

$$\forall \mathcal{S} \in \mathsf{F}(\mathcal{U}), \tilde{G}_U(\mathcal{S}; h, z) \geq C_{\mathcal{T}(Y,Z;h)}(\mathcal{S}; h, z), \text{ a.e. } z \in \mathcal{Z}. \tag{6.1}$$

there exist random variables $\left(\tilde{U}, \tilde{Y}\right)$ defined on $(\Omega, \mathcal{F}, \mathbb{P})$ such that (i) for almost every $z \in \mathcal{Z}$, the conditional distribution of $\tilde{Y}$ given $Z = z$ is identical to that of Y given $Z = z$, (ii) with probability one, $h\left(\tilde{Y}, Z, \tilde{U}\right) = 0$.

Proof. *Conditional on $Z = z$, $\mathcal{T}(Y, Z; h)$ has finite support given by $\{\mathcal{T}(y, z; h) : y \in \mathcal{Y}\}$. Since any subset of a finite set, in this case $\{\mathcal{T}(y, z; h) : y \in \mathcal{Y}\}$, is closed in the discrete topology, $\mathcal{T}(Y, z; h)$ with Y conditionally distributed $f^0_{Y|Z}(Y|z)$ is a random closed set.*

We now show in similar fashion to that of Theorem 2.1 of Beresteanu, Molchanov, and Molinari (2012), which applies to random compact sets, and as also shown in Chesher, Rosen, and Smolinski (2011) that (6.1) is equivalent to the conditional capacity functional inequalities

$$\forall \mathcal{S} \in \mathsf{K}(\mathcal{U}), \tilde{G}_U(\mathcal{S}; h, z) \leq \mathbb{P}[\mathcal{T}(Y, Z; h) \cap \mathcal{S} \neq \emptyset|Z = z], \text{ a.e. } z \in \mathcal{Z}, \tag{6.2}$$

where $K(\mathcal{U})$ denotes the collection of all compact subsets of $\mathcal{U}$. This is because for any $z \in \mathcal{Z}$,

$$\tilde{G}_U(\mathcal{S}; h, z) = 1 - \tilde{G}_U(\mathcal{S}^c; h, z),$$

and

$$C_{\mathcal{T}(Y,Z;h)}(\mathcal{S}; h, z) \equiv \mathbb{P}[\mathcal{T}(Y, Z; h) \subseteq \mathcal{S}|Z = z] = 1 - \mathbb{P}[\mathcal{T}(Y, Z; h) \cap \mathcal{S}^c \neq \emptyset|Z = z],$$

where $\mathcal{S}^c$ denotes the complement of $\mathcal{S}$ in $\mathcal{U}$. Therefore (6.1) is equivalent to

$$\forall \mathcal{S} \in \mathsf{F}(\mathcal{U}), \tilde{G}_U(\mathcal{S}^c; h, z) \leq \mathbb{P}[\mathcal{T}(Y, Z; h) \cap \mathcal{S}^c \neq \emptyset|Z = z], \text{ a.e. } z \in \mathcal{Z},$$

equivalently

$$\forall \mathcal{S} \in \mathsf{G}(\mathcal{U}), \tilde{G}_U(\mathcal{S}; h, z) \leq \mathbb{P}[\mathcal{T}(Y, Z; h) \cap \mathcal{S} \neq \emptyset|Z = z], \text{ a.e. } z \in \mathcal{Z},$$

where $\mathsf{G}(\mathcal{U})$ *denotes the collection of all open subsets of* $\mathcal{U}$. *By Corollary 1.4.44 of Molchanov (2005) this is in turn equivalent to (6.2).*

Next, it follows from this collection of inequalities and Molchanov (2005, Corollary 1.4.44), see also Artstein (1983) and Norberg (1992), that for each $z \in \mathcal{Z}$, *the probability measure* $\tilde{G}_U(\cdot; h, z)$ *is* $\mathbb{P}[\cdot|Z=z]$*-selectionable. This implies that there exists a random variable* $\tilde{U}$ *and a random set* $\tilde{\mathcal{T}}$ *both defined on* $(\Omega, \mathcal{F}, \mathbb{P})$ *such that for any* $z \in \mathcal{Z}$, *(a)* $\mathbb{P}\left[\tilde{U} \in \tilde{\mathcal{T}}|Z=z\right] = 1$, *(b) The distribution of* $\tilde{U}$ *conditional on* $Z=z$ *is the same as that of* U *conditional on* $Z=z$, *and (c) The distribution of* $\tilde{\mathcal{T}}$ *conditional on* $Z=z$ *is the same as the of* $\mathcal{T}(Y,Z;h)$ *conditional on* $Z=z$. *From (b) it follows that conditional on* $Z=z$ *the random sets* $\mathcal{Y}\left(\tilde{U}, Z; h\right)$ *and* $\mathcal{Y}(U,Z;h)$ *are identically distributed where as in the main text,*

$$\mathcal{Y}(u,z;h) \equiv \{y : h(y,z,u) = 0\}.$$

Furthermore by $\mathbb{P}[h(Y,Z,U)=0]=1$,

$$\mathbb{P}[Y \in \mathcal{Y}(U,Z;h)|Z=z] = \mathbb{P}[h(Y,Z,U)=0|Z=z] = 1, \text{ a.e. } z \in \mathcal{Z},$$

that is Y *is a measurable selection of* $\mathcal{Y}(U,Z;h)$. *The equivalence in distribution of* $\mathcal{Y}\left(\tilde{U}, Z; h\right)$ *and* $\mathcal{Y}(U,Z;h)$ *given* $Z=z$ *then implies that there exists a random variable* $\tilde{Y}$ *with the same distribution as* Y, *conditional on* $Z=z$, *such that* $\mathbb{P}\left[\tilde{Y} \in \mathcal{Y}\left(\tilde{U}, Z; h\right)|Z=z\right] = 1$ *for almost every* $z \in \mathcal{Z}$. *This establishes (i), and (ii) follows because by definition* $\tilde{Y} \in \mathcal{Y}\left(\tilde{U}, Z; h\right)$ *holds if and only if* $h\left(\tilde{Y}, Z, \tilde{U}\right) = 0$, *completing the proof.* ■

Appendix B: Proofs for Section 3

Proof of Theorem 1. *By arguments in the text we have that every* (h, G_U) *in the identified set satisfies the inequalities*

$$\forall \mathcal{S} \in \mathsf{F}(\mathcal{U}), \; G_U(\mathcal{S}) \geq C_{\mathcal{T}_1(Y,Z;h)}(\mathcal{S}|z), \text{ a.e. } z \in \mathcal{Z}.$$

Consider any $(h, G_U) \in \mathcal{H} \times \mathcal{G}_U$ *that satisfies these inequalities, equivalently any* $(h, G_U) \in \mathcal{D}_0(\mathcal{Z})$. *By application of Lemma 3 with* $\tilde{G}_U(\mathcal{S}; h, z) = G_U(S)$, *there exist random variables* $\left(\tilde{U}, \tilde{Y}\right)$ *on* $(\Omega, \mathcal{F}, \mathbb{P})$ *such that* $\forall z \in \mathcal{Z}$, *the conditional distribution of* $\tilde{Y}$ *given* $Z=z$ *is identical to that of* Y *given* $Z=z$ *with* $\mathbb{P}\left[h\left(\tilde{Y}, Z, \tilde{U}\right) = 0\right]$. *Thus every* $(h, G_U) \in D_0(\mathcal{Z})$ *is admissible, which establishes that* $\mathcal{D}_0(\mathcal{Z})$ *is the identified set.* ■

Proof of Theorem 2. *Consider the set*

$$\mathcal{D}^0(\mathcal{Z}) \equiv \left\{(h, G_U) \in \mathcal{H} \times G_U : \forall \mathcal{S} \in \mathsf{F}(\mathcal{U}), \; G_U(\mathcal{S}|\mathcal{Y}(z,U;h) \neq \emptyset) \geq C_{\mathcal{T}_2(Y,Z;h)}(\mathcal{S}|z), \text{ a.e. } z \in \mathcal{Z}\right\},$$

and any $(h, G_U) \in \mathcal{D}_0(\mathcal{Z})$. *By application of Lemma 3 with* $\tilde{G}_U(\mathcal{S}; h, z) = G_U(\mathcal{S}|\mathcal{Y}(z, U; h) \neq \emptyset)$, *and following the same steps as in the proof of Theorem 1 above, it follows that* $\mathcal{D}_0(\mathcal{Z})$ *is the identified set.* ■

Proof of Theorem 3. *The proof is identical to that of Theorem 1, but with* $C_{\mathcal{T}_3(Y,Z;h)}(\mathcal{S}|z)$ *the conditional capacity functional for* $\mathcal{T}_3(Y, Z; h)$ *defined in (3.18).* ■

Proof of Theorem 4. *Using (3.23) from the main text, the identified set is given by*

$$\mathcal{D}^0(\mathcal{Z}) = \left\{(h, G_U) \in \mathcal{H} \times \mathcal{G}_U : \forall \mathcal{S} \in \mathsf{F}(\mathcal{U}), G_U(\mathcal{S}|\#\mathcal{Y}(z, U; h) = 1) \geq C_{\mathcal{T}_4(Y,Z;h)}(\mathcal{S}|z), \text{ a.e. } z \in \mathcal{Z}\right\}.$$

For any $(h, z) \in \mathcal{H} \times \mathcal{Z}$, $\mathsf{S}(h, z) = \{\mathcal{T}(y, z; h) : y \in \mathcal{Y}\}$ *is a partition of* $\{u : \#\mathcal{Y}(z, u; h) = 1\}$ *so that the sets* $\mathcal{T}_4(y, z; h)$ *and* $\mathcal{T}_4(y', z; h)$ *are disjoint, for any* $y \neq y'$. *Thus Theorem 5 implies that*

$$\mathcal{D}^0(\mathcal{Z}) = \left\{\begin{array}{c} \{h, G_U\} \in \mathcal{H} \times \mathcal{G}_U : \\ \forall y \in \mathcal{Y}, G_U(\mathcal{T}(y, z; h)|\#\mathcal{Y}(z, U; h) = 1) \geq C_{\mathcal{T}(Y,Z;h)}(\mathcal{T}(y, z; h)|z), \text{ a.e. } z \in \mathcal{Z} \end{array}\right\}.$$

Proceeding with the same steps as in the proof of Theorem 6 from (6.11) on completes the proof. ■

Appendix C: Proofs for Sections 4 and 5

Proof of Lemma 1. $\mathcal{M}_{\mathcal{S}}(h, z)$ *is a union of sets contained in* $\mathcal{S}$, *so that* $\mathcal{M}_{\mathcal{S}}(h, z) \subseteq \mathcal{S}$, *and therefore*

$$\tilde{G}_U(\mathcal{S}) \geq \tilde{G}_U(\mathcal{M}_{\mathcal{S}}(h, z)).$$

Then $\tilde{G}_U(\mathcal{M}_{\mathcal{S}}(h, z)) \geq C_{\mathcal{T}(Y,Z;h)}(\mathcal{M}_{\mathcal{S}}(h, z)|z)$ *implies that* $\tilde{G}_U(\mathcal{S}) \geq C_{\mathcal{T}(Y,Z;h)}(\mathcal{M}_{\mathcal{S}}(h, z)|z)$. *The latter is*

$$\begin{aligned} C_{\mathcal{T}(Y,Z;h)}(\mathcal{M}_{\mathcal{S}}(h, z)|z) &\equiv \mathbb{P}\{\mathcal{T}(Y, Z; h) \subseteq \mathcal{M}_{\mathcal{S}}(h, z)|Z = z\} \\ &= \sum_{y \in \mathcal{Y}} 1[\mathcal{T}(y, z; h) \subseteq \mathcal{M}_{\mathcal{S}}(h, z)] f^0_{Y|Z}(y|z) \\ &= \sum_{y \in \mathcal{Y}} 1[\mathcal{T}(y, z; h) \subseteq \mathcal{S}] f^0_{Y|Z}(y|z) \\ &= C_{\mathcal{T}(Y,Z;h)}(\mathcal{S}|z), \end{aligned}$$

where the second line follows by the law of total probability, and the third by the definition of $M_{\mathcal{S}}(h, z)$ *in (4.2). This with* $\tilde{G}_U(\mathcal{S}) \geq C_{\mathcal{T}(Y,Z;h)}(\mathcal{M}_{\mathcal{S}}(h, z)|z)$ *establishes* $\tilde{G}_U(\mathcal{S}) \geq C_{\mathcal{T}(Y,Z;h)}(\mathcal{S}|z)$, *completing the proof.* ■

Proof of Lemma 2. *We have that*

$$\tilde{G}_U(\mathcal{S}) = \sum_{j=1}^{J} \tilde{G}_U(\mathcal{S}_j) \geq \sum_{j=1}^{J} C_{\mathcal{T}(Y,Z;h)}(\mathcal{S}_j|z) \geq C_{\mathcal{T}(Y,Z;h)}(\mathcal{S}|z),$$

where the first equality follows from condition (iii) in the statement of the lemma, the first inequality from (i), and second inequality from (ii) and (4.3) taken with $\mathbb{P}\{\cdot|z\}$ *in place of* $\mathbb{P}\{\cdot\}$. ■

Proof of Theorem 5. *Lemma 1 establishes that for any* $\mathcal{S} \subseteq \mathcal{U}$,

$$\tilde{G}_U(\mathcal{M}_{\mathcal{S}}(h,z)) \geq C_{\mathcal{T}(Y,Z;h)}(\mathcal{M}_{\mathcal{S}}(h,z)|z) \Rightarrow \tilde{G}_U(\mathcal{S}) \geq C_{\mathcal{T}(Y,Z;h)}(\mathcal{S}|z).$$

To prove the theorem it therefore suffices to show that

$$\forall \mathcal{S}' \in \mathsf{Q}(h,z), \tilde{G}_U(\mathcal{S}') \geq C_{\mathcal{T}(Y,Z;h)}(\mathcal{S}'|z) \tag{6.3}$$

implies that for any $\mathcal{S} \subseteq \mathcal{U}$,

$$\tilde{G}_U(\mathcal{M}_{\mathcal{S}}(h,z)) \geq C_{\mathcal{T}(Y,Z;h)}(\mathcal{M}_{\mathcal{S}}(h,z)|z), \tag{6.4}$$

where recall that $\mathcal{M}_{\mathcal{S}}(h,z) = \cup\left\{\mathcal{T} \in \mathsf{T}^{\mathcal{S}}(h,z)\right\}$. *The proof proceeds by induction on* $\#\mathsf{T}^{\mathcal{S}}(h,z) = \#\{\mathcal{T} \in \mathsf{S}(h,z) : \mathcal{T} \subseteq \mathcal{S}\}$.

Inductive Hypothesis: *Suppose that* (6.3) *holds. Then for any positive integer* k, *if* $\mathcal{S} \subseteq \mathcal{U}$ *such that* $\#\mathsf{T}_{\mathcal{S}}(h,z) \leq k$, *then* (6.4) *holds.*

Base Case: *Take* $k = 1$ *and suppose that* (6.3) *holds. That* (6.4) *holds for all* $\mathcal{S}$ *with* $\#\mathsf{T}^{\mathcal{S}}(h,z) \leq k$ *is immediate from* $\mathsf{S}(h,z) \subseteq \mathsf{Q}(h,z)$.

Inductive Step: *Suppose that* (6.3) *holds, and that for some positive integer* k *we have that for all* $\mathcal{S}$ *with* $\#\mathsf{T}^{\mathcal{S}}(h,z) \leq k$, (6.4) *holds. We need to prove that this implies that* (6.4) *also holds for all* $\mathcal{S}$ *with* $\#\mathsf{T}^{\mathcal{S}}(h,z) = k+1$.

Consider an arbitrary set $\mathcal{S} \subseteq \mathcal{U}$ *with* $\#\mathsf{T}_{\mathcal{S}}(h,z) = k+1$. *Suppose that* $\mathcal{M}_{\mathcal{S}}(h,z) \notin \mathsf{Q}(h,z)$, *as otherwise* (6.4) *is immediate. Then by the definition of* $\mathsf{Q}(h,z)$, *there exists some nonempty set* $\tilde{\mathcal{Y}} \subseteq \mathsf{T}^{\mathcal{S}}(h,z)$ *such that*

$$\tilde{G}_U\left(\mathcal{T}\left(\tilde{\mathcal{Y}}, z; h\right) \cap \mathcal{T}\left(\tilde{\mathcal{Y}}^c, z; h\right)\right) = 0, \tag{6.5}$$

where $\tilde{\mathcal{Y}}^c \equiv \mathsf{T}^{\mathcal{S}}(h,z)/\mathcal{Y}'$.

We now verify conditions (i)-(iii) of Lemma 2 for the sets $\mathcal{S}_1 = \mathcal{T}\left(\tilde{\mathcal{Y}}, z; h\right)$ *and* $\mathcal{S}_2 = \mathcal{T}\left(\tilde{\mathcal{Y}}^c, z; h\right)$. *Since* $\mathcal{M}_{\mathcal{S}}(h,z) = \mathcal{S}_1 \cup \mathcal{S}_2$, *application of the lemma will imply* (6.4), *as desired. Condition (i) follows from* (6.4) *holding for all* $\mathcal{S}$ *with* $\#\mathsf{T}^{\mathcal{S}}(h,z) \leq k$. *Condition (ii) holds because*

$$\mathcal{T}(Y,Z;h) \subseteq \mathcal{M}_{\mathcal{S}}(h,z) \Rightarrow Y \in \mathsf{T}^{\mathcal{S}}(h,z) \Rightarrow \mathcal{T}(Y,Z;h) \subseteq \mathcal{S}_1 \text{ or } \mathcal{T}(Y,Z;h) \subseteq \mathcal{S}_2.$$

Condition (iii) is precisely (6.5) *above. Thus we can apply Lemma 2 to conclude that* (6.4) *holds for* $\mathcal{M}_{\mathcal{S}}(h,z)$. *Since* k *is arbitrary this completes the proof.* ■

Proof of Corollary 1. *Consider any* $\mathcal{S} \in \mathsf{Q}^E(h,z)$. *Note that*

$$C_{\mathcal{T}(Y,Z;h)}(\mathcal{S}^c|z) = \mathbb{P}(\mathcal{T}(Y,Z;h) \subseteq \mathcal{S}^c|z) = \mathbb{P}(\mathcal{T}(Y,Z;h) \cap \mathcal{S} = \emptyset|z).$$

$\mathcal{S} \in \mathsf{Q}^E(h,z)$ *implies that for all* $y \in \mathcal{Y}$, *either* $\mathcal{T}(y,z;h) \subseteq \mathcal{S}$ *or* $\mathcal{T}(y,z;h) \cap \mathcal{S} = \emptyset$. *Thus*

$$C_{\mathcal{T}(Y,Z;h)}(\mathcal{S}|z) + C_{\mathcal{T}(Y,Z;h)}(\mathcal{S}^c|z) = \mathbb{P}(\mathcal{T}(Y,Z;h) \subseteq \mathcal{S}|z) + \mathbb{P}(\mathcal{T}(Y,Z;h) \cap \mathcal{S} = \emptyset|z) = 1. \quad (6.6)$$

The inequalities of Theorem **5** *imply that*

$$\tilde{G}_U(\mathcal{S}) \geq C_{\mathcal{T}(Y,Z;h)}(\mathcal{S}|z) \text{ and } \tilde{G}_U(\mathcal{S}^c) \geq C_{\mathcal{T}(Y,Z;h)}(\mathcal{S}^c|z).$$

Then $\tilde{G}_U(\mathcal{S}) + \tilde{G}_U(\mathcal{S}^c) = 1$ *and* (6.6) *imply that both weak inequalities hold with equality.* ■

Proof of Validity of the Algorithm for Core-Determining Sets. *Fix* $(h,z) \in \mathcal{H} \times \mathcal{Z}$. *Let* $\mathsf{Q}(h,z)$ *denote the core-determining sets characterized by Theorem 5 as in the main text, and let* $\mathsf{Q}^A(h,z)$ *denote the collections of sets produced by the algorithm. We show first that* $\mathsf{Q}(h,z) = \mathsf{Q}^A(h,z)$. *Next we show that the sets* $\mathsf{Q}^E(h,z)$ *from the algorithm coincide with those of Corollary 1.*

Step 1: *We first show that* $\mathcal{S} \in \mathsf{Q}(h,z) \Rightarrow \mathcal{S} \in \mathsf{Q}^A(h,z)$.

Consider a set $\mathcal{S} \in \mathsf{Q}(h,z)$. *Suppose that* $\mathcal{S} \notin \mathsf{S}(h,z)$, *as otherwise* $\mathcal{S} \in \mathsf{Q}^A(h,z)$ *is trivial. Then* $\mathcal{S} = \mathcal{T}(\mathcal{Y}_{\mathcal{S}}, z; h)$ *for some* $\mathcal{Y}_{\mathcal{S}} \subseteq \mathcal{Y}$ *such that for all nonempty* $\tilde{\mathcal{Y}}_{\mathcal{S}} \subset \mathcal{Y}_{\mathcal{S}}$,

$$\tilde{G}_U\left(\mathcal{T}\left(\tilde{\mathcal{Y}}_{\mathcal{S}}, z; h\right) \cap \mathcal{T}\left(\mathcal{Y}_{\mathcal{S}} \backslash \tilde{\mathcal{Y}}_{\mathcal{S}}, z; h\right)\right) > 0. \quad (6.7)$$

Starting with any $y_1 \in \tilde{\mathcal{Y}}_{\mathcal{S}}$, *we can then order the elements of* $\tilde{\mathcal{Y}}_{\mathcal{S}}$ *as* $y_1, ..., y_J$ *such that:*

$$\tilde{G}_U(\mathcal{T}(\mathcal{Y}_k, z; h) \cap \mathcal{T}(y_{k+1}, z; h)) > 0, \text{ where } \forall k = 1, ..., J-1,\ \mathcal{Y}_k \equiv \bigcup_{j=1}^{k} y_j. \quad (6.8)$$

Note that if this were not possible then (6.7) *would be violated with* $\tilde{\mathcal{Y}}_{\mathcal{S}}$ *equal to one of the sets* $\mathcal{Y}_k$, *which would be a contradiction. Thus, starting with* $\mathcal{T}(y_1, z; h)$ *in the first iteration, successive iterations of the algorithm will add the sets* $\mathcal{T}(\mathcal{Y}_k, z; h)$ *to* $\mathsf{Q}^*(h,z)$, *and after no more than J iterations,* $\mathcal{S} = \mathcal{T}(\mathcal{Y}_{\mathcal{S}}, z; h)$ *will be added to* $\mathsf{Q}^*(h,z)$, *guaranteeing that* $\mathcal{S} \in \mathsf{Q}^A(h,z)$.

Step 2: *We now establish that* $\mathcal{S} \notin \mathsf{Q}(h,z) \Rightarrow \mathcal{S} \notin \mathsf{Q}^A(h,z)$.

Let $\mathcal{S} \notin \mathsf{Q}(h,z)$ *be such that for some* $\mathcal{Y}_{\mathcal{S}} \subseteq \mathcal{Y}$, $\mathcal{S} = \mathcal{T}(\mathcal{Y}_{\mathcal{S}}, z; h)$, *as otherwise* $\mathcal{S} \notin \mathsf{Q}(h,z)$ *is immediate. Because* $\mathcal{S} \notin \mathsf{Q}(h,z)$, $\#\mathcal{Y}_{\mathcal{S}} \geq 2$, *and there exists* $\tilde{\mathcal{Y}}_{\mathcal{S}} \subset \mathcal{Y}_{\mathcal{S}}$ *such that*

$$\tilde{G}_U\left(\mathcal{T}\left(\tilde{\mathcal{Y}}_{\mathcal{S}}, z; h\right) \cap \mathcal{T}\left(\mathcal{Y}_{\mathcal{S}} \backslash \tilde{\mathcal{Y}}_{\mathcal{S}}, z; h\right)\right) = 0. \quad (6.9)$$

Therefore, for any y_1 and y_2 with $y_1 \in \tilde{\mathcal{Y}}$ and $y_2 \in \mathcal{Y}_{\mathcal{S}} \backslash \tilde{\mathcal{Y}}_{\mathcal{S}}$,

$$\tilde{G}_U \left(\mathcal{T}\left(y_1, z; h\right) \cap \mathcal{T}\left(y_2, z; h\right) \right) = 0. \tag{6.10}$$

It follows that there is no ordering $y_1, ..., y_J$ of the elements of $\mathcal{Y}_{\mathcal{S}}$ such that (6.8) *holds. This implies that $\mathcal{S} \notin \mathsf{Q}^A\left(h, z\right)$, which completes step 2.*

Step 3: *The sets $\mathsf{Q}^E\left(h, z\right)$ produced by the algorithm coincide with those of Corollary 1.*
This holds because the conditions of the corollary coincide with the condition $\overline{\mathsf{T}}^{\mathcal{S}}\left(h, z\right) = \emptyset$, which is explicitly checked for each $\mathcal{S} \in \mathsf{Q}^*\left(h, z\right)$ in step 2 of the algorithm. ■

Proof of Theorem 6. *Because the model is complete, we have that for any $(h, z) \in \mathcal{H} \times \mathcal{Z}$, $\mathsf{S}\left(h, z\right)$ constitutes a partition of $\mathcal{U}$. Thus the sets $\mathcal{T}\left(y, z; h\right)$ and $\mathcal{T}\left(y', z; h\right)$ are disjoint, for any $y \neq y'$. It follows from Theorem 5 that $\mathsf{S}\left(h, z\right)$ is core-determining for each $(h, z) \in \mathcal{H} \times \mathcal{Z}$, so that the representation of $\mathcal{D}^0(Z)$ given by 5.1 is equivalent to*

$$\mathcal{D}^0(\mathcal{Z}) = \left\{ \{h, G_U\} \in \mathcal{H} \times G_U : \forall y \in \mathcal{Y},\, \tilde{G}_U \left(\mathcal{T}\left(y, z; h\right); h, z \right) \geq C_{\mathcal{T}(Y,Z;h)} \left(\mathcal{T}\left(y, z; h\right) | z \right),\ \text{a.e. } z \in \mathcal{Z} \right\}. \tag{6.11}$$

Using again that for any $y \neq y'$ the sets $\mathcal{T}\left(y, z; h\right)$ and $\mathcal{T}\left(y', z; h\right)$, $y \neq y'$ are disjoint we have that

$$C_{\mathcal{T}(Y,Z;h)} \left(\mathcal{T}\left(y, z; h\right) | z \right) = \mathbb{P}[Y = y | Z = z],$$

and application of Corollary 1 implies that the inequalities of the characterization (6.11) must in fact hold with equality, completing the proof. ■

References

Amemiya, T. (1974): "Multivariate Regression and Simultaneous Equation Models when the Dependent Variables are Truncated Normal," *Econometrica*, 42(6), 999–1012.

Andrews, D. W. K., S. T. Berry, and P. Jia (2004): "Confidence Regions for Parameters in Discrete Games with Multiple Equilibria, with an Application to Discount Chain Store Location," working paper, Yale University.

Andrews, D. W. K., and X. Shi (2009): "Inference for Parameters Defined by Conditional Moment Inequalities," working paper, Cowles Foundation.

Aradillas-Lopez, A., and E. Tamer (2008): "The Identification Power of Equilibrium in Simple Games," *Journal of Business and Economic Statistics*, 26(3), 261–310.

Artstein, Z. (1983): "Distributions of Random Sets and Random Selections," *Israel Journal of Mathematics*, 46(4), 313–324.

Bajari, P., H. Hong, and S. P. Ryan (2010): "Identification and Estimation of Discrete Games of Complete Information," *Econometrica*, 78(5), 1529–1568.

Beresteanu, A., I. Molchanov, and F. Molinari (2011): "Sharp Identification Regions in Models with Convex Moment Predictions," *Econometrica*, 79(6), 1785–1821.

——— (2012): "Partial Identification Using Random Set Theory," *Journal of Econometrics*, 166(1), 17–32.

Bjorn, P., and Q. Vuong (1984): "Simultaneous Equations Models for Dummy Endogenous Variables: A Game Theoretic Formulation with an Application to Labor Force Participation," CIT working paper, SSWP 537.

Blundell, R., and R. J. Smith (1994): "Coherency and Estimation in Simultaneous Models with Censored or Qualitative Dependent Variables," *Journal of Econometrics*, 64, 355–373.

Bresnahan, T. F., and P. J. Reiss (1990): "Entry in Monopoly Markets," *Review of Economic Studies*, 57, 531–553.

——— (1991): "Empirical Models of Discrete Games," *Journal of Econometrics*, 48(1-2), 57–81.

Chernozhukov, V., S. Lee, and A. Rosen (2009): "Intersection Bounds, Estimation and Inference," CeMMAP working paper CWP19/09.

Chesher, A. (2005): "Nonparametric Identification Under Discrete Variation," *Econometrica*, 73, 1525–1550.

Chesher, A., A. Rosen, and K. Smolinski (2011): "An Instrumental Variable Model of Multiple Discrete Choice," CeMMAP working paper CWP39/11.

Ciliberto, F., and E. Tamer (2009): "Market Structure and Multiple Equilibria in Airline Markets," *Econometrica*, 77(6), 1791–1828.

Dagenais, M. (1997): "A Simultaneous Probit Model," working paper, University of Montreal.

Engers, M., and S. Stern (2002): "Family Bargaining and Long Term Care," *International Economic Review*, 43, 73–114.

Galichon, A., and M. Henry (2011): "Set Identification in Models with Multiple Equilibria," *Review of Economic Studies*, 78(4), 1264–1298.

Gourieroux, C., J. J. Laffont, and A. Monfort (1980): "Coherency Conditions in Simultaneous Linear Equation Models with Endogenous Switching Regimes," *Econometrica*, 48(3), 675–695.

Hajivassiliou, V. (2008): "Novel Approaches to Coherency Conditions in LDV Models," working paper, LSE.

Heckman, J. J. (1978): "Dummy Endogenous Variables in a Simultaneous Equation System," *Econometrica*, 46, 931–959.

Jovanovic, B. (1989): "Observable Implications of Models with Multiple Equilibria," *Econometrica*, 161(57), 1431–1437.

Jun, S. J., J. Pinkse, and H. Xu (2011): "Tighter Bounds in Triangular Systems," *Journal of Econometrics*, 161(2), 122–128.

Kline, B., and E. Tamer (2012): "Bounds on Best Response Functions in Binary Games," *Journal of Econometrics*, 166(1), 92–105.

Kooreman, P. (1994): "Estimation of Econometric Models of Some Discrete Games," *Journal of Applied Econometrics*, 9, 255–268.

Lewbel, A. (2007): "Coherency and Completeness of Structural Models Containing a Dummy Endogenous Variable," *International Economic Review*, 48(4), 1379–1392.

Maddala, G. (1983): *Limited-Dependent and Qualitative Variables in Econometrics*. Cambridge University Press, Cambridge, U.K.

Molchanov, I. S. (2005): *Theory of Random Sets*. Springer Verlag, London.

Norberg, T. (1992): "On the Existence of Ordered Couplings of Random Sets – with Applications," *Israel Journal of Mathematics*, 77, 241–264.

Schmidt, P. (1981): "Constraints on Parameters in Simultaneous Tobit and Probit Models," in *Structural Analysis of Discrete Data with Econometric Applications*, ed. by C. F. Manski, and D. L. McFadden, pp. 422–434. MIT Press.

Tamer, E. (2003): "Incomplete Simultaneous Discrete Response Models with Multiple Equilibria," *Review of Economic Studies*, 70(1), 147–167.

Figure 1: This figure depicts the level sets $\mathcal{L}(y) = \mathcal{L}(y, z; h)$ for a single value of z in the simultaneous binary model considered in Example 1. Each panel illustrates a different case for the signs of the interaction parameters δ_1 and δ_2. The dark blue regions indicate places where these sets overlap: in the top left panel this region is the intersection of $\mathcal{L}((0,1))$ and $\mathcal{L}((1,0))$, and in the top right panel this is the intersection of $\mathcal{L}((0,0))$ and $\mathcal{L}((1,1))$. The regions $\mathcal{L}(\emptyset)$ in the bottom panels indicate values of U such that the structural relation $h(y, z, u) = 0$ holds for no possible y.

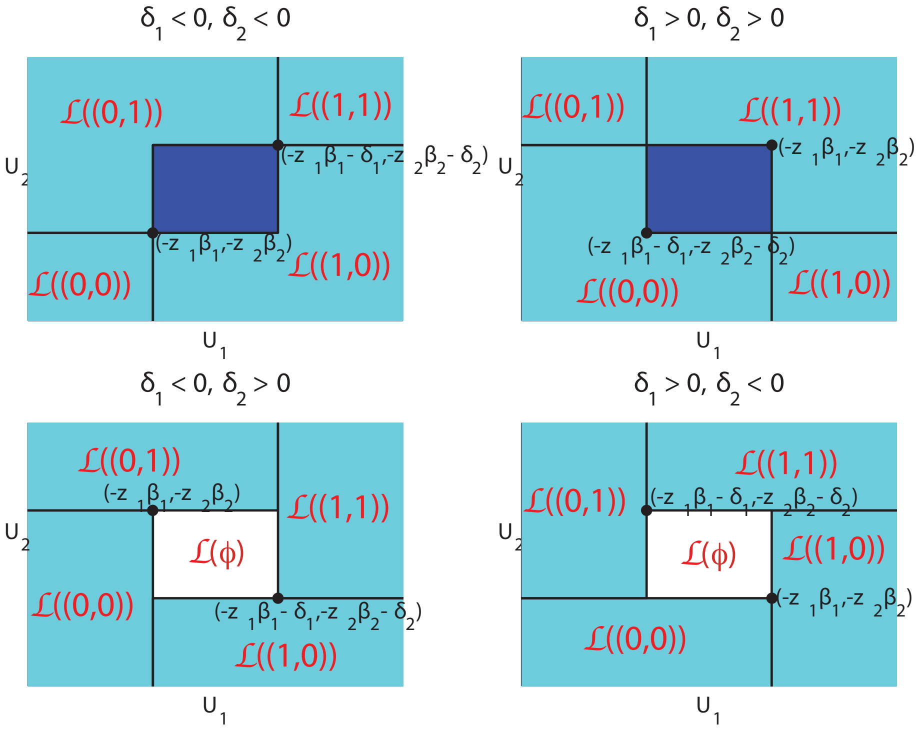

Figure 2: This figure depicts the level sets $\mathcal{L}(y) = \mathcal{L}(y, z; h)$ for a single value of z in the triangular binary model considered in Example 2. Here the model is proper and the level sets partition the support of the unobservable U.