cemmap
centre for microdata methods and practice

Generalized instrumental variable models

Andrew Chesher
Adam Rosen

The Institute for Fiscal Studies
Department of Economics, UCL

cemmap working paper CWP43/13

E·S·R·C ECONOMIC & SOCIAL RESEARCH COUNCIL
An ESRC Research Centre

Generalized Instrumental Variable Models*

Andrew Chesher† UCL and CeMMAP

Adam M. Rosen‡ UCL and CeMMAP

August 29, 2013

Abstract

The ability to allow for flexible forms of unobserved heterogeneity is an essential ingredient in modern microeconometrics. In this paper we extend the application of instrumental variable (IV) methods to a wide class of problems in which multiple values of unobservable variables can be associated with particular combinations of observed endogenous and exogenous variables. In our Generalized Instrumental Variable (GIV) models, in contrast to traditional IV models, the mapping from unobserved heterogeneity to endogenous variables need not admit a unique inverse. The class of GIV models allows unobservables to be multivariate and to enter non-separably into the determination of endogenous variables, thereby removing strong practical limitations on the role of unobserved heterogeneity. Important examples include models with discrete or mixed continuous/discrete outcomes and continuous unobservables, and models with excess heterogeneity where many combinations of different values of multiple unobserved variables, such as random coefficients, can deliver the same realizations of outcomes. We use tools from random set theory to study identification in such models and provide a sharp characterization of the identified set of structures admitted. We demonstrate the application of our analysis to a continuous outcome model with an interval-censored endogenous explanatory variable.

Keywords: instrumental variables, endogeneity, excess heterogeneity, limited information, set identification, partial identification, random sets, incomplete models.

JEL classification: C10, C14, C24, C26.

*We thank participants for comments at seminar and conference presentations given at Boston College, Chicago, Paris School of Economics, Vanderbilt, the April 2013 conference on Mathematical Statistics of Partially Identified Objects at Oberwolfach, and the 2013 Cowles summer econometrics conference. 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 from the European Research Council (ERC) grant ERC-2009-StG-240910-ROMETA is gratefully acknowledged.

†Address: Andrew Chesher, Department of Economics, University College London, Gower Street, London WC1E 6BT, andrew.chesher@ucl.ac.uk.

‡Address: Adam Rosen, Department of Economics, University College London, Gower Street, London WC1E 6BT, adam.rosen@ucl.ac.uk.

1 Introduction

In this paper we extend the application of instrumental variable (IV) methods to a wide class of problems in which multiple values of unobservable variables can be associated with particular combinations of observed endogenous and exogenous variables.

We study models that admit a residual *correspondence* ρ such that

$$U \in \rho(Y, Z), \tag{1.1}$$

where Y, Z denote observed endogenous and exogenous variables, respectively, and where U denotes unobserved heterogeneity. The endogenous variables $Y \equiv (Y_1, Y_2)$ include variables Y_1 modeled as the outcome of an economic process and variables Y_2 that enter into their determination, i.e. "right-hand-side" endogenous variables which are allowed to be stochastically dependent with unobserved heterogeneity. The exogenous variables $Z \equiv (Z_1, Z_2)$ are restricted in the degree to which they are stochastically dependent with U. Variation in variables Z_1, sometimes referred to as "exogenous covariates", may affect $\rho(Y, Z)$, while variables Z_2 are excluded instruments with respect to which $\rho(Y, Z)$ is restricted to be invariant.

The residual correspondence will be defined through a structural relation given by

$$h(Y, Z, U) = 0, \tag{1.2}$$

relating values of observed and unobserved variables. Then $\rho(Y, Z)$ is precisely the set of values of U such that (1.2) holds:

$$\rho(Y, Z) = \mathcal{U}(Y, Z; h) \equiv \{u \in \mathcal{R}_U : (1.2) \text{ holds when } U = u.\}, \tag{1.3}$$

where $\mathcal{R}_U$ denotes the support of U.

On the other hand, IV models typically require that the residual $\rho(Y, Z)$ be unique, so that

$$U = \rho(Y, Z), \tag{1.4}$$

for some *function* $\rho(Y, Z)$. This requires that the structural relation (1.2) produces a unique value of U for almost every realization of (Y, Z). The Generalized Instrumental Variable (GIV) models that give rise to only (1.1) are typically partially identifying although they include as special cases point identifying models, such as classical IV models, for which the solution to (1.2) in U is guaranteed unique. We provide a sharp characterization of the identified set of admissible structures delivered by GIV models, and examples of the sets obtained in particular cases.

The extension of IV methods to models with (1.1) allows for unobservables to be multivariate and to enter nonseparably into the determination of Y. This can be important in practice, as

failure to allow for these possibilities places strong limitations on the role of unobserved heterogeneity, ruling out for example random coefficients and discrete-valued outcomes. Restrictions on the structural relation h that guarantee the existence of a residual function as in (1.4) include additive separability in U, or strict monotonicity of h in scalar U, so that h has a unique inverse. Important cases in which invertibility fails and where GIV models deliver new results include models with discrete or mixed continuous/discrete outcomes and continuous unobservables, and models with excess heterogeneity where many combinations of values of multiple unobserved variables, such as random coefficients, can deliver the same realizations of outcomes. Even in fully parametric non-linear models invertibility can fail, and there may be only set identification.[1]

In this paper we relax this invertibility restriction, thereby substantially increasing the range of problems to which IV models can be applied. In the previous papers Chesher (2010), Chesher and Smolinski (2012), Chesher, Rosen, and Smolinski (2013), Chesher and Rosen (2012, 2013a, 2013b) we have given some results for particular cases in which outcomes are discrete. Here we present a complete development and results for a general case which includes problems in which outcomes may be continuous or discrete. Unlike our previous analyses, we consider conditional mean and conditional quantile restrictions on unobservables given exogenous variables, in addition to stochastic independence restrictions. We also provide a novel result allowing for characterization of the identified set for structural function h when U and Z are independent, but the distribution of U is completely unrestricted.

A simple and perhaps leading example of an econometric model where there is invertibility and that can be written in the more restrictive form (1.4) is the linear model with a single endogenous covariate, where $\rho(Y, Z) = Y_1 - Y_2\beta$. The instrument Z is excluded from the structural relation and hence from ρ, and is assumed exogenous, for example through the conditional mean restriction $E[U|Z] = 0$. Given this, the parameter β, and hence the distribution of U, are identified under the classical rank condition set out by Koopmans, Rubin, and Leipnik (1950).

Many model specifications, both linear and non-linear, as well as parametric, semiparametric, and nonparametric, yield residual functions of the form (1.4). The recent literature on nonparametric IV models achieves identification from (1.4) in conjunction with *completeness* conditions, as in Newey and Powell (2003), Chernozhukov and Hansen (2005), Hall and Horowitz (2005), Chernozhukov, Imbens, and Newey (2007), Blundell, Chen, and Kristensen (2007), Darolles, Fan, Florens, and Renault (2011), and Chen, Chernozhukov, Lee, and Newey (2011). These completeness conditions on the conditional distribution of endogenous covariates Y_2 given exogenous variables Z can be viewed as nonparametric analogs of the classical rank condition, and have been a central focus in recent papers by D'Haultfoeuille (2011), Andrews (2011), and Canay, Santos, and Shaikh (2012).

[1] Examples of parametric non-linear models where there is a residual correspondence as in (1.1) but not a residual function as in (1.4), include those of Ciliberto and Tamer (2009) and Section 4 of Chesher, Rosen, and Smolinski (2013).

The literature on nonparametric IV models has significantly advanced applied researchers' ability to deal with endogeneity beyond the classical linear setup. Although many models produce residual functions of the form (1.4), the requirement of a unique U associated with observed (Y, Z) is not innocuous. When there is no such residual function whereby residual U satisfies some independence restrictions with respect to instrument Z, for example conditional mean independence or stochastic independence, then the usual derivation of identification from completeness conditions does not go through. The requirement that U belong to some residual correspondence as in (1.1) is more generally applicable, but completeness conditions are generally insufficient for point identification, even if they are thought to hold.

In this paper we provide a sharp characterization of the identified set of structures admitted by GIV models where (1.1) holds but (1.4) may not, equivalently IV models where the mapping between outcomes Y and variables (Z, U) need not admit an inverse in U. Like classical IV models however, our models embody two types of restrictions:

1. Restrictions on the joint distribution of unobserved variables and exogenous covariates. In linear models there is typically a conditional mean independence restriction and in nonlinear models there is often a stochastic independence restriction.

2. Restrictions on the way in which exogenous covariates feature in structural relationships. These include exclusion restrictions and possibly other restrictions on structural relationships, for example index or parametric restrictions.

IV models often leave the determination of some endogenous variables unspecified, in which case they are incomplete, limited information models.

A central object in our identification analysis is the random set $\mathcal{U}(Y, Z; h)$ defined in (1.3). We use random set theory methods reviewed in Molchanov (2005) and introduced into econometric identification analysis by Beresteanu, Molchanov, and Molinari (2011), with particular use of a result known as Artstein's Inequality from Artstein (1983). We extend our analysis employing the distribution of random sets in the space of unobserved heterogeneity as in for example Chesher, Rosen, and Smolinski (2013) and Chesher and Rosen (2012b) to the much more general class of GIV models considered here.

The paper proceeds as follows. In Section 2 we provide an informal overview of our results and present some leading examples of GIV models to which our analysis applies. In Section 3 we lay out the formal restrictions of GIV models and provide identification analysis. This includes a new generalization of the classical notion of observational equivalence, e.g. Koopmans and Reiersøl (1950), Hurwicz (1950), Rothenberg (1971), and Bowden (1973) to models where a structure need not generate a unique distribution of outcomes given other observed variables. We use the notion of selectionability from random set theory, and in particular show how it can be applied in the space of unobserved heterogeneity to incorporate restrictions on unobserved variables of the sort

commonly used in econometric models. We demonstrate how this is done in models invoking stochastic independence, conditional mean, and conditional quantile restrictions. In Section 4 we illustrate the set identifying power of GIV models though application of our results to a continuous outcome model with an interval-censored endogenous explanatory variable. Section 5 concludes. All proofs are provided in the Appendix.

Notation: We use capital Roman letters A to denote random variables and lower case letters a to denote particular realizations. For probability measure $\mathbb{P}$, $\mathbb{P}(\cdot|a)$ is used to denote the conditional probability measure given $A = a$. We write $\mathcal{R}_A$ to denote the support of random vector A, and $\mathcal{R}_{A_1 \cdots A_m}$ to denote the joint support of random vectors $A_1, ..., A_m$. $\mathcal{R}_{A_1|a_2}$ denotes the support of random vector A_1 conditional on $A_2 = a_2$. $q_{A|B}(\tau|b)$ denotes the τ conditional quantile of A given $B = b$. $A \perp\!\!\!\perp B$ means that random vectors A and B are stochastically independent. $\emptyset$ denotes the empty set. Script font ($\mathcal{S}$) is reserved for sets, and sans serif font (S) is reserved for collections of sets. The sign $\subseteq$ is used to indicate nonstrict inclusion so "$\mathcal{A} \subseteq \mathcal{B}$" includes $\mathcal{A} = \mathcal{B}$, while "$\mathcal{A} \subset \mathcal{B}$" means $\mathcal{A} \subseteq \mathcal{B}$ but $\mathcal{A} \neq \mathcal{B}$. $\mathcal{A}/\mathcal{B}$ denotes elements of the set $\mathcal{A}$ that do not belong to $\mathcal{B}$, $\mathcal{A}/\mathcal{B} \equiv \{a \in \mathcal{A} : a \notin \mathcal{B}\}$. $\text{cl}(\mathcal{A})$ denotes the closure of $\mathcal{A}$. $C_h(\mathcal{S}|z)$ denotes the containment functional of random set $\mathcal{U}(Y, Z; h)$ conditional on $Z = z$, defined in Section 3.2. The notation $F \lesssim \mathcal{A}$ is used to indicate that the distribution F of a random vector is *selectionable* with respect to the distribution of random set $\mathcal{A}$, as defined in Section 3.1. $1[\mathcal{E}]$ denotes the indicator function, taking the value 1 if the event $\mathcal{E}$ occurs and 0 otherwise. For any real number c, $|c|_+$ and $|c|_-$ denote the positive and negative parts of c, respectively. $\mathbb{R}^m$ denotes m dimensional Euclidean space, and for any vector $v \in \mathbb{R}^m$, $\|v\|$ indicates the Euclidean norm: $\|v\| = \sqrt{v_1^2 + \cdots + v_m^2}$. Finally, in order to deal with sets of measure zero and conditions required to hold almost everywhere, we use the "sup" and "inf" operators to denote "essential supremum" and "essential infimum" with respect to the underlying measure when these operators are applied to functions of random variables (e.g. conditional probabilities, expectations, or quantiles). Thus $\sup_{z \in \mathcal{Z}} f(z)$ denotes the smallest value of $c \in \mathbb{R}$ such that $\mathbb{P}[f(Z) > c] = 0$ and $\inf_{z \in \mathcal{Z}} f(z)$ denotes the largest value of $c \in \mathbb{R}$ such that $\mathbb{P}[f(Z) < c] = 0$.

2 Examples and Informal Overview

The structural function $h : \mathcal{R}_{YZU} \to \mathbb{R}$ such that (1.2) holds with probability one determines the feasible values of endogenous variables Y when $(Z, U) = (z, u)$. We define the zero level sets of h for each $(y, z) \in \mathcal{R}_{YZ}$ and $(z, u) \in \mathcal{R}_{ZU}$, respectively as

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

These level sets are dual to each other in that for all z and all functions h, a value u^* lies in $\mathcal{U}(y^*, z; h)$ if and only if y^* lies in $\mathcal{Y}(u^*, z; h)$. Duality will be exploited to good effect.

In contrast to conventional IV models, GIV models admit structural functions h for which the level sets $\mathcal{U}(y,z;h)$ may have cardinality *exceeding* one. However, in both IV and GIV models $\mathcal{Y}(u,z;h)$ may have cardinality greater than one, in which case the model is incomplete. When this happens there are values y and y' such that the level sets $\mathcal{U}(y,z;h)$ and $\mathcal{U}(y',z;h)$ have non-empty intersection. This occurs in particular when the model does not specify the way in which endogenous explanatory variables Y_2 are determined, even if Y_1 is uniquely determined by (Y_2,Z,U). Some leading examples follow. Each of these may be combined with alternative restrictions on the joint distribution of (U,Z), for example $U \perp\!\!\!\perp Z$, $E[U|Z]=0$, $q_{U|Z}(\tau|Z)=0$, and/or parametric restrictions on the distribution of U.

2.1 Examples

Example 1. *A classical linear IV model with an instrument exclusion restriction has structural function*

$$h(y,z,u) = y_1 - \alpha - \beta y_2 - u,$$

in which case $\mathcal{Y}(u,z;h) = \{((\alpha+\beta y_2+u), y_2) : y_2 \in \mathcal{R}_{Y_2}\}$. *The level set* $\mathcal{U}(y,z;h)$ *is the singleton set* $\{(y_1-\alpha-\beta y_2)\}$. *In this instance the realization of exogenous variables* z *does not enter into* h *and so* $z = z_2$.

Example 2. *A binary outcome, threshold crossing GIV model with* $Y_1 = 1[g(Y_2,Z_1) < U]$ *and* U *normalized uniformly distributed on* $[0,1]$, *as studied in Chesher (2010) and Chesher and Rosen (2013), has structural function*

$$h(y,z,u) = y_1 |u - g(y_2,z_1)|_- + (1-y_1)|u-g(y_2,z_1)|_+,$$

where $y_1 \in \{0,1\}$. *The corresponding level sets are pairs of values of* (y_1,y_2),

$$\mathcal{Y}(u,z;h) = \{(1[g(y_2,z_1) < u], y_2) : y_2 \in \mathcal{R}_{Y_2}\},$$

and intervals

$$\mathcal{U}(y,z;h) = \begin{array}{c} [0, g(y_2,z_1)] \text{ if } y_1 = 0, \\ (g(y_2,z_1), 1] \text{ if } y_1 = 1. \end{array}$$

Example 3. *Multiple discrete choice with endogenous explanatory variables as studied in Chesher, Rosen, and Smolinski (2013). The structural function is*

$$h(y,z,u) = \left| \min_{k \in \{1,\ldots,M\}} \left(\pi_{y_1}(y_2,z_1,u_j) - \pi_k(y_2,z_1,u_k) \right) \right|_-,$$

where $\pi_j(y_2,z_1,u_j)$ *is the random utility associated with choice* $j \in \mathcal{J} \equiv \{1,\ldots,M\}$ *and* $u =$

$(u_1, \ldots, u_M)$ *is a vector of unobserved preference heterogeneity.* Y_1 *is the outcome or choice variable and* Y_2 *are endogenous explanatory variables.* Z_1 *are exogenous variables allowed to enter the utility functions* $\pi_1, ...\pi_M$*, while* Z_2 *are excluded exogenous variables, or instruments. The level sets are thus*

$$\mathcal{Y}(u, z; h) = \left\{ \left(\arg\max_{j \in \mathcal{J}} \pi_j (y_2, z_1, u_j), y_2 \right) : y_2 \in \mathcal{R}_{Y_2} \right\},$$

and

$$\mathcal{U}(y, z; h) = \left\{ u \in \mathcal{R}_U : y_1 = \arg\max_{j \in \mathcal{J}} \pi_j (y_2, z_1, u_j) \right\}.$$

Example 4. *A continuous outcome random coefficients model with endogeneity has structural function*

$$h(y, z, u) = y_1 - (\beta_2 + u_2) y_2 - (\beta_1 + u_1).$$

The random coefficients are $(\beta_2 + U_2)$ *and* $(\beta_1 + U_1)$*, with mean* (β_1, β_2)*. The level sets are*

$$\mathcal{Y}(u, z; h) = \{((\beta_2 + u_2) y_2 + (\beta_1 + u_1), y_2) : y_2 \in \mathcal{R}_{Y_2}\},$$

and

$$\mathcal{U}(y, z; h) = \{(y_1 - \beta_1 - \beta_2 y_2 - u_2 y_2, u_2) : u_2 \in \mathcal{R}_{U_2}\}.$$

Example 5. *Interval censored endogenous explanatory variables. Let* $g(\cdot, \cdot, \cdot) : \mathbb{R} \times \mathbb{R}^k \times \mathbb{R} \to \mathbb{R}$ *be increasing in its first argument and strictly increasing in its third argument such that*

$$Y_1 = g(Y_2^*, Z_1, U),$$

where endogenous variable $Y_2^* \in \mathbb{R}$ *is interval censored with*

$$P[Y_{2l} \leq Y_2^* \leq Y_{2u}] = 1,$$

for observed variables Y_{2l}, Y_{2u}*. There is no restriction on the process determining the realizations of the censoring variables* Y_{2l}, Y_{2u}*. The structural function is*

$$h(y, z, u) = |y_1 - g(y_{2l}, z_1, u)|_- + |g(y_{2u}, z_1, u) - y_1|_+,$$

with $y \equiv (y_1, y_{2l}, y_{2u})$*, and* $y_{2l} \leq y_{2u}$*. The resulting level sets are*

$$\mathcal{Y}(u, z; h) = \{y \in \mathcal{R}_Y : g(y_{2l}, z_1, u) \leq y_1 \leq g(y_{2u}, z_1, u) \wedge y_{2l} \leq y_{2u}\},$$

and

$$\mathcal{U}(y, z; h) = \left[g^{-1}(y_{2u}, z_1, y_1), g^{-1}(y_{2l}, z_1, y_1) \right],$$

where the function $g^{-1}(\cdot,\cdot,\cdot)$ *is the inverse of* $g(\cdot,\cdot,\cdot)$ *with respect to its third argument, so that for all* y_2, z_1, *and* u,

$$g^{-1}(y_2, z_1, g(y_2, z_1, u)) = u.$$

GIV models have numerous applications and many other incomplete models that include restrictions on the distribution of unobserved variables and exogenous covariates fall in this class of models. Examples include models with structural relationships defined by a system of inequalities such as the Haile and Tamer (2003) model of ascending English auctions, simultaneous equations models admitting multiple solutions such as the oligopoly entry game of Bresnehan and Reiss (1989, 1991) and Tamer (2003), and models with a nonadditive scalar unobservable and no monotonicity restriction as studied by Hoderlein and Mammen (2007).

2.2 The approach and the main results

A model $\mathcal{M}$ comprises a collection of admissible structures. A structure is a pair $\left(h, \mathcal{G}_{U|Z}\right)$ where h is a structural function previously defined and $\mathcal{G}_{U|Z}$ is shorthand for the collection of conditional distributions,

$$\mathcal{G}_{U|Z} \equiv \left\{G_{U|Z}(\cdot|z) : z \in \mathcal{R}_Z\right\},$$

where $G_{U|Z}(\mathcal{S}|z)$ denotes the probability mass placed on any set $\mathcal{S} \subseteq \mathcal{R}_{U|z}$ when $U \sim G_{U|Z}(\cdot|z)$. The model $\mathcal{M}$ can restrict both the structural function h and the distributions $\mathcal{G}_{U|Z}$. For example $\mathcal{M}$ may admit structures $\left(h, \mathcal{G}_{U|Z}\right)$ such that h is nonparametrically specified and sufficiently smooth and collections $G_{U|Z}$ of unknown form but such that U and Z are independently distributed. Alternatively the model may restrict h to be known up to a finite dimensional parameter vector with an index structure, or require for example that $E[U|Z] = 0$, or that $G_{U|Z}(\cdot|z)$ be parametrically specified.

The sampling process identifies a collection of probability distributions denoted

$$\mathcal{F}_{Y|Z} \equiv \left\{F_{Y|Z}(\cdot|z) : z \in \mathcal{R}_Z\right\}$$

where $F_{Y|Z}(\mathcal{T}|z)$ denotes the probability mass placed on any set $\mathcal{T} \subseteq \mathcal{R}_{Y|z}$. Our identification analysis answers the question: precisely which admissible structures $\left(h, \mathcal{G}_{U|Z}\right)$ are capable of generating the collection of conditional distributions $\mathcal{F}_{Y|Z}$?

We begin by revisiting the notion of observational equivalence. Classical definitions of observational equivalence, e.g. those of Koopmans and Reiersøl (1950), Hurwicz (1950), Rothenberg (1971), and Bowden (1973) require that a given structure produce for each $z \in \mathcal{R}_Z$ a unique conditional distribution of endogenous variables $F_{Y|Z}(\cdot|z)$. In our framework a structure $\left(h, \mathcal{G}_{U|Z}\right)$ produces a collection of random sets $\left\{\mathcal{Y}(U, z; h) : U \sim G_{U|Z}(\cdot|z), z \in \mathcal{R}_Z\right\}$. Conditional on $Z = z$,

the random set $\mathcal{Y}(U, z; h)$ can be viewed as a collection of its selections, where the selection of a random set $\mathcal{Y}$ is any random variable Y such that $\Pr[Y \in \mathcal{Y}] = 1$. Thus a structure $\left(h, \mathcal{G}_{U|Z}\right)$ admitted by our model generally produces collections of compatible distributions for $F_{Y|Z}(\cdot|z)$.

We therefore generalize the classical notion of observational equivalence to allow structures that permit collections of conditional distributions $F_{Y|Z}(\cdot|z)$, each $z \in \mathcal{R}_Z$. This formalizes the logic underlying previous identification analyses in incomplete models, while introducing some interesting subtleties. For example, the collection of conditional distributions produced by two different structures may be different yet overlap, so that whether or not the two structures are observationally distinct depends on the identified conditional distributions $\mathcal{F}_{Y|Z}$. This cannot happen when structures produce unique collections $\mathcal{F}_{Y|Z}$. The identified set of structures $\left(h, \mathcal{G}_{U|Z}\right)$ are then precisely those such that $F_{Y|Z}(\cdot|z)$ is *selectionable* with respect to the distribution of $\mathcal{Y}(U, z; h)$ given $Z = z$, for almost every $z \in \mathcal{R}_Z$.[2] This set may be empty, allowing for the possibility that model $\mathcal{M}$ of Restriction A4 is misspecified.

In Theorem 1 of Section 3, we show that because of the dual relationship between the two level sets of the structural function h, $\mathcal{U}(y, z; h)$ and $\mathcal{Y}(u, z; h)$, the probability distribution $F_{Y|Z}(\cdot|z)$ is selectionable with respect to the distribution of $\mathcal{Y}(U, z; h)$ when $U \sim G_{U|Z}(\cdot|z)$ if and only if $G_{U|Z}(\cdot|z)$ is selectionable with respect to the distribution of the random set $\mathcal{U}(Y, z; h)$ when $Y \sim F_{Y|Z}(\cdot|z)$. Using this result we show that we can completely characterize observational equivalence, and hence identified sets of model structures, through selectionability of $G_{U|Z}(\cdot|z)$ with respect to the conditional distribution of $\mathcal{U}(Y, z; h)$. This result, formalized in Theorem 2, facilitates the imposition of restrictions directly on the joint distribution of U and Z, i.e. restrictions on the distribution of unobserved heterogeneity, common to econometric modeling.

The selectionability criteria provide a widely applicable characterization of the identified set of structures $\left(h, \mathcal{G}_{U|Z}\right)$. Selectionability can be characterized in a variety of ways, and depending on which restrictions are imposed on the conditional distributions of unobserved heterogeneity $\mathcal{G}_{U|Z}$, different characterizations may prove more or less convenient. One such way is through the use of the containment functional of the random set $\mathcal{U}(Y, Z; h)$ defined in (1.3), as we show at the end of Section 3.2 in Corollary 1. This produces a generally applicable characterization comprising a collection of conditional moment inequalities. The collection of implied moment inequalities is potentially extremely large. In Section 3.3 we employ the notion of core-determining test sets to exploit the underlying geometry of possible realizations of random sets $\mathcal{U}(Y, Z; h)$ and thereby produce a reduction in the number of inequality restrictions necessary to characterize the identified set. We refine the containment functional characterization of the identified set by reducing the *number* of conditional moment restrictions required, and also by providing conditions whereby some of these inequalities must in fact hold with *equality*.

[2] A distribution F of a random variable is said to be selectionable with respect to the distribution of random set $\mathcal{Y}$ if there exists a random variable, say W, with distribution F, and a random set, say $\tilde{\mathcal{Y}}$, with the same distribution as $\mathcal{Y}$, such that $W \in \tilde{\mathcal{Y}}$ with probability one. See Definition 1, Section 3.1 for the formal definition.

Up to this point, our identification analysis holds no matter what restrictions are imposed on $\mathcal{G}_{U|Z}$. In Section 3.4 we consider three types of restrictions on the joint distribution of U and Z common to the econometrics literature, namely stochastic independence, mean independence, and quantile independence restrictions. Under stochastic independence we show how the containment functional characterization of Corollary 1 simplifies. We provide characterizations allowing for the distribution of unobserved heterogeneity to be either restricted to some known family, or completely unrestricted. With only mean independence imposed rather than stochastic independence, we employ the Aumann expectation to characterize the identified set, and show how, when applicable, the support function of random set $\mathcal{U}(Y, Z; h)$ can be used for computational gains. The ideas here follow closely those of Beresteanu, Molchanov, and Molinari (2011), although unlike their analysis we continue to consider characterizations through random sets in $\mathcal{R}_U$. We then provide a characterization of identified sets in models with conditional quantile restrictions and interval-valued random sets $\mathcal{U}(Y, Z; h)$, such as those used in models with censored variables. In Section 4 we illustrate the identifying power of each of these restrictions on $\mathcal{G}_{U|Z}$ in the context of Example 5 above, featuring a censored endogenous explanatory variable.

3 Identified sets for GIV models

We impose the following restrictions throughout.

Restriction A1: (Y, Z, U) are random vectors defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, endowed with the Borel sets on Ω. The support of (Y, Z, U) is a subset of Euclidean space. □

Restriction A2: A collection of conditional distributions

$$\mathcal{F}_{Y|Z} \equiv \left\{F_{Y|Z}(\cdot|z) : z \in \mathcal{R}_Z\right\}$$

is identified by the sampling process. □

Restriction A3: There is an $\mathcal{F}$-measurable function $h(\cdot, \cdot, \cdot) : \mathcal{R}_{YZU} \to \mathbb{R}$ such that

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

and there is a collection of conditional distributions

$$\mathcal{G}_{U|Z} \equiv \left\{G_{U|Z}(\cdot|z) : z \in \mathcal{R}_Z\right\},$$

where for all $\mathcal{S} \subseteq \mathcal{R}_{U|z}$, $G_{U|Z}(\mathcal{S}|z) \equiv \mathbb{P}[U \in \mathcal{S}|z]$ denotes a conditional distribution of U given $Z = z$. □

Restriction A4: The pair $\left(h, \mathcal{G}_{U|Z}\right)$ belongs to a known set of admissible structures $\mathcal{M}$ such that $h(\cdot, \cdot, \cdot) : \mathcal{R}_{YZU} \to \mathbb{R}$ is continuous in its first and third arguments. □

Restriction A1 defines the probability space on which (Y, Z, U) reside and restricts their support to Euclidean space. Restriction A2 requires that, for each $z \in \mathcal{R}_Z$, the conditional distribution of Y given $Z = z$ is identified. Random sampling of observations from $\mathbb{P}$ is sufficient, but not required. Restriction A3 posits the existence of structural relation h and provides notation for the collection of conditional distributions $\mathcal{G}_{U|Z}$ of U given Z induced by population measure $\mathbb{P}$. Restrictions A1-A3 place restrictions on neither the properties of h nor $\mathcal{G}_{U|Z}$. These restrictions are maintained throughout.

Restriction A4 imposes model $\mathcal{M}$, the collection of admissible structures $\left(h, \mathcal{G}_{U|Z}\right)$. Unlike restrictions A1-A3, restriction A4 is refutable based on knowledge of $\mathcal{F}_{Y|Z}$. Our characterizations of identified sets given admissible structures $\mathcal{M}$ entail those structures $\left(h, \mathcal{G}_{U|Z}\right) \in \mathcal{M}$ that under Restrictions A1-A3 could possibly deliver the identified conditional distributions $\mathcal{F}_{Y|Z}$. It is possible of course that there is no $\left(h, \mathcal{G}_{U|Z}\right)$ belonging to $\mathcal{M}$ such that $\mathbb{P}\left[h\left(Y, Z, U\right) = 0\right] = 1$ for some random variable U with conditional distributions belonging to $\mathcal{G}_{U|Z}$. We allow for this possibility, noting that in such cases the identified set of structures delivered by our results is empty, which would indicate that the model is misspecified. Continuity of $h\left(y, z, u\right)$ in y and u guarantees that $\mathcal{Y}\left(U, Z; h\right)$ and $\mathcal{U}\left(Y, Z; h\right)$ are random closed sets, but can be relaxed.[3]

In places we will find it convenient to refer separately to collections of admissible structural functions and distributions $\mathcal{G}_{U|Z}$. Notationally these are defined as the following projections of $\mathcal{M}$.

$$\begin{aligned} \mathcal{H} &\equiv \left\{h : \left(h, \mathcal{G}_{U|Z}\right) \in \mathcal{M} \text{ for some } \mathcal{G}_{U|Z}\right\}, \\ \mathsf{G}_{U|Z} &\equiv \left\{\mathcal{G}_{U|Z} : \left(h, \mathcal{G}_{U|Z}\right) \in \mathcal{M} \text{ for some } h\right\}. \end{aligned}$$

The model $\mathcal{M}$ could, but does not necessarily, consist of the full product space $\mathcal{H} \times \mathsf{G}_{U|Z}$.

Restrictions A1-A4 are noticeably weak. This is intentional, and not without consequence. The identification analysis built up in Sections 3.1 and 3.2 below is set out with only these restrictions in place, and is thus extremely general. The analysis characterizes the identified set of model structures, and the level of generality allows for the possibility that these sets are either large or small, for example the entire admissible space at one extreme, or a singleton point at the other.

The identifying power of any particular model manifests through three different mechanisms: (i) restrictions on the class of functions h; (ii) restrictions on the joint distribution of (U, Z), embodied through the admitted collections $\mathcal{G}_{U|Z}$, and (iii) the joint distribution of (Y, Z). The first two mechanisms are part of the model specification, with $\mathcal{M}$ constituting the set of structures $\left(h, \mathcal{G}_{U|Z}\right)$ deemed admissible for the generation of (Y, Z). A researcher may restrict these to belong to more or less restrictive classes, parametric, semiparametric, or nonparametric. For example, h could be allowed to be any sufficiently smooth function, or it could be restricted to a parametric family as

[3]Note that continuity of the function h does not rule out either Y or U being discrete-valued. Continuity itself is not essential, and may be relaxed as long as the relevant random sets can be shown to be closed in some topology.

in a linear index model. Likewise $\mathcal{G}_{U|Z}$ could be collections of conditional distributions such that $E[U|Z] = 0$, $q_{U|Z}(\tau|z) = 0$, $U \perp\!\!\!\perp Z$, or $U \perp\!\!\!\perp Z$ with $G_{U|Z}(\cdot|z)$ restricted to a parametric family. We consider various types of restrictions on $\mathcal{G}_{U|Z}$ in Section 3.4 below.

The third source of identifying power, the joint distribution of (Y, Z), is identified and hence left completely unrestricted. This may initially appear at odds with results requiring rank or completeness conditions, but in fact is not. Such conditions are used to ensure point identification. We allow for set identification, and so these are not required. Rather, we provide characterizations of identified sets in the class of models we study. If there is "sufficient variation" in the distribution of (Y, Z) to achieve point identification, such as the usual rank condition in a linear IV model, our characterization reduces to a singleton set. Thus, such conditions are not required in our set identification analysis, but are still very much of interest in consideration of which qualities of observed data may result in identified sets that are singleton points, or more generally affect the size of the identified set.

More broadly, the generality of our identification analysis offers a formal framework for the consideration of which models $\mathcal{M}$ and qualities of the joint distribution of (Y, Z) may be usefully applied. Given a joint distribution for (Y, Z), the use of less restrictive models will logically result in larger identified sets than will more restrictive models. A sufficiently general model may have so little identifying power as to be uninformative. This is useful for researchers to know when considering which models to employ in practice, and can be used to motivate the incorporation of further restrictions that may be deemed credible.

3.1 Observational Equivalence and Selectionability in Outcome Space

The standard definition of observational equivalence in the econometrics literature presumes the existence of a unique joint distribution of observed variables, $F_{YZ}(\cdot, \cdot; m)$, for any structure $m \equiv \left(h, \mathcal{G}_{U|Z}\right) \in \mathcal{M}$, equivalently a unique conditional distribution $F_Y(\cdot|z; m)$ for each $z \in \mathcal{R}_Z$, given identification of F_Z.[4] Two structures m, m' are then observationally equivalent if they produce the same conditional distribution a.e. $z \in \mathcal{R}_Z$, that is if

$$F_Y(\cdot|z; m) = F_Y\left(\cdot|z; m'\right) \text{ a.e. } z \in \mathcal{R}_Z.$$

As such, the classical notion of observational equivalence of two structures is a property for which the identified conditional distributions $\mathcal{F}_{Y|Z}$ are irrelevant.

As previously discussed, the requirement that each m produce a unique $F_Y(\cdot|z; m)$ for each z is important to relax when working with an incomplete model, since a given structure can generate

[4]For classical treatments of observational equivalence, see for example Koopmans and Reiersøl (1950), Hurwicz (1950), Rothenberg (1971), and Bowden (1973). For definitions in a fully nonparametric setting see e.g. Matzkin (2007, 2008).

a *collection* of possible $F_Y(\cdot|z;m)$ for each z.[5] For such models, the property of observational equivalence of two structures m, m' in contrast generally depends upon the collection of conditional distributions $\mathcal{F}_{Y|Z}$. This is because for any $z \in \mathcal{R}_Z$ there may be some $F_{Y|Z}(\cdot|z)$ that belong to the *sets* of conditional distributions generated by both m and m', and yet there may be other $F_{Y|Z}(\cdot|z)$ that belong to the set of conditional distributions generated by one structure, but not the other. Thus, in the following development, we define observational equivalence with respect to the (identified) collection of population distributions $\mathcal{F}_{Y|Z}$, and a corresponding notion of *potential* observational equivalence, which is a property of two structures that is irrespective of $\mathcal{F}_{Y|Z}$.

To provide formal definitions of these properties we begin by recalling the definition of a random selection from a random set as set out for example by Molchanov (2005, Definition 2.2, p. 26). We then use the related notion of selectionability in our subsequent discussion of observational equivalence.

Definition 1 *Let W and $\mathcal{W}$ denote a random vector and random set defined on the same probability space. W is a **selection** of $\mathcal{W}$, denoted $W \in \mathrm{Sel}(\mathcal{W})$, if $W \in \mathcal{W}$ with probability one. The distribution F_W of random vector W is **selectionable** with respect to the distribution of random set $\mathcal{W}$, which we abbreviate $F_W \preccurlyeq \mathcal{W}$, if there exists random variable $\tilde{W}$ distributed F_W and a random set $\tilde{\mathcal{W}}$ with the same distribution as $\mathcal{W}$ such that $\tilde{W} \in \tilde{\mathcal{W}}$ with probability one.*

A given structure $m = \left(h, \mathcal{G}_{U|Z}\right)$ induces a distribution for the random outcome set $\mathcal{Y}(U, Z; h)$ conditional on $Z = z$, for all $z \in \mathcal{R}_Z$. If $\mathcal{Y}(U, Z; h)$ is a singleton set with probability one, then the model is complete, and the conditional distribution of $\mathcal{Y}(U, Z; h)$ given $Z = z$ is simply that of $\{Y\}$ given $Z = z$ for each $z \in \mathcal{R}_Z$. In this case, again for each $z \in \mathcal{R}_Z$, $F_{Y|Z}(\cdot|z)$ is the *only* conditional distribution of Y given $Z = z$ that is selectionable with respect to the conditional distribution of $\mathcal{Y}(U, Z; h)$, and our definition of observational equivalence below simplifies to the classical one. If, on the other hand, the model is incomplete, so that $\mathcal{Y}(U, Z; h)$ is non-singleton with positive probability, then $h(Y, Z, U) = 0$ dictates only that $Y \in \mathcal{Y}(U, Z; h)$, which is insufficient to uniquely determine the conditional distributions $F_{Y|Z}$. That is, there are for at least some $z \in \mathcal{R}_Z$, multiple $F_{Y|Z}(\cdot|z)$ satisfying $F_{Y|Z}(\cdot|z) \preccurlyeq \mathcal{Y}(U, Z; h)$ given $Z = z$. The very definition of selectionability of $F_{Y|Z}(\cdot|z)$ from the distribution of $\mathcal{Y}(U, Z; h)$ given $Z = z$ for almost every $z \in \mathcal{R}_Z$ characterizes those distributions for which $h(Y, Z, U) = 0$ can hold with probability one for the given $\left(h, \mathcal{G}_{U|Z}\right)$. Those distributions $F_{Y|Z}(\cdot|z)$ that are selectionable with respect to the conditional distribution of $\mathcal{Y}(U, Z; h)$ when $U \sim G_{U|Z}(\cdot|z)$ are precisely those conditional distributions that can be generated by the structure $\left(h, \mathcal{G}_{U|Z}\right)$.

[5] In our formulation of observational equivalence and characterizations of identified sets, we continue to work with conditional distributions of endogenous and latent variables, $F_Y(\cdot|z)$ and $G_U(\cdot|z)$, respectively, for almost every $z \in \mathcal{R}_Z$. Knowledge of the distribution of Z, F_Z, combined with $F_Y(\cdot|z)$ or $G_U(\cdot|z)$ a.e. $z \in \mathcal{Z}$ is equivalent to knowledge of the joint distribution of (Y, Z) denoted F_{YZ}, or that of (U, Z), denoted G_{UZ}, respectively. We show formally in Appendix B that our characterizations using selectionability conditional on $Z = z$, a.e. $z \in \mathcal{R}_Z$, are indeed equivalent to using analogous selectionability criteria for the joint distributions F_{YZ} or G_{UZ}.

This leads to the following formal definitions of potential observational equivalence and observational equivalence with respect to a particular collection of conditional distributions $\mathcal{F}_{Y|Z}$.

Definition 2 *Under Restrictions A1-A3, two structures* $\left(h, \mathcal{G}_{U|Z}\right)$ *and* $\left(h', \mathcal{G}'_{U|Z}\right)$ *are* ***potentially observationally equivalent*** *if there exists a collection of conditional distributions* $\mathcal{F}_{Y|Z}$ *such that* $F_{Y|Z}(\cdot|z) \preccurlyeq \mathcal{Y}(U, z; h)$ *when* $U \sim G_{U|Z}(\cdot|z)$ *and* $F_{Y|Z}(\cdot|z) \preccurlyeq \mathcal{Y}(U, z; h')$ *when* $U \sim G'_{U|Z}(\cdot|z)$ *for almost every* $z \in \mathcal{R}_Z$*. Two structures* $\left(h, \mathcal{G}_{U|Z}\right)$ *and* $\left(h', \mathcal{G}'_{U|Z}\right)$ *are* ***observationally equivalent*** with respect to $\mathcal{F}_{Y|Z} = \left\{F_{Y|Z}(\cdot|z) : z \in \mathcal{R}_Z\right\}$ *if* $F_{Y|Z}(\cdot|z) \preccurlyeq \mathcal{Y}(U, z; h)$ *when* $U \sim G_{U|Z}(\cdot|z)$ *and* $F_{Y|Z}(\cdot|z) \preccurlyeq \mathcal{Y}(U, z; h')$ *when* $U \sim G'_{U|Z}(\cdot|z)$ *for almost every* $z \in \mathcal{R}_Z$*.*

The closely related definition of the identified set of structures $\left(h, \mathcal{G}_{U|Z}\right)$ is as follows.[6]

Definition 3 *Under Restrictions A1-A4, the* ***identified set*** *of structures* $\left(h, \mathcal{G}_{U|Z}\right)$ *with respect to the collection of distributions* $\mathcal{F}_{Y|Z}$ *are those structures such that the conditional distributions* $F_{Y|Z}(\cdot|z) \in \mathcal{F}_{Y|Z}$ *are selectionable with respect to the conditional distributions of random set* $\mathcal{Y}(U, z; h)$ *when* $U \sim G_{U|Z}(\cdot|z)$*, a.e.* $z \in \mathcal{R}_Z$*:*

$$\mathcal{M}^* \equiv \left\{\left(h, \mathcal{G}_{U|Z}\right) \in \mathcal{M} : F_{Y|Z}(\cdot|z) \preccurlyeq \mathcal{Y}(U, z; h) \text{ when } U \sim G_{U|Z}(\cdot|z), \text{ a.e. } z \in \mathcal{R}_Z\right\}. \tag{3.1}$$

Selectionability of observed conditional distributions from the random outcome set $\mathcal{Y}(U, z; h)$ provides a convenient and extremely general characterization of identified sets in a broad class of econometric models. A task that remains is how, in any particular model, to characterize the set of structures for which the given selectionability criteria holds. Any characterization of selectionability will suffice. Beresteanu, Molchanov, and Molinari (2011) show for example how one can cast selectionability in terms of the support function of the Aumann Expectation of the random outcome set in order to characterize identified sets in a broad class of econometric models.

Given Definition 3 for the identified set of model structures, we can now define set identification of structural features. As is standard, we define a structural feature $\psi(\cdot, \cdot)$ as any functional of a structure $\left(h, \mathcal{G}_{U|Z}\right)$. Examples include the structural function h, $\psi\left(h, \mathcal{G}_{U|Z}\right) = h$, the distributions of unobserved heterogeneity, $\psi\left(h, \mathcal{G}_{U|Z}\right) = \mathcal{G}_{U|Z}$, and counterfactual probabilities such as the probability that a component of Y is guaranteed to exceed a given threshold conditional on $Z = z$.

Definition 4 *The* ***identified set of structural features*** $\psi(\cdot, \cdot)$ *under Restrictions A1-A4 is*

$$\Psi \equiv \left\{\psi\left(h, \mathcal{G}_{U|Z}\right) : \left(h, \mathcal{G}_{U|Z}\right) \in \mathcal{M}^*\right\}.$$

Depending on the context, a variety of different features may be of interest. The identified set of structures $\mathcal{M}^*$ can be used to ascertain the identified set of any such feature. We thus take the

[6] The identified set $\mathcal{M}^*$ depends upon the collection of conditional distributions $\mathcal{F}_{Y|Z}$, although we do not make this dependence explicit in our notation.

identified set of structures $\mathcal{M}^*$ as the main object of interest, and unless we specify a particular feature of interest, reference to only the "identified set" without qualification refers to $\mathcal{M}^*$.[7]

A key component of econometric models are restrictions on the joint distribution of U and Z. The use of the Aumann Expectation of random outcome set $\mathcal{Y}(U, z; h)$ and associated support function dominance criteria can be convenient in models with conditional mean restrictions, as discussed by Beresteanu, Molchanov, and Molinari (2012). In models with $G_{U|Z}(\cdot|z)$ parametrically specified, this approach or a capacity functional characterization of selectionability can be used, see e.g. Beresteanu, Molchanov, and Molinari (2011) or the related characterization of Galichon and Henry (2011). Proposed estimation and inference strategies based on these approaches entail simulation of admissible distributions of $G_{U|Z}(\cdot|z)$, which are then plugged into the outcome correspondence $\mathcal{Y}(\cdot, \cdot; h)$ to generate simulated conditional distributions of random set $\mathcal{Y}(U, Z; h)$, see also Henry, Meango, and Queyranne (2011). In this class of models, that is with the same independence restriction and $G_{U|Z}(\cdot|z)$ parametrically specified, our characterization using selectionability from random sets in U-space provides characterizations of $\mathcal{M}^*$ comprising inequalities that can be checked via either numerical computation or simulation.

In the following Section we prove the equivalence of a characterization of $\mathcal{M}^*$ and indeed of observational equivalence put in terms of selectionability of $G_{U|Z}(\cdot|z)$ from the random residual set $\mathcal{U}(Y, Z; h)$. Working in the space of unobserved heterogeneity enables direct and immediate consideration of any conceivable alternative restrictions on the joint distribution of $G_{U|Z}(\cdot|z)$.

For example, we show in Section 3.4 that when U and Z are assumed independent but the form of $G_{U|Z}(\cdot|z)$ is left unspecified, characterization of $\mathcal{H}^*$, the identified set of structural functions h, can be reduced to a collection of inequality restrictions from which $G_{U|Z}(\cdot|z)$ is absent. Characterizations of identified sets using the selectionability criteria of $F_{Y|Z}(\cdot|z)$ from the random outcome set $\mathcal{Y}(U, z; h)$, on the other hand, do explicitly refer to $G_{U|Z}(\cdot|z)$. To check then whether any given h can be a component of an element $\left(h, \mathcal{G}_{U|Z}\right)$ of the identified set $\mathcal{M}^*$, one must then devise a way to check the selectionability criteria over a nonparametric infinite dimensional class of functions $G_{U|Z}(\cdot|z)$ for each $z \in \mathcal{R}_Z$.

3.2 Set Identification via Selectionability in U-Space

In this section a dual relation is derived between random outcome sets $\mathcal{Y}(U, Z; h)$ and random residual sets $\mathcal{U}(Y, Z; h)$. This is then used to relate selectionability of $F_{Y|Z}(\cdot|z)$ with respect to $\mathcal{Y}(U, Z; h)$ and selectionability of $G_{U|Z}(\cdot|z)$ with respect to $\mathcal{U}(Y, Z; h)$. In our setup the conditional distributions of the random sets $\mathcal{Y}(U, Z; h)$ and $\mathcal{U}(Y, Z; h)$ are completely determined by the conditional (on $Z = z$) distributions of U and Y, $G_{U|Z}(\cdot|z)$ and $F_{Y|Z}(\cdot|z)$, respectively.

[7] The identified set of structural features Ψ depends on on both $\mathcal{M}$ and the conditional distributions $\mathcal{F}_{Y|Z}$, but for ease of notation we suppress this dependence.

Theorem 1 *Let Restrictions A1-A3 hold. Then for any $z \in \mathcal{R}_Z$, $F_{Y|Z}(\cdot|z)$ is selectionable with respect to the conditional distribution of $\mathcal{Y}(U, Z; h)$ given $Z = z$ if and only if $G_{U|Z}(\cdot|z)$ is selectionable with respect to the conditional distribution of $\mathcal{U}(Y, Z; h)$ given $Z = z$.*

With Theorem 1 established, we now characterize the identified set in terms of random variables and sets in the space of unobserved heterogeneity. As previously expressed, a key advantage to doing this is the ability to impose restrictions directly on $G_{U|Z}$ through specification of the class $\mathcal{G}_{U|Z}$ admitted by the model $\mathcal{M}$. One can then check whether any such $G_{U|Z} \in \mathcal{G}_{U|Z}$ are selectionable with respect to the identified conditional distributions of random set $\mathcal{U}(Y, Z; h)$, given identification of the conditional distributions $\mathcal{F}_{Y|Z}$ under Restriction A2. That is, in the context of any particular model, events concerning this random set can be expressed as events involving observable variables, as we illustrate in the examples of Section 4.

Theorem 2 *Let Restrictions A1-A3 hold. Then (i) structures $\left(h, \mathcal{G}_{U|Z}\right)$ and $\left(h^*, \mathcal{G}^*_{U|Z}\right)$ are observationally equivalent with respect to $\mathcal{F}_{Y|Z}$ if and only if $G_{U|Z}(\cdot|z)$ and $G^*_{U|Z}(\cdot|z)$ are selectionable with respect to the conditional (on $Z = z$) distributions of random sets $\mathcal{U}(Y, Z; h)$ and $\mathcal{U}(Y, Z; h^*)$, respectively, a.e. $z \in \mathcal{R}_Z$; and (ii) the identified set of structures $\left(h, \mathcal{G}_{U|Z}\right)$ are those such that $G_{U|Z}(\cdot|z)$ is selectionable with respect to the conditional (on $Z = z$) distribution of random set $\mathcal{U}(Y, Z; h)$.*

Theorem 2 uses duality to express observational equivalence and characterization of the identified set of structures $\left(h, \mathcal{G}_{U|Z}\right)$ in terms of selectionability from the conditional distribution of $\mathcal{U}(Y, Z; h)$. Thus, any conditions that characterize the set of $\left(h, \mathcal{G}_{U|Z}\right)$ such that $G_{U|Z}$ is selectionable with respect to the conditional distribution of $\mathcal{U}(Y, Z; h)$ will suffice for characterization of the identified set.

One such characterization, used in previous papers allowing only discrete outcomes, e.g. Chesher, Rosen, and Smolinski (2013) and Chesher and Rosen (2013a, 2013b), uses Artstein's Inequality, see e.g. Artstein (1983), Norberg (1992), and Molchanov (2005, Section 1.4.8). This result allows us to characterize the identified set $\mathcal{M}^*$ through the conditional containment functional of random set $\mathcal{U}(Y, Z; h)$, defined as

$$C_h(\mathcal{S}|z) \equiv \mathbb{P}\left[\mathcal{U}(Y, Z; h) \subseteq \mathcal{S}|z\right].$$

Characterization via the containment functional produces an expression for $\mathcal{M}^*$ in the form of conditional moment inequalities, as given in the following Corollary.

Corollary 1 *Under the restrictions of Theorem 2, the identified set can be written*

$$\mathcal{M}^* \equiv \left\{\left(h, \mathcal{G}_{U|Z}\right) \in \mathcal{M} : \forall \mathcal{S} \in \mathsf{F}(\mathcal{R}_U),\ C_h(\mathcal{S}|z) \leq G_{U|Z}(\mathcal{S}|z),\ \text{a.e. } z \in \mathcal{R}_Z\right\},$$

where $\mathsf{F}(\mathcal{R}_U)$ denotes the collection of all closed subsets of $\mathcal{R}_U$.

Corollary 1 translates the selectionability requirement for characterization of the identified set to a collection of conditional moment inequalities. The inequalities in this characterization are for almost every value of the instrument $z \in \mathcal{R}_Z$ as well as all closed *test sets* $\mathcal{S}$ on $\mathcal{R}_U$. The containment functional inequality $C_h(\mathcal{S}) \leq G_{U|Z}(\mathcal{S}|z)$ follows immediately from the fact that U is, by virtue of $h(Y,Z,U) = 0$, a selection of $\mathcal{U}(Y,Z;h)$. Artstein's inequality establishes that the inequality holding for all $\mathcal{S} \in \mathsf{F}(\mathcal{R}_U)$ guarantees selectionability of $G_{U|Z}$ from the conditional distribution of $\mathcal{U}(Y,Z;h)$, a.e. $z \in \mathcal{R}_Z$.

3.3 Core Determining Test Sets

We now characterize a smaller collection $\mathsf{Q}(h,z)$ of *core-determining* test sets $\mathcal{S}$ for any h, and any $z \in \mathcal{R}_Z$, such that if

$$\forall \mathcal{S} \in \mathsf{Q}(h,z),\ C_h(\mathcal{S}|z) \leq G_{U|Z}(\mathcal{S}|z), \tag{3.2}$$

then the same inequality holds for all $\mathcal{S} \in \mathsf{F}(\mathcal{R}_U)$. Thus, characterization of the identified set is reduced to those such that (3.2) holds for all $\mathcal{S} \in \mathsf{Q}(h,z)$.[8] Galichon and Henry (2011) initially introduced core-determining sets for identification analysis in consideration of sets in outcome space, characterizing such classes of sets for incomplete models that satisfy a monotonicity requirement, which is not needed here. We extend their definition of such collections by devising core-determining sets for our characterizations in U-space, and by allowing them to be specific to the structural relation h and covariate value z.[9]

Our construction builds on ideas from Chesher, Rosen, and Smolinski (2013), but is much more widely applicable. We do not require independence of U and Z, and can thus accommodate other restrictions on $\mathcal{G}_{U|Z}$. We also establish conditions whereby for some of the core-determining sets belonging to $\mathsf{Q}(h,z)$, the inequality (3.2) must in fact hold with *equality*. Under these conditions the initial characterization via inequalities is sharp, but when coupled with the law of total probability some of these inequalities can be strengthened to equalities. Much is known about observable implication for models with conditional moment *equalities*, and recognition that some of the inequalities must in fact hold with equality can potentially be helpful for estimation and in consideration of conditions that may feasibly lead to point identification.

For this development we first define the support of the random set $\mathcal{U}(Y,Z;h)$ conditional on $Z = z$, and the collection of sets comprising unions of such sets. These objects both play important roles. In this definition and the subsequent analysis we employ the following notation at the cost

[8] Earlier versions of some results in this section appeared in the 2012 version of the CeMMAP working paper Chesher and Rosen (2012b), which concerned models with only discrete endogenous variables. In this Section we provide more general results that cover the broader class of models studied here. In revisions of Chesher and Rosen (2012b) we refer to the more general results here.

[9] Through (3.4) in Theorem 3 below, the core-determining set may also be dependent upon $G_{U|Z}(\cdot|z)$, but we suppress this from the notation.

of some slight abuse of notation:

$$\forall \mathcal{Y} \subseteq \mathcal{R}_{Y|z}, \qquad \mathcal{U}(\mathcal{Y}, z; h) \equiv \bigcup_{y \in \mathcal{Y}} \mathcal{U}(y, z; h).$$

That is, $\mathcal{U}(\mathcal{Y}, z; h)$ is the union of sets $\mathcal{U}(y, z; h)$ such that $y \in \mathcal{Y}$.

Definition 5 *Under Restrictions A1-A3, the* ***conditional support of random set*** $\mathcal{U}(Y, Z; h)$ *given* $Z = z$ *is*

$$\mathsf{U}(h, z) \equiv \left\{\mathcal{U} \subseteq \mathcal{R}_U : \exists y \in \mathcal{R}_{Y|z} \text{ such that } \mathcal{U} = \mathcal{U}(y, z; h)\right\}.$$

The collections of all sets that are unions of elements of $\mathsf{U}(h, z)$ *is denoted*

$$\mathsf{U}^*(h, z) \equiv \left\{\mathcal{U} \subseteq \mathcal{R}_U : \exists \mathcal{Y} \subseteq \mathcal{R}_{Y|z} \text{ such that } \mathcal{U} = \mathcal{U}(\mathcal{Y}, z; h)\right\}.$$

Lemma 1 below establishes that for any $(h, z) \in \mathcal{H} \times \mathcal{Z}$, in order for (3.2) to hold for all closed $\mathcal{S} \subseteq \mathcal{R}_U$, it suffices to show only that (3.2) holds for those sets $\mathcal{S} \in \mathsf{U}^*(h, z)$. To state the result, we define some additional notation. For any set $\mathcal{S} \subseteq \mathcal{R}_U$ and any $(h, z) \in \mathcal{H} \times \mathcal{R}_Z$, define

$$\mathsf{U}^{\mathcal{S}}(h, z) \equiv \{\mathcal{U} \in \mathsf{U}(h, z) : \mathcal{U} \subseteq \mathcal{S}\}, \qquad \mathsf{U}_{\mathcal{S}}(h, z) \equiv \left\{\mathcal{U} \in \mathsf{U}(h, z) : G_{U|Z}(\mathcal{U} \cap \mathcal{S}|z) = 0\right\},$$

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

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

which comprises the sets $\mathcal{U} \in \mathsf{U}(h, z)$ that belong to neither $\mathsf{U}^{\mathcal{S}}(h, z)$ nor $\mathsf{U}_{\mathcal{S}}(h, z)$. For ease of reference, Table 1 provides a summary of the collections of sets $\mathsf{U}(h, z)$, $\mathsf{U}^*(h, z)$, $\mathsf{U}^{\mathcal{S}}(h, z)$, $\mathsf{U}_{\mathcal{S}}(h, z)$, and $\overline{\mathsf{U}}^{\mathcal{S}}(h, z)$ used to establish the following Lemma and Theorem 3.

Lemma 1 *Let Restrictions A1-A3 hold. Let* $z \in \mathcal{R}_Z$, $h \in \mathcal{H}$, *and* $\mathcal{S} \subseteq \mathcal{R}_U$. *Let* $\mathcal{U}_{\mathcal{S}}(h, z)$ *denote the union of all sets in* $\mathsf{U}^{\mathcal{S}}(h, z)$,

$$\mathcal{U}_{\mathcal{S}}(h, z) \equiv \bigcup_{\mathcal{U} \in \mathsf{U}^{\mathcal{S}}(h,z)} \mathcal{U}. \tag{3.3}$$

If

$$C_h(\mathcal{U}_{\mathcal{S}}(h, z)|z) \leq G_{U|Z}(\mathcal{U}_{\mathcal{S}}(h, z)|z),$$

then

$$C_h(\mathcal{S}|z) \leq G_{U|Z}(\mathcal{S}|z).$$

Lemma 1 establishes that if (3.2) holds for all unions of sets in $\mathsf{U}(h, z)$, then for that (h, z) it must hold for all closed test sets $\mathcal{S} \subseteq \mathcal{R}_U$.

Collection	Description
$\mathsf{U}(h,z)$	Support of $\mathcal{U}(Y,Z;h)$ conditional on $Z=z$.
$\mathsf{U}^*(h,z)$	Sets that are unions of sets in $\mathsf{U}(h,z)$.
$\mathsf{U}^{\mathcal{S}}(h,z)$	Sets in $\mathsf{U}(h,z)$ that are contained in $\mathcal{S}$.
$\mathsf{U}_{\mathcal{S}}(h,z)$	Sets in $\mathsf{U}(h,z)$ with $G_{U\|Z}$ measure zero intersection with $\mathcal{S}$.
$\overline{\mathsf{U}}^{\mathcal{S}}(h,z)$	Sets in $\mathsf{U}(h,z)$ contained in neither $\mathsf{U}^{\mathcal{S}}(h,z)$ nor $\mathsf{U}_{\mathcal{S}}(h,z)$.

Table 1: Notation for collections of subsets of $\mathcal{R}_U$ and sets used in the development of core determining sets.

Theorem 3 below provides our collection of core-determining test sets $\mathsf{Q}(h,z)$, which is a refinement of $\mathsf{U}^*(h,z)$. That is, all sets in the collection of core-determining sets are also unions of sets in $\mathsf{U}(h,z)$. However not all such unions lie in the core-determining collection. Those elements of $\mathsf{U}^*(h,z)$ that can be excluded from the core-determining collection have the property that each one can be partitioned into two members of the collection $\mathsf{U}^*(h,z)$ such that (i) each is itself a member of the core-determining collection, and (ii) the two sets are disjoint relative to the probability measure $G_{U|Z}(\cdot|z)$.

Theorem 3 *Let Restrictions A1-A3 hold. For any $(h,z)\in\mathcal{H}\times\mathcal{R}_Z$, let $\mathsf{Q}(h,z)\subseteq\mathsf{U}^*(h,z)$, such that for any $\mathcal{S}\in\mathsf{U}^*(h,z)$ with $\mathcal{S}\notin\mathsf{Q}(h,z)$, there exist nonempty collections $\mathsf{S}_1,\mathsf{S}_2\in\mathsf{U}^{\mathcal{S}}(h,z)$ with $\mathsf{S}_1\cup\mathsf{S}_2=\mathsf{U}^{\mathcal{S}}(h,z)$ such that*

$$\mathcal{S}_1\equiv\bigcup_{\mathcal{T}\in\mathsf{S}_1}\mathcal{T},\ \mathcal{S}_2\equiv\bigcup_{\mathcal{T}\in\mathsf{S}_2}\mathcal{T},\ \text{and } G_{U|Z}(\mathcal{S}_1\cap\mathcal{S}_2|z)=0, \tag{3.4}$$

with $\mathcal{S}_1,\mathcal{S}_2\in\mathsf{Q}(h,z)$. Then $C_h(\mathcal{S}|z)\leq G_{U|Z}(\mathcal{S}|z)$ for all $\mathcal{S}\in\mathsf{Q}(h,z)$ implies that $C_h(\mathcal{S}|z)\leq G_{U|Z}(\mathcal{S}|z)$ holds for all $\mathcal{S}\subseteq\mathcal{R}_U$, and in particular for $\mathcal{S}\in\mathsf{F}(\mathcal{R}_U)$, so that the collection of sets $\mathsf{Q}(h,z)$ is core-determining.

Note that all sets of the form $\mathcal{U}(y,z;h)$ with $y\in\mathcal{R}_Y$ are contained in $\mathsf{Q}(h,z)$, so that $\mathsf{U}(h,z)\subseteq\mathsf{Q}(h,z)$. Theorem 3 implies that the identified sets of Theorem 2 are characterized by the set of $\left(h,\mathcal{G}_{U|Z}\right)$ that satisfy the containment functional inequalities of Corollary 1, but with $\mathsf{Q}(h,z)$ replacing $\mathsf{F}(\mathcal{R}_U)$. If, as is the case in many models, the sets in $\mathsf{U}(h,z)$ are each connected with boundary of Lebesgue measure zero, and $G_{U|Z}(\cdot|z)$ is absolutely continuous with respect to Lebesgue measure, then the condition $G_{U|Z}(\mathcal{S}_1\cap\mathcal{S}_2|z)=0$ in (3.4) is implied if the sets $\mathcal{S}_1$ and $\mathcal{S}_2$ have non-overlapping interiors. This implication was used for the construction of core-determining sets in Chesher, Rosen, and Smolinski (2013), in which all elements of $\mathsf{U}(h,z)$ were indeed connected.

To illustrate the results of Theorem 3 in a relatively simple context we refer back to Example 2 of Section 2.1, also studied in Chesher and Rosen (2013). In that model recall that $\mathcal{U}(y,z;h)=$

$[0, g(y_2, z_1)]$ when $y_1 = 0$ and $\mathcal{U}(y, z; h) = (g(y_2, z_1), 1]$ when $y_1 = 0$. Consider a fixed z and a conjectured structural function h, equivalently g. From Lemma 1 it thus follows that for the containment function inequality characterization of $\mathcal{M}^*$ in Corollary 1 we need only consider test sets that are unions of sets of the form $[0, g(y_2, z_1)]$ or $(g(y_2, z_1), 1]$, for $y_2 \in \mathcal{R}_{Y_2}$. The union of any collection of sets $\{[0, g(y_2, z_1)] : y_2 \in \mathcal{Y}_2 \subseteq \mathcal{R}_{Y_2}\}$ is simply $[0, \max_{y_2 \in \mathcal{Y}_2} g(y_2, z_1)]$. Likewise, the union of any collection of sets $\{(g(y_2, z_1), 1] : y_2 \in \mathcal{Y}_2 \subseteq \mathcal{R}_{Y_2}\}$ is $(\min_{y_2 \in \mathcal{Y}_2} g(y_2, z_1), 1]$. Thus, all unions of sets of the form $[0, g(y_2, z_1)]$ or $(g(y_2, z_1), 1]$ can be expressed as

$$\mathcal{S} = [0, g(y_2, z_1)] \cup \left(g\left(y_2', z_1\right), 1\right], \text{ for some } y_2, y_2' \in \mathcal{R}_{Y_2}. \tag{3.5}$$

Consider test sets $\mathcal{S}$ as in (3.5). If $g(y_2, z_1) \geq g(y_2', z_1)$, then $\mathcal{S} = \mathbb{R}$. This test set may be trivially discarded because in this case (3.2) is simply $1 \leq G_{U|Z}(\mathcal{R}_U|z)$, which holds by virtue of $G_{U|Z}(\cdot|z)$ being a probability measure on $\mathcal{R}_U$. If instead $g(y_2, z_1) < g(y_2', z_1)$, then $\mathcal{S} = [0, g(y_2, z_1)] \cup (g(y_2', z_1), 1]$ is a union of disconnected sets. Hence $G_{U|Z}(\mathcal{S}_1 \cap \mathcal{S}_2|z) = 0$, and we can apply Theorem 3 with $\mathcal{S}_1 = [0, g(y_2, z_1)]$ and $\mathcal{S}_2 = (g(y_2', z_1), 1]$ to conclude that as long as $\mathcal{S}_1$ and $\mathcal{S}_2$ are included in the collection of core-determining sets $\mathsf{Q}(h, z)$, $\mathcal{S}$ need not be included in $\mathsf{Q}(h, z)$. Thus it suffices to consider all $\mathcal{S} \in \mathsf{Q}(h, z)$ given by the collection of half-open intervals with endpoint $g(y_2, z_1)$ for some $y_2 \in \mathcal{R}_{Y_2}$.

The following Corollary shows that in some models some of the containment functional inequalities for core-determining sets can be replaced by equalities.[10] Then the identified set can be written as a collection of conditional moment inequalities and equalities. The strengthening of inequality (3.2) to an equality occurs for test sets $\mathcal{S} \in \mathsf{Q}(h, z)$ such that any $\mathcal{U}(y, z; h)$ not contained in $\mathcal{S}$ lies fully outside of $\mathcal{S}$. In this case each set $\mathcal{U}(y, z; h)$ is either contained in $\mathcal{S}$ or contained in $\mathcal{S}^c$, and we have that $C_h(\mathcal{S}|z) + C_h(\mathcal{S}^c|z) = 1$. Likewise $G_{U|Z}(\mathcal{S}|z) + G_{U|Z}(\mathcal{S}^c|z) = 1$, and this combined with the inequalities (3.2) for both $\mathcal{S}$ and $\mathcal{S}^c$ imply that the weak inequality must hold with equality. The formal statement of this result follows.

Corollary 2 *Define*

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

Then, under the conditions of Theorem 3, the collection of equalities and inequalities

$$\begin{aligned} C_h(\mathcal{S}|z) &= G_{U|Z}(\mathcal{S}|z), \text{ all } \mathcal{S} \in \mathsf{Q}^E(h, z), \\ C_h(\mathcal{S}|z) &\leq G_{U|Z}(\mathcal{S}|z), \text{ all } \mathcal{S} \in \mathsf{Q}^I(h, z) \equiv \mathsf{Q}(h, z) \setminus \mathsf{Q}^E(h, z). \end{aligned}$$

holds if and only if $C_h(\mathcal{S}|z) \leq G_{U|Z}(\mathcal{S}|z)$ for all $\mathcal{S} \in \mathsf{Q}(h, z)$.

[10] This is however not the case in the model studied in Example 2.

It is worth noting in particular that the conditions of the Corollary apply to *every* test set $\mathcal{S}$ in models where $\mathcal{U}(Y, Z; h)$ is a singleton set with probability one. Models with this property include traditional IV models with additive unobservables such as the classical linear IV model of Example 1 and the nonparametric IV model of Newey and Powell (2003), as well as IV models with structural function monotone in a scalar unobservable, for example the quantile IV model studied by Chernozhukov and Hansen (2005). In such models, the characterization delivered by the Corollary delivers a collection of *only* conditional moment *equalities*.

In general the collection of core-determining sets from Theorem 2 and Corollary 2 may be infinite. However, in models in which all endogenous variables are discrete and finite, the sets $\mathsf{Q}^E(h, z)$ and $\mathsf{Q}^I(h, z)$ are finite. In Chesher and Rosen (2012b) we provide an algorithm based on the characterization of core-determining sets in Theorem 2 and Corollary 2 to construct the collections $\mathsf{Q}^E(h, z)$ and $\mathsf{Q}^I(h, z)$ in such models.

3.4 Restrictions on the Joint Distribution of (U, Z)

Given a model $\mathcal{M}$, Theorem 2 provides a characterization of which structures $\left(h, \mathcal{G}_{U|Z}\right)$ belong to the identified set. A key element of econometric models are restrictions on the conditional distributions of unobserved heterogeneity, i.e. restrictions on the collections $\mathcal{G}_{U|Z}$ that are admitted by $\mathcal{M}$. The generality of Theorem 2 allows for its application whatever the specification of $\mathcal{M}$, though of course the size of the identified set $\mathcal{M}^*$ will depend crucially on the restrictions that $\mathcal{M}$ embodies.

In this section we consider particular restrictions on admissible collections of conditional distributions $\mathcal{G}_{U|Z}$, illustrating the use of such restrictions in further characterizing the identified set. The restrictions we consider are well-known in the literature, namely independence, conditional mean, conditional quantile, and parametric restrictions, although Theorem 2 can also be applied with other restrictions.

Stochastic Independence

We begin by considering the implications of stochastic independence of unobservables and exogenous variables set out in the following restriction.

Restriction SI: For all collections $\mathcal{G}_{U|Z}$ of conditional distributions admitted by $\mathcal{M}$, $U \perp\!\!\!\perp Z$. □

With this restriction in place, the conditional distributions $G_{U|Z}(\cdot|z)$ cannot vary with z, and we can simply write G_U in place of the collection $\mathcal{G}_{U|Z}$, where for each z, $G_{U|Z}(\cdot|z) = G_U(\cdot)$, and $\mathcal{M}$ is then denoted by a collection of structures (h, G_U). Let $\mathcal{G}_U \equiv \{G_U : \exists h \text{ s.t. } (h, G_U) \in \mathcal{M}\}$ denote the collection of distributions of unobserved heterogeneity admitted by model $\mathcal{M}$.

It follows from Theorem 2 that a given structure $(h, G_U) \in \mathcal{M}$ belongs to $\mathcal{M}^*$ if and only if G_U is selectionable with respect to the conditional (on $Z = z$) distribution of the random set $\mathcal{U}(Y, Z; h)$ induced by $F_{Y|Z}(\cdot|z)$ a.e. $z \in \mathcal{R}_Z$. A characterization of such structures is succinctly given through

the conditional containment inequality representation, as set out in the following Theorem.

Theorem 4 *Let Restrictions A1-A4 and SI hold. Then*

$$\begin{aligned}\mathcal{M}^* &= \left\{(h, G_U) \in \mathcal{M} : F_{Y|Z}(\cdot|z) \preccurlyeq \mathcal{Y}(U, z; h) \text{ when } U \sim G_U(\cdot),\ a.e.\ z \in \mathcal{R}_Z\right\} \quad (3.6)\\ &= \left\{(h, G_U) \in \mathcal{M} : G_U(\cdot) \preccurlyeq \mathcal{U}(Y, z; h) \text{ when } Y \sim F_{Y|Z}(\cdot|z),\ a.e.\ z \in \mathcal{R}_Z\right\}, \quad (3.7)\end{aligned}$$

equivalently,

$$\begin{aligned}\mathcal{M}^* &= \left\{ \begin{array}{c} (h, G_U) \in \mathcal{M} : \forall \mathcal{S}_I \in \mathsf{Q}^I(h, z),\ \forall \mathcal{S}_E \in \mathsf{Q}^E(h, z), \\ C_h(\mathcal{S}_I|z) \leq G_U(\mathcal{S}_I),\ C_h(\mathcal{S}_E|z) = G_U(\mathcal{S}_E),\ a.e.\ z \in \mathcal{R}_Z \end{array} \right\}. \quad (3.8)\\ &= \left\{ \begin{array}{c} (h, G_U) \in \mathcal{M} : \forall \mathcal{K} \in \mathsf{K}(\mathcal{Y}), \\ F_{Y|Z}(\mathcal{K}|z) \leq G_U\{\mathcal{Y}(U, z; h) \cap \mathcal{K} \neq \emptyset\},\ a.e.\ z \in \mathcal{R}_Z \end{array} \right\}, \quad (3.9)\end{aligned}$$

where $\mathsf{K}(\mathcal{Y})$ *denotes the collection of compact sets in* $\mathcal{R}_Y$*.*

Theorem 4 presents various representations of the identified set under restriction SI. The first two characterizations, (3.6) and (3.7) are direct applications of this restriction to Definition 3.1 and Theorem 2, respectively. The characterization (3.8) applies Theorem 3 and Corollary 2 to provide a characterization of the identified set through the conditional containment functional of $\mathcal{U}(Y, Z; h)$. The representation makes use of core-determining sets to reduce the required number of moment conditions in the characterization, and to distinguish which must hold as equalities and inequalities.

The last characterization, (3.9), characterizes the identified set through conditional moment inequalities implied by the capacity functional applied to random set $\mathcal{Y}(U, z; h)$. These inequalities closely coincide with characterizations provided by Beresteanu, Molchanov, and Molinari (2011, Appendix D.2) and Galichon and Henry (2011) in incomplete models of games. These inequalities must hold applied to all compact sets $\mathcal{K} \subseteq \mathcal{R}_Y$. Galichon and Henry (2011) provide core determining sets for this characterization in $\mathcal{R}_Y$ when a certain monotonicity condition holds. There are however many models where the required monotonicity condition does not hold. Nonetheless, the representation (3.8), and in particular the reduction in moment conditions achieved via the use of core determining sets on $\mathcal{R}_U$ given by Theorem 3, still holds.

A further difference between characterizations (3.8) and (3.9) is how they incorporate restrictions placed on the distribution of unobserved heterogeneity. Given an admissible distribution G_U, the use of characterization (3.9) computationally requires that one compute for each compact set $\mathcal{K}$ the probability that the random outcome set $\mathcal{Y}(U, z; h)$ hits $\mathcal{K}$. This has typically been achieved by means of simulation from each conjectured distribution G_U, see e.g. Beresteanu, Molchanov, and Molinari (2011, Appendix D.2) and Henry, Meango, and Queyranne (2011). $F_{Y|Z}(\mathcal{K}|z)$ is observed directly. On the other hand, characterization (3.9) requires, for each conjectured distribution G_U and each core-determining set $\mathcal{S}$, computation of $G_U(\mathcal{S})$. This can be done again via simulation, or

by means of numerical integration. The term $\mathbb{P}\left[\mathcal{U}\left(Y,Z;h\right)\subseteq\mathcal{S}|z\right]$ is, for any fixed h, the probability of an event concerning only the variables (Y,Z), which is point-identified and can be computed or estimated directly.

It is important to understand that with Restriction SI imposed, Theorem 4 applies with the admissible distributions $\mathcal{G}_U$ either parametrically or non-parametrically specified. For example, if admissible G_U were parameterized by $\lambda\in\Lambda\subseteq\mathbb{R}^d$ we could write $\mathcal{G}_U=\{G_U\left(\mathcal{S};\lambda\right):\lambda\in\Lambda\}$, and $\mathcal{M}^*$ could be represented by a collection of (h,λ) satisfying the equalities and inequalities of (3.8), with $G_U\left(\mathcal{S};\lambda\right)$ replacing $G_U\left(\mathcal{S}\right)$, for all $\mathcal{S}=\mathcal{S}_I\in\mathsf{Q}^I\left(h,z\right)$ and all $\mathcal{S}=\mathcal{S}_E\in\mathsf{Q}^E\left(h,z\right)$. Less stringent conditions on the class $\mathcal{G}_U$, all else equal, will of course result in larger identified sets.

If G_U is left completely unrestricted, then the task of checking the containment functional inequality for all admissible G_U is more difficult than with G_U parametrically specified. Indeed, it is not clear how to do this for all such G_U.

Fortunately, manipulation of the conditional containment functional inequality representation in U-space affords a representation of the identified set of structural functions h that does not explicitly involve the distribution G_U. The ability to characterize distribution-free identified sets using random set theory with statistical independence restrictions is new to the literature. The formal result follows.

Corollary 3 *Let Restrictions A1-A4 and SI hold, and let $\mathcal{G}_{U|Z}$ be otherwise unrestricted. Then the identified set of structural functions h is*

$$\mathcal{H}^*=\left\{\begin{array}{l} h\in\mathcal{H}:\forall\mathcal{S}_I\in\mathsf{Q}^I\left(h,z\right),\ \forall\mathcal{S}_E\in\mathsf{Q}^E\left(h,z\right), \\ \quad\sup\limits_{z\in\mathcal{R}_Z}C_h\left(\mathcal{S}_I|z\right)\leq\inf\limits_{z\in\mathcal{R}_Z}\left(1-C_h\left(\mathcal{S}_I^c|z\right)\right),\ \textit{and} \\ \quad\sup\limits_{z\in\mathcal{R}_Z}C_h\left(\mathcal{S}_E|z\right)=\inf\limits_{z\in\mathcal{R}_Z}\left(1-C_h\left(\mathcal{S}_E^c|z\right)\right),\ \textit{a.e. } z\in\mathcal{R}_Z \end{array}\right\}, \tag{3.10}$$

and the identified set of structures is

$$\mathcal{M}^*=\left\{\begin{array}{l} (h,G_U)\in\mathcal{M}:\forall\mathcal{S}_I\in\mathsf{Q}^I\left(h,z\right),\ \forall\mathcal{S}_E\in\mathsf{Q}^E\left(h,z\right), \\ \quad\sup\limits_{z\in\mathcal{R}_Z}C_h\left(\mathcal{S}_I|z\right)\leq G_U\left(\mathcal{S}_I\right)\leq\inf\limits_{z\in\mathcal{R}_Z}\left(1-C_h\left(\mathcal{S}_I^c|z\right)\right),\ \textit{and} \\ \quad\sup\limits_{z\in\mathcal{R}_Z}C_h\left(\mathcal{S}_E|z\right)=G_U\left(\mathcal{S}_E\right)=\inf\limits_{z\in\mathcal{R}_Z}\left(1-C_h\left(\mathcal{S}_E^c|z\right)\right),\ \textit{a.e. } z\in\mathcal{R}_Z \end{array}\right\}. \tag{3.11}$$

Corollary 3 shows that with a distribution-free specification, the probability $G_U\left(\mathcal{S}\right)$ can be profiled out of the containment functional inequality. This holds because for any set $\mathcal{S}$ we have

$$C_h\left(\mathcal{S}|z\right)\leq G_U\left(\mathcal{S}\right),\text{ and }C_h\left(\mathcal{S}^c|z\right)\leq G_U\left(\mathcal{S}^c\right). \tag{3.12}$$

The second equality is equivalent to

$$G_U(\mathcal{S}) \leq 1 - C_h(\mathcal{S}^c|z),$$

where $1 - C_h(\mathcal{S}^c|z)$ is the *capacity* functional of $\mathcal{U}(Y, Z; h)$ applied to argument $\mathcal{S}$ conditional on $Z = z$.[11] Rearranging and combining with (3.12) we have for almost every $z \in \mathcal{R}_Z$.

$$C_h(\mathcal{S}|z) \leq G_U(\mathcal{S}) \leq 1 - C_h(\mathcal{S}^c|z),$$

which in combination with Corollary 2 produces the characterizations of Corollary 3.

Mean Independence

The following restriction limits the collection $\mathcal{G}_{U|Z}$ to those such that U has conditional mean zero.[12]

Restriction MI: $\mathsf{G}_{U|Z}$ comprises all collections $\mathcal{G}_{U|Z}$ of conditional distributions for U given Z satisfying $E[U|Z = z] = 0$, a.e. $z \in \mathcal{R}_Z$. □

With this restriction imposed on the conditional distributions of unobserved heterogeneity, it is convenient to characterize the selectionability criterion of Theorem 2 by making use of the Aumann expectation. Also referred to as the selection expectation, the Aumann expectation of a random set $\mathcal{A}$ is the set of values that are the expectation of some random variable A that is a selection of $\mathcal{A}$. For clarity we provide the formal definition, repeated from Molchanov (2005, p. 151).

Definition 6 *The **Aumann expectation** of $\mathcal{A}$ is*

$$\mathbb{E}[\mathcal{A}] \equiv \text{cl}\{E[A] : A \in \text{Sel}(\mathcal{A}) \text{ and } E[A] < \infty\}.$$

The Aumann expectation of $\mathcal{A}$ conditional on $B = b$ is

$$\mathbb{E}[\mathcal{A}|b] \equiv cl\{E[A|b] : A \in \text{Sel}(\mathcal{A}) \text{ and } E[A|b] < \infty\}.$$

The resulting characterization of the identified sets for structural function h and for the structure $\left(h, \mathcal{G}_{U|Z}\right)$ employing the Aumann expectation with Restriction MI is given in the following Theorem.

[11] In other words

$$1 - C_h(\mathcal{S}^c|z) = \mathbb{P}[\mathcal{U}(Y, Z; h) \cap \mathcal{S} \neq \emptyset|z].$$

We use $1 - C_h(\mathcal{S}^c|z)$ for the capacity functional rather than introduce further notation or explicitly write out the longer expression $\mathbb{P}[\mathcal{U}(Y, Z; h) \cap \mathcal{S} \neq \emptyset|z]$ repeatedly.

[12] The use of zero in the restriction is simply a location normalization. The restriction $E[U|Z = z] = 0$ can be replaced by $E[U|Z = z] = c$ a.e. $z \in \mathcal{R}_Z$ for any known vector $c \in \mathcal{R}_U$, and Theorem 5 and Corollary 4 go through with c in place of 0.

Theorem 5 *Let Restrictions A1-A4 and MI hold and suppose that* $(\Omega, \mathcal{F}, \mathbb{P})$ *is non-atomic. Then the identified set for structural function* h *are those functions* h *such that* 0 *is an element of the Aumann expectation of* $\mathcal{U}(Y, Z; h)$ *conditional on* $Z = z$*, a.e.* $z \in \mathcal{R}_Z$*:*

$$\mathcal{H}^* = \{h \in \mathcal{H} : 0 \in \mathbb{E}[\mathcal{U}(Y, Z; h) | z], \text{ a.e. } z \in \mathcal{R}_Z\}.$$

The identified set for $\left(h, \mathcal{G}_{U|Z}\right)$ *is:*

$$\mathcal{M}^* = \left\{\left(h, \mathcal{G}_{U|Z}\right) \in \mathcal{M} : h \in \mathcal{H}^* \text{ and } G_{U|Z}(\cdot|z) \lesssim \mathcal{U}(Y, Z; h) \text{ conditional on } Z = z, \text{ a.e. } z \in \mathcal{R}_Z\right\}.$$

Theorem 5 succinctly characterizes $\mathcal{H}^*$ as those h such that $0 \in \mathbb{E}[\mathcal{U}(Y, Z; h) | z]$ a.e. $z \in \mathcal{R}_Z$. The identified set for $\left(h, \mathcal{G}_{U|Z}\right)$ is then simply those pairs of $\left(h, \mathcal{G}_{U|Z}\right)$ such that $0 \in \mathbb{E}[\mathcal{U}(Y, Z; h) | z]$ and $G_{U|Z}(\cdot|z)$ is selectionable with respect to $\mathcal{U}(Y, Z; h)$ conditional on $Z = z$, a.e. $z \in \mathcal{R}_Z$.

Properties of the random set $\mathcal{U}(Y, Z; h)$ can be helpful in characterizing its Aumann expectation, and consequently in determining whether any particular h is in $\mathcal{H}^*$. For example, if $\mathcal{U}(Y, Z; h)$ is integrably bounded, that is if

$$E \sup \{\|U\| : U \in \mathcal{U}(Y, Z; h)\} < \infty, \tag{3.13}$$

then from Molchanov (2005, Theorem 2.1.47-iv, p. 171), $0 \in \mathbb{E}[\mathcal{U}(Y, Z; h) | z]$ if and only if

$$\inf_{v \in \mathcal{R}_Z : \|v\| = 1} E[m(v, \mathcal{U}(Y, Z; h)) | z] \geq 0, \tag{3.14}$$

where for any set $\mathcal{S}$,

$$m(v, \mathcal{S}) \equiv \sup \{v \cdot s : s \in \mathcal{S}\}$$

denotes the support function of $\mathcal{S}$ evaluated at v. Beresteanu, Molchanov, and Molinari (2011) previously employed Molchanov (2005, Theorem 2.1.47-iv, p. 171) in consideration of the conditional Aumann expectation of random outcome set $\mathcal{Y}(Z, U; h)$ in characterizing its selections for identification analysis, and it can likewise lead to simplifications in determining whether $0 \in \mathbb{E}[\mathcal{U}(Y, Z; h) | z]$. Indeed, if structural function h is additively separable in Y, the two representations are equivalent, differing only by a trivial location shift.

More generally, Theorem 5 does not require the sets $\mathcal{U}(Y, Z; h)$ to be integrably bounded, but only integrable, which is important for dealing with cases where the support of unobserved heterogeneity $\mathcal{R}_U$ is unbounded, e.g. when $\mathcal{R}_U$ is some finite dimensional Euclidean space.

In some commonly occurring models, including all those of Examples 1-5 in Section 2.1, the random set $\mathcal{U}(Y, Z; h)$ is convex with probability one. In such cases the characterization of $\mathcal{H}^*$ can be simplified further as in the following Corollary. Unlike the simplification afforded by the support function characterization (3.14), it does not require that $\mathcal{U}(Y, Z; h)$ be integrably bounded.

Corollary 4 *Let the restrictions of Theorem 5 hold and suppose $\mathcal{U}(Y,Z;h)$ is convex with probability one. Then*

$$\mathcal{H}^* = \left\{ \begin{array}{l} h \in \mathcal{H} : E\left[u\left(Y,Z\right)|z\right] = 0 \text{ a.e. } z \in \mathcal{R}_Z, \\ \text{for some function } u : \mathcal{R}_{YZ} \to \mathcal{R}_U \text{ with } \mathbb{P}\left[u\left(Y,Z\right) \in \mathcal{U}\left(Y,Z;h\right)\right] = 1 \end{array} \right\}.$$

Finally, Theorem 5 can be generalized to characterize $\mathcal{H}^*$ under more general forms of conditional mean restrictions than Restriction MI, as expressed in Restriction MI*.

Restriction MI*: $\mathsf{G}_{U|Z}$ comprises all collections $\mathcal{G}_{U|Z}$ of conditional distributions for U given Z such that for some known function $d(\cdot,\cdot) : \mathcal{R}_U \times \mathcal{R}_Z \to \mathbb{R}^{k_d}$, $E\left[d\left(U,Z\right)|z\right] = 0$ a.e. $z \in \mathcal{R}_Z$, where $d(u,z)$ is continuous in u for all $z \in \mathcal{R}_Z$. □

Restriction **MI*** requires that the conditional mean of some function taking values in $\mathbb{R}^{k_d}$, namely $d(U,Z)$, $E\left[d\left(U,Z\right)|z\right] = 0$. This restriction can accommodate models that impose conditional mean restrictions on functions of unobservables U, and nests Restriction MI upon setting $d(U,Z) = U$. To express the identified set delivered under such a restriction, we define

$$\mathcal{D}\left(y,z;h\right) \equiv \left\{d\left(u,z\right) : u \in \mathcal{U}\left(y,z;h\right)\right\}.$$

Thus the random set $\mathcal{D}(Y,Z;h)$ is the set of feasible values for $d(U,Z)$ given observed (Y,Z). Given the requirement of Restriction MI* that $d(\cdot,z)$ is continuous for each z, the set $\mathcal{D}(Y,Z;h)$ is a random *closed* set. The same logic as that used for Theorem 5 then yields the following result.

Corollary 5 *Let Restrictions A1-A4 and MI* hold and suppose that $(\Omega,\mathcal{F},\mathbb{P})$ is non-atomic. Then the identified set for structural function h are those such that* 0 *is an element of the Aumann expectation of $\mathcal{D}(Y,Z;h)$ conditional on $Z = z$, a.e. $z \in \mathcal{R}_Z$:*

$$\mathcal{H}^* = \left\{h \in \mathcal{H} : 0 \in \mathbb{E}\left[\mathcal{D}\left(Y,Z;h\right)|z\right], \text{ a.e. } z \in \mathcal{R}_Z\right\}.$$

The identified set for $\left(h,\mathcal{G}_{U|Z}\right)$ is:

$$\mathcal{M}^* = \left\{\left(h,\mathcal{G}_{U|Z}\right) \in \mathcal{M} : h \in \mathcal{H}^* \text{ and } G_{U|Z}\left(\cdot|z\right) \lesssim \mathcal{U}\left(Y,Z;h\right) \text{ conditional on } Z = z, \text{ a.e. } z \in \mathcal{R}_Z\right\}.$$

Quantile Independence

Our analysis can also accommodate conditional quantile restrictions on unobserved heterogeneity. To illustrate how in a relatively simple yet useful framework we restrict the analysis of this section to models with univariate unobserved heterogeneity $U \in \mathbb{R}$, where the sets $\mathcal{U}(y,z;h)$ are closed intervals with lower bound $\underline{u}(y,z;h)$ and upper bound $\overline{u}(y,z;h)$, formalized in Restriction IS (interval support) below. The lower and upper bounds may be $-\infty$ and $+\infty$, respectively. This

restriction can be fruitfully applied to GIV models with censored endogenous or exogenous variables. We illustrate how in a model with interval censored endogenous variables in Section 4.1.3 below.[13]

Restriction IS: $\forall (y,z) \in \mathcal{R}_{YZ}$,

$$\mathcal{U}(y,z;h) = [\underline{u}(y,z;h), \overline{u}(y,z;h)], \tag{3.15}$$

where possibly $\underline{u}(y,z;h) = -\infty$ or $\overline{u}(y,z;h) = +\infty$, in which case the corresponding endpoint of the interval (3.15) above is open. □

The conditional quantile restriction is formalized as follows.[14]

Restriction QI: For some known $\tau \in (0,1)$, $\mathsf{G}_{U|Z}$ comprises all collections $\mathcal{G}_{U|Z}$ of conditional distributions for U given Z that are continuous in a neighborhood of zero and satisfy the conditional quantile restriction $q_{U|Z}(\tau|z) = 0$, a.e. $z \in \mathcal{R}_Z$. □

For random sets $\mathcal{U}(Y,Z;h)$ with interval-valued realizations, it is easy to check whether there exists a random variable that is selectionable with respect to the distribution of $\mathcal{U}(Y,Z;h)$ conditional on $Z = z$. It is well known that the quantile of a generic random variable W distributed F_W is a parameter that respects stochastic dominance. That is, if $\tilde{W} \sim F_{\tilde{W}}$, and $F_{\tilde{W}}$ stochastically dominates F_W, then $q_W(\tau) \leq q_{\tilde{W}}(\tau)$ for any $\tau \in [0,1]$. The smallest and largest selections of random set $\mathcal{U}(Y,Z;h)$, in terms of stochastic dominance, are those distributions that place all mass on $\underline{u}(Y,Z;h)$ and $\overline{u}(Y,Z;h)$, respectively. Thus, intuitively, the conditional quantiles of $\underline{u}(Y,Z;h)$ and $\overline{u}(Y,Z;h)$ provide sharp bounds on the conditional quantiles of all selections of $\mathcal{U}(Y,Z;h)$. This is formalized with the following result.

Theorem 6 *Let Restrictions A1-A4, IS, and QI hold. Then (i) the identified set for structural function h is*

$$\mathcal{H}^* = \left\{ h \in \mathcal{H} : \sup_{z \in \mathcal{R}_Z} F_{Y|Z}[\overline{u}(Y,Z;h) \leq 0|z] \leq \tau \leq \inf_{z \in \mathcal{R}_Z} F_{Y|Z}[\underline{u}(Y,Z;h) \leq 0|z] \right\}. \tag{3.16}$$

(ii) If $\underline{u}(Y,Z;h)$ and $\overline{u}(Y,Z;h)$ are continuously distributed conditional on $Z = z$, a.e. $z \in \mathcal{R}_Z$, then an equivalent formulation of $\mathcal{H}^*$ is given by

$$\mathcal{H}^* = \left\{ h \in \mathcal{H} : \sup_{z \in \mathcal{R}_Z} \underline{q}(\tau,z;h) \leq 0 \leq \inf_{z \in \mathcal{R}_Z} \overline{q}(\tau,z;h) \right\}, \tag{3.17}$$

[13] The use of conditional quantile restrictions with non-interval $\mathcal{U}(y,z;h)$ and more generally multivariate unobserved heterogeneity is an interesting line of research. This is so even in models with exogenous covariates absent instrumental variable restrictions, and as such is logically distinct from the study of generalized instrumental variable models that is our focus here.

[14] It is straightforward to modify the conditional quantile restriction to $Q_{U|Z}(\tau|z) = c$, for any known $c \in \mathbb{R}$, i.e. the use of $c = 0$ is simply a scale normalization. The ensuing analysis then carries through replacing 0 with c.

where

$$\underline{q}(\tau, z; h) \equiv \tau\text{-quantile of } \underline{u}(Y, Z; h), \quad \overline{q}(\tau, z; h) \equiv \tau\text{-quantile of } \overline{u}(Y, Z; h).$$

(iii) The identified set for $\left(h, \mathcal{G}_{U|Z}\right)$ *is:*

$$\mathcal{M}^* = \left\{\left(h, \mathcal{G}_{U|Z}\right) \in \mathcal{M} : h \in \mathcal{H}^* \text{ and } G_{U|Z}(\cdot|z) \lesssim \mathcal{U}(Y, Z; h) \text{ conditional on } Z = z, \text{ a.e. } z \in \mathcal{R}_Z\right\}.$$

Under Restriction QI, the conditional distributions belonging to $\mathcal{G}_{U|Z}$ are continuous in a neighborhood of zero. Indeed, it is common for econometric models to impose that unobserved heterogeneity is continuously distributed on its entire support. Given the continuity condition, the conditional quantile restriction $q_{U|Z}(\tau|z) = 0$ is equivalent to

$$G_{U|Z}((-\infty, c]\,|z) = \tau \quad \Leftrightarrow \quad c = 0.$$

The inequalities comprising (3.16) then follow from $\underline{u}(Y, Z; h) \leq U \leq \overline{u}(Y, Z; h)$. These inequalities also comprise the containment functional inequality $C_h(\mathcal{S}|z) \leq G_{U|Z}(\mathcal{S}|z)$ applied to test sets $\mathcal{S} = (-\infty, 0]$ and $\mathcal{S} = [0, \infty)$. In the proof of Theorem 6 we show that for any h, if the containment functional inequalities hold for these two test sets, then we can find an admissible collection of conditional distributions $\mathcal{G}_{U|Z}$ such that it holds for all closed test sets in $\mathcal{R}_U$. From Corollary 1 it follows that the characterization (3.16) is sharp.

This result helps to illustrate the relative identifying power under Restriction QI as compared to Restrictions SI. Under Restriction SI, $U \perp\!\!\!\perp Z$, to characterize $\mathcal{H}^*$ it suffices to consider $C_h(\mathcal{S}|z) \leq G_U(\mathcal{S})$ where $G_{U|Z}(\mathcal{S}|z) = G_U(\mathcal{S})$ for all core-determining test sets delivered by Theorem 3, which imply that it will hold for all closed subsets of $\mathcal{R}_U$. Under restriction QI, it is enough to consider $C_h(\mathcal{S}|z) \leq G_{U|Z}(\mathcal{S}|z)$ for only two test sets, namely $\mathcal{S} = (-\infty, 0]$ and $\mathcal{S} = [0, \infty)$.

The second part of Theorem 6 follows because when $\underline{u}(Y, Z; h)$ and $\overline{u}(Y, Z; h)$ are continuous, the inequalities in (3.16) which involve cumulative distributions $F_{Y|Z}[\cdot|z]$ may be inverted. Then $\mathcal{H}^*$ may be equivalently expressed as inequalities involving the lower and upper envelopes, $\underline{q}(\tau, z; h)$ and $\overline{q}(\tau, z; h)$, respectively, of conditional quantile functions for selections of $\mathcal{U}(Y, Z; h)$. Finally, as was the case for identified sets $\mathcal{M}^*$ using conditional mean restrictions given in Theorem 5, the third part of Theorem 6 states that the identified set of structures $\left(h, \mathcal{G}_{U|Z}\right)$ are elements of $\mathcal{H}^*$ paired with distributions $G_{U|Z}(\cdot|z)$ that are selectionable with respect to the conditional distribution of $\mathcal{U}(Y, Z; h)$ given $Z = z$, a.e. $z \in \mathcal{R}_Z$.

4 Illustration: A Model with Interval Censored Endogenous Variable

In this Section we return to Example 5 from Section 2.1, a generalization of a single equation model with an interval censored exogenous variable studied by Manski and Tamer (2002). Like Manski and Tamer (2002) we impose no assumption on the censoring process or the realization of the censored variable relative to the observed interval, but we allow the interval censored explanatory variable to be endogenous. We consider the identifying power of both independence and conditional mean restrictions with respect to unobservable U and observed exogenous variables Z, and we provide numerical illustrations of identified sets given particular data generating structures.

4.1 Identified Sets

The continuously distributed outcome of interest Y_1 is determined by the realizations of endogenous $Y_2^* \in \mathbb{R}$, exogenous $Z = (Z_1, Z_2) \in \mathbb{R}^{k_z}$, and unobservable variable $U \in \mathbb{R}$ with strictly monotone CDF $\Lambda(\cdot)$, such that

$$Y_1 = g(Y_2^*, Z_1, U), \tag{4.1}$$

where the function $g(\cdot,\cdot,\cdot)$ is increasing in its first argument, and strictly increasing in its third argument.[15] The endogenous variable Y_2^* is not observed, but there are observed variables Y_{2l}, Y_{2u} such that

$$Y_2^* = Y_{2l} + W \times (Y_{2u} - Y_{2l}), \tag{4.2}$$

for some unobserved variable $W \in [0,1]$. There is no restriction on the distribution of W on the unit interval, and no restriction its stochastic relation to observed variables. Together (U, W) comprise a two-dimensional vector of unobserved heterogeneity.

Since nothing is assumed about the censoring process, it is convenient to suppress the unobserved variable W by replacing (4.2) with the equivalent formulation

$$\mathbb{P}[Y_{2l} \leq Y_2^* \leq Y_{2u}] = 1. \tag{4.3}$$

The researcher observes a random sample of observations of (Y_1, Y_{2l}, Y_{2u}, Z).

The structural function is

$$h(y, z, u) = |y_1 - g(y_{2l}, z_1, u)|_- + |g(y_{2u}, z_1, u) - y_1|_+ ,$$

[15] It is important to note here that U is *marginally* distributed with CDF $\Lambda(\cdot)$. At this point, we have yet to impose restrictions on the joint distribution of (U, Z), so that for any $z \in \mathcal{R}_U$, The conditional CDF of $U|Z = z$ need not be $\Lambda(\cdot)$. It is straightforward to allow $g(y_2^*, z_1, u)$ increasing or decreasing in y_2^* for all (z_1, u), but we maintain that $g(y_2^*, z_1, u)$ is increasing in this example to simplify the exposition.

and $\mathbb{P}[h(Y,Z,U)=0]=1$ is equivalent to equations (4.1) and (4.3). The level sets in Y-space and U-space, respectively, are

$$\mathcal{Y}(u,z;h)=\{y=(y_1,y_{2l},y_{2u})\in\mathcal{R}_Y : g(y_{2l},z_1,u)\leq y_1\leq g(y_{2u},z_1,u)\},$$

and

$$\mathcal{U}(y,z;h)=\left[g^{-1}(y_{2u},z_1,y_1),g^{-1}(y_{2l},z_1,y_1)\right], \tag{4.4}$$

where g^{-1} denotes the inverse of g in its last argument.

In some of the following developments and indeed in our numerical illustrations we further restrict h, employing the commonly used linear index structure with additive unobservable. To do so we let

$$g(y_2^*,z,u)=\beta y_2^*+z_1\gamma+u, \tag{4.5}$$

where the first element of z_1 is one, and g (and hence h) are now parameterized by $(\beta,\gamma')'\in\mathbb{R}^{\dim(z_1)+1}$.

We now consider some alternative restrictions for the collection of conditional distributions $\mathcal{G}_{U|Z}$.

4.1.1 Stochastic Independence

Consider the restriction $U\perp\!\!\!\perp Z$. Each set $\mathcal{U}(y,z;h)$ is a closed interval on $\mathbb{R}$ and hence connected. Using Theorem 3 we can express the identified set for h as

$$\mathbb{P}[\mathcal{U}(Y,Z;h)\subseteq\mathcal{S}|z]\leq G_U(\mathcal{S}) \tag{4.6}$$

for all $\mathcal{S}\in\mathsf{Q}(h,z)$, where $\mathsf{Q}(h,z)$ is the collection of intervals that can be formed as unions of sets of the form $\left[g^{-1}(y_{2u},z_1,y_1),g^{-1}(y_{2l},z_1,y_1)\right]$. If the components of y are continuously distributed with sufficiently rich support the required test sets may constitute *all* intervals on $\mathbb{R}$.[16] Unless g has very restricted structure, the conditions for (4.6) to hold with equality will in general not be satisfied for any test set $\mathcal{S}$, and hence $\mathsf{Q}^E(h,z)=\emptyset$ and $\mathsf{Q}^I(h,z)=\mathsf{Q}(h,z)$ is the collection of all intervals on $\mathbb{R}$, which we henceforth denote

$$\mathsf{Q}\equiv\left\{[a,b]\in\mathbb{R}^2 : a\leq b\right\}.$$

[16] If the support of Y_1 is limited, application of Theorem 3 may dictate that not all intervals of $\mathbb{R}$ need to be considered as test sets. Nonetheless, this smaller collection core-determining sets will differ for different (h,z). A characterization based on all intervals, although employing more test sets than necessary, has the advantage of being invariant to (h,z). Both characterizations - that using the core determining sets of Theorem 3, or that using all intervals of interval on $\mathbb{R}$ - are for the same identified set. That is, both characterizations are sharp.

Thus we have from Theorem 4 that the identified set is

$$\mathcal{M}^* = \{m \in \mathcal{M} : \forall\, [u_*, u^*] \in \mathsf{Q},\, \mathbb{P}\left[\mathcal{U}\left(Y, Z; h\right) \subseteq [u_*, u^*]\,|z\right] \leq \Lambda\left(u^*\right) - \Lambda\left(u_*\right),\, \text{a.e. } z \in \mathcal{R}_Z\}\,.$$

Given structural function h, the probability $\mathbb{P}\left[\mathcal{U}\left(Y, Z; h\right) \subseteq [u_*, u^*]\,|z\right]$ is an event concerning only observed variables, and is thus identified. Specifically, the containment functional inequality in the definition of $\mathcal{H}^*$ can be equivalently written

$$\mathbb{P}\left[u_* \leq g^{-1}\left(Y_{2u}, Z_1, Y_1\right) \wedge g^{-1}\left(Y_{2l}, Z_1, Y_1\right) \leq u^* | z\right] \leq \Lambda\left(u^*\right) - \Lambda\left(u_*\right),$$

or, using monotonicity of $g\left(y_2, z_1, u\right)$ in its third argument,

$$\mathbb{P}\left[g\left(Y_{2u}, Z_1, u_*\right) \leq Y_1 \leq g\left(Y_{2l}, Z_1, u^*\right)|z\right] \leq \Lambda\left(u^*\right) - \Lambda\left(u_*\right). \tag{4.7}$$

With the added linear index restriction from (4.5) this produces the following representation for the identified set, where the model $\mathcal{M}$ stipulates a collection of admissible parameters β, γ and CDFs $\Lambda\left(\cdot\right)$.

$$\mathcal{M}^* = \left\{ \begin{array}{l} m \in \mathcal{M} : \forall\, [u_*, u^*] \in \mathsf{Q}, \\ \quad \mathbb{P}\left[u_* + \beta Y_{2u} \leq Y_1 - Z_1\gamma \leq u^* + \beta Y_{2l} | z\right] \leq \Lambda\left(u^*\right) - \Lambda\left(u_*\right),\, \text{a.e. } z \in \mathcal{R}_Z \end{array} \right\}. \tag{4.8}$$

We now specialize this result for a model incorporating a parametric restriction for Λ, and for a model leaving Λ completely unspecified.

Gaussian Unobservables

Suppose in addition to the linear index restriction (4.5) we further restrict $\Lambda\left(\cdot\right)$ to be a Gaussian CDF with variance $\sigma > 0$ so that $\Lambda\left(u\right) = \Phi\left(\sigma^{-1}u\right)$, where $\Phi\left(\cdot\right)$ is the standard normal CDF. In this case the model is fully parameterized by $\theta \equiv \left(\beta, \gamma', \sigma\right)'$, and $\mathcal{M}$ can be represented as the parameter space Θ for admissible θ. Using (4.8) the identified set is now

$$\mathcal{M}^* = \left\{ \begin{array}{l} \theta \in \Theta : \forall\, [u_*, u^*] \in \mathsf{Q}, \\ \quad \mathbb{P}\left[u_* + \beta Y_{2u} \leq Y_1 - Z_1\gamma \leq u^* + \beta Y_{2l} | z\right] \leq \Phi\left(\sigma^{-1}u^*\right) - \Phi\left(\sigma^{-1}u_*\right),\, \text{a.e. } z \in \mathcal{R}_Z \end{array} \right\}. \tag{4.9}$$

Equivalently, we can employ the change of variables $t^* = \Phi\left(\sigma^{-1}u^*\right)$ and $t_* \equiv \Phi\left(\sigma^{-1}u_*\right)$ to produce the following.[17]

$$\mathcal{M}^* = \left\{ \begin{array}{l} \theta \in \Theta : \forall\, [t_*, t^*] \subseteq [0, 1], \\ \quad \mathbb{P}\left[t_* \leq \Phi\left(\frac{Y_1 - \beta Y_{2u} - Z_1\gamma}{\sigma}\right) \wedge \Phi\left(\frac{Y_1 - \beta Y_{2l} - Z_1\gamma}{\sigma}\right) \leq t^* | z\right] \leq t^* - t_*,\, \text{a.e. } z \in \mathcal{R}_Z \end{array} \right\}. \tag{4.10}$$

[17] This can also be derived by normalizing the distribution of unobserved heterogeneity U to be uniform on the unit interval and defining $g\left(y_2, z_1, u\right) = \beta y_2 + z_1\gamma + \sigma\Phi^{-1}\left(u\right)$.

Using (4.9), the identified set can be represented as the set of parameter values θ satisfying the collection of conditional moment inequalities

$$E\left[m\left(\theta;Y,Z,u_*,u^*\right)|z\right]\leq 0,\quad \text{all } u_*,u^*\in\mathbb{R} \text{ s.t. } u_*\leq u^*, \text{ a.e. } z\in\mathcal{R}_Z,$$

with moment function

$$m\left(\theta;Y,Z.t_*,t^*\right)\equiv 1\left[u_*+\beta Y_{2u}\leq Y_1-Z_1\gamma\leq u^*+\beta Y_{2l}\right]-\left(\Phi\left(\sigma^{-1}u^*\right)-\Phi\left(\sigma^{-1}u_*\right)\right).$$

Distribution-Free Unobservables

Suppose now that we impose the independence restriction $U\perp\!\!\!\perp Z$ and the same additive index structure for g, but without imposing a parametric restriction on unobserved heterogeneity. If Y_2^* were observed, we would require a location normalization for identification of the first component of γ, the intercept. Thus it will be prudent to incorporate a location normalization in our model with Y_2^* censored as well, for example that the median of $U|Z=z$ is zero. Since Y_1 is continuously distributed, there is no scale normalization to be made.

We apply Corollary 3 to obtain the identified set h, equivalently that for parameters $\theta\equiv(\beta,\gamma')'$. To do so, we start with

$$\mathbb{P}\left[u_*+\beta Y_{2u}\leq Y_1-Z_1\gamma\leq u^*+\beta Y_{2l}|z\right]\leq\Lambda\left(u^*\right)-\Lambda\left(u_*\right), \text{ a.e.}z\in\mathcal{R}_Z, \tag{4.11}$$

for all $[u_*,u^*]\in\mathsf{Q}$ from (4.8) above. Noting that $G_U\left(\mathcal{S}\right)=\Lambda\left(u^*\right)-\Lambda\left(u_*\right)$ for any set $\mathcal{S}=[u_*,u^*]$ and following Corollary 3 we also have for all $-\infty<u_*\leq u^*<\infty$ and a.e. $z\in\mathcal{R}_Z$,

$$\begin{aligned}\Lambda\left(u^*\right)-\Lambda\left(u_*\right) &\leq 1-C_h\left(\mathcal{S}^c|z\right) \\ &= 1-\mathbb{P}\left[Y_1-\beta Y_{2l}-Z_1\gamma<u_*\vee Y_1-\beta Y_{2u}-Z_1\gamma>u^*|z\right] \\ &= \mathbb{P}\left[u_*+\beta Y_{2l}\leq Y_1-Z_1\gamma\leq u^*+\beta Y_{2u}|z\right].\end{aligned} \tag{4.12}$$

Define now

$$\begin{aligned}\underline{G}\left(\theta,u_*,u^*\right) &\equiv \sup_{z\in\mathcal{R}_Z}\mathbb{P}\left[u_*+\beta Y_{2u}\leq Y_1-Z_1\gamma\leq u^*+\beta Y_{2l}|z\right], \\ \overline{G}\left(\theta,u_*,u^*\right) &\equiv \inf_{z\in\mathcal{R}_Z}\mathbb{P}\left[u_*+\beta Y_{2l}\leq Y_1-Z_1\gamma\leq u^*+\beta Y_{2u}|z\right],\end{aligned}$$

each of which are identified for any parameter vector $\theta=(\beta,\gamma')'$ from knowledge of $\mathcal{F}_{Y|Z}$ under Restriction A2. Combining (4.11) and (4.12) as in Corollary 3, the identified set for parameters θ, where Θ denotes values admitted by model $\mathcal{M}$, is given by

$$\Theta^*=\left\{\theta\in\Theta:\forall\left[u_*,u^*\right]\in\mathsf{Q},\,\underline{G}\left(\theta,u_*,u^*\right)\leq\overline{G}\left(\theta,u_*,u^*\right)\right\}.$$

The identified set for $(\theta, \Lambda(\cdot))$ is

$$\mathcal{M}^* = \left\{(\theta, \Lambda(\cdot)) \in \mathcal{M} : \forall [u_*, u^*] \in \mathsf{Q},\, \underline{G}(\theta, u_*, u^*) \leq \Lambda(u^*) - \Lambda(u_*) \leq \overline{G}(\theta, u_*, u^*)\right\}.$$

4.1.2 Mean Independence

Now suppose we continue to assume the linear index structure (4.5) but replace the restriction $U \perp\!\!\!\perp Z$ with the conditional mean restriction $E[U|Z=z] = 0$ a.e. $z \in \mathcal{R}_Z$, equivalently Restriction MI from Section 3.4.

The random set $\mathcal{U}(Y, Z; h)$ in this model is given by the interval

$$\mathcal{U}(Y, Z; h) = [Y_1 - Z_1\gamma - \beta Y_{2u}, Y_1 - Z_1\gamma - \beta Y_{2l}],$$

rendering application of Theorem 5 and Corollary 4 particularly simple. This is because there exists a function $u(\cdot,\cdot)$ satisfying the conditions of Corollary 4, namely that (i) $E[u(Y,Z)|z] = 0$ a.e. $z \in \mathcal{R}_Z$, and (ii) $\mathbb{P}[u(Y,Z) \in \mathcal{U}(Y,Z;h)] = 1$ if and only if

$$E[Y_1 - Z_1\gamma - \beta Y_{2u}|z] \leq 0 \leq E[Y_1 - Z_1\gamma - \beta Y_{2l}|z] \text{ a.e. } z \in \mathcal{R}_Z.$$

Thus, applying Corollary 4, the identified set for $\theta \equiv (\beta, \gamma')'$, where again Θ denotes values admitted by model $\mathcal{M}$, is

$$\Theta^* = \left\{\theta \in \Theta : \underline{\mathbb{E}}(\theta) \leq 0 \leq \overline{\mathbb{E}}(\theta)\right\},$$

where

$$\underline{\mathbb{E}}(\theta) \equiv \sup_{z \in \mathcal{R}_Z} E[Y_1 - Z_1\gamma - \beta Y_{2u}|z], \quad \overline{\mathbb{E}}(\theta) \equiv \inf_{z \in \mathcal{R}_Z} E[Y_1 - Z_1\gamma - \beta Y_{2l}|z].$$

4.1.3 Quantile Independence

Finally, we consider the linear index structure (4.5) coupled with Restriction QI. That is, we now assert $q_{U|Z}(\tau|z) = 0$, a.e $z \in \mathcal{R}_Z$.

Again we have under (4.5) that

$$\mathcal{U}(Y, Z; h) = [Y_1 - Z_1\gamma - \beta Y_{2u}, Y_1 - Z_1\gamma - \beta Y_{2l}],$$

and the identified set for h is isomorphic to that of $\theta \equiv (\beta, \gamma')'$. As in Section 4.1.2 we again denote the parameter space and identified set for θ as Θ and Θ^*, respectively. Applying Theorem 6 the identified set of structural functions h is

$$\Theta^* = \left\{\theta \in \Theta : \sup_{z \in \mathcal{R}_Z} F_{Y|Z}[Y_1 \leq Z_1\gamma + \beta Y_{2l}|z] \leq \tau \leq \inf_{z \in \mathcal{R}_Z} F_{Y|Z}[Y_1 \leq Z_1\gamma + \beta Y_{2u}|z]\right\},$$

equivalently,

$$\Theta^* = \left\{ \theta \in \Theta : \sup_{z \in \mathcal{R}_Z} q_{\underline{V_\theta}|Z}(\tau|z) \leq z_1\gamma \leq \inf_{z \in \mathcal{R}_Z} q_{\overline{V_\theta}|Z}(\tau|z) \right\},$$

where $\underline{V_\theta} \equiv Y_1 - \beta Y_{2u}$ and $\overline{V_\theta} \equiv Y_1 - \beta Y_{2l}$. The identified set of structures $\mathcal{M}^*$ is then pairs of structural functions h parameterized by $\theta \in \Theta^*$ coupled with collections of conditional distributions $\mathcal{G}_{U|Z}$ satisfying the required conditional quantile restriction, and such that $G_{U|Z}(\cdot|z)$ is selectionable with respect to $\mathcal{U}(Y, Z; h)$ conditional on $Z = z$, a.e. $z \in \mathcal{R}_Z$.

4.2 Numerical Illustrations

In this section we provide illustrations of identified sets obtained for the interval censored endogenous variable model with the linear index restriction of (4.5). We consider the identified set obtained under the restriction that $U \sim N(0, \sigma)$ and $U \perp\!\!\!\perp Z$, i.e. the Gaussian unobservable case above with identified set given by (4.9).

To generate probability distributions $\mathcal{F}_{Y|Z}$ for observable variables (Y, Z) we employ a triangular Gaussian structure as follows.

$$\begin{aligned} Y_1 &= \gamma_0 + \gamma_1 Y_2^* + U, \\ Y_2^* &= \delta_0 + \delta_1 Z + V. \end{aligned}$$

with $(U, V) \perp\!\!\!\perp Z$, $\mathcal{R}_Z = \{-1, 1\}$, and

$$\begin{bmatrix} U \\ V \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_{11} & \sigma_{1v} \\ \sigma_{1v} & \sigma_{vv} \end{bmatrix} \right).$$

In this model there are no exogenous covariates Z_1, equivalently $Z = Z_2$.

We specify a censoring process that in place of Y_2^* reveals only to which of a collection of mutually exclusive intervals Y_2^* belongs. Such censoring processes are common in practice, for instance when interval bands are used for income in surveys. Specifically, we assume a sequence of J intervals, $I_1, I_2, \ldots, I_J$ with $I_j \equiv (c_j, c_{j+1}]$ and $c_j < c_{j+1}$ for all $j \in \{1, \ldots, J\}$. The censoring process is such that

$$\forall j \in \{1, \ldots, J\}, \quad (Y_{2l}, Y_{2u}) = (c_j, c_{j+1}) \Leftrightarrow Y_2^* \in I_j.$$

As above, the researcher observes a random sample of observations of (Y_1, Y_{2l}, Y_{2u}, Z).

In our examples we consider two data generation processes denoted DGP1 and DGP2, each with parameter values

$$\gamma_0 = 0, \quad \gamma_1 = 1, \quad \delta_0 = 0, \quad \delta_0 = 1, \quad \sigma_{11} = 0.5, \quad \sigma_{1v} = 0.25, \quad \sigma_{vv} = 0.5,$$

	c_1	c_2	c_3	c_4	c_5	c_6	c_7	c_8	c_9	c_{10}	c_{11}	c_{12}	c_{13}
DGP1	$-\infty$	-1.15	-0.67	-0.32	0.00	0.32	0.67	1.15	$+\infty$	-	-	-	-
DGP2	$-\infty$	-1.38	-0.97	-0.67	-0.43	-0.21	0.00	0.21	0.43	0.67	0.97	1.38	$+\infty$

Table 2: Endpoints of censoring process intervals in numerical illustrations.

and interval censoring endpoints $c_1, ..., c_J$ listed in Table 2. In DGP1, Y_2^* was censored into 8 intervals $I_j = (c_j, c_{j+1}]$ with endpoints given by the normal quantile function evaluated at 9 equally spaced values in $[0, 1]$, inclusive of 0 and 1. In DGP2, Y_2^* was censored into 12 such intervals with endpoints given by the normal quantile function evaluated at 13 equally spaced values.

Given these data generation processes, the distribution of $Y \equiv (Y_1, Y_{2l}, Y_{2u})$ conditional on Z is easily obtained as the product of the conditional distribution of (Y_{2l}, Y_{2u}) given Y_1 and Z and the distribution of Y_1 given Z. Combining these probabilities and the inequalities of (4.9), the conditional containment functional for random set $\mathcal{U}(Y, Z; h)$ applied to test set $\mathcal{S} = [u_*, u^*]$ is given by

$$C_\theta([u_*, u^*]|z) = \sum_j \mathbb{P}[g_1 c_{j+1} + u_* \leq Y_1 - g_0 \leq g_1 c_j + u^* | z, [Y_{2l}, Y_{2u}) = I_j] * \mathbb{P}[[Y_{2l}, Y_{2u}) = I_j | z], \tag{4.13}$$

where $\theta = (g_0, g_1, s)$ is used to denote generic parameter values for $(\gamma_0, \gamma_1, \sigma_1)$. C_θ replaces C_h for the containment functional, since in this model the structural function h is a known function of θ.[18] The identified set of structures $\left(h, \mathcal{G}_{U|Z}\right)$ is completely determined by the identified set for θ, which, following (4.10), is given by

$$\Theta^* = \left\{ \begin{array}{l} \theta \in \Theta : \forall [t_*, t^*] \subseteq [0, 1], \\ C_\theta\left(\left[s\Phi^{-1}(t_*), s\Phi^{-1}(t^*)\right] | z\right) \leq t^* - t_*, \text{ a.e. } z \in \mathcal{R}_Z \end{array} \right\}. \tag{4.14}$$

The set Θ^* comprises parameter values (g_0, g_1, s) such that the given conditional containment functional inequality holds for almost every z and *all* intervals $[t_*, t^*] \subseteq [0, 1]$. This collection of test sets is uncountable. For the purpose of illustration we used various combinations of collections Q_M of intervals from the full set of all possible $[t_*, t^*] \subseteq [0, 1]$. Each collection of intervals Q_M comprises the $M(M+1)/2 - 1$ *super-diagonal* elements of the following $(M+1) \times (M+1)$ array

[18] Computational details for the conditional containment probability $C_\theta([u_*, u^*]|z)$ are provided in Appendix C.

of intervals, excluding the interval $[0,1]$, where $m \equiv 1/M$.

$$\begin{bmatrix} [0,0] & [0,m] & [0,2m] & [0,3m] & \cdots & \cdots & \cdots & [0,1] \\ - & [m,m] & [m,2m] & [m,3m] & \cdots & \cdots & \cdots & [m,1] \\ - & - & [2m,2m] & [2m,3m] & \cdots & \cdots & \cdots & [2m,1] \\ - & - & - & [3m,3m] & \ddots & & & \vdots \\ \vdots & \vdots & \vdots & \vdots & & \ddots & & \vdots \\ - & - & - & - & & & [(M-1)m,(M-1)m] & [(M-1)m,1] \\ - & - & - & - & & & - & [1,1] \end{bmatrix}$$

The inequalities of (4.14) applied to those intervals of any collections of test sets Q_M defines an outer region for the identified set, with larger collections of test sets providing successively better approximations of the identified set.

Figure 1 shows three dimensional (3D) plots of outer regions for (g_0, g_1, s). Outer regions using $M \in \{5,7,9\}$ are noticeably smaller than those using only $M = 5$.[19] There was a noticeable reduction in the size of the outer region in moving from $M = 5$ to $M = \{5,7\}$, but hardly any change on including also the inequalities obtained with $M = 9$. Thus, only the outer regions obtained using $M = 5$ and $M \in \{5,7,9\}$ are shown. Figure 2 shows two dimensional projections of the outer region using $M \in \{5,7,9\}$ for each pair of the three parameter components. The surfaces of these sets were drawn as convex hulls of those points found to lie inside the outer regions and projections considered.[20] We have no proof of the convexity of the outer regions in general, but careful investigation of points found to lie in the outer regions strongly suggested that in the cases considered the sets are convex.

Figure 3 shows the 3D outer region for DGP2 employing 12 bins for the censoring of Y_2^* and $M \in \{5,7,9\}$. Compared to Figure 1, this outer region is smaller, as expected given the finer granularity of intervals with 12 rather than 8 bins. Figure 4 shows two dimensional projections for this outer region, again projecting onto each pair of parameter components. These projections help to further illustrate the extent of the reduction in the size of the outer region for DGP2 relative to DGP1.

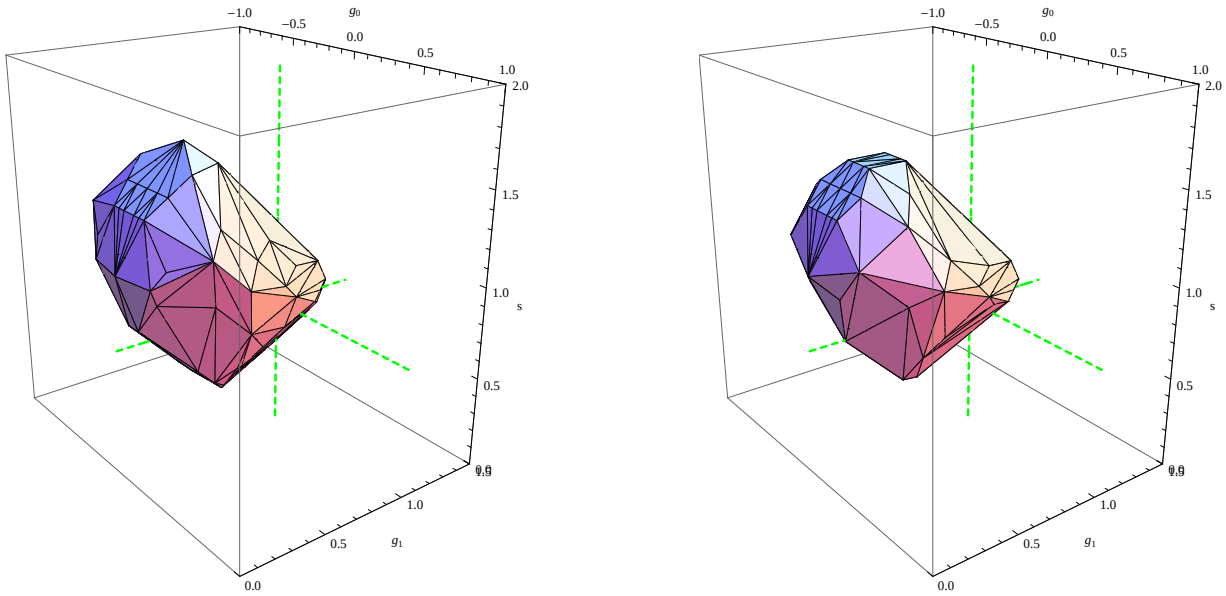
5 Conclusion

In this paper we have studied a broad class of Generalized Instrumental Variable (GIV) models, extending the use of instrumental variable restrictions to models with more flexible specifications for unobserved heterogeneity than previously allowed. In particular, our analysis allows for the

[19] The notation $M \in \{m_1, m_2, ..., m_R\}$ corresponds to the use of test sets $\mathsf{Q}_{m_1} \cup \mathsf{Q}_{m_2} \cdots \cup \cdots \cup \mathsf{Q}_{m_R}$.

[20] The 3D figures were produced using the `TetGenConvexHull` function available *via* the `TetGenLink` package in `Mathematica 9`. 2D figures below were drawn using `Mathematica`'s `ConvexHull` function.

Figure 1: Outer regions for parameters (g_0, g_1, s) for DGP1 with 8 bins using the 14 inequalities generated with $M = 5$ (left pane) and the 85 inequalities generated with $M \in \{5, 7, 9\}$ (right pane).



application of instrumental variable restrictions to models in which the structural mapping from unobserved heterogeneity to observed endogenous variables is not invertible. Thus, these models permit general forms of multivariate, nonadditive unobserved heterogeneity, as often appears for example in nonlinear models with discrete or mixed discrete/continuous outcomes, random coefficient models, and models with censoring. Without the existence of a unique value of unobserved heterogeneity, or "generalized residual", given values of observed variables, rank or more generally completeness conditions do not guarantee point identification of structural functions. We provided a comprehensive framework for characterizing identified sets for model structures in such contexts.

Using tools from random set theory, relying in particular on the concept of *selectionability*, we formally extended the classical notion of observational equivalence to models whose structures need not deliver a unique conditional distribution for endogenous variables given exogenous variables. We showed that the closely related definition of a model's identified set of structures may be equivalently formulated in terms of selectionability criteria in the space of unobserved heterogeneity. This formulation enables direct incorporation of restrictions on conditional distributions of unobserved heterogeneity, of the sort typically employed in econometric models, as we demonstrated by characterizing identified sets under stochastic independence, mean independence, and quantile independence restrictions. We specialized these characterizations to a model with interval censored endogenous explanatory variables, with a censoring and linear index structure following Manski and Tamer (2002), but where we relaxed the requirement that censored variables be exogenous. We provided numerical illustrations of identified sets in such models under a stochastic independence

Figure 2: Outer region projections for DGP1 onto the (g_0, g_1), (g_0, s), and (g_1, s) planes, respectively, with 8 bins using inequalities generated with $M \in \{5, 7, 9\}$. The red point marks the data generating value.

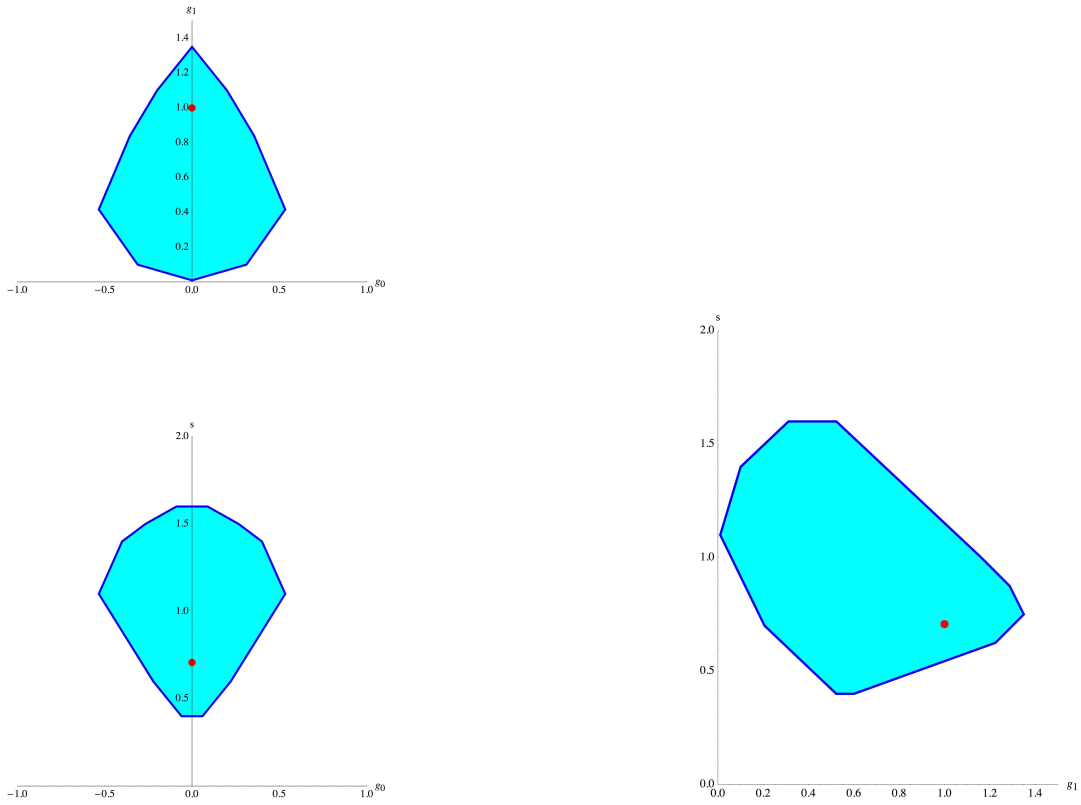


Figure 3: Outer region for DGP2 with 12 bins calculated using inequalities generated with $M \in \{5, 7, 9\}$. Dashed green lines intersect at the data generating value of the parameters.

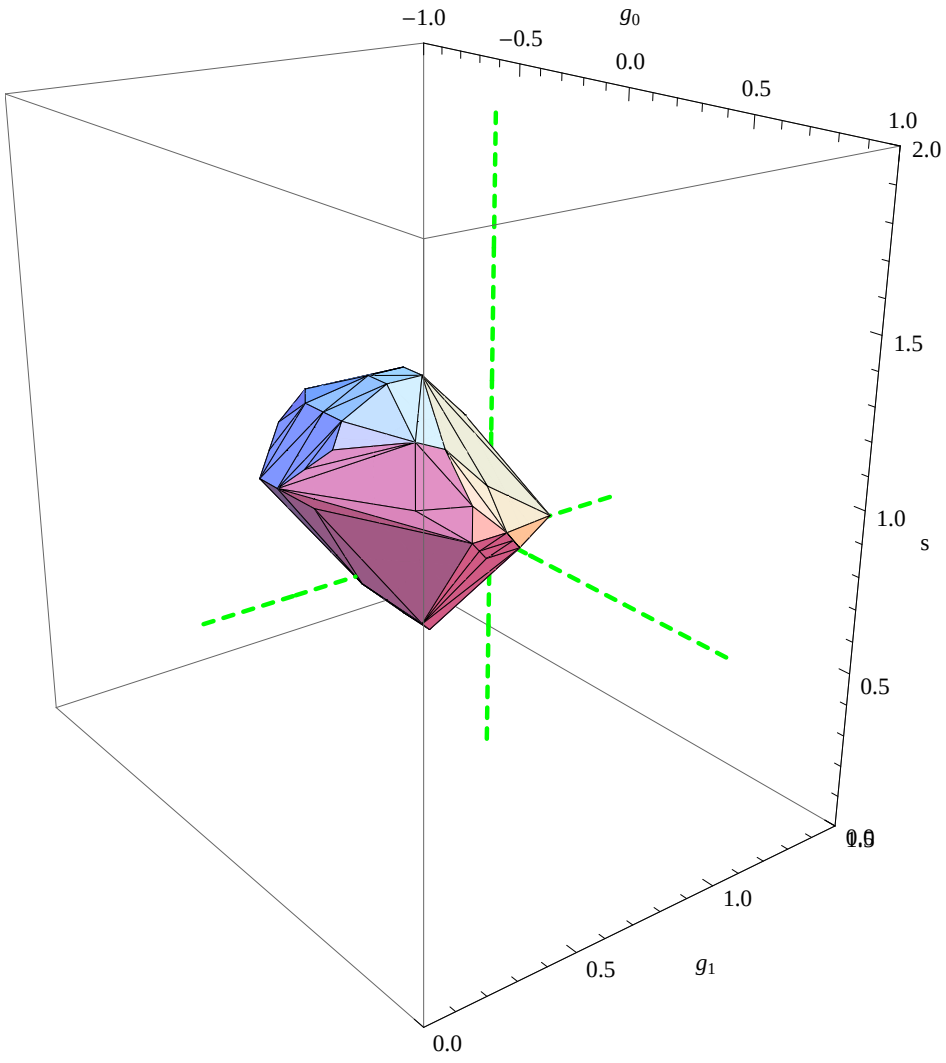
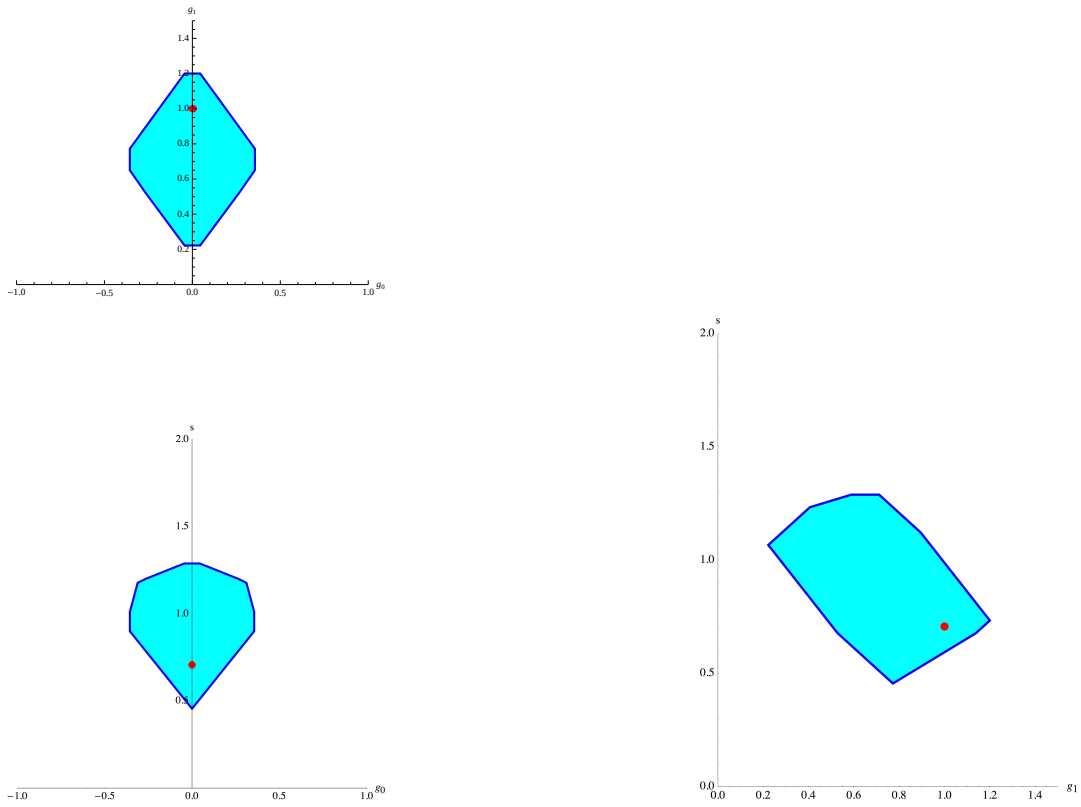


Figure 4: Outer region projections for DGP2 onto the (g_0, g_1), (g_0, s), and (g_1, s) planes, respectively, with 12 bins using inequalities generated with $M \in \{5, 7, 9\}$. The red point marks the data generating value.



restriction between instrumental variables and unobserved heterogeneity.

Our characterizations of identified sets can all be written in the form of conditional (moment or quantile) inequality restrictions. Estimation and inference based on such restrictions is a continuing topic of research in the econometrics literature. In many models, for example those employing conditional mean or conditional quantile restrictions on unobserved heterogeneity, existing methods such as those of Andrews and Shi (2013) and Chernozhukov, Lee, and Rosen (2013) are directly applicable. In other models, in particular those with stochastic independence restrictions where the number of core determining sets can be extremely large, the number of conditional inequality restrictions characterizing the identified set may be vast relative to sample size. This raises complications both in terms of computation and the quality of asymptotic approximations in finite samples for inference methods based on inequality restrictions, thus motivating future research on inference methods based on extremely large numbers of inequalities relative to sample size, or based on altogether different characterizations of the required selectionability criteria for identified sets.

References

Andrews, D. W. K. (2011): "Examples of L2-complete and Boundedly-Complete Distributions," Cowles Foundation Discussion Paper.

Andrews, D. W. K., and X. Shi (2013): "Inference for Parameters Defined by Conditional Moment Inequalities," *Econometrica*, 81(2), 609–666.

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

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.

Blundell, R., X. Chen, and D. Kristensen (2007): "Semi-Nonparametric IV Estimation of Shape-Invariant Engel Curves," *Econometrica*, 75(6), 1613–1669.

Bowden, R. (1973): "The Theory of Parametric Identification," *Econometrica*, 41(6), 1069–1074.

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.

Canay, I., A. Santos, and A. Shaikh (2012): "On the Testability of Identification in Some Nonparametric Models with Endogeneity," *Econometrica*, forthcoming. Formerly CeMMAP working paper CWP18/12.

Chen, X., V. Chernozhukov, S. Lee, and W. Newey (2011): "Local Identification of Nonparametric and Semiparametric Models," CWP17/11.

Chernozhukov, V., and C. Hansen (2005): "An IV Model of Quantile Treatment Effects," *Econometrica*, 73, 245–261.

Chernozhukov, V., G. Imbens, and W. Newey (2007): "Instrumental Variable Estimation of Nonseparable Models," *Journal of Econometrics*, 139, 4–14.

Chernozhukov, V., S. Lee, and A. Rosen (2013): "Intersection Bounds: Estimation and Inference," *Econometrica*, 81(2), 667–737.

Chesher, A. (2010): "Instrumental Variable Models for Discrete Outcomes," *Econometrica*, 78(2), 575–601.

Chesher, A., and A. Rosen (2012a): "An Instrumental Variable Random Coefficients Model for Binary Outcomes," *Econometrics Journal*, forthcoming. Formerly CeMMAP working paper CWP34/12.

——— (2012b): "Simultaneous Equations Models for Discrete Outcomes: Coherence, Completeness, and Identification," CeMMAP working paper CWP21/12.

Chesher, A., A. Rosen, and K. Smolinski (2013): "An Instrumental Variable Model of Multiple Discrete Choice," *Quantitative Economics*, 4(2), 157–196.

Chesher, A., and A. M. Rosen (2013): "What Do Instrumental Variable Models Deliver With Discrete Dependent Variables?," *American Economic Reivew: Papers and Proceedings*, 1(3), 557–562.

Chesher, A., and K. Smolinski (2012): "IV Models of Ordered Choice," *Journal of Econometrics*, 166(1), 33–48.

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

Darolles, S., Y. Fan, J.-P. Florens, and E. Renault (2011): "Nonparametric Instrumental Regression," *Econometrica*, 79(5), 1541–1565.

D'Haultfoeuille, X. (2011): "On the Completeness Condition in Nonparametric Instrumental Problems," *Econometric Theory*, 27, 460–471.

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

Haile, P. A., and E. Tamer (2003): "Inference with an Incomplete Model of English Auctions," *Journal of Political Economy*, 111(1), 1–51.

Hall, P., and J. Horowitz (2005): "Nonparametric Methods for Inference in the Presence of Instrumental Variables," *Annals of Statistics*, 33(6), 2904–2929.

Henry, M., R. Meango, and M. Queyranne (2011): "Combinatorial Bootstrap Inference in Partially Identified Incomplete Structural Models," working paper, University of Montreal.

Hoderlein, S., and E. Mammen (2007): "Identification of Marginal Effects in Nonseparable Models Without Monotonicity," *Econometrica*, 75(5), 1513–1518.

Hurwicz, L. (1950): "Generalization of the Concept of Identification," in *Statistical Inference in Dynamic Economic Models*, ed. by C. C. for Research in Economics, pp. 245–257. John Wiley and Sons.

Koopmans, T., and O. Reiersøl (1950): "The Identification of Structural Characteristics," *Annals of Mathematical Statistics*, 21(165-181), 1–27.

Koopmans, T., H. Rubin, and R. Leipnik (1950): "Measuring the equation systems of dynamic economics," in *Statistical Inference in Dynamic Economic Models*, ed. by C. C. for Research in Economics, pp. 53–237. John Wiley and Sons.

Manski, C. F., and E. Tamer (2002): "Inference on Regressions with Interval Data on a Regressor or Outcome," *Econometrica*, 70(2), 519–546.

Matzkin, R. (2007): "Nonparametric Identification," in *The Handbook of Econometrics*, ed. by J. J. Heckman, and E. E. Leamer, vol. 6b, pp. 5307–5368. North-Holland.

——— (2008): "Identification in Nonparametric Simultaneous Equations Models," *Econometrica*, 76, 945–978.

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

Newey, W. K., and J. L. Powell (2003): "Instrumental Variable Estimation of Nonparametric Models," *Econometrica*, 71, 1565–1578.

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

Rothenberg, T. J. (1971): "Identification in Parametric Models," *Econometrica*, 39(3), 577–591.

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

A Proofs

Proof of Theorem 1. *Fix $z \in \mathcal{R}_Z$ and suppose that $F_{Y|Z}(\cdot|z)$ is selectionable with respect to the conditional distribution of $\mathcal{Y}(U, Z; h)$ given $Z = z$. By Restriction A3, U is conditionally distributed $G_{U|Z}(\cdot|z)$ given $Z = z$, and thus selectionability implies that there exist random variables $\tilde{Y}$ and $\tilde{U}$ such that*

(i) $\tilde{Y}|Z = z \sim F_{Y|Z}(\cdot|z)$,

(ii) $\tilde{U}|Z = z \sim G_{U|Z}(\cdot|z)$,

(iii) $\mathbb{P}\left[\tilde{Y} \in \mathcal{Y}\left(\tilde{U}, Z; h\right)|Z = z\right] = 1$.

By Restriction A3, $\tilde{Y} \in \mathcal{Y}\left(\tilde{U}, Z; h\right)$ if and only if $h\left(\tilde{Y}, Z, \tilde{U}\right) = 0$, equivalently $\tilde{U} \in \mathcal{U}\left(\tilde{Y}, Z; h\right)$. Condition (iii) is therefore equivalent to

$$\mathbb{P}\left[\tilde{U} \in \mathcal{U}\left(\tilde{Y}, Z; h\right)|Z = z\right] = 1. \tag{A.1}$$

Thus there exist random variables $\tilde{Y}$ and $\tilde{U}$ satisfying (i) and (ii) such that (A.1) holds, equivalently such that $G_{U|Z}(\cdot|z)$ is selectionable with respect to the conditional distribution of $\mathcal{U}(Y, Z; h)$ given $Z = z$. The choice of z was arbitrary, and the argument thus follows for all $z \in \mathcal{R}_Z$. ■

Proof of Theorem 2. *This follows directly from application of Theorem **1** to Definitions 2 and 3, respectively.* ■

Proof of Corollary 1. *From the selectionability characterization of $\mathcal{M}^*$ in U-space in Theorem 2, we have that*

$$\mathcal{M}^* = \left\{(h, G_U) \in \mathcal{M} : G_U(\cdot|z) \preccurlyeq \mathcal{U}(Y, z; h) \text{ when } Y \sim F_{Y|Z}(\cdot|z), \text{ a.e. } z \in \mathcal{R}_Z\right\}.$$

Fix $z \in \mathcal{R}_Z$ and suppose $Y \sim F_{Y|Z}(\cdot|z)$. From Artstein's Inequality, see Artstein (1983), Norberg (1992), or Molchanov (2005, Section 1.4.8.), $G_U(\cdot|z) \preccurlyeq \mathcal{U}(Y, z; h)$ if and only if

$$\forall \mathcal{K} \in \mathsf{K}(\mathcal{R}_U), G_U(\mathcal{K}|z) \leq F_{Y|Z}[\mathcal{U}(Y, z; h) \cap \mathcal{K} \neq \emptyset|z],$$

where $\mathsf{K}(\mathcal{R}_Z)$ denotes the collection of all compact sets on $\mathcal{R}_U$. By Corollary 1.4.44 of Molchanov (2005) this is equivalent to

$$\forall \mathcal{S} \in \mathsf{G}(\mathcal{R}_U), G_U(\mathcal{S}|z) \leq F_{Y|Z}[\mathcal{U}(Y, z; h) \cap \mathcal{S} \neq \emptyset|z],$$

where $\mathsf{G}(\mathcal{R}_U)$ *denotes the collection of all open subsets of* $\mathcal{R}_U$*. Because* $G_U(\mathcal{S}|z) = 1 - G_U(\mathcal{S}^c|z)$ *and*

$$F_{Y|Z}\left[\mathcal{U}(Y,z;h) \subseteq \mathcal{S}^c|z\right] = 1 - F_{Y|Z}\left[\mathcal{U}(Y,z;h) \cap \mathcal{S} \neq \emptyset|z\right],$$

this is equivalent to

$$\forall \mathcal{S} \in \mathsf{G}(\mathcal{R}_U),\ F_{Y|Z}\left[\mathcal{U}(Y,z;h) \subseteq \mathcal{S}^c|z\right] \leq G_U(\mathcal{S}^c|z).$$

The collection of $\mathcal{S}^c$ *such that* $\mathcal{S} \in \mathsf{G}(\mathcal{R}_U)$ *is precisely the collection of closed sets on* $\mathcal{R}_U$*,* $\mathsf{F}(\mathcal{R}_U)$*, completing the proof.* ■

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

$$G_{U|Z}\left(\mathcal{U}_{\mathcal{S}}(h,z)\,|z\right) \leq G_{U|Z}(\mathcal{S}|z). \tag{A.2}$$

By supposition we have

$$C_h\left(\mathcal{U}_{\mathcal{S}}(h,z)\,|z\right) \leq G_{U|Z}\left(\mathcal{U}_{\mathcal{S}}(h,z)\,|z\right). \tag{A.3}$$

The result then holds because $C_h(\mathcal{S}|z) = C_h\left(\mathcal{U}_{\mathcal{S}}(h,z)\,|z\right)$*, since*

$$\begin{aligned} C_h\left(\mathcal{U}_{\mathcal{S}}(h,z)\,|z\right) &\equiv \mathbb{P}\left[\mathcal{U}(Y,Z;h) \subseteq \mathcal{U}_{\mathcal{S}}(h,z)\,|Z=z\right] \\ &= \int_{y\in\mathcal{R}_Y} 1\left[\mathcal{U}(y,z;h) \subseteq \mathcal{U}_{\mathcal{S}}(h,z)\right] dF_{Y|Z}(y|z) \\ &= \int_{y\in\mathcal{R}_Y} 1\left[\mathcal{U}(y,z;h) \subseteq \mathcal{S}\right] dF_{Y|Z}(y|z) \\ &= C_h(\mathcal{S}|z), \end{aligned}$$

where the second line follows by the law of total probability, and the third by the definition of $\mathcal{U}_{\mathcal{S}}(h,z)$ *in (3.3). Combining* $C_h\left(\mathcal{U}_{\mathcal{S}}(h,z)\,|z\right) = C_h(\mathcal{S}|z)$ *with (A.2) and (A.3) completes the proof.* ■

Proof of Theorem 3. *Fix* (h,z)*. Suppose that*

$$\forall \mathcal{U} \in \mathsf{Q}(h,z),\ C_h(\mathcal{U}|z) \leq G_{U|Z}(\mathcal{U}|z). \tag{A.4}$$

Let $\mathcal{S} \in \mathsf{U}^*(h,z)$ *and* $\mathcal{S} \notin \mathsf{Q}(h,z)$*. Since* $\mathcal{S} \notin \mathsf{Q}(h,z)$ *there exist nonempty collections of sets* $\mathsf{S}_1, \mathsf{S}_2 \in \mathsf{U}^{\mathcal{S}}(h,z)$ *with* $\mathsf{S}_1 \cup \mathsf{S}_2 = \mathsf{U}^{\mathcal{S}}(h,z)$ *such that*

$$\mathcal{S}_1 \equiv \bigcup_{\mathcal{T}\in\mathsf{S}_1} \mathcal{T} \in \mathsf{Q}(h,z),\ \mathcal{S}_2 \equiv \bigcup_{\mathcal{T}\in\mathsf{S}_2} \mathcal{T} \in \mathsf{Q}(h,z),$$

and

$$G_{U|Z}\left(\mathcal{S}_1 \cap \mathcal{S}_2|z\right) = 0. \tag{A.5}$$

Since $\mathcal{S}_1, \mathcal{S}_2 \in \mathsf{Q}\left(h,z\right)$ *we also have that*

$$C_h\left(\mathcal{S}_1|z\right) \leq G_{U|Z}\left(\mathcal{S}_1|z\right) \text{ and } C_h\left(\mathcal{S}_2|z\right) \leq G_{U|Z}\left(\mathcal{S}_2|z\right). \tag{A.6}$$

Because $\mathsf{S}_1 \cup \mathsf{S}_2 = \mathsf{U}^{\mathcal{S}}\left(h,z\right)$,

$$\mathcal{U}\left(Y,z;h\right) \subseteq \mathcal{S} \Rightarrow \left\{\mathcal{U}\left(Y,z;h\right) \subseteq \mathcal{S}_1 \text{ or } \mathcal{U}\left(Y,z;h\right) \subseteq \mathcal{S}_2\right\}. \tag{A.7}$$

Using (A.7), (A.6), and (A.5) in sequence we then have

$$C_h\left(\mathcal{S}|z\right) \leq C_h\left(\mathcal{S}_1|z\right) + C_h\left(\mathcal{S}_2|z\right) \leq G_{U|Z}\left(\mathcal{S}_1|z\right) + G_{U|Z}\left(\mathcal{S}_2|z\right) = G_{U|Z}\left(\mathcal{S}|z\right).$$

Combined with (A.4) this implies $C_h\left(\mathcal{S}|z\right) \leq G_{U|Z}\left(\mathcal{S}|z\right)$ *for all* $\mathcal{S} \in \mathsf{U}^*\left(h,z\right)$ *and hence all closed* $\mathcal{S} \subseteq \mathcal{R}_U$ *by Lemma 1, completing the proof.* ■

Proof of Corollary 2. *Consider any* $\mathcal{S} \in \mathsf{Q}^E\left(h,z\right)$. *Note that*

$$C_h\left(\mathcal{S}^c|z\right) = \mathbb{P}\left[\mathcal{U}\left(Y,Z;h\right) \subseteq \mathcal{S}^c|z\right] = \mathbb{P}\left[\mathcal{U}\left(Y,Z;h\right) \cap \mathcal{S} = \emptyset|z\right].$$

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

$$C_h\left(\mathcal{S}|z\right) + C_h\left(\mathcal{S}^c|z\right) = \mathbb{P}\left[\mathcal{U}\left(Y,Z;h\right) \subseteq \mathcal{S}|z\right] + \mathbb{P}\left[\mathcal{U}\left(Y,Z;h\right) \cap \mathcal{S} = \emptyset|z\right] = 1. \tag{A.8}$$

The inequalities of Theorem ***3*** *imply that*

$$G_{U|Z}\left(\mathcal{S}|z\right) \geq C_h\left(\mathcal{S}|z\right) \text{ and } G_{U|Z}\left(\mathcal{S}^c|z\right) \geq C_h\left(\mathcal{S}^c|z\right).$$

Then $G_{U|Z}\left(\mathcal{S}|z\right) + G_{U|Z}\left(\mathcal{S}^c|z\right) = 1$ *and* (A.8) *imply that both inequalities hold with equality.* ■

Proof of Theorem 4. *Under Restriction SI,* $G_{U|Z}\left(\cdot|z\right) = G_U\left(\cdot\right)$ *a.e.* $z \in \mathcal{R}_Z$. *(3.6) and (3.7) follow from (3.1) and Theorem 2, respectively, upon substituting* $G_U\left(\cdot\right)$ *for* $G_{U|Z}\left(\cdot|z\right)$. *(3.8) follows from Corollary 2, again by replacing* $G_{U|Z}\left(\cdot|z\right)$ *with* $G_U\left(\cdot\right)$. *The equivalence of (3.7) and (3.9) with* $G_{U|Z}\left(\cdot|z\right) = G_U\left(\cdot\right)$ *holds by Artstein's inequality, see e.g. Molchanov (2005, pp. 69-70, Corollary 4.44).* ■

Proof of Corollary 3. *By Theorem 4,* $\left(h, G_U\right) \in \mathcal{M}^*$ *if and only if*

$$C_h\left(\mathcal{S}_I|z\right) \leq G_U\left(\mathcal{S}_I\right), \qquad C_h\left(\mathcal{S}_E|z\right) = G_U\left(\mathcal{S}_E\right),$$

for all $\mathcal{S}_I \in \mathsf{Q}^I(h,z)$ *and* $\mathcal{S}_E \in \mathsf{Q}^E(h,z)$ *and a.e.* $z \in \mathcal{R}_Z$. *These inequalities imply that* $C_h(\mathcal{S}|z) \leq G_U(\mathcal{S})$ *for all* $\mathcal{S} \subseteq \mathcal{R}_U$ *and so in particular, for all* $\mathcal{S}_I \in \mathsf{Q}^I(h,z)$,

$$C_h(\mathcal{S}_I^c|z) \leq G_U(\mathcal{S}_I^c), \text{ a.e. } z \in \mathcal{R}_Z.$$

Also, by arguments in the proof of Corollary 2, for all $\mathcal{S}_E \in \mathsf{Q}^E(h,z)$,

$$C_h(\mathcal{S}_E^c|z) = G_U(\mathcal{S}_E^c), \text{ a.e. } z \in \mathcal{R}_Z.$$

Hence, using $G_U(\mathcal{S}) = 1 - G_U(\mathcal{S}^c)$, *we have that for all* $\mathcal{S}_I \in \mathsf{Q}^I(h,z)$ *and all* $\mathcal{S}_E \in \mathsf{Q}^E(h,z)$,

$$\begin{aligned} C_h(\mathcal{S}_I|z) &\leq G_U(\mathcal{S}_I) \leq 1 - C_h(\mathcal{S}_I^c|z), \text{ a.e. } z \in \mathcal{R}_Z, & \text{(A.9)} \\ C_h(\mathcal{S}_E|z) &= G_U(\mathcal{S}_E) = 1 - C_h(\mathcal{S}_E^c|z), \text{ a.e. } z \in \mathcal{R}_Z, & \text{(A.10)} \end{aligned}$$

which are precisely the conditions characterizing $\mathcal{M}^*$ *given in (3.11). The relations (A.9) and (A.10) thus imply and are implied by (3.8) in the statement of Theorem 4 and hence provide an equivalent characterization of* $\mathcal{M}^*$ *as given in (3.11).*

Furthermore, (A.9) and (A.10) yield the conditions that define $\mathcal{H}^*$ *in (3.10), namely that for all* $\mathcal{S}_I \in \mathsf{Q}^I(h,z)$ *and all* $\mathcal{S}_E \in \mathsf{Q}^E(h,z)$

$$\begin{aligned} C_h(\mathcal{S}_I|z) &\leq 1 - C_h(\mathcal{S}_I^c|z), \text{ a.e. } z \in \mathcal{R}_Z, & \text{(A.11)} \\ C_h(\mathcal{S}_E|z) &= 1 - C_h(\mathcal{S}_E^c|z), \text{ a.e. } z \in \mathcal{R}_Z. & \text{(A.12)} \end{aligned}$$

By monotonicity of the containment functional it follows that under (A.11) and (A.12) there exists some probability distribution G_U *such that (A.9) and (A.10) hold for all* $\mathcal{S}_I \in \mathsf{Q}^I(h,z)$ *and all* $\mathcal{S}_E \in \mathsf{Q}^E(h,z)$, *so that* $(h, G_U) \in \mathcal{M}^*$, *completing the proof.* ■

Proof of Theorem 5. *Restrictions A3 and A4 guarantee that* $\mathcal{U}(Y,Z;h)$ *is integrable and closed. In particular integrability holds because by Restriction A3 first* $G_{U|Z}(\mathcal{S}|z) \equiv \mathbb{P}[U \in \mathcal{S}|z]$ *so that* $E[U|z] = 0$ *a.e.* $z \in \mathcal{R}_Z$, *and second* $\mathbb{P}[h(Y,Z,U) = 0] = 1$ *so that*

$$U \in \mathcal{U}(Y,Z;h) \equiv \{u \in \mathcal{R}_U : h(Y,Z,u) = 0\},$$

implying that $\mathcal{U}(Y,Z;h)$ *has an integrable selection, namely* U. *From Definition 6,* $0 \in \mathbb{E}[\mathcal{U}(Y,Z;h)|z]$ *a.e.* $z \in \mathcal{R}_Z$ *therefore holds if and only if there exists a random variable* $\tilde{U} \in \mathrm{Sel}(\mathcal{U}(Y,Z;h))$ *such that* $E\left[\tilde{U}|z\right] = 0$ *a.e.* $z \in \mathcal{R}_Z$, *and hence* $\mathcal{H}^*$ *is the identified set for* h. *The representation of the identified set of structures* $\mathcal{M}^*$ *then follows directly from Theorem 2.* ■

Proof of Corollary 4. *Fix $z \in \mathcal{R}_Z$. The conditional Aumann expectation $\mathbb{E}[\mathcal{U}(Y,Z;h)|z]$ is the set of values for*

$$\int_{\mathcal{R}_{Y|z}} \int_{\mathcal{U}(y,z;h)} u dF_{U|YZ}(u|y,z)\, dF_{Y|Z}(y|z),$$

such that there exists for each $y \in \mathcal{R}_{Y|z}$ a conditional distribution $F_{U|YZ}(u|y,z)$ with support on $\mathcal{U}(y,z;h)$. Since each $\mathcal{U}(y,z;h)$ is convex, the inner integral

$$\int_{\mathcal{U}(y,z;h)} u dF_{U|YZ}(u|y,z)$$

can take any value in $\mathcal{U}(y,z;h)$, and hence $\mathbb{E}[\mathcal{U}(Y,Z;h)|z]$ is the set of values of the form

$$\int_{\mathcal{R}_{Y|z}} u(y,z)\, dF_{Y|Z}(y|z)$$

for some $u(y,z) \in \mathcal{U}(y,z;h)$, each $y \in \mathcal{R}_{Y|z}$. Since the choice of z was arbitrary, this completes the proof. ■

Proof of Corollary 5. *Restrictions A3 and A4 and the continuity requirement of Restriction MI* guarantee that $\mathcal{D}(Y,Z;h)$ is integrable and closed. From Definition 6, $0 \in \mathbb{E}[\mathcal{D}(Y,Z;h)|z]$ a.e. $z \in \mathcal{R}_Z$ therefore holds if and only if there exists a random variable $D \lesssim \mathcal{D}(Y,Z;h)$ such that $E[D|z] = 0$ a.e. $z \in \mathcal{R}_Z$. $D \lesssim \mathcal{D}(Y,Z;h)$ ensures that*

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

Define

$$\widetilde{\mathcal{U}}(D,Y,Z;h) \equiv \{u \in \mathcal{U}(Y,Z;h) : D = d(u,Z)\}.$$

By the definition of $\mathcal{D}(Y,Z;h)$, $D \in \mathcal{D}(Y,Z;h)$ implies that $\widetilde{\mathcal{U}}(D,Y,Z;h)$ is nonempty. Hence there exists a random variable $\tilde{U}$ such that with probability one $\tilde{U} \in \widetilde{\mathcal{U}}(D,Y,Z;h) \subseteq \mathcal{U}(Y,Z;h)$ where $D = d\left(\tilde{U},Z\right)$. Thus $\tilde{U}$ is a selection of $\mathcal{U}(Y,Z;h)$ and $E\left[d\left(\tilde{U},Z\right)|z\right] = 0$ a.e. $z \in \mathcal{R}_Z$, and therefore $\mathcal{H}^*$ is the identified set for h, and the given characterization of $\mathcal{M}^*$ follows. ■

Proof of Theorem 6. *Using Corollary 1 and Definition 4 with $\psi\left(h,\mathcal{G}_{U|Z}\right) = h$, the identified set of structural functions h is*

$$\mathcal{H}^{**} = \left\{h \in \mathcal{H} : \exists \mathcal{G}_{U|Z} \in \mathsf{G}_{U|Z} \text{ s.t. } \forall \mathcal{S} \in \mathsf{F}(\mathcal{R}_U),\, C_h(\mathcal{S}|z) \leq G_{U|Z}(\mathcal{S}|z) \text{ a.e. } z \in \mathcal{R}_Z\right\}. \quad \text{(A.13)}$$

Consider any $h \in \mathcal{H}^{**}$. We wish to show first that h belongs to the set $\mathcal{H}^*$ given in (3.16). Set $\mathcal{S} = (-\infty, 0]$, and fix $z \in \mathcal{R}_Z$. Then $\mathcal{G}_{U|Z} \in \mathsf{G}_{U|Z}$ and Restriction QI imply that

$$C_h\left((-\infty, 0]\,|z\right) \leq G_{U|Z}\left((-\infty, 0]\,|z\right) = \tau, \tag{A.14}$$

and because of Restriction IS, $\mathcal{U}(Y, Z; h) = [\underline{u}(Y, Z; h), \overline{u}(Y, Z; h)]$,

$$C_h\left((-\infty, 0]\,|z\right) = F_{Y|Z}\left[\overline{u}(Y, Z; h) \leq 0|z\right]. \tag{A.15}$$

Now consider $\mathcal{S} = [0, \infty)$. We have by monotonicity of the containment functional $C_h(\cdot|z)$ and from $C_h(\mathcal{S}|z) \leq G_{U|Z}(\mathcal{S}|z)$ in (A.13) that

$$C_h\left((0, \infty)\,|z\right) \leq C_h\left([0, \infty)\,|z\right) \leq G_{U|Z}\left([0, \infty)\,|z\right) = 1 - \tau, \tag{A.16}$$

where the equality holds by continuity of the distribution of $U|Z = z$ in a neighborhood of zero. Again using Restriction IS,

$$C_h\left((0, \infty)\,|z\right) = 1 - F_{Y|Z}\left[\underline{u}(Y, Z; h) \leq 0|z\right]. \tag{A.17}$$

Combining this with (A.16) and also using (A.14) and (A.15) above gives

$$F_{Y|Z}\left[\overline{u}(Y, Z; h) \leq 0|z\right] \leq \tau \leq F_{Y|Z}\left[\underline{u}(Y, Z; h) \leq 0|z\right]. \tag{A.18}$$

The choice of z was arbitrary and so we have that the above holds a.e. $z \in \mathcal{R}_Z$, implying that $h \in \mathcal{H}^*$.

Now consider any $h \in \mathcal{H}^*$. We wish to show that $h \in \mathcal{H}^{**}$. It suffices to show that for any such h under consideration there exists a collection of conditional distributions $\mathcal{G}_{U|Z}$ such that for almost every $z \in \mathcal{R}_Z$ (1) $G_{U|Z}(\cdot|z)$ has τ-quantile equal to zero, and (2) $\forall \mathcal{S} \in \mathsf{F}(\mathcal{R}_U)$, $C_h(\mathcal{S}|z) \leq G_{U|Z}(\mathcal{S}|z)$.

To do so we fix an arbitrary $z \in \mathcal{R}_Z$ and construct $G_{U|Z}(\cdot|z)$ such that (1) and (2) hold. Namely let $G_{U|Z}(\cdot|z)$ be such that for each $\mathcal{S} \in \mathsf{F}(\mathcal{R}_U)$,

$$G_{U|Z}(\mathcal{S}|z) = \lambda(z)\, C_h(\mathcal{S}|z) + (1 - \lambda(z))\left(1 - C_h(\mathcal{S}^c|z)\right), \tag{A.19}$$

where $\lambda(z)$ is chosen to satisfy

$$\lambda(z)\, F_{Y|Z}\left[\overline{u}(Y, Z; h) \leq 0|z\right] + (1 - \lambda(z))\, F_{Y|Z}\left[\underline{u}(Y, Z; h) \leq 0|z\right] = \tau. \tag{A.20}$$

The left hand side of equation (A.20) is precisely (A.19) with $\mathcal{S} = (-\infty, 0]$. *Because* $h \in \mathcal{H}^*$, *(A.18) holds, which guarantees that* $\lambda(z) \in [0,1]$. *(A.20) and (A.19) deliver*

$$G_{U|Z}((-\infty, 0]\,|z) = \tau,$$

so that (1) holds. Moreover, it easy to verify that for any $\mathcal{S}$,

$$C_h(\mathcal{S}|z) \leq 1 - C_h(\mathcal{S}^c|z),$$

since $C_h(\cdot|z)$ *is the conditional containment functional of* $\mathcal{U}(Y, Z; h)$ *and* $1 - C_h(\mathcal{S}^c|z)$ *is the conditional capacity functional of* $\mathcal{U}(Y, Z; h)$. *Hence* $C_h(\mathcal{S}|z) \leq G_{U|Z}(\mathcal{S}|z)$. *Thus (2) holds, and since the choice* z *was arbitrary,* $h \in \mathcal{H}^{**}$ *as desired. This verifies claim (i) of the Theorem.*

Claim (ii) of the Theorem holds because with $\overline{u}(Y, Z; h)$ *and* $\underline{u}(Y, Z; h)$ *continuously distributed given* $Z = z$, *a.e.* $z \in \mathcal{R}_Z$, *their conditional quantile functions are invertible at* τ. *Thus for any* $z \in \mathcal{R}_Z$,

$$\underline{q}(\tau, z; h) \leq 0 \leq \overline{q}(\tau, z; h) \Leftrightarrow F_{Y|Z}[\overline{u}(Y, Z; h) \leq 0|z] \leq \tau \leq F_{Y|Z}[\underline{u}(Y, Z; h) \leq 0|z].$$

Claim (iii) of the Theorem follows directly from Theorem 2. ■

B Equivalence of Selectionability of Conditional and Joint Distributions

In this section we prove that selectionability statements in the main text required for observational equivalence and characterization of identified sets conditional on $Z = z$ for almost every $z \in \mathcal{R}_Z$ are in fact equivalent to unconditional selectionability statements inclusive of Z. Intuitively this holds because knowledge of a conditional distribution of a random set or random vector given $Z = z$, a.e. $z \in \mathcal{R}_Z$, is logically equivalent to knowledge of the joint distribution of that given random vector or random set and Z.

Proposition 1 *(i)* $F_{Y|Z}(\cdot|z) \preccurlyeq \mathcal{Y}(U, Z; h)\,|Z = z$ *a.e.* $z \in \mathcal{R}_Z$ *if and only if* $F_{YZ}(\cdot) \preccurlyeq \mathcal{Y}(U, Z; h) \times Z$. *(ii)* $G_{U|Z}(\cdot|z) \preccurlyeq \mathcal{U}(Y, Z; h)\,|Z = z$ *a.e.* $z \in \mathcal{R}_Z$ *if and only if* $G_{UZ}(\cdot) \preccurlyeq \mathcal{U}(Y, Z; h) \times Z$.

Proof of Proposition 1. *Note that since the choice of z in the above Theorem is arbitrary the statement holds because*

$$\begin{aligned}\mathbb{P}\left[\left(\tilde{Y},Z\right)\in\mathcal{Y}\left(U,Z;h\right)\times Z\right] &= \int\limits_{z\in\mathcal{R}_Z}\mathbb{P}\left[\left(\tilde{Y},Z\right)\in\mathcal{Y}\left(\tilde{U},Z;h\right)\times Z|Z=z\right]dF_Z\left(z\right)\\ &= \int\limits_{z\in\mathcal{R}_Z}\mathbb{P}\left[\tilde{Y}\in\mathcal{Y}\left(\tilde{U},Z;h\right)|Z=z\right]dF_Z\left(z\right),\end{aligned}$$

which is equal to one if and only if $\mathbb{P}\left[\tilde{Y}\in\mathcal{Y}\left(\tilde{U},Z;h\right)|Z=z\right]=1$ *for almost every* $z\in\mathcal{R}_Z$. *By identical reasoning,* $G_{U|Z}\left(\cdot|z\right)$ *is selectionable with respect to the conditional distribution of* $\mathcal{U}\left(Y,Z;h\right)$ *given* $Z=z$ *for almost every* $z\in\mathcal{R}_Z$ *if and only if* $G_{UZ}\left(\cdot\right)$ *is selectionable with respect to the distribution of* $\mathcal{U}\left(Y,Z;h\right)\times Z$. ■

C Computational Details for Numerical Illustrations of Section 4.2

In this Section we describe computation of the conditional containment functional $C_\theta\left([u_*,u^*]\,|z\right)$ in (4.13). Computations were carried out in `Mathematica 9`.

Given the structure specified for DGP1 and DGP2 in Section 4.2, the conditional distribution of Y_2^* given $Y_1=y_1$ and $Z=z$ is for any (y_1,z)

$$\mathcal{N}\left(a\left(z\right)+\frac{\sigma_{1v}+\gamma_1\sigma_{vv}}{\sigma_{11}+2\gamma_1\sigma_{1v}+\gamma_1^2\sigma_{vv}}\left(y_1-\left(\gamma_0+\gamma_1 a\left(z\right)\right)\right),\sigma_{vv}-\frac{\left(\sigma_{1v}+\gamma_1\sigma_{vv}\right)^2}{\sigma_{11}+2\gamma_1\sigma_{1v}+\gamma_1^2\sigma_{vv}}\right),$$

where $a\left(z\right)\equiv\delta_0+\delta_1 z$. From this it follows that the conditional (discrete) distribution of (Y_{2l},Y_{2u}) given Y_1 and Z is:

$$\begin{aligned}\mathbb{P}\left[\left[Y_{2l},Y_{2u}\right)=I_j|y_1,z\right]=\Phi&\left(\frac{c_{j+1}-\left(a(z)+\frac{\sigma_{1v}+\gamma_1\sigma_{vv}}{\sigma_{11}+2\gamma_1\sigma_{1v}+\gamma_1^2\sigma_{vv}}\left(y_1-\left(\gamma_0+\gamma_1 a(z)\right)\right)\right)}{\sqrt{\sigma_{vv}-\frac{\left(\sigma_{1v}+\gamma_1\sigma_{vv}\right)^2}{\sigma_{11}+2\gamma_1\sigma_{1v}+\gamma_1^2\sigma_{vv}}}}\right)\\ &-\Phi\left(\frac{c_j-\left(a(z)+\frac{\sigma_{1v}+\gamma_1\sigma_{vv}}{\sigma_{11}+2\gamma_1\sigma_{1v}+\gamma_1^2\sigma_{vv}}\left(y_1-\left(\gamma_0+\gamma_1 a(z)\right)\right)\right)}{\sqrt{\sigma_{vv}-\frac{\left(\sigma_{1v}+\gamma_1\sigma_{vv}\right)^2}{\sigma_{11}+2\gamma_1\sigma_{1v}+\gamma_1^2\sigma_{vv}}}}\right).\end{aligned}$$

The distribution of Y_1 given $Z=z$ is

$$Y_1|Z=z\sim\mathcal{N}\left(\gamma_0+\gamma_1 a(z),\sigma_{11}+2\gamma_1\sigma_{1v}+\gamma_1^2\sigma_{vv}\right). \tag{C.1}$$

The conditional containment functional can thus be written

$$\begin{aligned} C_\theta\left([u_*, u^*]\,|z\right) &= \sum_j \mathbb{P}[(g_0 + g_1 c_{j+1} + u_* \leq Y_1 \leq g_0 + g_1 c_j + u^*) \wedge (Y_2, Y_3) = I_j | z] \\ &= \sum_j \max\left(0, \int_{\gamma_0+\gamma_1 c_{j+1}+u_*}^{\gamma_0+\gamma_1 c_j+u^*} f_{Y_1|Z}\left(y_1|z\right) \times \mathbb{P}\left[[Y_{2l}, Y_{2u}) = I_j | y_1, z\right] dy_1\right). \end{aligned}$$

where $f_{Y_1|Z}\left(\cdot|z\right)$ is the normal pdf with mean and variance given in (C.1).

In the calculations performed in `Mathematica` we used the following equivalent formulation employing a *single* numerical integration for computation of $C_\theta\left([u_*, u^*]\,|z\right)$.

$$C_\theta\left([u_*, u^*]\,|z\right) \equiv$$
$$\int_{-\infty}^{\infty} \left(\sum_j 1[g_0 + g_1 c_{j+1} + u_* < y_1 < g_0 + g_1 c_j + u^*] \times f_{Y_1|Z}\left(y_1|z\right) \times \mathbb{P}\left[[Y_{2l}, Y_{2u}) = I_j | y_1, z\right]\right) dy_1.$$