Set identifying models with discrete outcomes and endogenous variables

ANDREW CHESHER AND KONRAD SMOLINSKI*
CEMMAP & UCL

February 5th 2010

ABSTRACT. Nonparametric IV models for discrete outcomes with discrete endogenous variables set identify structural functions. The sets are the union of a number of convex sets, each of which is a projection onto the space in which a structural function resides of a convex set that resides in a much higher dimensional space. Calculation of these projections is infeasible unless the support of the outcomes and endogenous variables is quite limited. The paper develops a relatively easy-to-compute set which bounds the identified set in the sense that the latter is a subset of the former, derives some properties of the new set and suggests how the set can be used in practical econometric analysis when outcomes and endogenous variavbles are discrete.

KEYWORDS: Discrete outcomes, Discrete endogenous variables, Endogeneity, Incomplete models, Instrumental variables, Set Identification, Threshold Crossing Models.

JEL CODES: C10, C14, C50, C51.

1. INTRODUCTION

This paper considers single equation instrumental variable (SEIV) models in which both the outcome of interest and potentially endogenous explanatory variables are discrete. These models generally set rather than point identify structural features of central interest.[1] The paper presents new results on properties of the identified sets when there are no parametric restrictions.

The fully discrete case studied here arises frequently in applied econometrics practice and the case is of technical interest too. The identified sets in this case are unions of convex sets each of which is an intersection of linear half-spaces. Characterization and computation of the identified set is challenging and presents interesting problems which are similar to those arising in the study of large scale linear programmes.

In this paper we derive an outer set which we show contains the identified set. We show that when neither the outcome nor the endogenous variables are binary the new set is a subset of outer sets previously obtained and we present an example in which it is a proper subset.

Examples of settings in which the results of the paper are useful include situations in which a probit, or a logit or a count data analysis or some semiparametric or nonparametric alternative would be considered and explanatory variables are endogenous. The analysis of this paper deals with nonparametric models but characterizations of identified sets for nonparametric models are very helpful in the construction

*We gratefully acknowledge the financial support of the UK Economic and Social Research Council through a grant (RES-589-28-0001) to the ESRC Centre for Microdata Methods and Practice (CeMMAP).

[1] See Chesher (2008).

of identified sets in parametric cases as has been shown in Chesher and Smolinski (2009).

The commonly adopted, relatively restrictive, control function approach to identification in problems involving discrete outcomes and endogeneity[2] fails to deliver point identification when endogenous variables are discrete so the SEIV models studied here are leading contenders for application in practice.

The restrictions of the SEIV model are now set out and then the results given here are set in the context of earlier work.

1.1. The single equation instrumental variable model. In the SEIV model a scalar discrete outcome, Y, is determined by a structural function h as follows.

$$Y = h(X, U) \tag{1}$$

Here U is a scalar unobservable continuously distributed random variable and X is a list of explanatory variables. These explanatory variables may be endogenous in the sense that U and X may not be independently distributed. We focus on the identification of the structural function h.

In practice there may be exogenous variables appearing in h and the results of the paper are easily extended to accommodate these but for simplicity we proceed with the structural function specified as in equation (1).

The structural function h is restricted to be monotone in U for all values of X. It is normalized weakly increasing in what follows and the marginal distribution of U is normalized uniform on the unit interval. The support of X is denoted by $\mathcal{X}$.

The discrete outcome Y has M fixed points of support and without loss of generality these are taken to be the integers $1, \ldots, M$. There is the following threshold crossing representation of the structural function: for $m \in \{1, \ldots, M\}$:

$$h(x, u) = m \text{ if and only if } h_{m-1}(x) < u \leq h_m(x)$$

with $h_0(x) = 0$ and $h_M(x) = 1$ for all $x \in \mathcal{X}$. The value of $h(x, 0)$ is defined to be 1 for all x.[3]

In this set-up a conventional parametric probit model for $Y \in \{1, 2\}$ would have threshold functions as follows:

$$h_0(x) = 0 \qquad h_1(x) = \Phi\left(\alpha_0 + \alpha_1 x\right) \qquad h_2(x) = 1$$

and a conventional logit model would have $h_1(x) = (1 + \exp(\alpha_0 + \alpha_1 x))^{-1}$. Here Φ denotes the standard normal distribution function.

If X was exogenous then the threshold functions would be identified because in that case $\Pr[Y \leq m | X = x] = h_m(x)$. The SEIV model does not require X to be exogenous but admits instrumental variables, one or many, discrete or continuous, arranged in a vector Z which takes values in a set $\mathcal{Z}$. U and Z are independently distributed and Z is excluded from the structural function. The model set identifies the structural function. In this paper we study the case in which X is discrete.

[2] See for example Blundell and Powell (2003, 2004), Chesher (2003), Imbens and Newey (2009).

[3] So we are requiring $h(x, u) = 1$ if $0 \leq u \leq h_1(x)$ with a *weak* left hand inequality.

1.2. Relation to earlier work. The SEIV model studied here is an example of the sort of nonseparable model studied in Chernozhukov and Hansen (2005), Chesher (2003) and Imbens and Newey (2009).

The last two of these papers study complete models which specify structural equations for endogenous explanatory variables as well as for the outcome of interest. Both of those papers study triangular equations systems. When endogenous variables are continuous these models can point identify structural functions but when endogenous variables are discrete they do not. Dealing with the discrete endogenous variable case, Chesher (2005) introduces an additional restriction on the nature of the dependence amongst unobservables providing a set identifying triangular model with discrete endogenous variables. Jun, Pinkse and Xu (2009) deliver some refinements.

Discreteness of endogenous variables is not a problem for SEIV models, indeed it brings some simplifications - for example eliminating the "ill posed inverse problem" which arises when endogenous variables are continuous. This is shown clearly in Das (2005) where *additive error* nonparametric models with discrete endogenous variables and instrumental variable restrictions are considered. However because of the additive error restrictions this construction is not well suited to modelling discrete outcomes which sit more comfortably in the nonseparable setting studied here.

Chernozhukov and Hansen (2005) study a nonadditive-error SEIV model like that considered here, focussing on the case in which the outcome is *continuous*. The identification results of that paper are built around the following equality which holds, when Y is continuous, for all $\tau \in (0,1)$ and all $z \in \mathcal{Z}$.

$$\Pr[Y = h(X,\tau)|Z = z] = \tau \tag{2}$$

Additional (completeness) conditions are provided under which the model point identifies the structural function.

The condition (2) does not hold when Y is discrete. Instead, as shown in Chernozhukov and Hansen (2001), there are the following inequalities which hold for all $\tau \in (0,1)$ and $z \in Z$.

$$\Pr[Y < h(X,\tau)|Z = z] < \tau \leq \Pr[Y \leq h(X,\tau)|Z = z]$$

These imply that the inequalities:

$$\max_{z\in\mathcal{Z}} \Pr[Y < h(X,\tau)|Z = z] < \tau \leq \min_{z\in\mathcal{Z}} \Pr[Y \leq h(X,\tau)|Z = z] \tag{3}$$

hold for all $\tau \in (0,1)$. The result is that the SEIV model generally fails to point identify the structural function when the *outcome* Y is discrete. However the model is informative about the structural function as long as $\mathcal{Z}$ is not a singleton.

To see this suppose that for some value m and two values in $\mathcal{Z}$, z_1 and z_2, $\Pr[Y \leq m|Z = z_1] \neq \Pr[Y \leq m|Z = z_2]$. The restrictions of the model imply that in this case $h_m(x)$ is *not constant* for variations in x in admissible structures which generate the probability distribution under consideration. This is so because if $h_m(x)$ were constant, equal say to h_m^*, then for all $z \in \mathcal{Z}$, $\Pr[Y \leq m|Z = z] = h_m^*$

so *any* variation in $\Pr[Y \leq m|Z = z]$ with z rules out the possibility that $h_m(x)$ is constant for variations in x.[4]

The set identifying power of the SEIV model when the outcome is discrete is studied in Chesher (2008). Let $\mathcal{H}(\mathcal{Z})$ denote the identified set of structural functions associated with some probability distribution $F_{YX|Z}$ for Y and X given $Z = z \in \mathcal{Z}$.[5] Chesher (2008) develops a set, denoted here by $\mathcal{C}(\mathcal{Z})$, using the inequalities (3). That paper shows this to be, in general, an outer set in the sense that $\mathcal{H}(\mathcal{Z}) \subseteq \mathcal{C}(\mathcal{Z})$. It is shown that, when Y is binary and X is continuous, $\mathcal{H}(\mathcal{Z}) = \mathcal{C}(\mathcal{Z})$ and $\mathcal{C}(\mathcal{Z})$ provides tight set identification.

Chesher (2009) studies the binary outcome case, proving the tightness of $\mathcal{C}(\mathcal{Z})$ when endogenous variables are discrete, considering the impact of parametric restrictions and shape restrictions, and giving some results on estimation under shape restrictions using results on inference using intersection bounds given in Chernozhukov, Lee and Rosen (2009).

In Chesher and Smolinski (2009) a refinement[6] to $\mathcal{C}(\mathcal{Z})$, denoted $\tilde{\mathcal{D}}(\mathcal{Z})$, is developed. This delivers the identified set when there is a single binary endogenous variable no matter how many points of support the outcome Y has. The results are used in an investigation of the nature of the reduction in extent of the identified set as the number of points of support of Y increases in an endogenous parametric ordered probit example.

This paper develops a further refinement[7] to $\mathcal{C}(\mathcal{Z})$, denoted $\mathcal{E}(\mathcal{Z})$ and shows that $\mathcal{H}(\mathcal{Z}) \subseteq \mathcal{E}(\mathcal{Z})$. It is not clear at this point whether $\mathcal{E}(\mathcal{Z})$ is equal to $\mathcal{H}(\mathcal{Z})$. This is currently under investigation.

1.3. Plan of the paper. Section 2 defines the set identified by the SEIV models and reviews its characteristics.

Section 3 develops the new set, $\mathcal{E}(\mathcal{Z})$, shows that it contains the identified set, and in a series of sub-sections derives other properties of the set.

Section 3.1 sets out properties of the set $\mathcal{E}(\mathcal{Z})$ when the outcome is binary. Section 3.2 shows how the set $\mathcal{E}(\mathcal{Z})$ is related to the set defined in Chesher (2008). Section 3.3 gives alternative expressions for the inequalities defining the new set which help clarify its relationship to the set defined by the inequalities (3). Section 3.4 gives some illustrative calculations.

Section 4 concludes.

---

[4]

$$\Pr[Y \leq m|Z = z] = \sum_k \Pr[U \leq h^*_m|X = x_k, Z = z] \Pr[X = x_k|Z = z] = \Pr[U \leq h^*_m|Z = z] = h^*_m$$

[5] It would be clearer to give a distinctive symbol to the probability distribution under consideration, e.g. $F^0_{YX|Z}$ and label the various sets accordingly thus: $\mathcal{H}^0(\mathcal{Z})$, $\mathcal{C}^0(\mathcal{Z})$ and so forth. We do not do this here because the notation quickly becomes cumbersome. However it is important to keep in mind that each of the sets under discussion is associated with a particular probability distribution.

[6] By a refinement we mean that $\tilde{\mathcal{D}}(\mathcal{Z}) \subseteq \mathcal{C}(\mathcal{Z})$.

[7] Here, by a refinement we mean that $\mathcal{E}(\mathcal{Z}) \subseteq \tilde{\mathcal{D}}(\mathcal{Z}) \subseteq \mathcal{C}(\mathcal{Z})$.

2. THE IDENTIFIED SET

In this Section the set identified by the SEIV model is defined and notation is introduced.

We consider situations in which X, which may be a scalar or a vector, is discrete and takes values in the set $\mathcal{X} = \{x_k\}_{k=1}^K$. In this case the structural function h is characterized by $N \equiv K(M-1)$ parameters as follows,

$$\gamma_{mk} \equiv h_m(x_k), \qquad m \in \{1, \ldots, M-1\}, \quad k \in \{1, \ldots, K\}$$

which are arranged in a vector $\boldsymbol{\gamma}$, as follows.[8]

$$\boldsymbol{\gamma} \equiv [\gamma_{11}, \ldots, \gamma_{1K}, \gamma_{21}, \ldots, \gamma_{2K}, \ldots, \gamma_{M-1,1}, \ldots, \gamma_{M-1,K}]$$

It is the identification of the vector $\boldsymbol{\gamma}$ that is the focus of this paper. Each element of $\boldsymbol{\gamma}$ lies in the unit interval so each value of $\boldsymbol{\gamma}$ is a point in the unit N-cube. The identified set is a subset of the unit N-cube. There are the restrictions $\gamma_{lk} < \gamma_{mk}$ for all k and $l < m$.[9]

Consider a particular probability distribution for Y and X given $Z = z \in \mathcal{Z}$. The identified set of values of $\boldsymbol{\gamma}$ associated with this distribution contains all and only values of $\boldsymbol{\gamma}$ for which there exist admissible conditional probability distributions of U and X given Z for all values of Z in $\mathcal{Z}$ such that the resulting structures deliver the probability distribution under consideration. Notation for that probability distribution is now introduced.

For values $z \in \mathcal{Z}$, for $m \in \{1, \ldots, M-1\}$ and $k \in \{1, \ldots, K\}$, there are the following point probabilities:

$$\alpha_{mk}(z) \equiv \Pr[Y = m | X = x_k, Z = z] \qquad \delta_k(z) \equiv \Pr[X = x_k | Z = z]$$

and cumulative probabilities

$$\bar{\alpha}_{mk}(z) \equiv \Pr[Y \leq m | X = x_k, Z = z]$$

and it is convenient to have a notation for the following mixed cumulative and point probabilities.

$$\bar{\rho}_{mk}(z) \equiv \Pr[Y \leq m \wedge X = x_k | Z = z] = \bar{\alpha}_{mk}(z)\delta_k(z)$$

For all k define: $\gamma_{0k} = 0$, $\gamma_{Mk} = 1$, $\bar{\alpha}_{0k}(z) = \bar{\rho}_{0k}(z) = 0$, $\bar{\alpha}_{Mk}(z) = 1$, $\bar{\rho}_{Mk}(z) = \delta_k(z)$. In what follows "for all m" means for $m \in \{0, \ldots, M\}$.

Associated with a particular value of $\boldsymbol{\gamma}$ and each value $z \in \mathcal{Z}$, define a piecewise uniform conditional distribution for U given X and Z, such that for all m, k and k':

$$\beta_{mkk'}(z) \equiv \Pr[U \leq \gamma_{mk} | X = x_{k'}, Z = z]$$

[8] If Γ is a matrix with (m, k) element equal to γ_{mk} then $\gamma \equiv \text{vec}(\Gamma')$. Considering γ_n, the nth element of γ, there are the following relationships.

$$n = (m-1)K + k$$

$$k = n \text{ modulo } K \qquad m = (n-k)/K + 1$$

[9] Henceforth "for all k" means for $k \in \{1, \ldots, K\}$.

and let $\boldsymbol{\beta}(z)$ denote the complete list of $(M-1)K^2$ such terms.[10]

A list of values of $(\boldsymbol{\gamma}, \boldsymbol{\beta}(z))$ produced as z varies in $\mathcal{Z}$ characterizes a *structure* which is *admissible* if it satisfies the following *independence* and *properness* conditions

[1]. Independence. For all $z \in \mathcal{Z}$ and for all m and k the following equalities hold.[11]

$$\sum_{k'} \beta_{mkk'}(z)\delta_{k'}(z) = \gamma_{mk}$$

[2]. Properness. For all $z \in \mathcal{Z}$ and for all j, k, l, m and k', $\beta_{ljk'}(z) \leq \beta_{mkk'}(z)$ if and only if $\gamma_{lj} \leq \gamma_{mk}$.

If in addition the following *observational equivalence* condition is satisfied then the structure generates the probability distribution under consideration.

[3]. Observational equivalence. For all $z \in \mathcal{Z}$ and for m and k the following equalities hold.

$$\beta_{mkk}(z) = \bar{\alpha}_{mk}(z)$$

All and *only* structures that obey conditions [1], [2] and [3] are in the set of structures identified by the model for the probabilities considered. Let $\mathcal{S}(\mathcal{Z})$ denote that set. The identified set of structural *functions*, $\mathcal{H}(\mathcal{Z})$, is the set of values of $\boldsymbol{\gamma}$ for which there are values of $\boldsymbol{\beta}(z)$ for $z \in \mathcal{Z}$ such that the resulting structure is in the identified set, $\mathcal{S}(\mathcal{Z})$. The identified set for $\boldsymbol{\gamma}$, $\mathcal{H}(\mathcal{Z})$, is the *projection* of $\mathcal{S}(\mathcal{Z})$ onto the unit N-cube within which all values of $\boldsymbol{\gamma}$ lie.

The geometry of these sets is considered in Chesher and Smolinski (2009). A brief account is given here. Because of the properness condition [2] the order in which the elements of $\boldsymbol{\gamma}$ lie is an important consideration. There are $T \equiv (K(M-1))!/\left((M-1)!\right)^K$ admissible[12] arrangements of the elements of $\boldsymbol{\gamma}$. In each arrangement, $t \in \{1, \ldots, T\}$, the set of admissible observationally equivalent structures defined by [1], [2] and [3], denoted by $\mathcal{S}_t(\mathcal{Z})$, is either empty or a convex polyhedron because it is an intersection of bounded linear half spaces. The identified set of structures is the union of the sets obtained under each admissible arrangement.

$$\mathcal{S}(\mathcal{Z}) = \bigcup_{t=1}^{T} \mathcal{S}_t(\mathcal{Z})$$

In each arrangement, t, the identified set of structural functions obtained by projecting away $\boldsymbol{\beta}(z)$ for $z \in \mathcal{Z}$, denoted $\mathcal{H}_t(\mathcal{Z})$, is also either empty or a convex

[10] Between each pair of adjacent knots γ_{mk} each conditional density function is uniform. The construction is justified in Chesher (2009). The conditional density functions have a histogram-like appearance.

[11] The left hand sides are expected values of U given $Z = z$ which the independence restriction requires to be free of z. The values γ_{mk} on the right hand sides arises because of the uniform distribution normalisation of the marginal distribution of U. See Chesher (2008).

[12] See Chesher and Smolinski (2009). Arrangements in which there is a pair of indices m and m' with $m > m'$ such that for some k, $\gamma_{mk} \leq \gamma_{m'k}$ are inadmissible.

polyhedron. The complete identified set of structural functions is the union of these convex sets. The result may not itself be convex, nor even connected.

$$\mathcal{H}(\mathcal{Z}) = \bigcup_{t=1}^{T} \mathcal{H}_t(\mathcal{Z})$$

In problems in which M or K are at all large computation of an identified set of structural functions is difficult. A head on attack would consider each admissible arrangement in turn and use the method of Fourier-Motzkin elimination[13] to project away the $(M-1)K^2$ elements $\beta_{mkk'}(z)$ for each $z \in \mathcal{Z}$ but this is computationally infeasible when M and/or K are large.

In the next Section we develop easy-to-compute sets in the unit N-cube which are shown to contain the identified set of structural functions, $\mathcal{H}(\mathcal{Z})$.

3. Bounds on the identified set

Conditions [1] - [3] place restrictions on values of $(\boldsymbol{\gamma}, \boldsymbol{\beta}(z))$ for z varying in $\mathcal{Z}$. They define the set $\mathcal{S}(\mathcal{Z})$. In this Section we derive some implications of these restrictions for admissible values of $\boldsymbol{\gamma}$, that is for values of $\boldsymbol{\gamma}$ that lie in the set $\mathcal{H}(\mathcal{Z})$. The result is a set denoted $\mathcal{E}(\mathcal{Z})$ within which the identified set $\mathcal{H}(\mathcal{Z})$ lies. There may be values of $\boldsymbol{\gamma}$ in $\mathcal{E}(\mathcal{Z})$ which are not in the identified set $\mathcal{H}(\mathcal{Z})$. In ongoing research we are investigating whether that is the case.

As has already been noted the ordering of the elements of $\boldsymbol{\gamma}$ is important. Just like $\mathcal{S}(\mathcal{Z})$ and $\mathcal{H}(\mathcal{Z})$ the set $\mathcal{E}(\mathcal{Z})$ is a union of convex sets, $\mathcal{E}_t(\mathcal{Z})$, one associated with each admissible arrangement, t, of $\boldsymbol{\gamma}$. Each set is defined as an intersection of linear half spaces.

$$\mathcal{E}(\mathcal{Z}) = \bigcup_{t=1}^{T} \mathcal{E}_t(\mathcal{Z})$$

We proceed to develop a definition of a set $\mathcal{E}_t(\mathcal{Z})$ obtained under a particular arrangement of the elements of $\boldsymbol{\gamma}$. Let $\gamma_{[n]}$ be the nth largest value in that arrangement.[14]

Define functions $m(n)$ and $k(n)$ such that $\gamma_{m(n)k(n)} = \gamma_{[n]}$. For example with $M = 3$ and $K = 3$, for which $N = 6$ and

$$\gamma = [\gamma_{11}, \gamma_{12}, \gamma_{13}, \gamma_{21}, \gamma_{22}, \gamma_{23}]$$

and for the arrangement

$$[\gamma_{11}, \gamma_{12}, \gamma_{21}, \gamma_{13}, \gamma_{23}, \gamma_{22}]$$

[13] See Zeigler (2007).

[14] Recall here are $N = (M-1)K$ elements in γ. We adopt the notation used in the literature on order statistics to denote the ordered values of γ.

$$\gamma_{[1]} \leq \gamma_{[2]} \cdots \leq \gamma_{[N-1]} \leq \gamma_{[N]}.$$

the functions $m(s)$ and $k(s)$ are as shown below. We will work with this example throughout this Section.

n	$m(n)$	$k(n)$
1	1	1
2	1	2
3	2	1
4	1	3
5	2	3
6	2	2

These functions are specific to the particular arrangement under consideration.

We consider a particular value $z \in \mathcal{Z}$ and develop a set $\mathcal{E}_t(z)$ defining $\mathcal{E}_t(\mathcal{Z})$ as the intersection of such sets for z varying in $\mathcal{Z}$.

$$\mathcal{E}_t(\mathcal{Z}) \equiv \bigcap_{z \in \mathcal{Z}} \mathcal{E}_t(z)$$

To avoid cumbersome notation we do not make dependence on the chosen value z explicit in the notation for the moment.

We use an abbreviated notation as follows: $\bar{\rho}_{[n]}$ denotes $\bar{\rho}_{m(n)k(n)}$, $\beta_{[n]k'}$ denotes $\beta_{m(n)k(n)k'}$. Define $\eta_{mkk'} \equiv \beta_{mkk'}\delta_{k'}$ and let $\eta_{[n]k'}$ denote $\eta_{m(n)k(n)k'}$.[15]

Table 1 relates to a case with $K = 3$ and shows the conditional distribution-mass function terms $\eta_{[n]k}$ for each value in $\mathcal{X} \equiv \{x_1, x_2, x_3\}$. We define $\gamma_{[0]} = 0$ and $\gamma_{[N]} = 1$.

Now consider Conditions [1] - [3]. If the value of $\boldsymbol{\gamma}$ that appears in the final column of the Table lies in the identified set then there exist values of the terms $\eta_{[n]k}$ such that: (i) the row sum of these terms is the value of $\boldsymbol{\gamma}$ indicated in each row and (ii) the entries are non decreasing as we pass down each column. In addition the observational equivalence condition [3] requires that in each row, n, there is a term $\bar{\rho}_{[n]k}$ for the value of k delivered by the function $k(n)$.

Table 2 gives an example for the case in which $K = 3$ and $M = 3$ and a particular arrangement of $\boldsymbol{\gamma}$, shown in the penultimate column of the Table. This is the arrangement used in the example above.

The smallest element in this arrangement is $\gamma_{[1]} = \gamma_{11}$ so $k(1) = 1$ (shown in the final column of the Table) and in the first row it is $\eta_{[1]1}$ which is equal to $\bar{\rho}_{[1]} = \bar{\rho}_{11}$. The largest element in the arrangement is $\gamma_{[6]} = \gamma_{22}$ so $k(6) = 2$ and in the row corresponding to $\gamma_{[6]}$ it is $\eta_{[6]2}$ that is equal to $\bar{\rho}_{[6]} = \bar{\rho}_{22}$.

Figure 1 shows a configuration of threshold functions that is consonant with this arrangement. Here $M = 3$ so there are two threshold functions, $h_1(x)$ and $h_2(x)$. In the Figure values of X are measured along the horizontal axis and three points of support, x_1, x_2 and x_3 are marked. Values of threshold functions are measured along the vertical axis which represents the unit interval $[0, 1]$. This axis also measures values of the unobservable variable U. At any value of x values of U falling below the lowest threshold function deliver the value 1 for Y, values of U falling between the two threshold functions deliver the value 2 for Y and values of U falling above the highest

[15] All these objects depend on the chosen value z. A more precise but more cumbersome notation would use $\bar{\rho}_{[s]}(z)$, $\beta_{[s]k'}(z)$ and so forth.

n	Value of x			Ordered values of $\boldsymbol{\gamma}$
	x_1	x_2	x_3	
0	0	0	0	$\gamma_{[0]}$
1	$\eta_{[1]1}$	$\eta_{[1]2}$	$\eta_{[1]3}$	$\gamma_{[1]}$
2	$\eta_{[2]1}$	$\eta_{[2]2}$	$\eta_{[2]3}$	$\gamma_{[2]}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
n	$\eta_{[n]1}$	$\eta_{[n]2}$	$\eta_{[n]3}$	$\gamma_{[n]}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$N-1$	$\eta_{[N-1]1}$	$\eta_{[N-1]2}$	$\eta_{[N-1]3}$	$\gamma_{[N-1]}$
N	δ_1	δ_2	δ_3	$\gamma_{[N]}$

Table 1: Conditional distribution-mass function values arranged by ordered values of γ and values of the conditioning variable X.

n	Value of x			Ordered values of $\boldsymbol{\gamma}$	$\boldsymbol{\gamma}$	$k(n)$
	x_1	x_2	x_3			
0	0	0	0	$\gamma_{[0]}$	0	
1	$\bar{\rho}_{[1]}$	$\eta_{[1]2}$	$\eta_{[1]3}$	$\gamma_{[1]}$	γ_{11}	1
2	$\eta_{[2]1}$	$\bar{\rho}_{[2]}$	$\eta_{[2]3}$	$\gamma_{[2]}$	γ_{12}	2
3	$\bar{\rho}_{[3]}$	$\eta_{[3]2}$	$\eta_{[3]3}$	$\gamma_{[3]}$	γ_{21}	1
4	$\eta_{[4]1}$	$\eta_{[4]2}$	$\bar{\rho}_{[4]}$	$\gamma_{[4]}$	γ_{13}	3
5	$\eta_{[5]1}$	$\eta_{[5]2}$	$\bar{\rho}_{[5]}$	$\gamma_{[5]}$	γ_{23}	2
6	$\eta_{[6]1}$	$\bar{\rho}_{[6]}$	$\eta_{[6]3}$	$\gamma_{[6]}$	γ_{22}	2
7	δ_1	δ_2	δ_3	$\gamma_{[7]}$	1	

Table 2: Conditional distribution-mass function values for the example with observational equivalence restrictions imposed.

threshold deliver the value 3 for Y. The regions in which Y takes its three values are shaded. The threshold functions have been drawn to reflect the arrangement shown in Table 2. Notice that the upper threshold function is not monotone in x reflecting the inequality $\gamma_{23} < \gamma_{22}$.

We obtain bounds on the elements of γ for the arrangement under consideration using the following argument.

Consider the nth largest element $\gamma_{[n]}$. If this lies in the identified set then there exist values of the terms $\eta_{[n]k}$ which sum to $\gamma_{[n]}$. We find lower and upper bounds for each term $\eta_{[n]k}$ in the sum and obtain a bound on $\gamma_{[n]}$ by summing the bounds. In fact, as will be seen, we can do better than this.

Each term $\eta_{[n]k}$ is bounded *below* by the *maximum* of the terms $\bar{\rho}_{[i]k}$ that appear in rows 1 through n of the column associated with x_k. That bound is λ_{nk} defined now. Recall we have defined $\bar{\rho}_{0k} = 0$ for all k.

$$\lambda_{nk} \equiv \max_{0 \leq i \leq n} \left\{ \bar{\rho}_{[i]} \text{ s.t. } k(i) = k \right\}$$

Each term $\eta_{[n]k}$ is bounded *above* by the *minimum* of the terms $\bar{\rho}_{[i]k}$ that appear in rows n through N of the column associated with x_k. That bound is π_{nk} defined now. Recall we have defined $\bar{\rho}_{Mk} = \delta_k$ for all k.

$$\pi_{nk} \equiv \min_{n \le i \le N} \left\{ \bar{\rho}_{[i]} \text{ s.t. } k(i) = k \right\}$$

Combining results there are the following bounds:

$$\lambda_{nk} \le \eta_{[n]k} \le \pi_{nk} \tag{4}$$

and on summing and noting that for $\boldsymbol{\gamma}$ in the identified set the independence condition holds so that $\gamma_{[n]} = \sum_k \eta_{[n]k}$ there are the following lower and upper bounds.

$$\lambda_n \equiv \sum_k \lambda_{nk} \le \gamma_{[n]} \le \sum_k \pi_{nk} \equiv \pi_n \tag{5}$$

These bounds, unlike $\gamma_{[n]}$, depend on the chosen value z. This dependence is made explicit in the notation at the next step where we obtain tighter bounds by *intersecting* the intervals given in (5), as follows.

$$\lambda_n(\mathcal{Z}) \equiv \max_{z \in \mathcal{Z}} (\lambda_n(z)) \le \gamma_{[n]} \le \min_{z \in \mathcal{Z}} (\pi_n(z)) \equiv \pi_n(\mathcal{Z}) \tag{6}$$

It will be shown shortly that these bounds are precisely the bounds given in Chesher (2008) which are built on the probability inequalities (3).

The inequalities (4) can also be used to place bounds on differences, $\gamma_{[b]} - \gamma_{[a]}$, as follows. For all a and b in $\{0, \ldots, N\}$ and for all k there are bounds on $\eta_{[b]k}$ and on $-\eta_{[a]k}$ as follows:

$$\lambda_{bk} \le \eta_{[b]k} \le \pi_{bk}$$

$$-\pi_{ak} \le -\eta_{[a]k} \le -\lambda_{ak}$$

and adding there is the following bound.

$$\lambda_{bk} - \pi_{ak} \le \eta_{[b]k} - \eta_{[a]k} \le \pi_{bk} - \lambda_{ak} \tag{7}$$

Summing across k we have the following.

$$\lambda_b - \pi_a = \sum_k (\lambda_{bk} - \pi_{ak}) \le \gamma_{[b]} - \gamma_{[a]} \le \sum_k (\pi_{bk} - \lambda_{ak}) = \pi_b - \lambda_a \tag{8}$$

This is a direct implication of (5) but the lower bound *can be improved upon* by exploiting the properness condition [2]. Thus, consider values a and b such that $b \ge a$. If $\boldsymbol{\gamma}$ is in the identified set then for all $b \ge a$ the inequality $\eta_{[b]k} - \eta_{[a]k} \ge 0$ holds. The lower bound in (7) can therefore be tightened as follows.

$$\max(0, \lambda_{bk} - \pi_{ak}) \le \eta_{[b]k} - \eta_{[a]k} \le \pi_{bk} - \lambda_{ak} \tag{9}$$

Summing across k we obtain the following bounds which hold for all $b \ge a \in \{0, \ldots, N\}$.

$$\sum_k \max(0, \lambda_{bk} - \pi_{ak}) \le \gamma_{[b]} - \gamma_{[a]} \le \pi_b - \lambda_a \tag{10}$$

Finally, making explicit the dependence on the value z of the terms in these bounds and intersecting the bounds across $z \in \mathcal{Z}$ gives the following inequalities which hold for all $N \geq b \geq a \geq 0$.

$$\phi^L_{ba}(\mathcal{Z}) \equiv \max_{z\in\mathcal{Z}} \left(\sum_k \max(0, \lambda_{bk}(z) - \pi_{ak}(z)) \right) \leq \gamma_{[b]} - \gamma_{[a]} \leq \min_{z\in\mathcal{Z}} (\pi_b(z) - \pi_a(z)) \equiv \phi^U_{ba}(\mathcal{Z}) \tag{11}$$

The lower bound here can be tighter than the lower bound in (8) if for any z and k, $\lambda_{bk}(z) < \pi_{ak}(z)$.

Consider the case in which $a = 0$. There is $\gamma_{[0]} = 0$, $\lambda_0 = 0$ and for all k, $\pi_{0k} = 0$ and so (11) reduces to the bounds (6), that is, for all b, $\phi^L_{b0}(\mathcal{Z}) = \lambda_b(\mathcal{Z})$.

The set $\mathcal{E}_t(\mathcal{Z})$ obtained under arrangement t is defined by the inequalities (11) generated as a and b vary with $N \geq b \geq a \geq 0$. We have shown in this Section that $\mathcal{H}_t(\mathcal{Z}) \subseteq \mathcal{E}_t(\mathcal{Z})$ from which it follows directly that $\mathcal{H}(\mathcal{Z}) \subseteq \mathcal{E}(\mathcal{Z})$.

3.1. Binary outcomes. When Y is binary there is just one threshold function and the parameters of interest are $\gamma_{11}, \gamma_{12}, \ldots, \gamma_{1K}$. We now show that in this case the lower bound in (11) is zero when $a > 0$ so these bounds place no restrictions on $\boldsymbol{\gamma}$ additional to those defined by (6).

Without loss of generality we consider an arrangement t in which the elements of $\boldsymbol{\gamma}$ are arranged in the order of the index k. The situation for a case in which $K = 6$ is as pictured in Table 3. Notice that with the given arrangement of $\boldsymbol{\gamma}$ for every value of n, $k(n) = n$, so the values $\bar{\rho}_{[n]}$ lie on the diagonal in Table 3. Because Y is binary there is only one such entry in each column.

We now show that for all indices $b > a > 0$ the terms $\lambda_{bk} - \pi_{ak}$ are zero or negative for all k from which it follows that the lower bound in (11) is zero.

Consider some value k and the difference $\lambda_{bk} - \pi_{ak}$ with $b \geq a$. Referring to Table 3 it can be seen that values taken by λ_{bk} and π_{ak} are as follows.

$$\lambda_{bk} = \begin{cases} 0 & , \quad b < k \\ \bar{\rho}_{[k]} & , \quad b \geq k \end{cases} \qquad\qquad \pi_{ak} = \begin{cases} \bar{\rho}_{[k]} & , \quad a \leq k \\ \delta_k & , \quad a > k \end{cases}$$

The resulting values of $\lambda_{bk} - \pi_{ak}$ are therefore as shown below.

Values of $\lambda_{bk} - \pi_{ak}$

	$b < k$	$b = k$	$b > k$
$a < k$	$-\bar{\rho}_{[k]}$	$-\bar{\rho}_{[k]}$	0
$a = k$	$*$	0	0
$a > k$	$*$	$*$	$\bar{\rho}_{[k]} - \delta_k$

All the values are zero or negative and the result is that the lower bound $\phi^L_{ba}(\mathcal{Z})$ is zero. Therefore in the binary Y case the restrictions imposed by the bounds (11) for $a \neq 0$ have no force. It is shown in the next Section that the bounds obtained from (11) setting $a = 0$, equivalently the bounds (6), are identical to the bounds given in Chesher (2008, 2009) which are shown in those papers to define the identified set $\mathcal{H}(\mathcal{Z})$.

	Value of X						Ordered		
n	x_1	x_2	x_3	x_4	x_5	x_6	values of $\boldsymbol{\gamma}$	$\boldsymbol{\gamma}$	$k(n)$
0	0	0	0	0	0	0	$\gamma_{[0]}$	0	
1	$\bar{\rho}_{[1]}$	$\eta_{[1]2}$	$\eta_{[1]3}$	$\eta_{[1]4}$	$\eta_{[1]5}$	$\eta_{[1]6}$	$\gamma_{[1]}$	γ_{11}	1
2	$\eta_{[2]1}$	$\bar{\rho}_{[2]}$	$\eta_{[2]3}$	$\eta_{[2]4}$	$\eta_{[2]5}$	$\eta_{[2]6}$	$\gamma_{[2]}$	γ_{12}	2
3	$\eta_{[3]1}$	$\eta_{[3]2}$	$\bar{\rho}_{[3]}$	$\eta_{[3]4}$	$\eta_{[3]5}$	$\eta_{[3]6}$	$\gamma_{[3]}$	γ_{13}	3
4	$\eta_{[4]1}$	$\eta_{[4]2}$	$\eta_{[4]3}$	$\bar{\rho}_{[4]}$	$\eta_{[4]5}$	$\eta_{[4]6}$	$\gamma_{[4]}$	γ_{14}	4
5	$\eta_{[5]1}$	$\eta_{[5]2}$	$\eta_{[5]3}$	$\eta_{[5]4}$	$\bar{\rho}_{[5]}$	$\eta_{[5]6}$	$\gamma_{[5]}$	γ_{15}	5
6	$\eta_{[6]1}$	$\eta_{[6]2}$	$\eta_{[6]3}$	$\eta_{[6]4}$	$\eta_{[6]5}$	$\bar{\rho}_{[6]}$	$\gamma_{[6]}$	γ_{16}	6
7	δ_1	δ_2	δ_3	δ_4	δ_5	δ_6	$\gamma_{[7]}$	1	

Table 3: Conditional distribution-mass function values for the binary case example with observational equivalence restrictions imposed.

3.2. Relationship to earlier results. It is shown in Chesher (2008, 2009) that all structural functions h that lie in the set identified by a SEIV model given a particular probability distribution $F_{YX|Z}$ with $Z = z \in \mathcal{Z}$ satisfy the following inequalities for all $\tau \in (0, 1)$.

$$\max_{z\in\mathcal{Z}} \Pr[Y < h(X,\tau)|Z = z] < \tau \leq \min_{z\in\mathcal{Z}} \Pr[Y \leq h(X,\tau)|Z = z] \tag{12}$$

Here probabilities are calculated using the distribution $F_{YX|Z}$.

The inequalities generated by (12) as τ varies over $(0, 1)$ define a set of structural functions referred to as $\mathcal{C}(\mathcal{Z})$ in Chesher and Smolinski (2009). When X is discrete and characterized by a vector $\boldsymbol{\gamma}$ as in the previous discussion the set $\mathcal{C}(\mathcal{Z})$ is a union of convex sets, $\mathcal{C}_t(\mathcal{Z})$, one associated with each arrangement, t, of $\boldsymbol{\gamma}$.

$$\mathcal{C}(\mathcal{Z}) = \bigcup_{t=1}^{T} \mathcal{C}_t(\mathcal{Z})$$

Each set $\mathcal{C}_t(\mathcal{Z})$ is an intersection of sets obtained as z varies within $\mathcal{Z}$.

$$\mathcal{C}_t(\mathcal{Z}) = \bigcap_{z\in\mathcal{Z}} \mathcal{C}_t(z)$$

We now show that the bounds (12) are identical to those generated by (6) as n varies over $\{1, \ldots, N\}$. Chesher and Smolinski (2009) show that in the discrete endogenous variable case considered here the bounds (12) hold for all $\tau \in (0, 1)$ if and only if the following inequalities hold for all $l \in \{1, \ldots, M-1\}$ and $s \in \{1, \ldots, K\}$.

$$\max_{z\in\mathcal{Z}} \sum_{k=1}^{K} \sum_{m=1}^{M-1} \delta_k(z)\alpha_{mk}(z)1(\gamma_{mk} \leq \gamma_{ls}) \leq \gamma_{ls} \leq \min_{z\in\mathcal{Z}} \sum_{k=1}^{K} \sum_{m=1}^{M} \delta_k(z)\alpha_{mk}(z)1(\gamma_{m-1,k} < \gamma_{ls})$$

Consider a particular arrangement of $\gamma_{[n]}$ and its nth largest element, $\gamma_{[n]}$. Substituting $\gamma_{[n]}$ for γ_{ls} above gives the following.

$$\max_{z\in\mathcal{Z}}\sum_{k=1}^{K}\sum_{m=1}^{M-1}\delta_k(z)\alpha_{mk}(z)1(\gamma_{mk}\leq\gamma_{[n]})\leq\gamma_{[n]}\leq\min_{z\in\mathcal{Z}}\sum_{k=1}^{K}\sum_{m=1}^{M}\delta_k(z)\alpha_{mk}(z)1(\gamma_{m-1,k}<\gamma_{[n]}) \tag{13}$$

Comparing this with (6) it can be concluded that the bounds are identical because both of the following equations are satisfied:

$$\max_{0\leq i\leq n}\left\{\bar{\rho}_{[i]}(z)\text{ s.t. }k(i)=k\right\}\equiv\lambda_{nk}(z)=\sum_{m=1}^{M-1}\delta_k(z)\alpha_{mk}(z)1(\gamma_{mk}\leq\gamma_{[n]}) \tag{14}$$

$$\min_{n\leq i\leq N}\left\{\bar{\rho}_{[i]}(z)\text{ s.t. }k(i)=k\right\}\equiv\pi_{nk}(z)=\sum_{m=1}^{M}\delta_k(z)\alpha_{mk}(z)1(\gamma_{m-1,k}<\gamma_{[n]}) \tag{15}$$

and on the right hand side here are the expressions summed over k to produce the bounds in (13).

3.3. Alternative expressions for the bounds. The objects $\lambda_{nk}(z)$ and $\pi_{nk}(z)$ can be expressed in terms of probabilities involving the structural function as follows.

$$\lambda_{nk}(z)=\Pr[Y<h(x_k,\gamma_{[n+1]})|X=x_k,Z=z]\Pr[X=x_k|Z=z] \tag{16}$$

$$\pi_{nk}(z)=\Pr[Y\leq h(x_k,\gamma_{[n]})|X=x_k,Z=z]\Pr[X=x_k|Z=z] \tag{17}$$

Therefore the bounds (6) can be written as follows:

$$\lambda_n(\mathcal{Z})\equiv\max_{z\in\mathcal{Z}}\left(\Pr[Y<h(X,\gamma_{[n+1]})|Z=z]\right)\leq\gamma_{[n]}\leq\min_{z\in\mathcal{Z}}\left(\Pr[Y\leq h(X,\gamma_{[n]})|Z=z]\right)\equiv\pi_n(\mathcal{Z})$$

and the lower bound in (10) can be written as

$$\phi^L_{ba}(\mathcal{Z})=\max_{z\in\mathcal{Z}}\left(\phi^L_{ba}(z)\right)$$

where there is the following expression for $\phi^L_{ba}(z)$.

$$\sum_k\max\left(0,\left\{\Pr[Y<h(x_k,\gamma_{[b+1]})|x_k,z]-\Pr[Y\leq h(x_k,\gamma_{[a]})|x_k,z]\right\}\right)\Pr[X=x_k|z]$$

An example may help at this point.

Table 4 shows the values of λ_{nk} and π_{nk} in the arrangement used in the example considered earlier in the paper. Dependence on the value z is no longer made explicit in the notation.

Table 5 shows the value of the structural function $h(x,u)$ for all the combinations of x and u that arise in (16) and (17) in this example. For example the entry for $n=3$ (row) and $k=2$ under the heading π_{n2} is $h(x_2,\gamma_{[3]})=h(x_2,\gamma_{21})=2$. The entries in this Table are easily verified by considering Figure 1.

			$k=1$		$k=2$		$k=3$	
n	$\gamma_{(n)}$	$\gamma_{(n+1)}$	λ_{n1}	π_{n1}	λ_{n2}	π_{n2}	λ_{n3}	π_{n3}
0	0	γ_{11}	0	$\bar{\rho}_{11}$	0	$\bar{\rho}_{12}$	0	$\bar{\rho}_{13}$
1	γ_{11}	γ_{12}	$\bar{\rho}_{11}$	$\bar{\rho}_{11}$	0	$\bar{\rho}_{12}$	0	$\bar{\rho}_{13}$
2	γ_{12}	γ_{21}	$\bar{\rho}_{11}$	$\bar{\rho}_{21}$	$\bar{\rho}_{12}$	$\bar{\rho}_{12}$	0	$\bar{\rho}_{13}$
3	γ_{21}	γ_{13}	$\bar{\rho}_{21}$	$\bar{\rho}_{21}$	$\bar{\rho}_{12}$	$\bar{\rho}_{22}$	0	$\bar{\rho}_{13}$
4	γ_{13}	γ_{23}	$\bar{\rho}_{21}$	$\bar{\rho}_{31}$	$\bar{\rho}_{12}$	$\bar{\rho}_{22}$	$\bar{\rho}_{13}$	$\bar{\rho}_{13}$
5	γ_{23}	γ_{22}	$\bar{\rho}_{21}$	$\bar{\rho}_{31}$	$\bar{\rho}_{12}$	$\bar{\rho}_{22}$	$\bar{\rho}_{23}$	$\bar{\rho}_{23}$
6	γ_{22}	1	$\bar{\rho}_{21}$	$\bar{\rho}_{31}$	$\bar{\rho}_{22}$	$\bar{\rho}_{22}$	$\bar{\rho}_{23}$	$\bar{\rho}_{33}$

Table 4: Values of λ_{nk} and π_{nk} in the arrangement used in the example.

			$k=1$		$k=2$		$k=3$	
n	$\gamma_{(n)}$	$\gamma_{(n+1)}$	A: λ_{n1}	B: π_{n1}	A: λ_{n2}	B: π_{n2}	A: λ_{n3}	B: π_{n3}
0	0	γ_{11}	1	1	1	1	1	1
1	γ_{11}	γ_{12}	2	1	1	1	1	1
2	γ_{12}	γ_{21}	2	2	2	1	1	1
3	γ_{21}	γ_{13}	3	2	2	2	1	1
4	γ_{13}	γ_{23}	3	3	2	2	2	1
5	γ_{23}	γ_{22}	3	3	2	2	3	2
6	γ_{22}	1	3	3	3	2	3	3

Table 5: For the arrangement used in the example these are the values of A: $h(x_k, \gamma_{[n+1]})$ appearing in the definition of λ_{nk} and of B: $h(x_k, \gamma_{[n]})$ appearing in the definition of π_{nk}.

Consider for example λ_{42}. From (16) we have

$$\lambda_{42} = \Pr[Y < h(x_2, \gamma_{[5]})|X = x_2, Z = z] \Pr[X = x_2|Z = z]$$

and since $\gamma_{[5]} = \gamma_{23}$ there is from Table 5 $h(x_2, \gamma_{23}) = 2$. Accordingly

$$\lambda_{42} = \Pr[Y \leq 1|X = x_2, Z = z] \Pr[X = x_2|Z = z]$$

which is equal to $\bar{\rho}_{12}$ as shown in Table 4 in the entry for $n = 4$ and $k = 2$.

Consider for example π_{33}. From (17) we have

$$\pi_{33} = \Pr[Y \leq h(x_3, \gamma_{[3]})|X = x_3, Z = z] \Pr[X = x_3|Z = z]$$

and since $\gamma_{[3]} = \gamma_{21}$ there is from Table 5 $h(x_3, \gamma_{21}) = 1$. Accordingly

$$\pi_{33} = \Pr[Y \leq 1|X = x_3, Z = z] \Pr[X = x_3|Z = z]$$

which is equal to $\bar{\rho}_{13}$ as shown in Table 4 in the entry for $n = 3$ and $k = 3$.

3.4. Illustrative calculations. This Section provides an example of the bounds for a case with $M = 3$ and $K = 3$ and the arrangement of $\boldsymbol{\gamma}$ considered throughout

n	γ_{mk}	$\gamma_{[n]}$	λ_n	π_n
1	γ_{11}	$\gamma_{[1]}$	$\frac{1}{16}$	$\frac{485}{3024}$
2	γ_{12}	$\gamma_{[2]}$	$\frac{59}{432}$	$\frac{1619}{3024}$
3	γ_{21}	$\gamma_{[3]}$	$\frac{221}{432}$	$\frac{2291}{3024}$
4	γ_{13}	$\gamma_{[4]}$	$\frac{1619}{3024}$	$\frac{155}{189}$
5	γ_{23}	$\gamma_{[5]}$	$\frac{1763}{3024}$	$\frac{164}{189}$
6	γ_{22}	$\gamma_{[6]}$	$\frac{2435}{3024}$	$\frac{26}{27}$

Table 6: Illustrative calculations: bounds on elements of γ.

the paper. There are the following values for probabilities for some value z.[16]

$$\alpha(z) \equiv \begin{pmatrix} \alpha_{11}(z) & \alpha_{12}(z) & \alpha_{13}(z) \\ \alpha_{21}(z) & \alpha_{22}(z) & \alpha_{23}(z) \\ \alpha_{31}(z) & \alpha_{32}(z) & \alpha_{33}(z) \end{pmatrix} = \begin{pmatrix} \frac{1}{8} & \frac{2}{9} & \frac{1}{7} \\ \frac{6}{8} & \frac{6}{9} & \frac{2}{7} \\ \frac{1}{8} & \frac{1}{9} & \frac{4}{7} \end{pmatrix}$$

$$\delta(z) = \begin{pmatrix} \delta_1(z) & \delta_2(z) & \delta_3(z) \end{pmatrix} = \begin{pmatrix} \frac{3}{6} & \frac{2}{6} & \frac{1}{6} \end{pmatrix}$$

Table 6 shows the lower and upper bounds on the elements of $\boldsymbol{\gamma}$ for the single value z under consideration delivered by the bounds (5). These would contribute to the construction of a set $\mathcal{C}_t(z)$ - see Section 3.2. The definition of that set also involves the restrictions imposed by the ordering in the arrangement, e.g. $\gamma_{11} \leq \gamma_{12}$ and $\gamma_{23} \leq \gamma_{22}$. This means that the set set $\mathcal{C}_t(z)$ will not admit $\gamma_{11} = \frac{485}{3024}$ and $\gamma_{12} = \frac{59}{432}$ even though these both lie in the intervals defined by the entries in Table 6 because $\frac{485}{3024} \simeq 0.16 > \frac{59}{432} \simeq 0.14$ which violates the restriction $\gamma_{11} \leq \gamma_{12}$.

Table 7 shows the values of lower bounds on $\gamma_{[b]} - \gamma_{[a]}$ for b (columns) and a (rows) varying in $\{0, \ldots, 7\}$. Because $\gamma_{[0]} = 0$ the first row of the Table shows the lower bounds on elements $\gamma_{[b]}$ and reproduces the "λ_n" column of Table 6. Similarly, because $\gamma_{[7]} = 1$ the final column of Table 7 shows the lower bound on $1 - \gamma_{[a]}$, so 1 minus the entries in that column are upper bounds on $\gamma_{[a]}$ and with that adjustment they reproduce the entries in the "π" column in Table 6.

The entries that lie in neither the first row ($\gamma_{[0]}$) nor the last column ($\gamma_{[7]}$) in Table 7 show the restrictions additional to those defining $\mathcal{C}_t(z)$ that arise in defining the set $\mathcal{E}_t(z)$. For example there are the restrictions $\gamma_{21} - \gamma_{11} \geq \frac{3}{8}$ and $\gamma_{22} - \gamma_{12} \geq \frac{17}{63}$.

4. Concluding remarks

We have considered identification of a nonparametrically specified structural function in a discrete outcome - discrete endogenous variable setting. The single equation instrumental variable (SEIV) model we have studied set identifies the structural function which is characterized by a point in the unit N-cube.

We have developed expressions for a set $\mathcal{E}(\mathcal{Z})$ which is easily calculated for any probability distribution of the outcome Y and endogenous variables X given instruments Z taking values in a set $\mathcal{Z}$. We have shown that the set identified by the SEIV model, $\mathcal{H}(\mathcal{Z})$, is a subset of $\mathcal{E}(\mathcal{Z})$. In no calculations done by us to date have we

[16] These values have been fabricated purely to provide an illustrative calculation. They are not generated by some structure at some actual value z.

		0	γ_{11}	γ_{12}	γ_{21}	γ_{13}	γ_{23}	γ_{22}	1
		$\gamma_{[0]}$	$\gamma_{[1]}$	$\gamma_{[2]}$	$\gamma_{[3]}$	$\gamma_{[4]}$	$\gamma_{[5]}$	$\gamma_{[6]}$	$\gamma_{[7]}$
0	$\gamma_{[0]}$	0	$\frac{1}{16}$	$\frac{59}{432}$	$\frac{221}{432}$	$\frac{1619}{3024}$	$\frac{1763}{3024}$	$\frac{2435}{3024}$	1
γ_{11}	$\gamma_{[1]}$	$\cdot$	0	0	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{71}{168}$	$\frac{325}{504}$	$\frac{2539}{3024}$
γ_{12}	$\gamma_{[2]}$	$\cdot$	$\cdot$	0	0	0	$\frac{1}{21}$	$\frac{17}{63}$	$\frac{1405}{3024}$
γ_{21}	$\gamma_{[3]}$	$\cdot$	$\cdot$	$\cdot$	0	0	$\frac{1}{21}$	$\frac{1}{21}$	$\frac{733}{3024}$
γ_{13}	$\gamma_{[4]}$	$\cdot$	$\cdot$	$\cdot$	$\cdot$	0	$\frac{1}{21}$	$\frac{1}{21}$	$\frac{34}{189}$
γ_{23}	$\gamma_{[5]}$	$\cdot$	$\cdot$	$\cdot$	$\cdot$	$\cdot$	0	0	$\frac{25}{189}$
γ_{22}	$\gamma_{[6]}$	$\cdot$	$\cdot$	$\cdot$	$\cdot$	$\cdot$	$\cdot$	0	$\frac{1}{27}$
1	$\gamma_{[7]}$	$\cdot$	$\cdot$	$\cdot$	$\cdot$	$\cdot$	$\cdot$	$\cdot$	0

Table 7: Illustrative calculations: values of lower bounds on $\gamma_{[b]} - \gamma_{[a]}$ for b (in columns) and a (in rows).

found structural functions in $\mathcal{E}(\mathcal{Z})$ which are not also in $\mathcal{H}(\mathcal{Z})$ but at this point we cannot claim that $\mathcal{E}(\mathcal{Z}) = \mathcal{H}(\mathcal{Z})$. Determining if this is the case is one of a number of items on our research agenda.

In cases in which the number of values of the outcome, M,or the number of points of support of the endogenous variable, K, are at all large direct calculation of the identified set $\mathcal{H}(\mathcal{Z})$ is infeasible. However $\mathcal{E}(\mathcal{Z})$ can be computed in many cases in which $\mathcal{H}(\mathcal{Z})$ cannot be computed so even if $\mathcal{E}(\mathcal{Z})$ is an outer set it can provide valuable information about structural functions. For example it can be instrumental in the rejection of hypotheses.

When M and K are both large even computation of $\mathcal{E}(\mathcal{Z})$ is challenging because of the large number of potential arrangements of $\boldsymbol{\gamma}$, that is of the K values of the $M-1$ threshold functions, that may arise. Here *shape restrictions* will be helpful in cutting down the scale of the problem. The use and benefit of restrictions on threshold functions such as monotonicity or single-peakedness for variations in endogenous variables is one of the topics we are currently studying.

An alternative way to proceed in this situation is to bring *parametric restrictions* on board as described and illustrated in Chesher and Smolinski (2009). With a reasonably parsimonious parametrization it is possible to work with a grid of values of parameters, $\boldsymbol{\theta}$. At each point in the grid one calculates the values of the parametrically specified threshold functions at the points of support of the endogenous variable, $\boldsymbol{\gamma}(\boldsymbol{\theta})$. The order in which these elements appear is determined at that stage so one knows which arrangement t to consider. One asks whether the value of $\boldsymbol{\gamma}(\boldsymbol{\theta})$ lies within the set $\mathcal{E}_t(\mathcal{Z})$ - if it does not the value $\boldsymbol{\theta}$ does not lie in the identified set of parameter values. As one passes across the grid knowledge of that identified set is built up. Efficient computational strategies for conducting this sort of search is another topic of current research.

References

Blundell Richard W., and James L. Powell (2003): “Endogeneity in Nonparametric and Semiparametric Regression Models,” in Dewatripont, M., L.P. Hansen, and S.J. Turnovsky, eds., *Advances in Economics and Econometrics: Theory and Applications, Eighth World Congress, Vol. II.* Cambridge: Cambridge University Press.

Blundell, Richard W., and James .L. Powell (2004): “Endogeneity in Semiparametric Binary Response Models,” *Review of Economic Studies*, 71, 655-679.

Chernozhukov, Victor and Christian Hansen (2001): “An IV Model of Quantile Treatment Effects,” Working Paper 02-06, Department of Economics, MIT. Available at http://ssrn.com/abstract=298879 or doi: 10.2139/ssrn.298879.

Chernozhukov, Victor And Christian Hansen (2005): “An IV Model of Quantile Treatment Effects,” *Econometrica*, 73, 245-261.

Chernozhukov, Victor, Sokbae Lee and Adam Rosen (2009): “Intersection Bounds: Estimation and Inference,” CeMMAP Working Paper 19/09.

Chesher, Andrew D., (2003): “Identification in Nonseparable Models,” *Econometrica*, 71, 1405-1441.

Chesher, Andrew D., (2005): “Nonparametric Identification under Discrete Variation,” *Econometrica*, 73, 1525-1550.

Chesher, Andrew D., (2008): “Instrumental Variable Models for Discrete Outcomes,” CeMMAP Working Paper 30/08, forthcoming in *Econometrica*.

Chesher, Andrew D., (2009): “Single Equation Endogenous Binary Response Models,” CeMMAP Working Paper 23/09.

Chesher, Andrew D., and Konrad Smolinski (2009): “IV Models of Ordered Choice,” CeMMAP Working Paper 37/09.

Das, Mitali (2005): "Instrumental Variables Estimators of Nonparametric Models with Discrete Endogenous Regressors," *Journal of Econometrics*, 124, 335-361.

Heckman, James J., (1978): “Dummy Endogenous Variables in a Simultaneous Equations System," *Econometrica*, 46, 931-959.

Imbens, Guido W., and Whitney K. Newey (2009): “Identification and Estimation of Triangular Simultaneous Equations Models Without Additivity,” *Econometrica*, 77, 1481–1512.

Jun, Sung Jae, Joris Pinkse and Haiqing Xu (2009): "Tighter Bounds in Triangular Systems,” unpublished paper.

Pemmaraju, Sriram and Steven Skienka (2003); *Combinatorial Discrete Mathematics.* New York: Cambridge University Press.

Ziegler, Gunter M., (2007): *Lectures on Polytopes.* (Updated seventh printing of the first edition*).* New York: Springer Science.

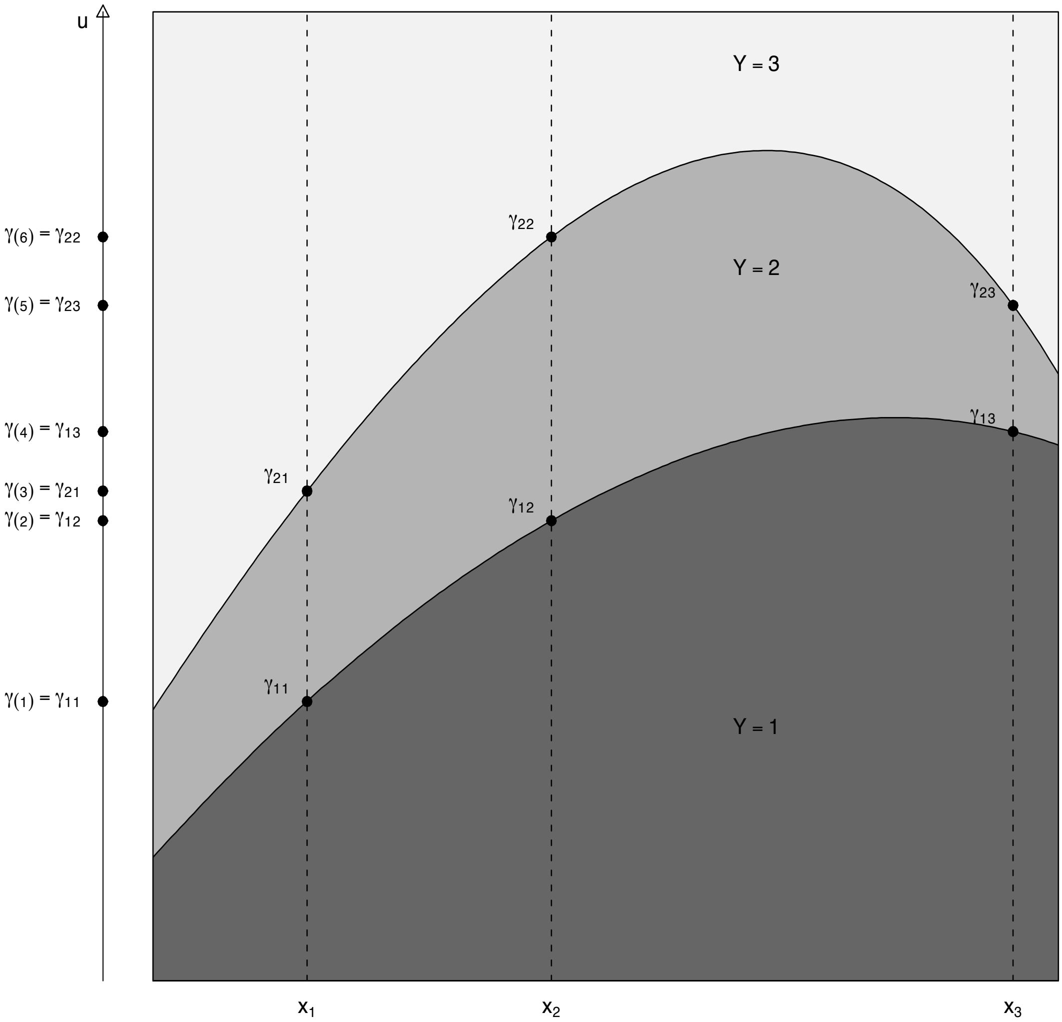

Figure 1: Examples of two threshold functions for the case $M = 3$ and $K = 3$ that are consonant with the arrangement of elements of γ shown on the vertical axis. The outcome Y takes the value 1 below the lowest threshold in the dark shaded region and the value 3 above the highest threshold in the light shaded region. The vertical scale is the unit interval $[0, 1]$.