cemmap

centre for microdata methods and practice

Characterizations of identified sets delivered by structural econometric models

Andrew Chesher
Adam Rosen

The Institute for Fiscal Studies
Department of Economics, UCL

cemmap working paper CWP63/15

An ESRC Research Centre

Institute for Fiscal Studies, 7 Ridgmount Street, London WC1E 7AE **Tel:** 44 (0)20 7291 4800 **Email:** info@cemmap.ac.uk **www.cemmap.ac.uk**

Characterizations of Identified Sets Delivered by Structural Econometric Models*

Andrew Chesher[†]
UCL and CeMMAP

Adam M. Rosen[‡]
UCL and CeMMAP

October 3, 2015

Abstract

This paper develops characterizations of identified sets of structures and structural features for complete and incomplete models involving continuous and/or discrete variables. Multiple values of unobserved variables can be associated with particular combinations of observed variables. This can arise when there are multiple sources of heterogeneity, censored or discrete endogenous variables, or inequality restrictions on functions of observed and unobserved variables. The models generalize the class of incomplete instrumental variable (IV) models in which unobserved variables are single-valued functions of observed variables. Thus the models are referred to as Generalized IV (GIV) models, but there are important cases in which instrumental variable restrictions play no significant role. The paper provides the first formal definition of observational equivalence for incomplete models. The development uses results from random set theory which guarantee that the characterizations deliver sharp bounds, thereby dispensing with the need for case-by-case proofs of sharpness. One innovation is the use of random sets defined on the space of unobserved variables. This allows identification analysis under mean and quantile independence restrictions on the distributions of unobserved variables conditional on exogenous variables as well as under a full independence restriction. It leads to a novel general characterization of identified sets of structural functions when the sole restriction on the

*This paper is a revised version of a paper initially circulated as CeMMAP working paper CWP43/13 under the title "Generalized Instrumental Variable Models." We thank Federico Bugni, Matt Masten, and Francesca Molinari for detailed comments and discussion. We have benefitted from discussion with participants 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, the 2013 Cowles summer econometrics conference, Yale, Virginia, Georgetown, Maryland, Johns Hopkins, Toulouse School of Economics, Oxford, Duke, Illinois, Reading, and Surrey. 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.

distribution of unobserved and observed exogenous variables is that they are independently distributed. Illustrations are presented for a parametric random coefficients linear model and for a model with an interval censored outcome, in both cases with endogenous explanatory variables, and for an incomplete nonparametric model of English auctions. Numerous other applications are indicated.
Keywords: instrumental variables, endogeneity, excess heterogeneity, limited information, set identification, partial identification, random sets, incomplete models.

JEL classification: C10, C14, C24, C26.

1 Introduction

This paper develops characterizations of identified sets – equivalently sharp bounds – for a wide class of complete and incomplete structural models admitting general forms of unobserved heterogeneity.[1] Multiple values of unobserved variables can be associated with particular combinations of values of observed endogenous and exogenous variables. This occurs in models admitting multiple sources of heterogeneity such as random coefficients, in models with discrete or censored outcomes, and in models in which observed and unobserved variables are constrained by inequality restrictions.

Leading examples of the models studied here are classical single equation instrumental variable (IV) models such as the linear model (e.g. Wright (1928), Theil (1953), Basmann (1959)) as well as semiparametric and nonparametric IV models (e.g. Newey and Powell (1989, 2003), and Chernozhukov and Hansen (2005)) and extensions of these models allowing random coefficients and/or discrete or censored outcomes. In these IV models there are restrictions on the influence of certain exogenous variables on the determination of outcomes and restrictions on the extent of dependence between observed exogenous and unobserved variables. We use the catch-all descriptor Generalized Instrumental Variable (GIV) models to describe the class of models studied in this paper. However, our results can also be applied to models in which instrumental variables play no significant role.

Let Y and Z denote, respectively, observed endogenous and exogenous variables, and let U denote unobserved heterogeneity. Lower case y, z and u denote realizations of these random vectors which may be continuous, discrete, or mixed continuous-discrete. The models studied in this paper place restrictions on a structural function $h(y, z, u)$ mapping the joint support of Y, Z, and U onto the real line. The structural function defines the combinations of values of Y, Z and U that may occur through the restriction that $h(Y, Z, U) = 0$ almost surely. For example, a classical linear IV model has $h(y, z, u) = y_1 - \alpha y_2 - z\beta - u$. More examples are given in Section 2.2.

The primary focus of this paper is on identification of structures. A structure $(h, \mathcal{G}_{U|Z})$ comprises

[1]The terms "sharp bounds" and "identified sets" are used interchangeably throughout. Non-sharp bounds are referred to as "outer sets".

a structural function h coupled with a family of conditional distributions of U given Z:

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

where $G_{U|Z}(\mathcal{S}|z)$ is the probability that U belongs to set $\mathcal{S}$ given $Z = z$, and $\mathcal{R}_Z$ is the support of exogenous Z. Identified sets for structural features, for example a structural function or some functional of it, are obtained as projections of identified sets of structures.

Level sets of the structural function $h(y, z, u)$ play a central role in the development. Let $\mathcal{R}_U$ and $\mathcal{R}_Y$ denote the support of U and Y, respectively. The random set

$$\mathcal{U}(Y, Z; h) \equiv \{u \in \mathcal{R}_U : h(Y, Z, u) = 0\} \tag{1.1}$$

has realizations $\mathcal{U}(y, z; h)$ which contain all values of U that can give rise to $Y = y$ when $Z = z$ according to structural function h. The random set

$$\mathcal{Y}(U, Z; h) \equiv \{y \in \mathcal{R}_Y : h(y, Z, U) = 0\} \tag{1.2}$$

has realizations $\mathcal{Y}(u, z; h)$ which contain all values of Y that can occur when $U = u$ and $Z = z$ according to the structural function h. Complete models require $\mathcal{Y}(U, Z; h)$ to be a singleton with probability one for all admissible h. Incomplete models admit structural functions h such that $\mathcal{Y}(U, Z; h)$ can have cardinality greater than one.[2] Models with multiple sources of heterogeneity, discrete or censored outcomes, or observed and unobserved variables restricted by inequality constraints have sets $\mathcal{U}(Y, Z; h)$ with realizations which may not be singleton sets. The GIV models studied here require neither of these sets to be singleton. Consequently, classical rank conditions or more generally nonparametric completeness conditions are typically not sufficient for point identification. GIV models are generally partially identifying.

This paper provides characterizations of identified sets of structures and structural features delivered by GIV models given distributions of observable variables. Examples of the sets obtained are supplied for particular cases. Previously the question of whether sharp bounds are obtained has been primarily handled on a case-by-case basis. The usual approach to proving sharpness is constructive, see for example Chesher (2010, 2013) and Rosen (2012). This approach requires one to show that every structure in the identified set can deliver the distribution of observed variables. This is often difficult to accomplish and sometimes, as in the auction model of Haile and Tamer (2003), it is infeasible. The methodology introduced here is shown to always deliver characterizations of *sharp* bounds. It is shown that these sets can be expressed as systems of moment inequalities

[2] In Chesher and Rosen (2012) we specialize our approach for identification analysis to simultaneous discrete outcome models, and there define incoherent models in which $\mathcal{Y}(U, Z; h)$ can be empty. There we discuss several ways in which incoherence can be addressed. For further details about incoherence and references to the larger literature on simultaneous discrete outcome models where this has been a pervasive issue, we refer to that paper.

and equalities to which recently developed inferential procedures are applicable. See for example Chesher and Rosen (2013a,b), and Aradillas-Lopez and Rosen (2013) for empirical applications using treatment effect and simultaneous ordered response models.

The results of this paper are obtained using random set theory, reviewed in Molchanov (2005) and introduced into econometric identification analysis by Beresteanu, Molchanov, and Molinari (2011), henceforth BMM11. The analysis there employs the random set $\mathcal{Y}(U, Z; h)$ in models where the identified set can be characterized through a finite number of conditional moment inequalities involving an unobservable and possibly infinite-dimensional nuisance function, such as an equilibrium selection mechanism in econometric models of games. BMM11 provides a novel and computationally tractable characterization of the identified set in the form of conditional moment inequalities in which this nuisance function does not appear. In Beresteanu, Molchanov, and Molinari (2012), it is further shown how the random set $\mathcal{Y}(U, Z; h)$ can be used to characterize sharp bounds on the distribution of counterfactual response functions in treatment effect models, such as those of Balke and Pearl (1997), Manski (1997), and Kitagawa (2009), and extensions thereof. In another related paper, Galichon and Henry (2011) take an alternative approach using optimal transportation theory to characterize sharp parameter bounds in parametrically specified incomplete models. An innovation in their analysis is the construction of core-determining sets, in their case subsets of the space on which $\mathcal{Y}(U, Z; h)$ is realized, for the purpose of simplifying computation of the identified set.

Instead of using the random Y level set $\mathcal{Y}(U, Z; h)$, the approach taken in this paper uses the random U-level sets $\mathcal{U}(Y, Z; h)$ to characterize identified sets for $(h, \mathcal{G}_{U|Z})$ in structural econometric models. The analysis does not require the existence of a representation of the identified set through a finite number of conditional moment equalities involving an unknown nuisance function as required in BMM11. This allows treatment of models with continuous endogenous variables and independence restrictions on the joint distribution U and Z. Nor does the analysis here require parametric specification for the structural function or the conditional distributions of unobserved heterogeneity as required in Galichon and Henry (2011). The use of random U-level sets allows consideration of a variety of restrictions on unobserved heterogeneity common in structural econometrics, including stochastic independence, conditional mean, conditional quantile, and parametric restrictions. The results are sufficiently general to allow the use of other restrictions on the distribution of U and Z through application of Theorem 2. In our previous papers employing random set theory such as Chesher, Rosen, and Smolinski (2013) and Chesher and Rosen (2012), unobserved U and exogenous Z were required to be independently distributed, and certain outcome variables were required to be discrete. This paper studies a much broader class of models in which each of the components of endogenous Y can be continuous, discrete or mixed.

The main result of the paper is as follows. Let θ be a structure, that is, an object with components which are a structural function h and a collection of conditional distributions $\mathcal{G}_{U|Z}$.

Let $F_{Y|Z}(\cdot|z)$ denote a conditional distribution of endogenous variables given $Z = z$. A random set $\mathcal{U}(Y, Z; h)$ is characterized by the collection of random variables that almost surely lie in the random set. These random variables are called the *selections* of the random set.[3] It is shown that the identified set of structures delivered by a model given distributions $F_{Y|Z}(\cdot|z)$, $z \in \mathcal{R}_Z$, comprises all θ admitted by the model such that for almost every $z \in \mathcal{R}_Z$, $G_{U|Z}(\cdot|z) \in \mathcal{G}_{U|Z}$ is the distribution of one of the selections of $\mathcal{U}(Y, Z; h)$ when Y given $Z = z$ has distribution $F_{Y|Z}(\cdot|z)$.

Alternative characterizations of this selectionability property deliver alternative characterizations of the identified set. One such characterization is delivered by Artstein's (1983) inequality, characterizing the identified set as those θ admitted by the model such that the inequality

$$G_{U|Z}(\mathcal{S}|z) \geq \mathbb{P}[\mathcal{U}(Y, Z; h) \subseteq \mathcal{S}|Z = z]$$

holds for all $\mathcal{S}$ in a collection of closed sets $\mathsf{Q}(h, z)$ and almost every $z \in \mathcal{R}_Z$. On the left hand side is the probability that U has a realization in set S given $Z = z$. On the right hand side is the conditional probability of the occurrence of one one of the values of Y such that $\mathcal{U}(Y, Z; h) \subseteq \mathcal{S}$, under structural function h. A definition of collections of sets $\mathsf{Q}(h, z)$ is provided and conditions are given under which certain inequalities can be replaced by equalities. Identified sets for structural features that are functionals of θ are obtained as projections of the identified set for θ. Characterizations employing the Aumann expectation of random sets $\mathcal{U}(Y, Z; h)$ are also provided.

The paper provides other new results, briefly reviewed now.

Observational equivalence. The concept of *observational equivalence* is formally defined for incomplete models. While identification analysis for incomplete models has been conducted in other papers, including some of our own, this appears to be the first formal definition of observational equivalence for such models. The definition extends the classical definition of observational equivalence, e.g. Koopmans (1949), Koopmans and Reiersøl (1950), Hurwicz (1950), Rothenberg (1971), Bowden (1973), and Matzkin (2007, 2008) to models admitting structures that may not generate a single distribution of endogenous variables given other observed variables.[4] In complete models observationally equivalent structures generate common distributions of endogenous outcomes, Y. In incomplete models structures can generate multiple distributions of outcomes so observational equivalence is naturally defined in terms of random outcome sets $\mathcal{Y}(U, Z; h)$.

Dual level sets. The analysis is based on a key duality result concerning the random level sets $\mathcal{Y}(U, Z; h)$ and $\mathcal{U}(Y, Z; h)$. This allows the development of characterizations of observational equivalence in terms of properties of $\mathcal{U}(Y, Z; h)$ rather than $\mathcal{Y}(U, Z; h)$. This dual characterization of observational equivalence makes plain the relationship between the characterizations of identified

[3]See Definition 1 and Molchanov (2005) for further details.

[4]Hurwicz (1950) does not explicitly use the term *observational equivalence* but employs the concept nonetheless. In the cited papers Hurwicz (1950) and Matzkin (2007, 2008) both consider observational equivalence from a nonparametric perspective.

sets in this paper and those using outcome sets $\mathcal{Y}(U, Z; h)$, as in for example BMM11 and Galichon and Henry (2011). The use of random U-level sets $\mathcal{U}(Y, Z; h)$ rather than random Y-level sets $\mathcal{Y}(U, Z; h)$ is what enables characterization of identified sets in models employing a variety of different restrictions on the distribution of U and Z. It also facilitates projection of identified sets onto the space in which structural functions are defined.

Projection and distribution free analysis. Theorem 4 Corollary 3 characterizes identified sets of structural functions in models which restrict unobserved U and exogenous Z to be independently distributed while placing no restrictions on the marginal distribution of U. This provides a characterization of the identified set for the structural function h that does not make explicit reference to the distribution of unobserved heterogeneity. Until now such results have not been available.

Conditional moment and quantile restrictions. Sections 5.2 and 5.3 provide new results characterizing identified sets of structures and structural features in econometric models incorporating conditional moment and conditional quantile restrictions on unobserved heterogeneity. Conditional mean restrictions on *outcome* variables Y given exogenous variables can be readily incorporated in an analysis employing random Y-level sets $\mathcal{Y}(U, Z; h)$ using techniques developed in BMM11 and Beresteanu, Molchanov, and Molinari (2012). The approach taken here additionally delivers identified sets in models featuring conditional moment restrictions on *unobserved* variables that may enter structural equations non-separably, that is non-linear moment condition models, that do not map directly to conditional mean restrictions on outcome variables. The characterizations of identified sets in models featuring conditional quantile restrictions are also novel.

Core determining sets. The conditional moment inequalities characterizing identified sets comprise inequalities concerning the probability of the event $\mathcal{U}(Y, Z; h) \subseteq \mathcal{S}$ for sufficiently rich collections of non-stochastic sets $\mathcal{S}$ defined on the support of U. Such collections are called core determining sets, and, as previously noted, were first introduced for characterizations involving Y sets by Galichon and Henry (2011). Here several new results on core-determining sets for characterizations involving $\mathcal{U}(Y, Z; h)$ are presented. The results are widely applicable, making use of the geometry of the random U-level sets $\mathcal{U}(Y, Z; h)$. A novel feature of these results is that we provide conditions whereby some combinations of conditional moment inequalities can be simplified to conditional moment equalities.[5]

Applications. The scope of application of the new results is demonstrated in three worked examples. Section 6.1 applies the results of this paper to obtain a characterization of the sharp identified set of distributions of valuations in the incomplete model of English auctions studied in Haile and Tamer (2003). This demonstrates the ability of the methods introduced in this paper to deliver sharp identified sets in cases where instrumental variable restrictions play no significant role and where characterizations of sharp identified sets have hitherto been unavailable. In Section 6.2,

[5]Earlier versions of some of the results on core determining sets appeared in the 2012 version of the CeMMAP working paper Chesher and Rosen (2012), which concerned models with only discrete endogenous variables. In this paper we provide more general results covering the broader class of models studied here.

identified sets are calculated for a single equation random coefficients linear IV model. Other IV treatments of continuous outcome random coefficient models have employed simultaneous equations models as in Masten (2014) or complete models that fully specify the mapping from exogenous to endogenous variables, as in Hoderlein, Holzmann, and Meister (2015). The online materials provide a characterization of identified sets in a linear model with an interval censored outcome and an endogenous explanatory variable, an extension of the model studied in Manski and Tamer (2002) in which explanatory variables are restricted to be exogenous.

1.1 Plan

The paper proceeds as follows. Section 2 formalizes the GIV model restrictions and provides some leading examples of GIV models. Section 3 provides our generalization of the classical notion of observational equivalence, our duality result, and some accompanying formal set identification characterizations, including a widely-applicable construction written in terms of conditional moment inequalities. Section 4 shows how to use the notion of core-determining sets to exploit the geometric structure of the random sets $\mathcal{U}(Y, Z; h)$ to reduce the collection of conditional moment inequalities without losing identifying power. Section 5 shows how restrictions on unobserved heterogeneity and exogenous variables, such as independence, conditional mean, conditional quantile, and parametric restrictions can then be incorporated to further refine characterization of the identified set. Section 6 demonstrates how the application of our results to IV random coefficient models and incomplete models of English auctions delivers novel characterizations of sharp bounds in models featuring multivariate unobserved heterogeneity. Section 7 concludes. All proofs are provided in the Appendix. Additional results not included in the paper are given in an on-line supplement. Section C.3 of this supplement further illustrates the set identifying power of GIV models through application to a continuous outcome model with an interval-censored endogenous explanatory variable, providing numerical illustrations of the resulting identified sets.

1.2 Notation

Capital Roman letters A denote point-valued random variables and lower case letters a denote particular point-valued realizations. For probability measure $\mathbb{P}$, $\mathbb{P}(\cdot|a)$ is used to denote the conditional probability measure given $A = a$. $\mathcal{R}_{A_1 \cdots A_m}$ denotes 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. Calligraphic font ($\mathcal{S}$) is reserved for sets, and sans serif font (S) is reserved for collections of sets. The symbol $\subseteq$ indicates nonstrict set inclusion, $\mathrm{cl}(\mathcal{A})$ 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 $F \lesssim \mathcal{A}$ indicates that the distribution F of a random vector is *selectionable* with respect to the distribution of random set $\mathcal{A}$, and $A = \text{Sel}(\mathcal{A})$ indicates that random variable A is a selection of random set $\mathcal{A}$, both 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. $\mathbb{R}^m$ denotes m dimensional Euclidean space, $\mathbb{R}^1$ is abbreviated to $\mathbb{R}$ and for any vector $v \in \mathbb{R}^m$, $\|v\|$ indicates the Euclidean norm: $\|v\| = \sqrt{v_1^2 + \cdots + v_m^2}$. 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 GIV Models

This section starts with a formal statement of the restrictions comprising GIV models. Examples of particular GIV models are then provided.

2.1 GIV Models

The following restrictions are employed.

Restriction A1: (Y, Z, U) are finite dimensional random vectors defined on a probability space $(\Omega, \mathsf{L}, \mathbb{P})$, endowed with the Borel sets on Ω. □

Restriction A2: The support of (Y, Z) is a subset of Euclidean space. 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, where for all $\mathcal{T} \subseteq \mathcal{R}_{Y|z}$, $F_{Y|Z}(\mathcal{T}|z) \equiv \mathbb{P}[Y \in \mathcal{T}|z]$. □

Restriction A3: The support of U is a subset of a locally compact second countable Hausdorff topological space. There is an L-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 $\mathcal{S} \subseteq \mathcal{R}_{U|z}$, $G_{U|Z}(\mathcal{S}|z) \equiv \mathbb{P}[U \in \mathcal{S}|z]$. □

Restriction A4: The pair $\left(h, \mathcal{G}_{U|Z}\right)$ belongs to a known set of admissible structures $\mathcal{M}$. □

Restriction A5: $\mathcal{U}(Y, Z; h)$ is closed almost surely $\mathbb{P}[\cdot|z]$, each $z \in \mathcal{R}_Z$. □

Restriction A6: $\mathcal{Y}(Z, U; h)$ is closed almost surely $\mathbb{P}[\cdot|z]$, each $z \in \mathcal{R}_Z$. □

Restriction A1 defines the probability space on which (Y, Z, U) reside. Restriction A2 restricts $\mathcal{R}_{YZ}$ to a subset of Euclidean space, and requires that for each $z \in \mathcal{R}_Z$, $F_{Y|Z}(\cdot|z)$ is identified. Random sampling of observations from $\mathbb{P}$ is sufficient, but not required.

The first part of restriction A3 requires that $\mathcal{R}_U$ is a subset of a locally compact second countable Hausdorff space. This is automatically satisfied if $\mathcal{R}_U$ is a subset of a finite dimensional Euclidean space, as is the case in the examples we cover in this paper.[6] The second part of 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.

Restriction A4 imposes model $\mathcal{M}$, the collection of admissible structures $\left(h, \mathcal{G}_{U|Z}\right)$. Unlike the previous restrictions, it is refutable based on knowledge of $\mathcal{F}_{Y|Z}$ in that it is possible that there is no $\left(h, \mathcal{G}_{U|Z}\right) \in \mathcal{M}$ such that $\mathbb{P}\left[h(Y, Z, U) = 0\right] = 1$. In such cases the identified set of structures is empty, indicating model misspecification.

Restrictions A5 and A6 restrict $\mathcal{U}(Y, Z; h)$ and $\mathcal{Y}(Z, U; h)$, respectively, to be random closed sets. The purpose of these restrictions is to enable use of results from random set theory characterizing the distributions of selections of random closed sets.[7] These restrictions are satisfied for example if $\mathcal{M}$ specifies that all admissible h are continuous in their first and third arguments, respectively, but can also hold more generally. A given econometric model can generally be represented through a variety of different but substantively equivalent structural functions h, and judicious choice of this function can often be made to ensure these requirements are satisfied.[8] See Section 2.2 for examples.[9]

Sometimes it is convenient to refer separately to collections of admissible structural functions and distributions $\mathcal{G}_{U|Z}$. 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 $\mathcal{H} \times \mathsf{G}_{U|Z}$.

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), and

[6] Molchanov (2005) p. 1 notes that the Euclidean space $\mathbb{R}^d$ is a generic example of a locally compact second countable Hausdorff space. Importantly, the requirement of local compactness allows $\mathcal{R}_U$ to be unbounded, requiring only that each point in $\mathcal{R}_U$ has a neighborhood with compact closure, see e.g. Molchanov (2005) p. 388.

[7] The definition of a selection of a random set is provided in Section 3.

[8] In many models U is restricted to be continuously distributed, in which case the requirement that h be specified such that $\mathcal{U}(y, z; h)$ is closed is not restrictive, since the difference G_U places on any set and its closure is zero.

[9] Importantly, the realizations of $\mathcal{Y}(Z, U; h)$ and $\mathcal{U}(Y, Z; h)$ may be unbounded, as closedness merely requires that they contain their limit points. Moreover, whether these sets are closed depends on the underlying topological space. We use the Euclidean topology on $\mathbb{R}^d$ throughout, but in some cases other topological spaces could be used to establish closedness. For instance, if $\mathcal{U}(Y, Z; h)$ can take only a finite number of realizations, then this set is closed in the discrete topology on $\{\mathcal{U}(Y, z; h) : z \in \mathcal{R}_Z\}$, see e.g. Sutherland (2009), page 94, Exercise 9.1.

(iii) the joint distribution of (Y, Z). The first two mechanisms are part of the model specification, with $\mathcal{M}$ comprising 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 function satisfying some particular smoothness restrictions, or it could be restricted to a parametric family as in a linear index model. Likewise $\mathcal{G}_{U|Z}$ could be collections of conditional distributions such that $E\left[U|Z\right] = 0$, or $q_{U|Z}\left(\tau|z\right) = 0$, or $U \perp\!\!\!\perp Z$, or satisfying parametric restrictions. Various types of restrictions on $\mathcal{G}_{U|Z}$ are considered in Section 5.

The third source of identifying power, the joint distribution of (Y, Z), is identified and hence left unrestricted. In many models, rank or completeness conditions are invoked to ensure point identification. We allow for set identification, so such conditions are not required here. Moreover, in models with nonsingleton sets $\mathcal{U}\left(Y, Z; h\right)$, as we allow here, such conditions are typically not sufficient to achieve point identification. 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 characterizations reduce to a singleton set.

2.2 Examples

Example 1. *A binary outcome, threshold crossing GIV model as studied in Chesher (2010) and Chesher and Rosen (2013b) has the following structural function:*

$$h(y, z, u) = y_1 \min\left\{u - g\left(y_2, z_1\right), 0\right\} + \left(1 - y_1\right) \max\left\{u - g\left(y_2, z_1\right), 0\right\},$$

with U normalized uniformly distributed on $[0, 1]$.[10] *The corresponding level sets are:*

$$\mathcal{Y}\left(u, z; h\right) = \left\{\left(y_1, y_2\right) \in \mathcal{R}_{Y_1 Y_2} : \left\{y_1 = 1 \wedge u \geq g\left(y_2, z_1\right)\right\} \vee \left\{y_1 = 0 \wedge u \leq g\left(y_2, z_1\right)\right\}\right\},$$

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

Example 2. *Multiple discrete choice with endogenous explanatory variables as studied in Chesher, Rosen, and Smolinski (2013). The structural function is*

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

where $\pi_j\left(y_2, z_1, u_j\right)$ is the utility associated with choice $j \in \mathcal{J} \equiv \{1, \ldots, M\}$ and $u = \left(u_1, \ldots, u_M\right)$ is a vector of unobserved preference heterogeneity. Y_1 is the outcome or choice variable and Y_2 are

[10] With U continuously distributed conditional on the other variables, $\mathbb{P}\left[g\left(Y_2, Z_1\right) = U\right] = 0$. Up to events of measure zero, this model is then equivalent one that specifies $Y_1 = 1\left[g\left(Y_2, Z_1\right) > 0\right]$.

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*

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

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

Example 3. *A continuous outcome random coefficients model has structural function*

$$h(y, z, u) = y_1 - z_1\gamma - (\beta_2 + u_2) y_2 - (\beta_1 + u_1). \tag{2.1}$$

The random coefficients are $(\beta_1 + U_1)$ *and* $(\beta_2 + U_2)$*, with means* β_1 and β_2*, respectively. The coefficient* γ *multiplying exogenous variables in* h *could also be random. The level sets are*

$$\mathcal{Y}(u, z; h) = \{(z_1\gamma + (\beta_2 + u_2) y_2 + (\beta_1 + u_1), y_2) : y_2 \in \mathcal{R}_{Y_2}\},$$

$$\mathcal{U}(y, z; h) = \{u \in \mathcal{R}_U : u_1 = y_1 - z_1\gamma - \beta_1 - \beta_2 y_2 - u_2 y_2\}. \tag{2.2}$$

Section 6.2 investigates by way of numerical illustration the identifiying power of instrumental variables with the single equation structural function 2.1. The example comprises a limited information single equation instrumental variable model, in contrast to the simultaneous equations random coefficient model studied by Masten (2014), for which point identification of the marginal distributions of the random coefficients was established.

Example 4. *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 and continuous 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* $\mathbb{P}[Y_{2l} \le Y_2^* \le Y_{2u}] = 1$ *for observed variables* Y_{2l}, Y_{2u}*. No further restriction is placed on the process determining the realizations of* Y_{2l}, Y_{2u}*. The structural function is*

$$h(y, z, u) = \max\{y_1 - g(y_{2u}, z_1, u), 0\} + \max\{g(y_{2l}, z_1, u) - y_1, 0\},$$

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

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

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

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

Example 5. *Endogenous censoring. Observed Y_1 is the minimum of observed Y_2 and a partially observed variable which is a function $g(Z,U)$ of observed and unobserved exogenous variables, continuous in its second argument, leading to the structural function*

$$h(y,z,u) = y_1 - \min(g(z,u), y_2)$$

which has level set $\mathcal{Y}(u,z;h) = \{(y_1 - \min(g(z,u), y_2), y_2) : y_2 \in \mathcal{R}_{Y_2}\}$. If u is scalar and $g(z,u)$ is monotone increasing in u then the level set $\mathcal{U}(y,z;h)$ is:

$$\mathcal{U}(y,z;h) = \begin{cases} \{g^{-1}(z,y_1)\} & , \quad y_1 < y_2 \\ [g^{-1}(z,y_2), \infty) & , \quad y_1 = y_2 \end{cases}$$

where $g^{-1}(\cdot,\cdot)$ is the inverse of $g(\cdot,\cdot)$ with respect to its second argument. In an example Y_1 is the amount of perishable fish sold in a day at a stall in a fish market, Y_2 is the amount on sale at the start of the day and $g(Z,U)$ is demand, only observed when it fails to exceed the amount stocked. The stall owner may choose Y_2 having some signal of the value of U thus rendering Y_2 endogenous.

Example 6. *Endogenous variables measured with error. Observed Y_1 is a function, $g(Y_2^*, Z, U_1)$, of latent endogenous Y_2^*, exogenous Z and unobserved U_1. Observed $Y_2 = Y_2^* + U_2$ where U_2 may be measurement error or transitory variation around a long run level. There is the structural function*

$$h(y,z,u) = y_1 - g(y_2 - u_2, z, u_1)$$

with level sets

$$\mathcal{Y}(u,z;h) = \{(y_1 - g(y_2 - u_2, z, u_1), y_2) : y_2 \in \mathcal{R}_{Y_2}\}$$

and, if u is scalar and $g(\cdot,\cdot,\cdot)$ is monotone increasing in its third argument with inverse function relative to that argument $g^{-1}(\cdot,\cdot,\cdot)$, as in Chernozhukov and Hansen (2005),

$$\mathcal{U}(y,z;h) = \{u \in \mathcal{R}_U : u_1 = g^{-1}(y_2 - u_2, z, y_1)\}.$$

GIV models have numerous applications and many additional structural econometric models fall in this class. The examples cited above are limited information instrumental variable models in which the determination of some endogenous variables is left completely unspecified, which are our main focus. Nonetheless, the methods of this paper also apply to simultaneous equations and triangular models, see for instance Chesher and Rosen (2012) for some examples involving discrete endogenous variables. In addition, by way of example we show in Section 6.1 how the incomplete model of English auctions in which bidders have independent private valuations introduced in Haile and Tamer (2003) can be set up as a GIV model, and we show how the results developed in this paper produce a characterization of the identified set of valuation distributions.

3 Observational Equivalence and Duality

3.1 Observational Equivalence and Selectionability in Outcome Space

The standard definition of observational equivalence found in the econometrics literature applies in contexts in which each structure, $m \in \mathcal{M}$, delivers a single collection of conditional distributions:

$$\mathcal{P}_{Y|Z}(m) \equiv \{P_{Y|Z}(\cdot|z;m) : z \in \mathcal{R}_Z\}$$

where $P_{Y|Z}(\cdot|z;m)$ is the conditional distribution of Y given $Z = z$ delivered by structure m.[11] Under this definition of observational equivalence, structures m and m' are observationally equivalent if $\mathcal{P}_{Y|Z}(m) = \mathcal{P}_{Y|Z}(m')$ almost surely. The point identified collection of conditional distributions $\mathcal{F}_{Y|Z}$ plays no role in determining whether structures are observationally equivalent.

In the incomplete models studied in this paper, a particular structure m may generate more than one collection of conditional distributions. The *set* of collections of conditional distributions that can be generated by a structure m is denoted by $\mathsf{P}_{Y|Z}(m)$. It is possible that collections $\mathsf{P}_{Y|Z}(m)$ and $\mathsf{P}_{Y|Z}(m')$ generated by distinct structures, m and m', have some but not all collections of conditional distributions in common. Observational equivalence of two structures m and m' may thus depend on the particular collection of conditional distributions $\mathcal{F}_{Y|Z}$ under consideration in identification analysis. This is so because there may be a collection, say $\mathcal{F}^{*}_{Y|Z}$, which lies in $\mathsf{P}_{Y|Z}(m)$ and in $\mathsf{P}_{Y|Z}(m')$ and a collection $\mathcal{F}^{**}_{Y|Z}$ which lies in $\mathsf{P}_{Y|Z}(m)$ but not in $\mathsf{P}_{Y|Z}(m')$. Structures m and m' are observationally equivalent in identification analysis employing $\mathcal{F}^{*}_{Y|Z}$ but not in identification analysis employing $\mathcal{F}^{**}_{Y|Z}$.

Consequently, in the following development, observational equivalence is defined with respect to the (identified) collection of distributions $\mathcal{F}_{Y|Z}$, and a corresponding notion of *potential* observational equivalence, which is a property which two structures can possess irrespective of the collection of conditional distributions $\mathcal{F}_{Y|Z}$ under consideration in identification analysis.[12] In order to give formal definitions of these properties we first provide the definition of a *selection* from a random set

[11] See for example Koopmans and Reiersøl (1950), Hurwicz (1950), Rothenberg (1971), Bowden (1973), and Matzkin (2007, 2008).

[12] 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. $z \in \mathcal{R}_Z$ 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}.

and the definition of *selectionability* as given by Molchanov (2005).[13] Note that a selection from a random set is itself a random variable. It can have arbitrary distribution – not necessarily a point mass – subject to the caveat that it is contained in the given random set with probability one.

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 $\mathcal{W}$, abbreviated $F_W \preccurlyeq \mathcal{W}$, if there exists a random variable $\tilde{W}$ distributed F_W and a random set $\tilde{\mathcal{W}}$ with the same distribution as $\mathcal{W}$ such that $\tilde{W} \in \mathrm{Sel}\left(\tilde{\mathcal{W}}\right)$.*

A given structure $\left(h, \mathcal{G}_{U|Z}\right)$ induces a distribution for the random outcome set $\mathcal{Y}(U, Z; h)$ conditional on $Z = z$, for all $z \in \mathcal{R}_Z$. This is because $h(Y, Z, U) = 0$ dictates only that $Y \in \mathcal{Y}(U, Z; h)$, which is in general insufficient to uniquely determine the conditional distributions $\mathcal{F}_{Y|Z}$. The 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 precisely those distributions for which $h(Y, Z, U) = 0$ can hold with probability one for the given $\left(h, \mathcal{G}_{U|Z}\right)$. In other words, these conditional distributions $F_{Y|Z}(\cdot|z)$ are exactly those that can be generated by the structure $\left(h, \mathcal{G}_{U|Z}\right)$, leading to the following definitions.[14]

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 $\mathcal{F}_{Y|Z}$, 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 $z \in \mathcal{R}_Z$.*

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 admissible 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. $z \in \mathcal{R}_Z$:*

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

A definition of set identification of structural features follows directly from Definition 3. A structural feature $\psi(\cdot, \cdot)$ is any functional of a structure $\left(h, \mathcal{G}_{U|Z}\right)$. Examples include the structural function h itself, $\psi\left(h, \mathcal{G}_{U|Z}\right) = h$, the distributions of unobserved heterogeneity, $\psi\left(h, \mathcal{G}_{U|Z}\right) = \mathcal{G}_{U|Z}$, and various counterfactual probabilities.

[13]Specifically, Definition 2.2 on page 26 and Definition 2.19 on page 34.

[14]The identified set $\mathcal{M}^*$ in Definition 3 depends upon the collection of conditional distributions $\mathcal{F}_{Y|Z}$, although we do not make this dependence explicit in our notation.

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

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

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 identified set of structures $\mathcal{M}^*$ as the focus of our analysis, and unless we specify a particular feature of interest, reference to only the "identified set" without qualification refers to $\mathcal{M}^*$.[15]

3.2 Observational Equivalence and Selectionability in U-Space

In the previous section, observational equivalence was defined in terms collections of distributions $\mathcal{F}_{Y|Z}(\cdot|z)$ selectionable with respect to $\mathcal{Y}(U, Z; h)$ conditional on $Z = z$. Equivalent, and in many cases more useful, characterizations are now developed in terms of (i) random sets $\mathcal{U}(Y, Z; h)$ whose distribution is determined by the structural function (h) along with a collection of distributions of outcomes $\mathcal{F}_{Y|Z}$ and (ii) selectionability relative to these sets of the distributions of unobservables ($\mathcal{G}_{U|Z}$). These dual representations flow directly from an elementary duality property of the two types of level sets of structural functions, namely that for all h and z:

$$u^* \in \mathcal{U}(y^*, z; h) \Longleftrightarrow y^* \in \mathcal{Y}(u^*, z; h).$$

The advantage of this new characterization is that it allows direct imposition of restrictions on the collection $\mathcal{G}_{U|Z}$ admitted by the model $\mathcal{M}$. Such restrictions - for example mean, quantile, full independence, and parametric restrictions are commonplace in econometrics. Further, for any structure under consideration, the distributions of the random residual sets to be considered are entirely determined by the identified distributions of the observed random variables (Y, Z). The characterization is set out in the following two theorems.

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 $Z = z$ when $U \sim G_{U|Z}(\cdot|z)$ if and only if $G_{U|Z}(\cdot|z)$ is selectionable with respect to the conditional distribution of $\mathcal{U}(Y, Z; h)$ given $Z = z$ when $Y \sim F_{Y|Z}(\cdot|z)$.*

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) if additionally Restriction A4 holds, then the identified set of*

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

structures $\left(h, \mathcal{G}_{U|Z}\right)$ *are those elements of* $\mathcal{M}$ *such that* $G_{U|Z}(\cdot|z)$ *is selectionable with respect to the conditional (on* $Z = z$*) distribution of* $\mathcal{U}(Y, Z; h)$*.*

Theorem 2 expresses observational equivalence and the characterization of the identified set of structures $\left(h, \mathcal{G}_{U|Z}\right)$ in terms of selectionability from the conditional distribution of the random residual set $\mathcal{U}(Y, Z; h)$. Any conditions that characterize the set of structures $\left(h, G_{U|Z}\right)$ 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 that applies when $\mathcal{U}(Y, Z; h)$ is a random closed set uses Artstein's Inequality, see e.g. Artstein (1983), Norberg (1992), and Molchanov (2005, Section 1.4.8), which delivers characterization of the identified set $\mathcal{M}^*$ using the conditional containment functional of $\mathcal{U}(Y, Z; h)$, defined as:

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

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 Restrictions A1-A5, the identified set can be written*

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

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

There are inequalities in this characterization for almost every value of the instrument $z \in \mathcal{R}_Z$ and for all closed subsets of $\mathcal{R}_U$. The next section is concerned with reducing the number of sets required to characterize $\mathcal{M}^*$.

4 Core Determining Test Sets

We now characterize a 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, for all $\mathcal{S}$ in $\mathsf{Q}(h, z)$

$$C_h(\mathcal{S}|z) \leq G_{U|Z}(\mathcal{S}|z), \tag{4.1}$$

then the same inequality holds for all $\mathcal{S} \subseteq \mathcal{R}_U$. This can substantially reduce the number of test sets required to characterize an identified set.

Galichon and Henry (2011) introduced core-determining sets for identification analysis considering sets in outcome space and characterizing core-determining sets for incomplete models that satisfy a certain monotonicity requirement. Here no monotonicity condition is imposed and their definition is extended by introducing core-determining sets for the characterizations in U-space developed in Section 3, and by allowing core-determining sets to be specific to the structural relation h and covariate value z.[16]

It is shown that every set in a core determining collection $\mathsf{Q}(h,z)$ is a union of realizations of $\mathcal{U}(Y,Z;h)$ that can arise when $Y \sim F_{Y|Z}(\cdot|z)$. Lemma 1 establishes this result. A particular class of unions can be excluded from the core-determining collection. Theorem 3 defines this class and sets out the composition of $\mathsf{Q}(h,z)$.[17]

In Corollary 2 conditions are established under which the inequality (4.1) must hold with *equality* for particular members of $\mathsf{Q}(h,z)$.[18] Since much is known about observable implications of models with conditional moment *equalities*, this can be helpful for estimation and in determining when there is point identification.[19]

To proceed we define $\mathsf{U}(h,z)$, the support of the random set $\mathcal{U}(Y,Z;h)$ conditional on $Z=z$, $\mathsf{U}^*(h,z)$, the collection of sets comprising unions of such sets, and three sub-collections of $\mathsf{U}^*(h,z)$ associated with a set $\mathcal{S} \subseteq \mathcal{R}_U$.[20] We employ the notation $\mathcal{U}(\mathcal{Y},z;h)$ for the union of the sets $\mathcal{U}(y,z;h)$ such that $y \in \mathcal{Y}$.

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

Definition 5 *The conditional support of random set $\mathcal{U}(Y,Z;h)$ given $Z=z$ is $\mathsf{U}(h,z)$:*

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

and $\mathsf{U}^*(h,z)$ is the collection of all sets that are unions of elements of $\mathsf{U}(h,z)$.

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

For any $\mathcal{S} \subseteq \mathcal{R}_U$, $(h,\mathcal{G}_{U|Z})$ and z there are the following sub-collections of $\mathsf{U}(h,z)$.

$$\mathsf{U}^{\mathcal{S}}(h,z) \equiv \{\mathcal{U} \in \mathsf{U}(h,z) : \mathcal{U} \subseteq \mathcal{S}\}$$

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

[16]Core-determining sets may also be dependent upon $G_{U|Z}(\cdot|z)$ as set out in Theorem 3 below, but this is not made explicit in the notation.

[17]The construction builds on ideas from Chesher, Rosen, and Smolinski (2013), but is much more widely applicable.

[18]Under these conditions the characterization *via* inequalities is sharp, but through use of the law of total probability some of these inequalities can be strengthened to equalities.

[19]For instance, Tamer (2003) shows that observable implications of a simultaneous binary outcome model constitute conditional moment equalities and inequalities, and proves point identification of parameters through use of the equalities. Aradillas-Lopez and Rosen (2013) provide conditions for point identification of a subset of model parameters in a simultaneous ordered entry model that also delivers conditional moment inequalities and equalities.

[20]The definitions and following analysis are easily extended to cases in which $\mathcal{R}_{U|z}$ varies with z.

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

The sets $\mathsf{U}^{\mathcal{S}}(h,z)$, $\mathsf{U}_{\mathcal{S}}(h,z)$ and $\overline{\mathsf{U}}^{\mathcal{S}}(h,z)$ are the sets $\mathcal{U} \in \mathsf{U}(h,z)$ that, respectively: are contained in $\mathcal{S}$, up to zero measure $G_{U|Z}(\cdot|z)$ do not intersect $\mathcal{S}$, and belong to neither of the first two collections.

Lemma 1 establishes that, for any $(h, \mathcal{G}_{U|Z})$ and z, to show that the inequality (4.1) holds for all $\mathcal{S} \subseteq \mathcal{R}_U$ it suffices to show that it holds for all sets $\mathcal{S}$ in the collection of unions $\mathsf{U}^*(h,z)$.

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 $\mathsf{U}^{\mathcal{S}}(h,z)$,*

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

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

All sets in the collection of core-determining sets are unions of sets in $\mathsf{U}(h,z)$, but not all such unions lie in the core-determining collection. Theorem 3 defines a collection of core-determining test sets $\mathsf{Q}(h,z)$, which is a refinement of $\mathsf{U}^*(h,z)$.

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 $\mathsf{S}_1, \mathsf{S}_2 \subseteq \mathsf{U}^{\mathcal{S}}(h,z)$ with $\mathsf{S}_1 \cup \mathsf{S}_2 = \mathsf{U}^{\mathcal{S}}(h,z)$ such that*

$$\mathcal{S}_1 \equiv \bigcup_{\mathcal{T} \in \mathsf{S}_1} \mathcal{T},\ \mathcal{S}_2 \equiv \bigcup_{\mathcal{T} \in \mathsf{S}_2} \mathcal{T},\ \text{and } G_{U|Z}(\mathcal{S}_1 \cap \mathcal{S}_2|z) = 0, \tag{4.3}$$

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 $\mathcal{S} \subseteq \mathcal{R}_U$, so that the collection of sets $\mathsf{Q}(h,z)$ is core-determining.

Theorem 3 implies that the identified sets of Theorem 2 are the set of structures $\left(h, \mathcal{G}_{U|Z}\right)$ that satisfy the containment functional inequalities of Corollary 1, but with $\mathsf{Q}(h,z)$ replacing $\mathsf{F}(\mathcal{R}_U)$. All sets in $\mathsf{U}(h,z)$ are elements of $\mathsf{Q}(h,z)$.

If, as is the case in many models, the sets in $\mathsf{U}(h,z)$ are each connected with boundary having 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 (4.3) is implied if the sets $\mathcal{S}_1$ and $\mathcal{S}_2$ have non-overlapping interiors. This implication was used in the construction of core-determining sets in Chesher, Rosen, and Smolinski (2013), in which all elements of $\mathsf{U}(h,z)$ were connected.

Example 1 *To illustrate the results of Theorem 3 consider the binary outcome IV model in Example 1 of Section 2.2 in which*

$$\mathsf{U}(h,z) = \left\{([0, g(y_2, z_1)], [g(y_2, z_1), 1]) : y_2 \in \mathcal{R}_{Y_2|z}\right\}.$$

In the collection of unions $\mathsf{U}^*(h,z)$ *there are three types of unions:*

$$\mathcal{J}_0(\mathcal{Y}_2) \equiv \bigcup_{y_2 \in \mathcal{Y}_2 \subseteq \mathcal{R}_{Y_2}} [0, g(y_2, z_1)] = \left[0, \max_{y_2 \in \mathcal{Y}_2} g(y_2, z_1)\right]$$

$$\mathcal{J}_1(\mathcal{Y}_2) \equiv \bigcup_{y_2 \in \mathcal{Y}_2 \subseteq \mathcal{R}_{Y_2}} [g(y_2, z_1), 1] = \left[\min_{y_2 \in \mathcal{Y}_2} g(y_2, z_1), 1\right]$$

$$\mathcal{J}_{01}(\mathcal{Y}_2', \mathcal{Y}_2'') \equiv \mathcal{J}_0(\mathcal{Y}_2') \cup \mathcal{J}_1(\mathcal{Y}_2'') = \left[0, g\left(y_2', z_1\right)\right] \cup \left[g\left(y_2'', z_1\right), 1\right]$$

where $y_2' \in \underset{y_2 \in \mathcal{Y}_2'}{\operatorname{argmax}}\, g(y_2, z_1)$ *and* $y_2'' \in \underset{y_2 \in \mathcal{Y}_2''}{\operatorname{argmax}}\, g(y_2, z_1)$. *Any union* $\mathcal{J}_{01}(\mathcal{Y}_2', \mathcal{Y}_2'')$ *such that* $g(y_2', z_1) \geq g(y_2'', z_1)$ *is equal to the unit interval* $[0,1]$. *Such unions can be excluded from* $\mathsf{Q}(h,z)$ *because with* $\mathcal{S} = [0,1]$ *the inequality in (4.1) is always satisfied. Any union* $\mathcal{J}_{01}(\mathcal{Y}_2', \mathcal{Y}_2'')$ *such that* $g(y_2', z_1) < g(y_2'', z_1)$ *is not connected and by Theorem 3 can be excluded from* $\mathsf{Q}(h,z)$. *It follows that, in the binary outcome IV model,* $\mathsf{Q}(h,z) = \mathsf{U}(h,z)$.

The following Corollary sets out cases in which certain of the containment functional inequalities can be replaced by equalities.

Corollary 2 *Define*

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

If $G_{U|Z}(\cdot|z)$ *is absolutely continuous with respect to Lebesgue measure then, under the conditions of Theorem 3, the collection of equalities and inequalities*

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

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

There are two classes of models in which *all* members of $\mathsf{Q}(h,z)$ belong to $\mathsf{Q}^E(h,z)$, so that the characterization of the identified set delivered by the Corollary comprises a collection of *only* conditional moment *equalities*, as follows.

1. *Models where* $\mathcal{U}(Y,Z;h)$ *is a singleton set with probability one.* This includes models with additive unobservables such as the classical linear IV model, the nonparametric IV model of Newey and Powell (2003), and IV models with structural function strictly monotone in a scalar unobservable, for example the quantile IV model studied by Chernozhukov and Hansen (2005).

2. *Complete models for which $\mathcal{Y}(U, Z; h)$ is a singleton with probability one.* In such models for any z and any $y \neq y'$ the sets $\mathcal{U}(y, z; h)$ and $\mathcal{U}(y', z; h)$ have measure zero intersection, as otherwise the model would produce non-singleton outcome sets $\mathcal{Y}(U, Z; h)$ with positive probability. Since $\mathsf{Q}(h, z)$ is a collection of sets that are unions of sets on the support of $\mathcal{U}(Y, Z; h)$ we then have that for all (y, z), any h, and all $\mathcal{S} \in \mathsf{Q}(h, z)$, either $\mathcal{U}(y, z; h) \subseteq \mathcal{S}$ or $\mathcal{U}(y, z; h) \subseteq \mathrm{cl}(\mathcal{S}^c)$.

Identified sets that are characterized by systems of moment equalities *and* inequalities typically arise when models are incomplete and $\mathcal{U}(Y, Z; h)$ is not required to be a singleton set with probability one. This includes incomplete discrete outcome models, models with high dimensional unobserved heterogeneity, and models where the relation between outcomes and unobservable variables is defined by inequalities.

The collection of core-determining sets from Theorem 2 and Corollary 2 may be uncountably infinite in models with continuous endogenous variables. However, in models in which all endogenous variables are discrete with finite support, the sets $\mathsf{Q}^E(h, z)$ and $\mathsf{Q}^I(h, z)$ are finite. In Chesher and Rosen (2012) 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.

The precise form and cardinality of $\mathsf{Q}(h, z)$ depends on the particular model under consideration. In some cases, such as Example 1 above, this collection is quite small and can be computed quickly. In other cases the collection can be extremely large, for example in the two examples studied in Section 6 and that featured in Appendix C, $\mathsf{Q}(h, z)$ comprises an uncountable infinity of tests sets. There is no single catch-all rule for picking out a finite number of sets from this collection in such cases. Nevertheless, the characterization can be helpful in providing guidance for intelligently selecting from among different potential collections of test sets in practice. We describe precisely how we have done this in the examples in Section 6 and in the supplementary material.

5 Identified Sets Under Restrictions on the Distribution of (U, Z)

Theorem 2 provides a characterization of the structures $\left(h, \mathcal{G}_{U|Z}\right)$ contained in the identified set delivered by a model $\mathcal{M}$ and a collection of distributions $\mathcal{F}_{Y|Z}$. A key element of econometric models are restrictions on the conditional distributions of unobserved variables. In this Section we show how some commonly employed restrictions on admissible collections of conditional distributions $\mathcal{G}_{U|Z}$ refine the characterization of an identified set. The restrictions considered are full stochastic independence, conditional mean and conditional quantile independence.

5.1 Stochastic Independence

Restriction SI: For all collections $\mathcal{G}_{U|Z}$ of conditional distributions admitted by $\mathcal{M}$, U and Z are stochastically independent. $\square$

Under Restriction SI conditional distributions $G_{U|Z}(\cdot|z)$ cannot vary with z and we write G_U in place of the collection $\mathcal{G}_{U|Z}$ where for all z, $G_U(\cdot) = G_{U|Z}(\cdot|z)$.

It follows from Theorem 2 that a structure $(h, G_U) \in \mathcal{M}$ belongs to $\mathcal{M}^*$ if and only if G_U is selectionable with respect to the conditional distribution of the random set $\mathcal{U}(Y, Z; h)$ induced by $F_{Y|Z}(\cdot|z)$ a.e. $z \in \mathcal{R}_Z$. Four characterizations of such structures are set out in Theorem 4.

Theorem 4 *Let Restrictions A1-A5 and SI hold. Then:*

$$\begin{aligned} \mathcal{M}^* &= \left\{(h, G_U) \in \mathcal{M} : G_U(\cdot) \preccurlyeq \mathcal{U}(Y, z; h) \ \text{ when } Y \sim F_{Y|Z}(\cdot|z), \ \text{a.e. } z \in \mathcal{R}_Z\right\} & (5.1) \\ &= \left\{ \begin{array}{c} (h, G_U) \in \mathcal{M} : \forall \mathcal{S}_I \in \mathsf{Q}^I(h, z), \forall \mathcal{S}_E \in \mathsf{Q}^E(h, z), \\ C_h(\mathcal{S}_I|z) \leq G_U(\mathcal{S}_I), \ C_h(\mathcal{S}_E|z) = G_U(\mathcal{S}_E), \ \text{a.e. } z \in \mathcal{R}_Z \end{array} \right\} & (5.2) \end{aligned}$$

If Restriction A6 also holds, then equivalently:

$$\begin{aligned} \mathcal{M}^* &= \left\{(h, G_U) \in \mathcal{M} : F_{Y|Z}(\cdot|z) \preccurlyeq \mathcal{Y}(U, z; h) \ \text{ when } U \sim G_U(\cdot), \ \text{a.e. } z \in \mathcal{R}_Z\right\}, & (5.3) \\ &= \left\{ \begin{array}{l} (h, G_U) \in \mathcal{M} : \forall \mathcal{K} \in \mathsf{K}(\mathcal{Y}), \\ \quad F_{Y|Z}(\mathcal{K}|z) \leq G_U\{\mathcal{Y}(U, z; h) \cap \mathcal{K} \neq \emptyset\}, \ \text{a.e. } z \in \mathcal{R}_Z \end{array} \right\}, & (5.4) \end{aligned}$$

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

Theorem 4 presents alternative representations of the identified set under Restriction SI. Characterizations (5.3) and (5.1) arise directly on application of the restriction to Definition 3.1 and Theorem 2, respectively. The characterization (5.2) applies Theorem 3 and Corollary 2 to define the identified set in terms of the conditional containment functional of the random set $\mathcal{U}(Y, Z; h)$. This representation employs core-determining sets to reduce the number of moment conditions in the characterization, and distinguishes which ones hold as equalities and inequalities.

The characterization (5.4) defines the identified set through conditional moment inequalities involving the capacity functional of $\mathcal{Y}(U, Z; h)$. This delivers the characterizations provided in Appendix D.2 of BMM11 and in Galichon and Henry (2011) when applied to incomplete models of games. In general this characterization using random sets in Y-space, $\mathcal{R}_Y$, requires the inequalities to hold for all compact sets $\mathcal{K} \subset \mathcal{R}_Y$. Simplification is sometimes possible: Galichon and Henry (2011) provide core-determining sets in $\mathcal{R}_Y$ when a certain monotonicity condition holds; BMM11 Appendix D.3 provides alternative conditions under which some inequalities are redundant.

In many cases, the representation (5.2) will be the simplest to use. This characterization uses the containment functional of $\mathcal{U}(Y, Z; h)$ which has support in U-space. This allows the use of core determining sets on $\mathcal{R}_U$ given by Theorem 3, which is in general a smaller collection of sets than all compact sets in Y-space. The ability to exploit the structure of sets $\mathcal{U}(Y,Z;h)$ for this purpose is a benefit of working in the space of unobserved heterogeneity. Our construction is based on core determining sets specific to each (h,z) pair, while the collections of core-determining sets working in Y-space characterized by Galichon and Henry (2011) under monotonicity are not.

A further difference between characterizations (5.2) and (5.4) is how they incorporate restrictions on the distribution of unobserved heterogeneity. Given an admissible distribution G_U, use of characterization (5.4) requires computation of the probability that $\mathcal{Y}(U,z;h)$ hits $\mathcal{K}$ for each compact set $\mathcal{K}$. This has typically been achieved by means of simulation from each conjectured distribution G_U, see e.g. Appendix D.2 of BMM11 and Henry, Meango, and Queyranne (2011).

The characterization (5.2) in U-space shows that there is an alternative to simulating draws from the distribution of unobservables. Computation using (5.2) requires computation of $G_U(\mathcal{S})$ for each conjectured distribution G_U and each core-determining set $\mathcal{S}$, which can be done either by simulation or by numerical integration. The term $\mathbb{P}[\mathcal{U}(Y,Z;h)\subseteq\mathcal{S}|z]$ is the probability of an event concerning only the observed variables (Y,Z), which is point-identified and can be computed or estimated directly.

Our characterization of identified sets employing random sets in U-space leads directly to Corollary 3 which characterizes the identified set for the structural function h under Restriction SI, in the absence of further restrictions on G_U.[21] This identified set is the projection of the identified set of structures (h,G_U) onto the space $\mathcal{H}$ in which the structural function h is restricted to lie.

Corollary 3 *If Restrictions A1-A5 and SI hold, and $\mathcal{G}_{U|Z}$ is otherwise unrestricted the identified set of structural functions h is*

$$\mathcal{H}^* = \left\{ h\in\mathcal{H} : \forall\mathcal{S}\in\mathsf{Q}^*(h),\quad \sup_{z\in\mathcal{R}_Z} C_h(\mathcal{S}|z) \le \inf_{z\in\mathcal{R}_Z}\left(1-C_h(\mathcal{S}^c|z)\right), \right\}, \tag{5.5}$$

where $\mathsf{Q}^*(h)$ is any collection of sets $\mathcal{S}\subseteq\mathcal{R}_U$ such that for all $z\in\mathcal{R}_Z$, $\mathsf{Q}(h,z)\subseteq\mathsf{Q}^*(h)$.

In (5.5) $1-C_h(\mathcal{S}^c|z) = \mathbb{P}[U(Y,Z;h)\cap S\neq\emptyset|Z=z]$ is the conditional capacity functional of $\mathcal{U}(Y,Z;h)$ given $Z=z$. The result is obtained using an upper bound on $G_U(\mathcal{S})$ produced by applying the containment functional inequalities in Theorem 4 to $\mathcal{S}^c$, the complement of $\mathcal{S}$, and using the monotonicity in $\mathcal{S}$ of the containment and capacity functionals of $U(Y,Z;h)$

This strikingly simple projection result is extremely useful in situations in which G_U is not parametrically specified, allowing information about the structural functions that could have delivered data without the need to simultaneously understand the various distributions of unobserved U that, coupled with these structural functions, could have delivered the data. Such a result is rare

[21] If there are further restrictions on G_U (e.g. symmetry) the identified set for the structural function will be a subset of $\mathcal{H}^*$ defined in Corollary 3.

in the partial identification literature and a notable benefit coming from working with random sets defined in the space of unobservables rather than the space of outcomes.

5.2 Mean Independence

Restriction MI: $\mathsf{G}_{U|Z}$ comprises all collections $\mathcal{G}_{U|Z}$ of conditional distributions for U given Z satisfying $E[U|z] = c$, a.e. $z \in \mathcal{R}_Z$, for some fixed, finite c belonging to a known set $\mathcal{C} \subseteq \mathcal{R}_U$. $\square$

This restriction limits the collection $\mathcal{G}_{U|Z}$ to those containing conditional distributions $G_{U|Z}(\cdot|z)$ such that $E[U|z]$ is equal to a constant c which does not vary with z. This covers cases where numerical values are provided for some components of c but not for others. For instance, in a model with bivariate U, Restriction MI with $\mathcal{C} = \{(c_1, c_2) : c_1 = 0, c_2 \in \mathbb{R}\}$ restricts $E[U_1|z] = 0$, which could be a normalization, and restricts $E[U_2|z]$ to be invariant with respect to z.

Under Restriction MI it is convenient to characterize the selectionability criterion of Theorem 2 using the Aumann expectation.

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

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

Molchanov (2005, p. 151). The Aumann expectation of random set $\mathcal{A}$ conditional on $B = b$ is

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

A characterization of the identified sets for structural function h and for the structure $(h, \mathcal{G}_{U|Z})$ under Restriction MI is given in the following Theorem.

Theorem 5 *Let Restrictions A1-A5 and MI hold and suppose that $(\Omega, \mathcal{F}, \mathbb{P})$ is non-atomic. Then the identified set for structural function h comprises those functions h such that some $c \in \mathcal{C}$ is an element of the Aumann expectation of $\mathcal{U}(Y, Z; h)$ conditional on $Z = z$ a.e. $z \in \mathcal{R}_Z$:*

$$\mathcal{H}^* = \{h \in \mathcal{H} : \exists c \in \mathcal{C} \text{ s.t. for almost every } z \in \mathcal{R}_Z,\ c \in \mathbb{E}[\mathcal{U}(Y, Z; h)|z]\}.$$

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

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

where by virtue of Restriction MI all structures $(h, \mathcal{G}_{U|Z}) \in \mathcal{M}^* \subseteq \mathcal{M}$ are such that for some $c \in \mathcal{C}$, $E[U|z] = c$ a.e. $z \in \mathcal{R}_Z$.

Knowledge of 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

$$E \sup \{\|U\| : U \in \mathcal{U}(Y,Z;h)\} < \infty, \tag{5.6}$$

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

$$\inf_{v \in \mathcal{R}_U : \|v\|=1} \left\{ E[m(v, \mathcal{U}(Y,Z;h))|z] - v'c \right\} \geq 0, \tag{5.7}$$

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. BMM11 employed Molchanov (2005, Theorem 2.1.47-iv, p. 171) when using the conditional Aumann expectation of random *outcome* set $\mathcal{Y}(Z,U;h)$ to characterize its selections. In our analysis using random sets in U-space the result simplifies the task of determining whether $c \in \mathbb{E}[\mathcal{U}(Y,Z;h)|z]$ for some $c \in \mathcal{C}$.[22] If the structural function h is additively separable in Y, the two representations are equivalent, differing only by a trivial location shift.

On the other hand, if h is not additively separable in U, the conditional mean restriction MI cannot generally be written as a conditional mean restriction on Y, and previous identification results using random sets in Y-space appear inapplicable. Theorem 5 provides a novel characterization for the identified set in such cases. Furthermore, Theorem 5 does not require $\mathcal{U}(Y,Z;h)$ to be integrably bounded, but only integrable, enabling its application when the support of unobserved heterogeneity $\mathcal{R}_U$ is unbounded.

In some commonly occurring models, including all those of Examples 1-5 in Section 2.2, $\mathcal{U}(Y,Z;h)$ 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 (5.7), 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*

$$\mathcal{H}^* = \left\{ \begin{array}{l} h \in \mathcal{H} : \exists c \in \mathcal{C} \textit{ s.t. for almost every } z \in \mathcal{R}_Z,\ E[u(Y,Z)|z] = c, \\ \textit{for some function } u : \mathcal{R}_{YZ} \to \mathcal{R}_U \textit{ with } \mathbb{P}[u(Y,Z) \in \mathcal{U}(Y,Z;h)] = 1 \end{array} \right\}.$$

Theorem 5 can be generalized to characterize $\mathcal{H}^*$ under more general forms of conditional mean

[22] For instance, when $\mathcal{C} = 0$, equivalently when $E[U|z] = 0$ a.e. $z \in \mathcal{R}_Z$ is imposed, the support function inequality (5.7) implies that the identified set for h are those $h \in \mathcal{H}$ such that

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

restriction as expressed in Restriction MI*.

Restriction MI*: *$G_{U|Z}$ comprises all collections $G_{U|Z}$ of conditional distributions for U given Z such that for some known function $d(\cdot,\cdot) : R_U \times R_Z \to \mathbb{R}^{k_d}$, $E[d(U,Z)|z] = c$ a.e. $z \in R_Z$, for some fixed c belonging to a known set $C \subseteq R_U$, where $d(u,z)$ is continuous in u for all $z \in R_Z$.* $\square$

Restriction **MI*** requires that the conditional mean given $Z = z$ of some function $d(U,Z)$ taking values in $\mathbb{R}^{k_d}$ does not vary with respect to z. This restriction can accommodate models that impose conditional mean restrictions on functions of unobservables U, for example homoskedasticity restrictions or restrictions on covariances of elements of U. To express the identified set delivered under restriction **MI*** define

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

Then $\mathcal{D}(Y,Z;h)$ is a random set of feasible values for $d(U,Z)$ given observed (Y,Z). This set is closed under the requirement of Restriction MI* that $d(\cdot,z)$ is continuous for each z. The arguments that deliver Theorem 5 yield the following result.

Corollary 5 *Let Restrictions A1-A5 and MI* hold and suppose that $(\Omega,\mathcal{F},\mathbb{P})$ is non-atomic. Then the identified set for structural function h are those h such that there exists at least one $c \in \mathcal{C}$ that is an element of the Aumann expectation of $\mathcal{D}(Y,Z;h)$ conditional on $Z = z$ a.e. $z \in \mathcal{R}_Z$:*

$$\mathcal{H}^* = \{h \in \mathcal{H} : \exists c \in \mathcal{C} \text{ s.t. for almost every } z \in \mathcal{R}_Z,\ c \in \mathbb{E}[\mathcal{D}(Y,Z;h)|z]\}.$$

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

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

where, by Restriction MI, all structures $(h,\mathcal{G}_{U|Z}) \in \mathcal{M}^* \subseteq \mathcal{M}$ are such that for some $c \in \mathcal{C}$, $E[d(U,Z)|z] = c$, a.e. $z \in \mathcal{R}_Z$.*

5.3 Quantile Independence

Conditional quantile restrictions on the distribution of unobserved U can also be accommodated. This is illustrated in a simple setting under Restriction IS.

Restriction IS *(interval support): $U \in \mathbb{R}$ and for all $(y,z) \in R_{YZ}$,*

$$\mathcal{U}(y,z;h) = [\underline{u}(y,z;h), \overline{u}(y,z;h)], \tag{5.8}$$

where possibly $\underline{u}(y,z;h) = -\infty$ or $\overline{u}(y,z;h) = +\infty$, in which case the corresponding endpoint of the interval (5.8) is open. $\square$

This restriction requires U to be scalar and the sets $\mathcal{U}(y, z; h)$ to be closed intervals. This is a case that often arises in practice when quantile independence restrictions are employed.

This restriction can be fruitfully applied to GIV models with censored endogenous or exogenous variables, as we show in a model with interval censored endogenous variables in Section C.1.3 of the on-line supplement.

The conditional quantile restriction is as follows.

Restriction QI: *For some known $\tau \in (0,1)$ and some known set $\mathcal{C} \subseteq \mathbb{R}$, $G_{U|Z}$ comprises all collections $G_{U|Z}$ of conditional distributions for U given Z that are continuous in a neighborhood of their τ-quantile and satisfy the conditional quantile restriction $q_{U|Z}(\tau|z) = c$, a.e. $z \in R_Z$ for some $c \in \mathcal{C}$.* $\square$

Theorem 6 *Let Restrictions A1-A5, IS, and QI hold. Then (i) the identified set for structural function h is*

$$\mathcal{H}^* = \left\{ h \in \mathcal{H} : \exists c \in \mathcal{C} \ s.t. \ \sup_{z \in \mathcal{R}_Z} F_{Y|Z}\left[\overline{u}(Y, Z; h) \leq c|z\right] \leq \tau \leq \inf_{z \in \mathcal{R}_Z} F_{Y|Z}\left[\underline{u}(Y, Z; h) \leq c|z\right] \right\}. \tag{5.9}$$

(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

$$\mathcal{H}^* = \left\{ h \in \mathcal{H} : \exists c \in \mathcal{C} \ s.t. \ \sup_{z \in \mathcal{R}_Z} \underline{q}(\tau, z; h) \leq c \leq \inf_{z \in \mathcal{R}_Z} \overline{q}(\tau, z; h) \right\}, \tag{5.10}$$

where $\underline{q}(\tau, z; h)$ and $\overline{q}(\tau, z; h)$ are the τ-quantiles of respectively $\underline{u}(Y, Z; h)$ and $\overline{u}(Y, Z; h)$, (iii) The identified set for $\left(h, \mathcal{G}_{U|Z}\right)$ is:

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

where following from Restriction QI, all structures $\left(h, \mathcal{G}_{U|Z}\right) \in \mathcal{M}^* \subseteq \mathcal{M}$ are such that for some $c \in \mathcal{C}$, $q_{U|Z}(\tau|z) = c$, a.e. $z \in \mathcal{R}_Z$.

Under Restriction QI, the conditional distributions belonging to $\mathcal{G}_{U|Z}$ are continuous in a neighborhood of zero and therefore

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

The inequalities comprising (5.9) then follow from $\underline{u}(Y, Z; h) \leq U \leq \overline{u}(Y, Z; h)$. These inequalities also arise on applying the containment functional inequality $C_h(\mathcal{S}|z) \leq G_{U|Z}(\mathcal{S}|z)$ to test sets $\mathcal{S} = (-\infty, c]$ and $\mathcal{S} = [c, \infty)$. In the proof of Theorem 6 it is shown that for any h and any c, if the containment functional inequalities hold for these two test sets, then an admissible collection of conditional distributions $\mathcal{G}_{U|Z}$ can be found such that they hold for all closed test sets in $\mathcal{R}_U$. From Corollary 1 it follows that the characterization (5.9) is sharp.

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 (5.9) 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. $z\in\mathcal{R}_Z$.

6 Two Applications

This Section illustrates the application of the results of this paper to two particular models, both admitting multivariate unobserved heterogeneity.

First, identified sets of valuation distributions are characterized for an incomplete model of English Auctions with independent private values, previously studied by Haile and Tamer (2003), henceforth HT. The approach of this paper obviates the need for a constructive proof of sharpness, which, as noted in HT, is difficult to produce in the auction model. The new characterization of the identified set includes the inequalities derived in HT and refines the HT bounds with additional inequalities.

Second, identified sets of parameter values are characterized for continuous outcome random coefficient linear models with endogenous explanatory variables and instrumental variable restrictions.

6.1 An Incomplete Model of Auctions

This example revisits the incomplete model of an open outcry English ascending auction introduced in Haile and Tamer (2003), henceforth HT. In that IPV model M bidders have valuations which are independent realizations drawn from a common conditional distribution of valuations given observed auction characteristics $Z=z$, denoted $A_z(v)\equiv\mathbb{P}[V\leq v|Z=z]$. HT develop pointwise bounds on $A_z(v)$ which hold at each value v.

We show how the model can be set up as a GIV model. Applying Theorem 4 and the results in Section 3.3 on core determining sets delivers a characterization of the sharp bounds on valuation distributions supported by the HT model. The HT pointwise bounds appear in this characterization along with an uncountable infinity of additional inequalities which further restrict the shape of the distribution function. This resolves the sharpness question raised in the final paragraph of HT.

The complete analysis given in Chesher and Rosen (2015) is now summarized for the case in which there is no reserve price and no minimum bid increment.[23]

Realizations of random variable $Y = (Y_1, \ldots, Y_M)$ are ordered final bids made by M bidders. Realizations of $V = (V_1, \ldots, V_M)$ are ordered, continuously distributed, valuations of the bidders.[24] Let $\tilde{U} \in [0, 1]^M$ be M mutually independent uniform variates with $\tilde{U} \perp\!\!\!\perp Z$ and with order statistics $U = (U_1, \ldots, U_M)$. The elements of V, ordered valuations, can be expressed as functions of uniform order statistics as follows: $V_m = A_z^{-1}(U_m)$, $m \in \{1, \ldots, M\}$.

The HT model includes the restrictions: (i) the second highest valuation must be no larger than the winning bid which implies that almost surely $Y_M \geq V_{M-1}$, and (ii) no one bids more than their valuation, which implies that the inequality in order statistics $V_m \geq Y_m$ holds almost surely for all m. Applying the strictly monotone function A_z to both sides of these inequalities gives the structural function of the HT model.[25]

$$h(y, z, u) = \max\{u_{M-1} - A_z(y_M), 0\} + \sum_{m=1}^{M} \max\{A_z(y_m) - u_m, 0\} \tag{6.1}$$

The vector of M uniform order statistics, U, has constant density function equal to $M!$ supported on $\mathcal{R}_U$ which is the orthoscheme of the unit M-cube in which $U_1 \leq \cdots \leq U_M$, David and Nagaraja (2003). Let $G_U(\mathcal{S})$ denote the probability mass placed by this distribution on a set $\mathcal{S} \subseteq \mathcal{R}_U$. Structures in this model are pairs (h, G_U).

The U-level sets of the structural function (6.1) are as follows.

$$\mathcal{U}(y, z; h) = \left\{ u \in \mathcal{R}_U : (A_z(y_M) \geq u_{M-1}) \wedge \bigwedge_{m=1}^{M} (A_z(y_m) \leq u_m) \right\} \tag{6.2}$$

Lemma 1 states that core determining test sets which characterize the identified set for $A_z(\cdot)$ are unions of these U-level sets. There is an uncountable infinity of such unions, and in Chesher and Rosen (2015) we make a selection of unions of such sets, $\mathcal{S}(y', y''_M, z; h)$, of the following form.

$$\mathcal{S}(y', y''_M, z; h) \equiv \bigcup_{y_M \in [y'_M, y''_M]} \mathcal{U}((y'_1, y'_2, \ldots, y_M), z; h), \qquad y''_M \geq y'_M \geq \cdots \geq y'_1 \tag{6.3}$$

Applied to such sets the containment functional inequality given in (3.8) in Theorem 4 requires

[23] These restrictions are easily removed as shown in Chesher and Rosen (2015).

[24] Here and later, in M-element ordered lists, index M identifies the highest value.

[25] The elements in y and u in (6.2) are values of order statistics satisfying $y_1 \leq \cdots \leq y_M$ and $u_1 \leq \cdots \leq u_M$.

that all valuation distributions, A_z, in the identified set satisfy

$$M! \int_{A_z(y'_M)}^{1} \int_{A_z(y'_{M-1})}^{\min(u_M, A_z(y''_M))} \int_{A_z(y'_{M-2})}^{u_{M-1}} \cdots \int_{A_z(y'_1)}^{u_2} du_M du_{M-1} du_{M-2} \cdots du_1 \geq \mathbb{P}\left[(y''_M \geq Y_M \geq y'_M) \wedge \left(\bigwedge_{m=1}^{M-1} (Y_m \geq y'_m) \right) |z \right] \quad (6.4)$$

for all z, M, and y' and y''_M such that $y''_M \geq y'_M \geq \cdots \geq y'_1$. On the left hand side is $G_U(\mathcal{S}(y', y''_M, z; h))$; on the right hand side is the containment functional $\mathbb{P}[\mathcal{U}(Y, Z; h) \subseteq \mathcal{S}(y', y''_M, z; h)|z]$.[26]

Particular choices of y' and y''_M deliver the pointwise bounds in HT. Plugging $y''_M = +\infty$, $y'_m = -\infty$ for $m < n$ and $y'_m = v$ for $m \geq n$ into (6.4) delivers the following inequality.

$$\mathbb{P}[U_n \geq A_z(v)|Z = z] \geq \mathbb{P}[Y_n \geq v|z] \quad (6.5)$$

The marginal distribution of U_n, the nth of M uniform order statistics, is Beta$(n, M + 1 - n)$. Let $Q(\cdot, n, M)$ denote the associated quantile function. Transforming both sides of (6.5) expressed in terms of distribution functions using this quantile function gives

$$A_z(v) \leq Q(\mathbb{P}[Y_n \leq v|z], n, M)$$

which leads to directly to the pointwise upper bound in Theorem 1 of HT.

Plugging $y''_M = v$ and $y' = (-\infty, \cdots, -\infty)$ into (6.4) delivers, after a similar manipulation, the inequality

$$A_z(v) \geq Q(\mathbb{P}[Y_M \leq v|z], M - 1, M)$$

which leads directly to the pointwise lower bound in Theorem 3 of HT.

The choices of y' and y''_M in (6.4) used till now deliver pointwise bounds on $A_z(\cdot)$, that is, bounds on the value of the distribution function at a single value of its argument. Other choices of y' and y''_M lead to inequalities which restrict values of $A_z(\cdot)$ at two or more values of its argument. For example, plugging $y''_M = +\infty$ and $y' = (-\infty, \cdots, -\infty, v_2, v_1)$ into (6.4) delivers the inequality

$$1 - A_z(v_1)^M - MA_z(v_2)^{M-1} + MA_z(v_1)A_z(v_2)^{M-1} \geq \mathbb{P}[Y_M \geq v_1 \wedge Y_{M-1} \geq v_2|z] \quad (6.6)$$

[26] These expressions are derived in Chesher and Rosen (2015).

Table 1: Projections of the identified set onto each parameter axis in turn and value of the parameters in the structures that generate probability distributions of Y given Z.

Instrument Strength	Parameter	Value in Structures	Lower bound	Upper bound
	μ_1	0	-0.083	0.025
	μ_2	1	0.955	1.066
$d = 1.0$	σ_{11}	1	0.881	1.112
	σ_{12}	0	-0.036	0.050
	σ_{22}	0.25	0.050	0.364
	μ_1	0	-0.326	0.158
	μ_2	1	0.797	1.388
$d = 0.5$	σ_{11}	1	0.811	1.399
	σ_{12}	0	-0.189	0.275
	σ_{22}	0.25	0.000	1.612

and $y''_M = v_1$, $y' = (-\infty, \cdots, -\infty, v_2, v_2)$ inserted in (6.4) delivers the following inequality

$$\begin{aligned} A_z(v_1)^M - MA_z(v_1)A_z(v_2)^{M-1} + (M-1)A_z(v_2)^M + \\ M(1 - A_z(v_1))(A_z(v_1)^{M-1} - A_z(v_2)^{M-1}) \geq \\ \mathbb{P}[v_1 \geq Y_M \geq v_2 \wedge Y_{M-1} \geq v_2 | z] \end{aligned} \quad (6.7)$$

Chesher and Rosen (2015) gives examples of distributions of ordered final bids for which both of these inequalities are binding. In the examples considered further test sets defined by three points on the valuation distribution are found to achieve further refinement, though only slightly.

Consideration of all possible choices for y' and y''_M in the union of U-level sets (6.3) yields *via* (6.4) a dense system of inequalities involving values of the conditional distribution function of valuations at all choices of up to $M + 1$ values of its arguments. A complete characterization of the identified set of valuation distributions involves consideration as well of an uncountable infinity of unions of sets of the form (6.3) so there is no limit to the number of coordinates of valuation distributions simultaneously constrained by the HT model. The results of this paper applied to this problem deliver the first complete characterization of the identified set of valuation distributions delivered by the HT model and reveal its extraordinary complexity.

6.2 An Instrumental Variable Random Coefficients Model

This section employs an instance of the single equation instrumental random coefficient model of Section 2.2 Example 3. A parametric Gaussian restriction on the distribution of unobserved heterogeneity is imposed. For ease of illustration, there is a scalar instrumental variable Z subject

Table 2: Instrument power in the random coefficients model: $\mathbb{P}[Y_2 = y|Z = z]$.

		z				
d	y	-2	-1	0	1	2
1.0	-1	.01	.07	.31	.69	.93
	0	.06	.24	.38	.24	.06
	$+1$	.93	.69	.31	.07	.01
0.5	-1	.07	.16	.31	.50	.69
	0	.24	.34	.38	.34	.24
	$+1$	.69	.50	.31	.16	.07

to an exclusion restriction and there are no included exogenous variables present.

$$Y_1 = U_1 + U_2 Y_2 \qquad U \sim N_2(\mu, \Sigma) \qquad U \perp\!\!\!\perp Z \in \{-2, -1, 0, 1, 2\} = \mathcal{R}_Z.$$

The corresponding definitions of structural function h and level set $\mathcal{U}(y, z; h)$ are:

$$h(y, z, u) = y_1 - u_1 - u_2 y_2, \qquad \mathcal{U}(y, z; h) = \{u \in \mathcal{R}^2 : y_1 = u_1 + u_2 y_2\}.$$

We stay with the notation for U-level sets $\mathcal{U}((y_1, y_2), z; h)$ used elsewhere in the paper even though in this example, under the restrictions imposed by the IV model, these level sets do not vary with z, and the structural function is known. In this random coefficients linear model the unknown parameter vector θ, comprising the unique elements of μ and Σ, affects only the bivariate normal distribution of U, denoted here by $G_U(\cdot; \theta)$.

Allowing for additional values of Z or higher dimensional Z, as well as the presence of included exogenous variables, is conceptually straightforward, with the analysis proceeding as below with $\mathcal{U}(y, z; h)$ as defined in (2.2). If there are random coefficients on additional included variables, then $\mathcal{U}(y, z; h)$ remains a linear manifold on $\mathcal{R}_U$, of higher dimension.

Approximations to identified sets delivered by this IV model are calculated using two particular probability distributions of Y given Z when Y_2 is a discrete random variable with support $\{-1, 0, 1\}$.[27] In the two cases considered the instrumental variable has different predictive power for endogenous Y_2. The incomplete IV model is shown to be capable of delivering substantial partial identifying power even in the weaker instrument case.

To illustrate, conditional distributions for $Y \equiv (Y_1, Y_2)$ given Z are generated using a triangular structure in which Y_2 is delivered by an ordered probit model:

$$Y_2 = -1[d \times Z + U_3 < -0.5] + 1[d \times Z + U_3 > 0.5]$$

[27] These are approximations because we examine a finite number of the uncountable infinity of inequalities that characterize the identified set.

with parameter d equal to 1.0 and 0.5 as we study respectively stronger and weaker instrument situations and with $U^* \equiv (U_1, U_2, U_3)$ trivariate Gaussian with mean $(0, 1, 0)$ and variance matrix S, as follows.

$$S = \begin{bmatrix} 1 & 0 & 0.25 \\ 0 & 0.25 & -0.25 \\ 0.25 & -0.25 & 1 \end{bmatrix}$$

Core determining test sets for this model comprise certain unions of manifolds:

$$\mathcal{S}(\mathcal{A}(-1), \mathcal{A}(0), \mathcal{A}(1)) \equiv \bigcup_{y_2 \in \{-1,0,1\}} \left(\bigcup_{y_1 \in \mathcal{A}(y_2)} \mathcal{U}((y_1, y_2), z; h) \right) \tag{6.8}$$

defined by sets of values of Y_1, denoted $\mathcal{A}(y_2)$, whose membership may depend on the value y_2 of Y_2.

There are uncountably infinitely many unions of the form (6.8) and thus an equivalent number of conditional moment inequalities characterizing the identified set. Incorporating all of these inequalities is therefore computationally infeasible. We thus illustrate outer sets based on the collections of test sets described below.

In the calculations reported here the sets $\mathcal{A}(y_2)$ are intervals.[28] For each integer n define the sequence $(0, b, 2b, \ldots, (n-1)b, 1)$ where $b \equiv n^{-1}$, and define a triangular array of intervals:

$$B(n) \equiv \begin{bmatrix} [0,b] & [0,2b] & [0,3b] & \cdots & [0,(n-1)b] & [0,1] \\ & [b,2b] & [b,3b] & \cdots & [b,(n-1)b] & [b,1] \\ & & [2b,3b] & \cdots & [2b,(n-1)b] & [2b,1] \\ & & & \ddots & & \vdots \\ & & & & \ddots & \vdots \\ & & & & & [(n-1)b,1] \end{bmatrix}.$$

A collection of $n(n+1)/2$ intervals is obtained as the unique elements of this array. This is then transformed into a collection $\mathsf{C}(n)$ of intervals on the real line by replacing $[0, 1]$ with the empty interval and replacing all others with twice the value of the standard normal quantile function evaluated at the original endpoints.

The collection of test sets used in our calculations is

$$\{\mathcal{S}(\mathcal{A}(-1), \mathcal{A}(0), \mathcal{A}(1)) : \mathcal{A}(-1) \in \mathsf{C}(n), \mathcal{A}(0) \in \mathsf{C}(n), \mathcal{A}(1) \in \mathsf{C}(n)\}.$$

The inclusion of the empty interval in $\mathsf{C}(n)$ ensures that within the collection there are some test

[28] In the complete collection of core determining sets, the sets $\mathcal{A}(y_2)$ comprise all collections of subsets of the real line.

sets defined by a single value of Y_2, some defined by each selection of two of the three possible values of Y_2 as well as test sets defined by three values of Y_2. There are $(n(n+1)/2)^3$ test sets in this collection. In the calculations reported here $n = 9$ and the collection of test sets generates $91,125$ inequalities, each having the following form.

$$G_U(\mathcal{S}(\mathcal{A}(-1), \mathcal{A}(0), \mathcal{A}(1)); \theta) \geq \sup_{z \in \{-2,-1,0,1,2\}} \left\{ \sum_{y_2 \in \{-1,0,1\}} \mathbb{P}[Y_1 \in \mathcal{A}(y_2) \wedge Y_2 = y_2 | Z = z] \right\}$$

$G_U(\mathcal{S}; \theta)$ is the probability mass placed on a set $\mathcal{S}$ by a $N_2(\mu, \Sigma)$ probability distribution.[29] The probabilities on the right hand side are calculated using the triangular structure specified earlier.

Table 1 shows projections of the 5 dimensional identified set for θ onto each axis in turn.[30] The IV model is clearly informative about the magnitudes of the 5 parameters, much more so when $d = 1$ than when $d = 0.5$. Why do we see this substantial difference as the coefficient d varies? With the additional restriction that Y_2 is exogenous $\theta \equiv (\mu, \Sigma)$ is point identified. Therefore, if the instrumental variable Z were a perfect predictor of Y_2 the IV model would be point identifying.[31] With $d = 1$ the instrumental variable is a much better, though not perfect, predictor of Y_2. This can be seen by examining Table 2 which shows the conditional distributions of Y_2 given Z for $d \in \{1.0, 0.5\}$.

The identified interval for μ_2 has width 0.111 when $d = 1$ and width 0.591 when $d = 0.5$. In both cases the sign of μ_2 is identified. When the instrument is weaker, the identified interval for σ_{22}, the variance of the random coefficient on Y_2, includes zero. This means that the hypothesis of an additive scalar unobservable with fixed coefficient on Y_2 cannot be refuted. However, with the stronger instrument zero is excluded from the identified interval, implying that there are no admissible structures with a fixed coefficient on Y_2 in the identified set.

7 Conclusion

This paper provides characterizations of identified sets of structures and structural features for a very broad class of models. It delivers results for complete and incomplete models and for partially and point identifying models. The results apply to models in which the inverse of the structural mapping from unobserved heterogeneity to observed endogenous variables may not be

[29] These probabilities are calculated as numerical integrals using the `integrate` function in R, R Core Team (2014).

[30] Exploratory calculations reveal that the projections are all connected intervals. The projections are calculated as follows. A measure of the distance of a value of 5-element θ from the identified set is defined such that the measure is negative for θ in the identified set, zero on the boundary of the set and positive off the set. A value of a parameter $\theta_i = \theta_i^*$ is judged inside the projection for θ_i if the minimum of this measure with respect to θ subject to $\theta_i = \theta_i^*$ is nonpositive. Minimization is done using the `optim` function in R. Endpoints of projection intervals are determined using the `uniroot` function in R.

[31] This is so because if $U \perp\!\!\!\perp Z$ and Z is a perfect predictor of Y_2, then $U \perp\!\!\!\perp Y_2$.

unique-valued. Models with discrete and censored endogenous variables fall under this heading, as do models permitting general forms of multivariate unobserved heterogeneity, such as random coefficient models and models placing inequality constraints on unobserved and endogenous variables.

The results extend the scope of application of instrumental variables for use in structural econometric models, in view of which we have described the class of models covered here as Generalized Instrumental Variable models. It is straightforward to incorporate instrumental variable restrictions involving conditional mean or quantile independence of unobserved variables and instruments as well as full stochastic independence. However the coverage of the results is wider because in some of the models to which the results apply instrumental variable restrictions play no significant role.

The characterizations developed here always deliver sharp bounds, removing the need for case-by-case constructive proofs of sharpness. This is a great benefit since it is often difficult to formulate such proofs and the task is sometimes impossible as, for example, is the case in the model of English auctions of Section 6.1.

Using tools from random set theory, relying in particular on the concept of *selectionability*, the classical notion of observational equivalence has been formally extended to models whose structures may not be required to deliver unique conditional distributions for endogenous variables given exogenous variables. We have shown 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.

All of the characterizations of identified sets presented in this paper can be expressed as systems of conditional moment inequalities and equalities. These can be employed for estimation and inference using a variety of approaches from the recent literature. With a discrete conditioning variable the identified sets derived in Section 3 can be expressed using unconditional moment inequality representations, for example as in Chernozhukov, Hong, and Tamer (2007), Beresteanu and Molinari (2008), Romano and Shaikh (2008a,b), Rosen (2008), Andrews and Guggenberger (2009), Andrews and Soares (2010), Andrews and Jia-Barwick (2010), Bugni (2010), Canay (2010), and Romano, Shaikh, and Wolf (2013). With a continuous conditioning variable inference using conditional moment inequalities can be performed, see for example Andrews and Shi (2013a,b), Chernozhukov, Lee, and Rosen (2013), Lee, Song, and Whang (2013a,b), Armstrong(2011a,b), and Chetverikov (2011).

In some models the number of inequality restrictions fully characterizing an identified set can be very large relative to the sample size. This is a common problem, not unique to the models considered in this paper, and it is the subject of current research efforts, see for example: Menzel

(2009), Chernozhukov, Chetverikov, and Kato (2013) and Andrews and Shi (2015).

The complexity of such characterizations in this paper are a consequence of using complete characterizations of identified sets, that is sharp bounds, which the methods of this paper always deliver, rather than outer bounds. Compare for example the relative simplicity and ease of use of the pointwise bounds on valuation distributions in the English auction model of Section 6.1 and the complexity of the complete characterization of the identified set of distributions obtained using the results of this paper. The additional inequalities afforded by the sharp characterization will generally deliver tighter bounds, and so their use is beneficial. In practice, the benefit of incorporating additional inequalities must be weighed against computational cost. We have demonstrated approaches for selecting finite collections of inequalities from the uncountable infinity of those characterizing the identified set in the context of each of our examples. Formal prescriptions for this task, accounting for the scale of reduction of the resulting set given observable data, as well as for sampling variation of the moments involved, is beyond the scope of this paper but seems a useful avenue for future research.

References

Andrews, D. W. K., and P. Guggenberger (2009): "Validity of Subsampling and Plug-In Asymptotic Inference for Parameters Defined by Moment Inequalities," *Econometric Theory*, 25, 669–709.

Andrews, D. W. K., and P. Jia-Barwick (2010): "Inference for Parameters Defined by Moment Inequalities: A Recommended Moment Selection Procedure," *Econometrica*, 80(6), 28052826.

Andrews, D. W. K., and X. Shi (2013a): "Inference Based on Conditional Moment Inequalities," *Econometrica*, 81(2), 609–666.

——— (2013b): "Nonparametric Inference Based on Conditional Moment Inequalities," *Journal of Econometrics*, forthcoming.

——— (2015): "Inference Based on Many Conditional Moment Inequalities," Working Paper, Yale University and University of Wisconsin.

Andrews, D. W. K., and G. Soares (2010): "Inference for Parameters Defined by Moment Inequalities Using Generalized Moment Selection," *Econometrica*, 78(1), 119–157.

Aradillas-Lopez, A., and A. M. Rosen (2013): "Inference in Ordered Response Games with Complete Information," CEMMAP working paper CWP33/13.

Armstrong, T. B. (2014): "Weighted KS Statistics for Inference on Conditional Moment Inequalities," *Journal of Econometrics*, 181(2), 92–116.

——— (2015): "Asymptotically Exact Inference in Conditional Moment Inequality Models," *Journal of Econometrics*, 186(1), 51–65.

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

Balke, A., and J. Pearl (1997): "Bounds on Treatment Effects From Studies With Imperfect Compliance," *Journal of the American Statistical Association*, 92(439), 1171–1176.

Basmann, R. L. (1959): "A Generalized Classical Method of Linear Estimation of Coefficients in a Structural Equation," *Econometrica*, 25(1), 77–83.

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.

Beresteanu, A., and F. Molinari (2008): "Asymptotic Properties for a Class of Partially Identified Models," *Econometrica*, 76(4), 763–814.

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

Bugni, F. (2010): "Bootstrap Inference for Partially Identified Models Defined by Moment Inequalities: Coverage of the Identified Set," *Econometrica*, 78(2), 735–753.

Canay, I. (2010): "EL Inference for Partially Identified Models: Large Deviations Optimality and Bootstrap Validity," *Journal of Econometrics*, 156(2), 408–425.

Chernozhukov, V., D. Chetverikov, and K. Kato (2013): "Testing Many Moment Inequalities," CeMMAP working paper CWP65/13.

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

Chernozhukov, V., H. Hong, and E. Tamer (2007): "Estimation and Confidence Regions for Parameter Sets in Econometric Models," *Econometrica*, 75(5), 1243–1284.

Chernozhukov, V., S. Lee, and A. M. 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.

——— (2013): "Semiparametric Structural Models of Binary Response: Shape Restrictions and Partial Identification," *Econometric Theory*, 29, 231–266.

Chesher, A., and A. M. Rosen (2012): "Simultaneous Equations Models for Discrete Outcomes: Coherence, Completeness, and Identification," CeMMAP working paper CWP21/12.

——— (2013a): "Intersection Bound Inference on Partially Identified Average Treatment Effects," in preparation.

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

Chesher, A., and A. M. Rosen (2015): "Identification of the Distribution of Valuations in an Incomplete Model of English Auctions," CeMMAP working paper CWP30/15.

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

Chetverikov, D. (2011): "Adaptive Test of Conditional Moment Inequalities," CeMMAP working paper CWP36/12.

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.

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

Hoderlein, S., H. Holzmann, and A. Meister (2015): "The Triangular Model with Random Coefficients," CeMMAP working paper CWP33/15.

Hong, H., and E. Tamer (2003): "Inference in Censored Models with Endogenous Regressors," *Econometrica*, 71(3), 905–932.

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.

Khan, S., and E. Tamer (2009): "Inference on Endogenously Censored Regression Models Using Conditional Moment Inequalities," *Journal of Econometrics*, 152(2), 104–119.

Kitagawa, T. (2009): "Identification Region of the Potential Outcome Distributions Under Instrument Independence," CeMMAP Working Paper CWP30/09.

Koopmans, T. (1949): "Identification Problems in Economic Model Construction," *Econometrica*, 17(125-144), 2.

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

Lee, S., K. Song, and Y.-J. Whang (2013a): "Testing for a General Class of Functional Inequalities," working paper, Seoul National University.

——— (2013b): "Testing Functional Inequalities," *Journal of Econometrics*, 172(1), 14–32.

Manski, C. F. (1997): "Monotone Treatment Response," *Econometrica*, 65(6), 1311–1334.

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

Masten, M. (2014): "Random Coefficients on Endogenous Variables in Simultaneous Equations Models," CeMMAP working paper CWP01/14.

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.

Menzel, K. (2009): "Estimation and Inference with Many Weak Moment Inequalities," working paper, MIT.

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

Newey, W. K., and J. L. Powell (1989): "Nonparametric Instrumental Variable Estimation," working paper, Princeton University.

——— (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.

Romano, J., and A. Shaikh (2008): "Inference for Identifiable Parameters in Partially Identified Econometric Models," *Journal of Planning and Statistical Inference*, 138, 2786–2807.

Romano, J., A. Shaikh, and M. Wolf (2013): "A practical two-step method for testing moment inequalities with an application to inference in partially identified models.," working paper, Stanford University.

Romano, J. P., and A. M. Shaikh (2010): "Inference for the Identified Set in Partially Identified Econometric Models," *Econometrica*, 78(1), 169–211.

Rosen, A. M. (2008): "Confidence Sets for Partially Identified Parameters that Satisfy a Finite Number of Moment Inequalities," *Journal of Econometrics*, 146, 107–117.

——— (2012): "Set Identification via Quantile Restrictions in Short Panels," *Journal of Econometrics*, 166(1), 127–137.

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

Sutherland, W. A. (2009): *Introduction to Metric and Topological Spaces*. Oxford Univerity Press, New York.

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

Theil, H. (1953): *Repeated Least Squares Applied to Complete Equation Systems*. Central Planning Bureau, The Hague.

Wright, P. (1928): *The Tariff on Animal and Vegetable Oils*. Macmillan, New York.

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

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

this is equivalent to

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

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*

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

By supposition we have

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

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

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

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 (4.2). Combining $C_h(\mathcal{U}_{\mathcal{S}}(h,z)|z) = C_h(\mathcal{S}|z)$ 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

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

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

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

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

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

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

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

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

Proof of Corollary 2. *Consider any* $\mathcal{S} \in \mathsf{Q}^E(h,z)$*. Then for all* $y \in \mathcal{Y}$*, either* $\mathcal{U}(y,z;h) \subseteq \mathcal{S}$ *or* $\mathcal{U}(y,z;h) \subseteq \overline{\mathcal{S}^c}$*. Thus*

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

The inequalities of Theorem 3 imply that

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

Then absolute continuity of $G_{U|Z}(\cdot|z)$ *implies that* $G_{U|Z}(\mathcal{S}|z)+G_{U|Z}\left(\overline{\mathcal{S}^c}|z\right) = 1$*, which taken with* (A.8) *implies that both inequalities hold with equality.* ■

Proof of Theorem 4. *Under Restriction SI,* $G_{U|Z}(\cdot|z) = G_U(\cdot)$ *a.e.* $z \in \mathcal{R}_Z$*. (5.3) and (5.1) follow from (3.1) and Theorem 2, respectively, upon substituting* $G_U(\cdot)$ *for* $G_{U|Z}(\cdot|z)$*. (5.2) follows from Corollary 2, again by replacing* $G_{U|Z}(\cdot|z)$ *with* $G_U(\cdot)$*. The equivalence of (5.1) and (5.4) with* $G_{U|Z}(\cdot|z) = G_U(\cdot)$ *holds by Artstein's inequality, see e.g. Molchanov (2005, pp. 69-70, Corollary 4.44).* ■

Proof of Corollary 3. *It follows from the main text that for any structure* $(h, G_U) \in \mathcal{M}^*$*,* $h \in \mathcal{H}^*$*. Now consider an arbitrary structural function* $h \in \mathcal{H}^*$ *as defined in the statement of the Corollary. Then*

$$\forall \mathcal{S} \in \mathsf{Q}^*(h), \sup_{z \in \mathcal{R}_Z} C_h(\mathcal{S}|z) \leq \inf_{z \in \mathcal{R}_Z} \left(1 - C_h(\mathcal{S}^c|z)\right).$$

By monotonicity of the containment functional in the argument $\mathcal{S}$*, and the fact that* $C_h(\mathcal{S}|z) \in [0,1]$*, it follows that there exists some probability distribution function* $G_U(\cdot)$ *such that:*

$$\forall \mathcal{S} \in \mathsf{Q}^*(h), \sup_{z \in \mathcal{R}_Z} C_h(\mathcal{S}|z) \leq G_U(\mathcal{S}) \leq \inf_{z \in \mathcal{R}_Z} \left(1 - C_h(\mathcal{S}^c|z)\right). \tag{A.9}$$

Since $\mathsf{Q}(h,z) \subseteq \mathsf{Q}^*(h)$ *a.e.* $z \in \mathcal{R}_Z$*, it follows that for almost every* $z \in \mathcal{R}_Z$*:*

$$\forall \mathcal{S} \in \mathsf{Q}(h,z), \quad C_h(\mathcal{S}|z) \leq G_U(\mathcal{S}).$$

By Theorem 3 this implies that $C_h(\mathcal{S}|z) \leq G_U(\mathcal{S})$ *holds for almost every* $z \in \mathcal{R}_Z$ *and for all sets* $\mathcal{S}$ *in* $\mathcal{R}_U$*, so that by Corollary 1,* $(h, G_U) \in \mathcal{M}^*$*, completing the proof.* ■

Proof of Theorem 5. *Restrictions A3 and A5 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, for some finite* $c \in \mathcal{C}$*,* $E[U|z] = c$ *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,* $c \in \mathbb{E}[\mathcal{U}(Y,Z;h)|z]$ *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* $E\left[\tilde{U}|z\right] = c$ *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 A5 and the continuity requirement of Restriction MI* guarantee that* $\mathcal{D}(Y,Z;h)$ *is integrable and closed. From Definition 6, for any* $c \in \mathcal{C}$, $c \in \mathbb{E}[\mathcal{D}(Y,Z;h)|z]$ *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* $E[D|z] = c$ *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* $E\left[d\left(\tilde{U},Z\right)|z\right] = d$ *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*

$$\mathcal{H}^{**}=\left\{h\in\mathcal{H}:\exists\mathcal{G}_{U|Z}\in\mathsf{G}_{U|Z}\ s.t.\ \forall\mathcal{S}\in\mathsf{F}\left(\mathcal{R}_U\right),\ C_h\left(\mathcal{S}|z\right)\leq G_{U|Z}\left(\mathcal{S}|z\right)\ \ a.e.\ \ z\in\mathcal{R}_Z\right\}. \quad \text{(A.10)}$$

Consider any $h\in\mathcal{H}^{**}$*. We wish to show first that* h *belongs to the set* $\mathcal{H}^{*}$ *given in (5.9). Fix* $z\in\mathcal{R}_Z$ *and choose* c *such that* $G_{U|Z}\left(\left(-\infty,c\right]|z\right)=\tau$*, which can be done by virtue of the continuity condition of Restriction QI. Then*

$$C_h\left(\left(-\infty,c\right]|z\right)\leq G_{U|Z}\left(\left(-\infty,c\right]|z\right)=\tau, \quad \text{(A.11)}$$

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

$$C_h\left(\left(-\infty,c\right]|z\right)=F_{Y|Z}\left[\overline{u}\left(Y,Z;h\right)\leq c|z\right]. \quad \text{(A.12)}$$

Now consider $\mathcal{S}=\left[c,\infty\right)$*. We have by monotonicity of the containment functional* $C_h\left(\cdot|z\right)$ *and from* $C_h\left(\mathcal{S}|z\right)\leq G_{U|Z}\left(\mathcal{S}|z\right)$ *in (A.10) that*

$$C_h\left(\left(c,\infty\right)|z\right)\leq C_h\left(\left[c,\infty\right)|z\right)\leq G_{U|Z}\left(\left[c,\infty\right)|z\right)=1-\tau, \quad \text{(A.13)}$$

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

$$C_h\left(\left(c,\infty\right)|z\right)=1-F_{Y|Z}\left[\underline{u}\left(Y,Z;h\right)\leq c|z\right]. \quad \text{(A.14)}$$

Combining this with (A.13) and also using (A.11) and (A.12) above gives

$$F_{Y|Z}\left[\overline{u}\left(Y,Z;h\right)\leq c|z\right]\leq\tau\leq F_{Y|Z}\left[\underline{u}\left(Y,Z;h\right)\leq c|z\right]. \quad \text{(A.15)}$$

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)* $G_{U|Z}\left(\cdot|z\right)$ *has* τ*-quantile equal to* c*, and (2)* $\forall\mathcal{S}\in\mathsf{F}\left(\mathcal{R}_U\right)$*,* $C_h\left(\mathcal{S}|z\right)\leq G_{U|Z}\left(\mathcal{S}|z\right)$*.*

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

$$G_{U|Z}\left(\mathcal{S}|z\right)=\lambda\left(z\right)C_h\left(\mathcal{S}|z\right)+\left(1-\lambda\left(z\right)\right)\left(1-C_h\left(\mathcal{S}^c|z\right)\right), \quad \text{(A.16)}$$

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

$$\lambda(z) F_{Y|Z}[\overline{u}(Y,Z;h) \leq c|z] + (1-\lambda(z)) F_{Y|Z}[\underline{u}(Y,Z;h) \leq c|z] = \tau. \tag{A.17}$$

The left hand side of equation (A.17) is precisely (A.16) with $\mathcal{S} = (-\infty, c]$. *Because* $h \in \mathcal{H}^*$, *(A.15) holds, which guarantees that* $\lambda(z) \in [0,1]$. *(A.17) and (A.16) deliver*

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

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* $z \in \mathcal{R}_Z$,

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

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

Supplement to "Characterizations of Identified Sets Delivered by Structural Econometric Models"

Andrew Chesher* UCL and CeMMAP

Adam M. Rosen† UCL and CeMMAP

October 3, 2015

Abstract

This supplement provides appendices not included in the main text. Appendix B shows that consideration of selectionability criteria conditional on $Z = z$ for almost every z on the support of Z, as in the main text, is equivalent to using selectionability criteria for the *joint* distribution of the random variable in question and Z. Appendix C further illustrates the application of the paper's results to a continuous outcome model with an interval-censored endogenous explanatory variable, providing numerical illustrations of the resulting identified sets.

*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.

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*

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

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,* $G_{U|Z}(\cdot|z)$ *is selectionable with respect to the conditional distribution of* $\mathcal{U}(Y, Z; h)$ *given* $Z = z$ *for almost every* $z \in \mathcal{R}_Z$ *if and only if* $G_{UZ}(\cdot)$ *is selectionable with respect to the distribution of* $\mathcal{U}(Y, Z; h) \times \{Z\}$*.* ■

C A Model with an Interval Censored Endogenous Variable

In this Section, Example 4 from Section 2.2 is studied in detail. This is a generalization of a single equation model with an interval censored exogenous variable studied in Manski and Tamer (2002).[1] As in Manski and Tamer (2002), there is no restriction on the censoring process, but there are three main differences in the models we consider. First, the models here allow the interval censored explanatory variable as well as the endpoints of the censoring interval to be endogenous. Second, there is no analog of Manski and Tamer's (2002) Assumption MI, which stipulates that the conditional mean of the outcome variable given the censored variable and its observed interval

[1] Models allowing censored *outcome* variables with *uncensored* endogenous explanatory variables with sufficient conditions for point identification include those of Hong and Tamer (2003) and Khan and Tamer (2009).

endpoints does not vary with the values of the interval endpoints.[2] Third, Manski and Tamer (2002) focused exclusively on conditional mean restrictions when considering continuous outcomes. Here we consider the identifying power of conditional quantile and stochastic independence restriction in addition to conditional mean restrictions on the conditional distribution of unobservable U given exogenous variables Z. We provide numerical illustrations of identified sets delivered by particular data generating structures for several cases. Application of our results ensures sharpness of the the resulting bound characterizations directly, under all the different sets of restrictions considered, without need for constructive proofs of sharpness.

C.1 Identified Sets

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

$$Y_1 = g(Y_2^*, Z_1, U), \tag{C.1}$$

where the function $g(\cdot,\cdot,\cdot)$ is increasing in its first argument, and strictly increasing in its third argument.[3] The endogenous variable Y_2^* is not observed, but there are observed variables Y_{2l}, Y_{2u} such that

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

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

Since there is no restriction on the censoring process, it is convenient to suppress the unobserved variable W by replacing (C.2) with the equivalent formulation

$$\mathbb{P}[Y_{2l} \leq Y_2^* \leq Y_{2u}] = 1. \tag{C.3}$$

The researcher observes realizations of (Y_1, Y_{2l}, Y_{2u}, Z) under conditions which identify their joint distribution.

[2] This assumption, through use of the law of iterated expectations, plays a role in their analysis by establishing a direct relation between the expected value of outcome Y given the observed interval endpoints, and its expected value conditional on the unobserved censored variable.

[3] It is important to note here that $\Lambda(\cdot)$ is the *marginal* distribution of U. At this point no restrictions have been placed on the joint distribution of (U, Z), so that for any $z \in \mathcal{R}_U$, the conditional distribution of $U|Z = z$ need not be $\Lambda(\cdot)$. It is straightforward to allow $g(y_2^*, z_1, u)$ monotone increasing or decreasing in y_2^* for all (z_1, u). Indeed, it is also possible for the model specification to allow some functions g that are monotone increasing and others that are monotone decreasing in y_2^* for all (z_1, u), but the restriction that $g(y_2^*, z_1, u)$ is monotone increasing in y_2^* is maintained here to simplify the exposition.

The structural function

$$h(y, z, u) = \max\{g(y_{2l}, z_1, u) - y_1, 0\} + \max\{y_1 - g(y_{2u}, z_1, u), 0\}$$

and $\mathbb{P}[h(Y, Z, U) = 0] = 1$ is equivalent to equations (C.1) and (C.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

$$\mathcal{U}(y, z; h) = \left[g^{-1}(y_{2u}, z_1, y_1), g^{-1}(y_{2l}, z_1, y_1)\right], \tag{C.4}$$

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

In most of the following development and in the numerical illustrations the function h is further restricted by imposing a linear index structure on $g(y_2^*, z, u)$, thus:

$$g(y_2^*, z, u) = \beta y_2^* + z_1\gamma + u, \tag{C.5}$$

with the first element of z_1 normalized to one. With this restriction in place the functions g and h are parameterized by $(\beta, \gamma') \in \mathbb{R}^{\dim(z_1)+1}$.

The identified sets delivered by this model are now obtained under alternative restrictions on the collection of conditional distributions $\mathcal{G}_{U|Z}$. In each case it is shown how the identified sets can be characterized as a system of conditional moment inequalities that can be used as a basis for estimation and inference.

C.1.1 Stochastic Independence

Consider the restriction $U \perp\!\!\!\perp Z$. Using Theorem 3 the identified set is characterized as the admissible structures (h, Λ) such that the inequalities

$$\mathbb{P}[\mathcal{U}(Y, Z; h) \subseteq \mathcal{S}|z] \leq G_U(\mathcal{S}) \tag{C.6}$$

hold for all sets $\mathcal{S} \in \mathsf{Q}(h, z)$, where $\mathsf{Q}(h, z)$ is the collection of intervals that can be formed as unions of sets of the form $\left[g^{-1}(y_{2u}, z_1, y_1), g^{-1}(y_{2l}, z_1, y_1)\right]$. Here $G_U(\mathcal{S})$ is the probability mass placed on a set $\mathcal{S}$ by the distribution of U which has cumulative distribution function $\Lambda(\cdot)$.

If the components of Y are continuously distributed with rich support the collection of required test sets may comprise *all* intervals on $\mathbb{R}$.[4] Unless g has very restricted structure, the conditions for

[4]If the support of Y_1 is limited, application of Theorem 3 may indicate that not all intervals on $\mathbb{R}$ need be considered as test sets. This smaller collection of 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, and that using all closed intervals on $\mathbb{R}$

(C.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 is henceforth denoted

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

It follows from Theorem 4 that the identified set of structures (h,Λ) admitted by a model $\mathcal{M}$ embodying the restrictions so far introduced is as follows.

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

The containment functional inequality in this characterization can be written

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

equivalently, using monotonicity of $g(y_2, z_1, u)$ in its third argument, as follows.

$$\mathbb{P}[g(Y_{2u}, Z_1, u_*) \leq Y_1 \leq g(Y_{2l}, Z_1, u^*)\,|z] \leq \Lambda(u^*) - \Lambda(u_*) \tag{C.7}$$

With the added linear index restriction from (C.5) this produces the following representation for the identified set, where the model $\mathcal{M}$ defines a collection of admissible parameters β, γ and distribution functions $\Lambda(\cdot)$.

$$\mathcal{M}^* = \left\{ \begin{array}{l} (\beta,\gamma,\Lambda) \in \mathcal{M} : \forall [u_*, u^*] \in \mathsf{Q}, \\ \quad \mathbb{P}[u_* + \beta Y_{2u} \leq Y_1 - Z_1\gamma \leq u^* + \beta Y_{2l}|z] \leq \Lambda(u^*) - \Lambda(u_*),\ \text{a.e. } z \in \mathcal{R}_Z \end{array} \right\} \tag{C.8}$$

In some of the calculations reported later $\Lambda(\cdot)$ is restricted to be the distribution function of a $N(0,\sigma^2)$ random variable. In this case $\Lambda(u^*) - \Lambda(u_*)$ in (C.8) is replaced by $\Phi\left(\sigma^{-1}u^*\right) - \Phi\left(\sigma^{-1}u_*\right)$ and M is the admissible parameter space for (β, γ', σ). Using (C.8) the identified set is then as follows.

$$\mathcal{M}^* = \left\{ \begin{array}{l} \theta \in \Theta : \forall [u_*, u^*] \in \mathsf{Q}, \\ \quad \mathbb{P}[u_* + \beta Y_{2u} \leq Y_1 - Z_1\gamma \leq u^* + \beta Y_{2l}|z] \leq \Phi\left(\sigma^{-1}u^*\right) - \Phi\left(\sigma^{-1}u_*\right),\ \text{a.e. } z \in \mathcal{R}_Z \end{array} \right\} \tag{C.9}$$

Equivalently, the change of variables $t^* = \Phi\left(\sigma^{-1}u^*\right)$ and $t_* \equiv \Phi\left(\sigma^{-1}u_*\right)$ can be employed to

- define the same identified set and deliver sharp bounds.

produce the following representation.

$$\mathcal{M}^* = \left\{ \begin{array}{l} \theta \in \Theta : \forall [t_*, t^*] \subseteq [0,1], \\ \quad \mathbb{P}\left[t_* \leq \Phi\left(\frac{Y_1 - \beta Y_{2u} - Z_1\gamma}{\sigma}\right) \wedge \Phi\left(\frac{Y_1 - \beta Y_{2l} - Z_1\gamma}{\sigma}\right) \leq t^* | z \right] \leq t^* - t_*, \text{ a.e. } z \in \mathcal{R}_Z \end{array} \right\} \quad \text{(C.10)}$$

Under the stochastic independence restriction, $U \perp\!\!\!\perp Z$, and with no further restriction on the distribution function of U, applying Corollary 3 gives the identified set for the parameters $\theta \equiv (\beta, \gamma')$ alone. This involves the containment and capacity functionals of the random set $\mathcal{U}(Y, Z; h)$. For the case considered here the containment functional for sets $\mathcal{S} = [u_*, u^*]$ is

$$C_h(\mathcal{S}|z) = \mathbb{P}[u_* + \beta Y_{2u} \leq Y_1 - Z_1\gamma \leq u^* + \beta Y_{2l}|z]$$

and the capacity functional is

$$1 - C_h(\mathcal{S}^c|z) = \mathbb{P}[u_* + \beta Y_{2l} \leq Y_1 - Z_1\gamma \leq u^* + \beta Y_{2u}|z]$$

where the structural function h is characterized by the parameter vector θ.

Define

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

and let Θ denote parameter values admitted by model $\mathcal{M}$. Applying Corollary 3, the identified set for parameters θ is as follows.

$$\Theta^* = \left\{ \theta \in \Theta : \forall [u_*, u^*] \in \mathsf{Q}, \ \underline{G}(\theta, u_*, u^*) \leq \overline{G}(\theta, u_*, u^*) \right\} \quad \text{(C.11)}$$

Equivalent to (C.11), the identified set for θ are those $\theta \in \Theta$ satisfying the moment inequality representation:

$$\begin{aligned} E[m_1(\theta; Y, Z, u_*, u^*)|z] - E\left[m_2(\theta; Y, Z, u_*, u^*)|z'\right] &\leq 0, \\ \text{all } u_*, u^* &\in \mathbb{R} \text{ s.t. } u_* \leq u^*, \text{ a.e. } z, z' \in \mathcal{R}_Z \times \mathcal{R}_Z, \end{aligned}$$

where

$$\begin{aligned} m_1(\theta; Y, Z, u_*, u^*) &\equiv 1[u_* + \beta Y_{2u} \leq Y_1 - Z_1\gamma \leq u^* + \beta Y_{2l}], \\ m_2(\theta; Y, Z, u_*, u^*) &\equiv 1[u_* + \beta Y_{2l} \leq Y_1 - Z_1\gamma \leq u^* + \beta Y_{2u}]. \end{aligned}$$

C.1.2 Mean Independence

Now consider the case in which the linear index restriction (C.5) is retained and the stochastic independence restriction $U \perp\!\!\!\perp Z$ is replaced by the mean independence restriction $E[U|Z=z]=0$ a.e. $z \in \mathcal{R}_Z$. This is Restriction MI of Section 5 with $\mathcal{C}=\{0\}$.

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

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

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

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

where

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

C.1.3 Quantile Independence

Consider the case in which the linear index restriction (C.5) is coupled with the restriction $q_{U|Z}(\tau|z) = 0$, a.e $z \in \mathcal{R}_Z$ for some specified value of τ. This is Restriction QI of Section 5.3 with $\mathcal{C}=\{0\}$.

As before under the linear index restriction (C.5) there is

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

and the structural function h is determined by $\theta \equiv (\beta, \gamma')$. As in Section C.1.2 the parameter space and identified set for θ are denoted by Θ and Θ^*, respectively. Applying Theorem 6 the identified set of values of θ is

$$\Theta^* = \left\{\theta \in \Theta : \sup_{z\in\mathcal{R}_Z} F_{Y|Z}[Y_1 \le Z_1\gamma + \beta Y_{2l}|z] \le \tau \le \inf_{z\in\mathcal{R}_Z} F_{Y|Z}[Y_1 \le Z_1\gamma + \beta Y_{2u}|z]\right\}, \quad \text{(C.12)}$$

equivalently,

$$\Theta^* = \left\{ \theta \in \Theta : \sup_{z \in \mathcal{R}_Z} \left(q_{\underline{V_\theta}|Z} (\tau|z) - z_1\gamma \right) \leq 0 \leq \inf_{z \in \mathcal{R}_Z} \left(q_{\overline{V_\theta}|Z} (\tau|z) - z_1\gamma \right) \right\},$$

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. $z \in \mathcal{R}_Z$.

Using (C.12) the identified set Θ^* can be characterized via the moment inequalities

$$\begin{aligned} E[m_1(\theta; Y, Z)|z] &\leq 0, \text{ a.e. } z \in \mathcal{R}_Z \\ E[m_2(\theta; Y, Z)|z] &\leq 0, \text{ a.e. } z \in \mathcal{R}_Z \end{aligned}$$

where

$$\begin{aligned} m_1(\theta; Y, Z) &\equiv 1[Y_1 \leq Z_1\gamma + \beta Y_{2l}] - \tau \\ m_2(\theta; Y, Z) &\equiv \tau - 1[Y_1 \leq Z_1\gamma + \beta Y_{2u}]. \end{aligned}$$

C.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 (C.5). We first 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 (C.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 so $Z = Z_2$. The binary support of Z simplifies the calculations. Richer support would provide greater identifying power, equivalently smaller

	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 1: Endpoints of censoring process intervals in DGP1 and DGP2.

identified sets.

We specify a censoring process that reveals to which one of a collection of mutually exclusive intervals a realization of Y_2^* belongs. Such censoring processes are common in practice, for instance when interval bands are used to record income in surveys. We specify 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

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

The researcher observes realizations of (Y_1, Y_{2l}, Y_{2u}, Z) through a process such that their joint distribution is identified.

In the first set of examples we work with probability distributions generated by two structures denoted DGP1 and DGP2, both with parameter values

$$\gamma_0 = 0, \quad \gamma_1 = 1, \quad \delta_0 = 0, \quad \delta_1 = 1, \quad \sigma_{11} = 0.5, \quad \sigma_{1v} = 0.25, \quad \sigma_{vv} = 0.5, \tag{C.13}$$

and interval censoring endpoints $c_1, ..., c_J$ listed in Table 1. In DGP1, Y_2^* is 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^* is censored into 12 such intervals with endpoints given by the normal quantile function evaluated at 13 equally spaced values.

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 (C.9), the conditional containment functional for random set $\mathcal{U}(Y, Z; h)$ applied to test set $\mathcal{S} = [u_*, u^*]$ is given by

$$C_\theta([u_*, u^*] | z) = \sum_j \mathbb{P}[g_1 c_{j+1} + u_* \leq Y_1 - g_0 \leq g_1 c_j + u^* | z, [Y_{2l}, Y_{2u}) = I_j] * \mathbb{P}[[Y_{2l}, Y_{2u}) = I_j | z], \tag{C.14}$$

where $\theta = (g_0, g_1, s)$ is used to denote generic parameter values for $(\gamma_0, \gamma_1, \sigma_{11})$. C_θ replaces C_h for the containment functional, since in this model the structural function h is a known function of θ.[5] The identified set of structures $\left(h, \mathcal{G}_{U|Z}\right)$ is completely determined by the identified set for θ,

[5]Computational details for the conditional containment probability $C_\theta([u_*, u^*] | z)$ are provided in Appendix C.3.

which, following (C.10), is given by

$$\Theta^* = \left\{ \begin{array}{l} \theta \in \Theta : \forall \left[t_*, t^*\right] \subseteq [0,1], \\ C_\theta\left(\left[s\Phi^{-1}\left(t_*\right), s\Phi^{-1}\left(t^*\right)\right] |z\right) \leq t^* - t_*, \text{ a.e. } z \in \mathcal{R}_Z \end{array} \right\}. \tag{C.15}$$

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 super-diagonal elements of the following $(M+1) \times (M+1)$ array of intervals with the interval $[0,1]$ excluded. Here $m \equiv 1/M$ and there are $M(M+1)/2 - 1$ intervals in Q_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 (C.15) applied to the 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.

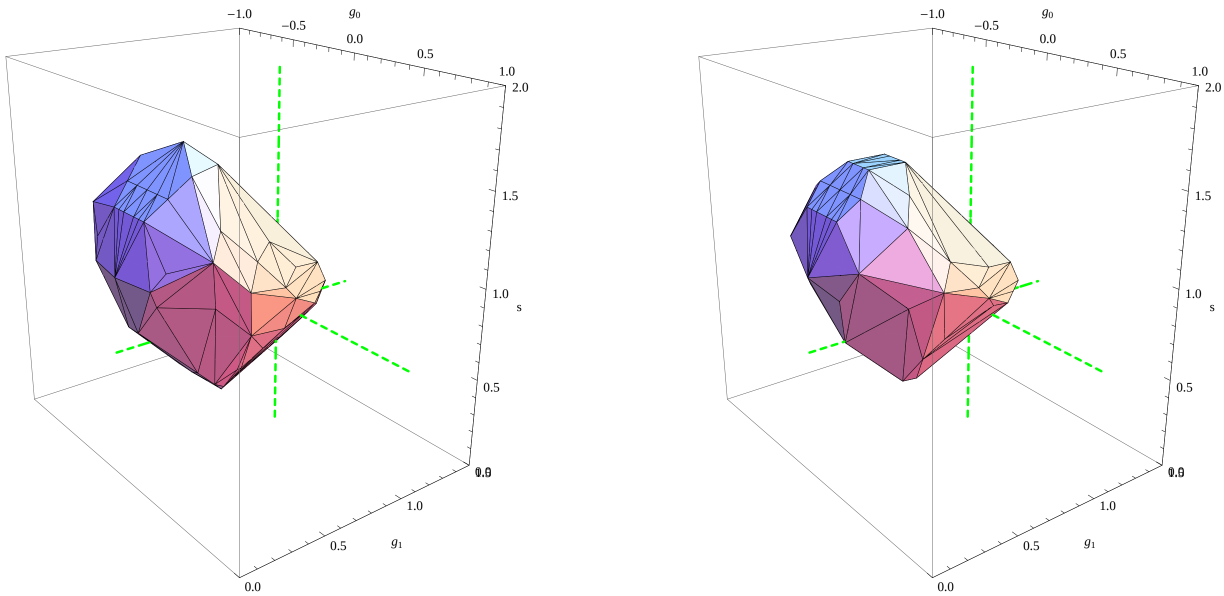
Figure 1 shows perspective plots of three dimensional outer regions for (g_0, g_1, s). Outer regions using $M \in \{5,7,9\}$ are noticeably smaller than those using only $M = 5$.[6] There was a noticeable reduction in the size of the outer region in moving from $M = 5$ to $M = \{5,7\}$, but hardly any 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.[7] 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 suggests that in the cases considered the sets are convex.

Figure 3 shows perspective plots of outer regions for DGP2 employing 12 bins for the censoring

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

[7]Perspective plots were produced using the `TetGenConvexHull` function available *via* the `TetGenLink` package in `Mathematica 9`, Wolfram Research, Inc. (2012). The projections 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).



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 further illustrate the extent of the reduction in the size of the outer region for DGP2 relative to DGP1.

The second set of numerical illustrations employs the same triangular Gaussian error structure for our DGPs with parameter values as specified in (C.13). However, we consider two alternative censoring processes, where Y_2^* is again observed only to lie in one of a fixed set of bins, but where now these bins are set to be of a fixed width. We consider fixed bins, first with width 0.4:

$$\ldots\ldots,\ (-0.8, -0.4],\ (-0.4, 0.0],\ (0, 0.4],\ (0.4, 0.8],\ \ldots\ldots$$

and then of width 0.2:

$$\ldots\ldots,\ (-0.4, -0.2],\ (-0.2, 0.0],\ (0, 0.2],\ (0.2, 0.4],\ \ldots\ldots$$

With this censoring structure in place, we now compare the identifying power of alternative

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.

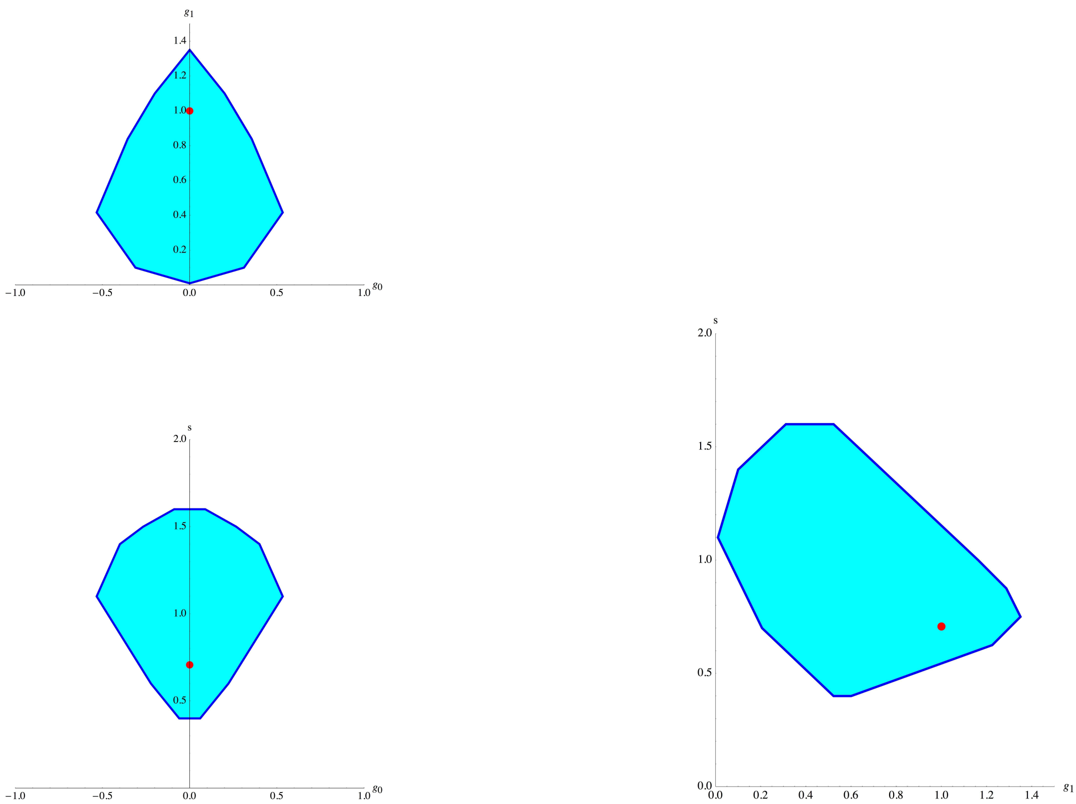


Figure 3: Outer region for DGP2 with 12 bins calculated using inequalities generated with $M \in \{5, 7, 9\}$. Dashed green lines intersect at the data generating value of the parameters.

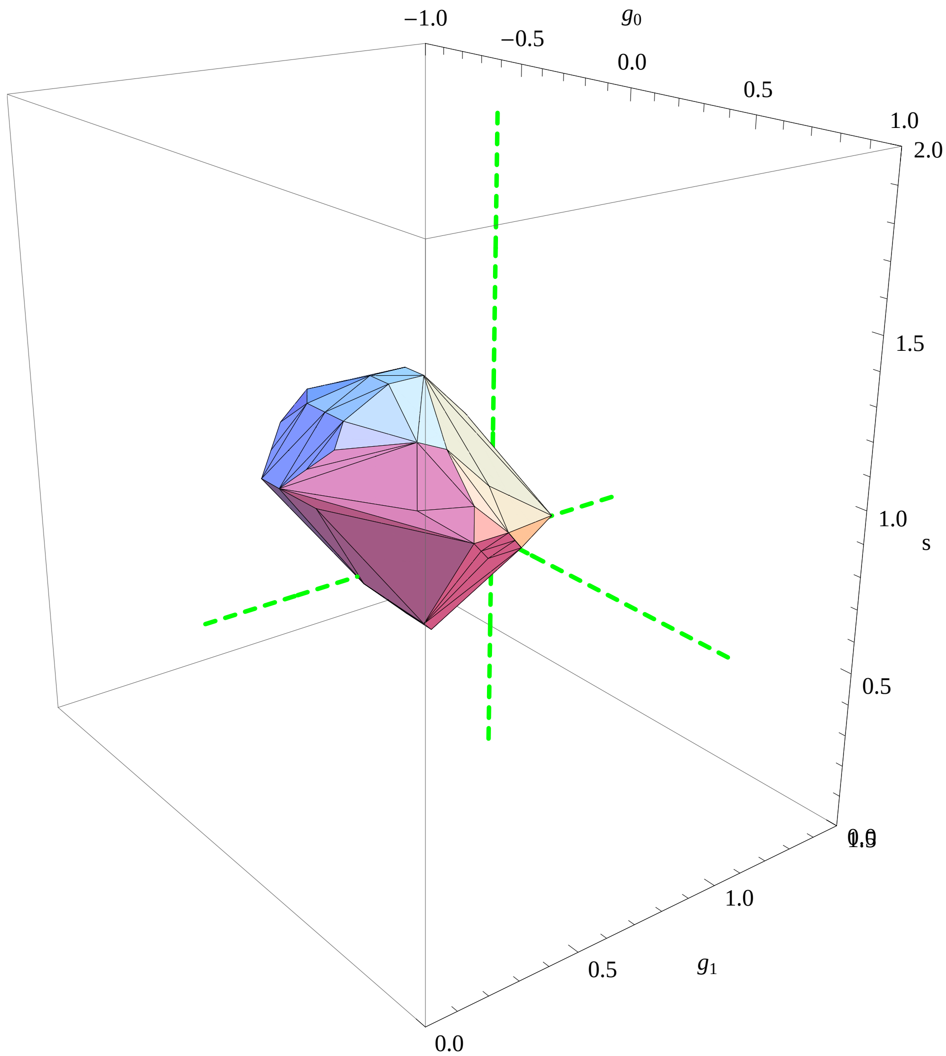
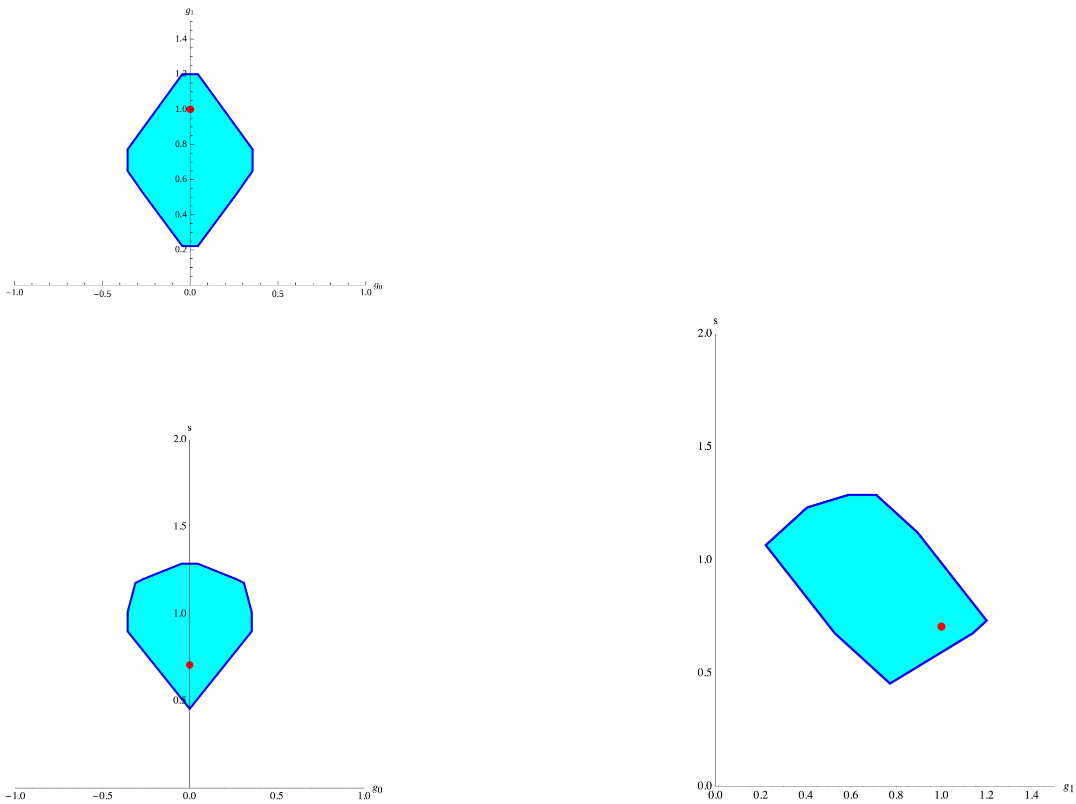


Figure 4: Outer region projections for DGP2 onto the (g_0, g_1), (g_0, s), and (g_1, s) planes, respectively, with 12 bins using inequalities generated with $M \in \{5, 7, 9\}$. The red point marks the data generating value.



restrictions on unobserved heterogeneity, in both cases imposing the linear functional form

$$Y_1 = \gamma_0 + \gamma_1 Y_2^* + U.$$

We again consider the parametric Gaussian restriction on unobserved heterogeneity that $U \sim N(0, \sigma)$ and $U \perp\!\!\!\perp Z$, and compare to the alternative restriction imposing only knowledge that $q_{U|Z}(0.5|z) = 0$, a.e. $z \in \mathcal{R}_Z$. This semiparametric specification has no scale parameter s, so we focus attention on the implied identified set for (γ_0, γ_1).

Figure 5 below illustrates the identified sets obtained for bin widths 0.4 (top panels) and 0.2 (bottom panels), as well as for $\delta_1 = 1$ (left panels) and $\delta_1 = 1.5$ (right panels). In the triangular structure employed to generate the actual distributions $\mathcal{F}_{Y|Z}$ the parameter δ_1 is the coefficient multiplying the instrument Z in the equation determining the value of the censored endogenous variable Y_2^*. With a higher value of δ_1 the value of this variable as well as the censoring points is more sensitive with respect to the instrument Z. As we might expect, identified sets when $\delta_1 = 1.5$ are smaller than those for the case $\delta_1 = 1$, as are sets obtained when the bin width is only 0.2 rather than 0.4.

Identified sets obtained from a model imposing independent Gaussian unobservable U (in light blue) are naturally contained in those obtained from a model only imposing the less restrictive zero conditional quantile restriction (in dark blue). However, the difference between the identified sets obtained under these different restrictions is not so great, at least under the particular data generating structures employed. In these cases, the use of the weaker conditional quantile restriction does not seem to lose much in the way of identifying power relative to the Gaussian distributional restriction.

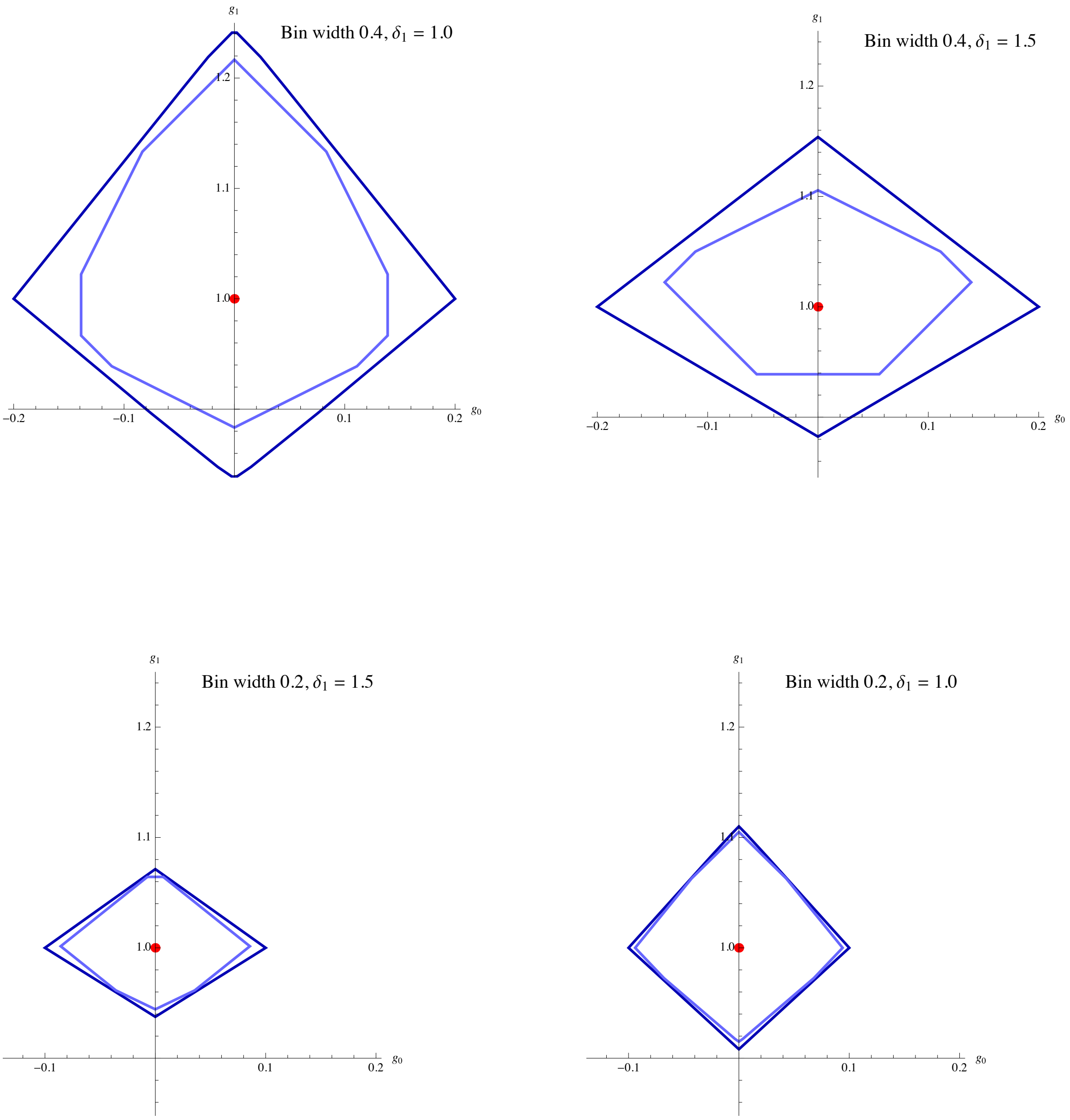
A partial explanation for this observation may be the fixed-width binning setup. Indeed, under this censoring process with the given triangular data-generating structure, it can be shown that under a distribution-free independence restriction - that is $U \perp\!\!\!\perp Z$ with the distribution of U otherwise unrestricted except for a zero median location normalization - the identified set is identical to that obtained under the conditional median restriction alone. This is not generally the case. For other censoring processes (not reported here) the identified set under the distribution-free independence restriction is a strict subset of that obtained under only the conditional quantile restriction.

C.3 Computational Details for Numerical Illustrations

In this Section we describe computation of the conditional containment functional $C_\theta([u_*, u^*]|z)$ in (C.14). Computations were carried out in `Mathematica 9`.

Given the structure specified for DGP1 and DGP2 in Section C.2, the conditional distribution

Figure 5: Identified sets for (γ_0, γ_1). The top panels display sets for bins of width 0.4 and the bottom panels display sets for bins of width 0.2. In the panels on the left $\delta_1 = 1$ and on the right $\delta_1 = 1.5$. The dark blue lines indicate boundaries of identified sets obtained with the conditional quantile restriction $q_{U|Z}(\tau|z) = 0$, while the inner light blue lines indicate boundaries of identified sets when U is restricted to be Gaussian, independent of Z.



of Y_2^* given $Y_1 = y_1$ and $Z = z$ is for any (y_1, z)

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

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

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

The distribution of Y_1 given $Z = z$ is

$$Y_1 | Z = z \sim \mathcal{N}\left(\gamma_0 + \gamma_1 a(z), \sigma_{11} + 2\gamma_1 \sigma_{1v} + \gamma_1^2 \sigma_{vv}\right). \tag{C.16}$$

The conditional containment functional can thus be written

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

where $f_{Y_1|Z}(\cdot|z)$ is the normal probability density function with mean and variance given in (C.16).

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

$$C_\theta\left([u_*, u^*] | z\right) \equiv$$

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

References

Hong, H., and E. Tamer (2003): "Inference in Censored Models with Endogenous Regressors," *Econometrica*, 71(3), 905–932.

Khan, S., and E. Tamer (2009): "Inference on Endogenously Censored Regression Models Using Conditional Moment Inequalities," *Journal of Econometrics*, 152(2), 104–119.

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

Wolfram Research, Inc. (2012): *Mathematica, Version 9.0.* Champaign, IL.