cemmap

centre for microdata methods and practice

IV methods for Tobit models

Andrew Chesher
Dongwoo Kim
Adam M. Rosen

The Institute for Fiscal Studies
Department of Economics, UCL

cemmap working paper CWP26/21

Economic
and Social
Research Council

IV Methods for Tobit Models*

Andrew Chesher[†]
UCL and CeMMAP

Dongwoo Kim[‡]
Simon Fraser University

Adam M. Rosen[§]
Duke University and CeMMAP

June 27, 2021

Abstract

This paper studies models of processes generating censored outcomes with endogenous explanatory variables and instrumental variable restrictions. Tobit-type left censoring at zero is the primary focus in the exposition. The models studied here are unrestrictive relative to others widely used in practice, so they are relatively robust to misspecification. The models do not specify the process determining endogenous explanatory variables and they do not embody restrictions justifying control function approaches. The models can be partially or point identifying. Identified sets are characterized and it is shown how inference can be performed on scalar functions of partially identified parameters when exogenous variables have rich support. In an application using data on UK household tobacco expenditures inference is conducted on the coefficient of an endogenous total expenditure variable with and without a Gaussian distributional restriction on the unobservable and compared with the results obtained using a point identifying complete triangular model.

*We thank Xinyue Bei for providing excellent research assistance. Financial support from the UK Economic and Social Research Council through grants (RES-589-28-0001, RES-589-28-0002 and ES/P008909/1) to the ESRC Centre for Microdata Methods and Practice (CeMMAP) is gratefully acknowledged.

[†]Address: Department of Economics, University College London, Gower Street, London WC1E 6BT, United Kingdom. Email: andrew.chesher@ucl.ac.uk.

[‡]Address: Department of Economics, Simon Fraser University, West Mall Centre, SFU, 8888 University Drive Burnaby, BC Canada, V5A 1S6. Email: dongwook@sfu.ca.

[§]Address: Department of Economics, Duke University, 213 Social Sciences, Box 90097 Durham, NC 27708, United States. Email: adam.rosen@duke.edu.

Keywords: censored models, endogeneity, incomplete models, instrumental variables, partial identification, random sets.

JEL classification: C10, C24, C26, C51.

1 Introduction

This paper develops and applies results on the identifying power of models for continuous censored outcomes, focussing on cases with left censoring at a fixed known value. In econometrics the Tobit model, Tobin (1958), is a well known example of such a model.

In contrast to the classical Tobit model, the focus is on cases in which one or more of the explanatory variables is endogenous and there are exogenous variables that satisfy exclusion restrictions that limit their degree of influence on the censored outcomes. These instrumental variables (IVs) are restricted to be distributed to some degree independently of the unobserved variables in the structural equation for the censored outcomes. A variety of independence restrictions are considered, including mean, quantile, and full stochastic independence restrictions, with both parametric and nonparametric distributional restrictions.

This paper is unlike most that allow for endogenous explanatory variables in situations in which outcomes are censored because there is no specification of the determination of the endogenous explanatory variables as either a deterministic or stochastic function of the exogenous explanatory variables and instruments. Consequently the models studied here are incomplete for the determination of endogenous explanatory variables. The IV approach explored here is an alternative to approaches that use complete models, which are prey to misspecification if they introduce an incorrect specification of the genesis of endogenous explanatory variables. STATA's `ivtobit` command,[1] despite its name, computes estimates of parameters of Tobit models with endogenous explanatory variables using a complete triangular model assuming Gaussian unobserved variables. That attack does not deliver consistent estimates when endogenous variables are discrete or affected by multiple sources of heteorgeneity. By contrast, the IV approach employed here can be used when endogenous variables are discrete or when the way in which endogenous variables' values are generated is unrestricted. The IV approach is a useful alternative to control function approaches which require endogeneity to be absent once there is conditioning on identifiable functions of observed variables. This paper shows how inferences can be drawn using censored data and weakly restrictive robust IV models.

[1] StataCorp (2019).

The models studied are single equation IV models for a censored outcome variable. They fall in the class of generalized instrumental variable models analyzed in Chesher and Rosen (2017), henceforth CR17. We use techniques developed in that paper to carry out identification analysis, and we show how to implement the high level results given in CR17 in models for censored outcomes. The models are in general partially identifying but they can be point identifying, and it is usually not possible to determine identification status using the data available in applications. We conduct inference on scalar functions or subvectors of partially identified parameters following the approach of Belloni et al. (2018) based on a self-normalized critical value as in Chernozhukov et al. (2019), which is appropriate for inference based on a very large number of moment inequalities. The approach applies regardless of identification status, and we propose a method for implementation when exogenous variables have rich support. We illustrate with an application to UK household survey data recording tobacco expenditures in which around 70% of households record zero expenditures.

The main focus in this paper is on IV Tobit models with left censoring at zero, with and without a Gaussian distributional restriction on the scalar unobserved variable. However it is straightforward to extend to cases with right censored outcomes and to cases in which the censoring value is stochastic, distributed independently of the unobserved variable in the censored outcome equation.[2] Leading examples with right censoring arise when the censored outcome is the time until some event occurs, so the methods developed here find application in models of durations admitting endogenous explanatory variables.[3]

As noted, much of the prior literature studying models with censored outcomes and endogenous explanatory variables relies on a complete model for the determination of endogenous variables. Examples include the fully parametric specifications studied in Heckman (1978), Nelson and Olson (1978), Amemiya (1979), Smith and Blundell (1986), Newey (1987), and Blundell and Smith (1989) that made early contributions to the study of limited dependent variable models (including Tobit models in particular) permitting endogenous explanatory variables and enabling consistent estimation by way of control function approaches and marginal or conditional maximum likelihood procedures.[4] Control function approaches for semiparametric triangular models for censored outcomes are provided in Das (2002), Blundell and Powell (2007), and Chernozhukov et al. (2015). These papers do not

[2]Cases in which the censoring variable is also endogenous are considered in ongoing research.

[3]See for example Lancaster and Chesher (1984), Lancaster (1985), Olsen and Farkas (1989), Frandsen (2015), and Wrenn et al. (2017).

[4]Comparisons of the efficiency of different procedures are provided in Newey (1987) and Blundell and Smith (1989).

require parametric distributional restrictions on unobservable heterogeneity, with Das (2002) employing symmetry restrictions and Blundell and Powell (2007) and Chernozhukov et al. (2015) using conditional quantile restrictions.

Our approach is in the spirit of Manski and Tamer (2002), which pioneered the use of incomplete models for censored outcomes or covariates and used partial identification analysis. That paper characterized identified sets and proposed consistent set estimators for a variety of models with censored variables, in which the censoring process is not specified. In the IV Tobit models studied here, the censoring process is specified, but endogenous explanatory variables are permitted and it is the lack of specification of their determination that renders the models incomplete. In this respect the models studied here have similarities with the model studied in Hong and Tamer (2003). That model employs conditional quantile restrictions with a censored outcome variable and does not impose a specification of the process determining values of endogenous explanatory variables. The model of Hong and Tamer (2003) is incomplete but the focus there is on settings in which the parameters of the model are point identified. Sufficient conditions for point identification are proposed along with a point estimator, and the asymptotic properties of the estimator are characterized. The support conditions shown to guarantee point identification are strong and there are many cases arising in practice in which they will not be satisfied. We characterize sharp identified sets for model parameters applicable when these support conditions are not guaranteed to hold, and we illustrate the use of inference that is robust to the possibility of partial identification.

There is also research that considers the different but important problems of *endogenous censoring* of explanatory variables. This includes Khan and Tamer (2009), Khan et al. (2011), and Section 7 of Chesher and Rosen (2020b).

This paper makes several contributions to the literature on models of censoring with endogenous explanatory variables and instrumental variable restrictions.

1. As already pointed out, the vast majority of the previous papers on this topic consider models that, unlike those considered here, require a complete specification for the determination of endogenous explanatory variables. Our analysis thus shows what can still be learned when the specification of the genesis of endogenous explanatory variables is dropped.

2. We consider the use of more or less demanding restrictions on the distribution of unobservable heterogeneity conditional on instruments. This can be used to assess

(for example) how robust empirical findings are to relaxation from a full stochastic independence restriction to selected conditional quantile restrictions.

3. Our analysis is robust to the possibility of partial identification, and is thus applicable when data are not compatible with conditions that are known to ensure point identification.

4. We show how to conduct inference on functions or subvectors of parameters partially identified by moment inequalities in these IV Tobit models, using recent developments in Chernozhukov et al. (2019) and Belloni et al. (2018) allowing for a large number of moment inequalities relative to the sample size, as encountered in our application.

5. We show how quantile independence restrictions at multiple quantiles can be incorporated. This enables the study of the increase in identifying power as one moves from invoking a single conditional quantile restriction to successively more quantile restrictions, approaching full stochastic independence as more such restrictions are imposed.

6. We show how to determine the identifying power of an IV model for censored outcomes under a stochastic independence restriction with a nonparametric specification of the distribution of the unobservable variable in the structural equation for the censored outcome.

The paper proceeds as follows. In the following section we present the class of IV Tobit models studied. In Section 3 we characterize the identified set of structures – combinations of structural functions and distributions of unobservable heterogeneity – that are compatible with the censored outcome model. The identified set is shown to comprise those structures that lie in the intersection of two sets, each defined by a collection of conditional moment inequalities. We show how under some circumstances certain subsets of the inequalities reduce to moment equalities, and we show how exclusion restrictions can be incorporated into the characterization of the identified set. In Section 4 we analyze the identifying content of various restrictions on the joint distribution of exogenous variables and unobservable heterogeneity, such as conditional mean, conditional quantile, and stochastic independence restrictions. In Section 5 we describe how inference is carried out using results from Chernozhukov et al. (2019) and Belloni et al. (2018), and we propose a practical approach for application of the inference method when an identified set is characterized by conditional moment inequalities with continuous conditioning variables. We illustrate with an application in which we focus on conducting inference on the effect of total household nondurable expenditure on the share of expenditure spent on tobacco, previously considered using a control function approach in Adams et al. (2019). All proofs and figures are provided in in the Appendix.

2 The IV Tobit model

Scalar endogenous outcome Y_1, possibly endogenous vector Y_2, exogenous vector $Z \in \mathcal{R}_Z$, and unobserved scalar $U \in \mathbb{R}$ satisfy:

$$Y_1 = \max(Y_1^*, 0), \qquad Y_1^* = m(Y_2, Z, U) \tag{1}$$

where the function m is strictly increasing in its third argument (U) and for all y_2 and z, $m(y_2, z, -\infty) \leq 0$. There is the inverse function $m^{-1}(y_2, z, y_1^*)$ such that for all y_1^*, y_2 and z

$$m(y_2, z, m^{-1}(y_2, z, y_1^*)) = y_1^*.$$

In a leading case of interest the function m is linear

$$m(y_2, z, u) = \alpha y_2 + \beta z + u,$$

as in the classical Tobit model which has $U \sim N(0, \sigma^2)$ independent of Z and no endogenous explanatory variables. In the linear case the inverse function is

$$m^{-1}(y_2, z, y_1^*) = y_1^* - \alpha y_2 - \beta z.$$

We consider cases in which the model embodies a parametric specification of the distribution of U, for example $U \sim N(0, \sigma^2)$, and cases in which there is no parametric specification. We consider models restricting U and Z to be stochastically independent, less restrictive specifications requiring quantile independence at specified quantiles, or, alternatively, mean independence restrictions.

We cast the problem into the Generalized Instrumental Variable framework set out in CR17 in which a structure, $(m, \mathcal{G}_{U|Z})$ comprises two components, namely (i) a structural function and (ii) a distribution of unobservable heterogeneity conditional on each possible value of exogenous variables. The first of these components, the structural function, determines which combinations of (Y, Z, U) can jointly occur. In the IV Tobit model this is fully determined by the function m.[5] The second component of a structure is a collection of conditional distributions of U given Z, denoted

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

where for any set $\mathcal{S} \subseteq \mathbb{R}$

$$G_{U|Z}(\mathcal{S}|z) \equiv \mathbb{P}[U \in \mathcal{S}|Z = z].$$

A model, A, comprises a list of restrictions on structures, defining a set of admissible structures, $\mathcal{M}_A$, which satisfy the restrictions. A model's restrictions may limit the dependence between U and Z and may require that the function m satisfies conditions additional to those so far imposed, for example functional form and exclusion restrictions. When U and Z are stochastically independent a collection $\mathcal{G}_{U|Z}$ is a singleton $\{G_U\}$ where G_U is the marginal distribution of U. A model can additionally impose parametric restrictions on the distribution of U.

In summary, throughout the paper a model is referred to as an *IV Tobit Model* if it satisfies the following definition.

Definition 1 *An* IV Tobit Model A *comprises a set of restrictions on the process generating observed variables* Y_1, Y_2, *and* Z *such that (1) holds for some unobservable variable* U *residing on the same probability space as* (Y_1, Y_2, Z). *The function* m *and conditional distributions of* U *given* Z *are restricted to belong to some set,* $\mathcal{M}_A$, *of admissible pairs* $(m, \mathcal{G}_{U|Z})$.

A variety of IV Tobit Models are considered, imposing different sets of restrictions on the pair $(m, \mathcal{G}_{U|Z})$. When convenient, notation $\tilde{G}_{U|Z}(t|z) \equiv G_{U|Z}((-\infty, t]\,|z)$ is used to denote the conditional cumulative distribution function of U given $Z = z$ associated with $G_{U|Z}(\cdot)$, and $\tilde{g}_{U|Z}(\cdot|z)$ is used to denote the corresponding conditional density when conditions are such that the density exists.

[5] In CR17 notation $h : \mathcal{R}_{YZU} \to \mathbb{R}$ is used to denote a structural function defined on the support of (Y, Z, U) such that $h(Y, Z, U) = 0$ with probability one. In the the IV Tobit model there is a unique mapping from values of (Y_2, Z, U) to values of Y and it is more natural to work with this function, m, directly. In the notation of CR17, one would define $h(Y, Z, U) \equiv Y_1 - \max(0, m(Y_2, Z, U))$.

3 Identification

3.1 Characterizations of identified sets

CR17 gives characterizations of identified sets of structures from which identified sets of structural features are obtained by projection. The characterizations make use of *residual sets* associated with a structure $(m, \mathcal{G}_{U|Z})$. A residual set is the set of values of unobserved U that, for a structural function m, can deliver the value y of endogenous Y when exogenous Z is equal to z. In the IV Tobit model the residual sets are singleton sets when $y_1 > 0$ and semi-infinite intervals when $y_1 = 0$, as follows.

$$\mathcal{U}(y, z, m) = \begin{cases} (-\infty, m^{-1}(y_2, z, 0)] & , \quad y_1 = 0 \\ \{m^{-1}(y_2, z, y_1)\} & , \quad y_1 > 0 \end{cases}$$

Let $\mathcal{F}_{Y|Z} \equiv \{F_{Y|Z}(\cdot|z) : z \in \mathcal{R}_Z\}$ denote the collection of conditional distributions of Y given Z, where for any set[6] $\mathcal{Y} \subseteq \mathcal{R}_Y$

$$F_{Y|Z}(\mathcal{Y}|z) \equiv \mathbb{P}[Y \in \mathcal{Y}|Z = z].$$

Data is informative about this collection of distributions, which is assumed to be identified. The following Proposition characterizes the identified set of structures given knowledge of $\mathcal{F}_{Y|Z}$.

Proposition 1 *The identified set of structures delivered by an IV Tobit model, A, and a collection of conditional distributions of Y given Z, $\mathcal{F}_{Y|Z} \equiv \{F_{Y|Z}(\cdot|z) : z \in \mathcal{R}_Z\}$ is*

$$\mathcal{I}(\mathcal{F}_{Y|Z}, \mathcal{R}_Z, A) = \mathcal{I}_1(\mathcal{F}_{Y|Z}, \mathcal{R}_Z, A) \cap \mathcal{I}_2(\mathcal{F}_{Y|Z}, \mathcal{R}_Z, A), \tag{2}$$

where

$$\mathcal{I}_1(\mathcal{F}_{Y|Z}, \mathcal{R}_Z, A) \equiv \left\{(m, \mathcal{G}_{U|Z}) \in \mathcal{M}_A : \forall t \in \mathbb{R}, G_{U|Z}((-\infty, t]|z) \geq C(z, t, m) \text{ a.e. } z \in \mathcal{R}_Z\right\}, \tag{3}$$

[6] To keep the notation uncluttered the support of Y is taken to be independent of z. This can easily be relaxed.

$$\mathcal{I}_2(\mathcal{F}_{Y|Z}, \mathcal{R}_Z, A) \equiv \big\{(m, \mathcal{G}_{U|Z}) \in \mathcal{M}_A : \forall [t_1, t_2] \subset \mathbb{R}, \\ G_{U|Z}([t_1, t_2]|z) \geq \Delta(z, t_1, t_2, m) \;\; a.e. \;\; z \in \mathcal{R}_Z \big\}, \quad (4)$$

$$\Delta(z, t_1, t_2, m) \equiv B(z, t_2, m) - B(z, t_1, m),$$

$$B(z, t, m) \equiv P[0 < Y_1 \leq m(Y_2, Z, t)|Z = z],$$

$$D(z, t, m) \equiv \mathbb{P}[Y_1 = 0 \wedge 0 \leq m(Y_2, Z, t)|Z = z],$$

and

$$C(z, t, m) \equiv P[Y_1 \leq m(Y_2, Z, t)|Z = z] = D(z, t, m) + B(z, t, m).$$

The proof of the Proposition follows from application of the inequality

$$G_{U|Z}(\mathcal{S}|z) \geq \mathbb{P}[\mathcal{U}(Y, Z, h) \subseteq \mathcal{S}|Z = z] \quad (5)$$

to sets $\mathcal{S}$ comprising certain intervals on the real line. The probability $\mathbb{P}[\mathcal{U}(Y, Z, h) \subseteq \mathcal{S}|Z = z]$, termed a containment probability, is the conditional probability given $Z = z$ of the occurrence of a value of Y that can occur only when unobserved U takes a value in the interval $\mathcal{S}$.[7] The expressions $\Delta(z, t_1, t_2, m)$, $B(z, t, m)$, $D(z, t, m)$, and $C(z, t, m)$ defined in the Proposition help to provide concise representations for these containment probabilities. Specifically, for intervals $(-\infty, t]$ that are unbounded below:

$$\mathbb{P}[\mathcal{U}(Y, Z, m) \subseteq (-\infty, t] \, |Z = z] = C(z, t, m). \quad (6)$$

For intervals $[t_1, t_2]$, with $t_1 > -\infty$ but with $t_2 \geq t_1$ unrestricted:

$$\mathbb{P}[\mathcal{U}(Y, Z, m) \subseteq [t_1, t_2]|Z = z] = \Delta(z, t_1, t_2, m), \quad (7)$$

from which it also follows that for intervals $[t_1, \infty)$ with $t_1 > -\infty$:

$$\mathbb{P}[\mathcal{U}(Y, Z, m) \subseteq [t_1, \infty)|Z = z] = P[Y_1 > 0|z] - B(z, t_1, m). \quad (8)$$

[7] Other characterizations of identified sets are available. One such will be employed when we consider the force of the restriction that unobserved U is mean independent of Z. All of the characterizations follow from the result that a structure $(m, G_{U|Z})$ is in the identified set if and only if for all z in the support of Z the distribution $G_{U|Z}(\cdot|z)$ is selectionable with respect to the distribution of the random set $\mathcal{U}(Y, Z; m)$ induced by the distribution of (Y, Z) delivered by the process under study. Definitions and details are in CR17.

In addition, it will be useful later to have an expression for the containment probability which applies for intervals $[t_1, t_2]$ with t_1 finite or infinite, as follows.

$$\mathbb{P}[\mathcal{U}(Y,Z,m) \subseteq [t_1,t_2]|Z=z] = 1[t_1=-\infty] \times \mathbb{P}[Y_1 = 0 \wedge 0 \leq m(Y_2,Z,t_2)|Z=z] + \\ \mathbb{P}[Y_1 > 0 \wedge m(Y_2,Z,t_1) \leq Y_1 \leq m(Y_2,Z,t_2)|Z=z] \quad (9)$$

3.2 Characterizations using singly-infinite systems of moment inequalities

The sets $\mathcal{I}_1(\mathcal{F}_{Y|Z}, \mathcal{R}_Z, A)$ and $\mathcal{I}_2(\mathcal{F}_{Y|Z}, \mathcal{R}_Z, A)$ are determined by systems of respectively singly- and doubly-infinite moment inequalities. Under additional restrictions that imply that $B(z,t,m)$ is everywhere differentiable in t the doubly-infinite system can be replaced with an equivalent singly-infinite system. This can have computational advantages.

With $\widetilde{G}_{U|Z}(\cdot|z)$ denoting the conditional cumulative distribution function of U given $Z=z$ the condition

$$G_{U|Z}([t_1,t_2]|z) \geq \Delta(z,t_1,t_2,m)$$

that appears in $\mathcal{I}_2(\mathcal{F}_{Y|Z}, \mathcal{R}_Z)$ can be expressed as

$$\widetilde{G}_{U|Z}(t_2|z) - \widetilde{G}_{U|Z}(t_1|z) \geq \Delta(z,t_1,t_2,m).$$

From this, with a differentiability restriction, Proposition 2 provides a singly-infinite moment inequality characterization for the set $\mathcal{I}_2(\mathcal{F}_{Y|Z}, \mathcal{R}_Z, A)$ originally defined in (4).

Proposition 2 *Suppose $m(y_2,z,t)$ is everywhere differentiable with respect to t for all values of (y_2,z) and that U is continuously distributed given Z with conditional density $\widetilde{g}_{U|Z}(\cdot|z)$. Then*

$$\mathcal{I}_2(\mathcal{F}_{Y|Z}, \mathcal{R}_Z, A) = \left\{(m, \mathcal{G}_{U|Z}) \in \mathcal{M}_A : \forall t \in \mathbb{R}, \widetilde{g}_{U|Z}(t|z) \geq b(z,t,m) \text{ a.e. } z \in \mathcal{R}_Z\right\}, \quad (10)$$

where $b(z,t,m) \equiv \nabla_t B(z,t,m)$ is the partial derivative of $B(z,t,m)$ with respect to t.

3.3 Upper and lower bounds and moment equalities

The containment inequality (5) used to produce Proposition 1 provides a lower bound on the distribution of U. It is shown in CR17 that applying the inequality in (5) to the complement

$\mathcal{S}^c$ of a set $\mathcal{S}$ delivers the following inequality satisfied by all structures in the identified set for all $z \in \mathcal{R}_Z$ and all closed sets $\mathcal{S}$ on the support of U.[8]

$$G_{U|Z}(\mathcal{S}|z) \leq \mathbb{P}[\mathcal{U}(Y,Z,m) \cap \mathcal{S} \neq \emptyset | Z = z] \tag{11}$$

So, for all structures in the identified set $\mathcal{I}(\mathcal{F}_{Y|Z}, \mathcal{R}_Z, A)$ the inequalities

$$\mathbb{P}[\mathcal{U}(Y,Z,m) \cap \mathcal{S} \neq \varnothing | Z = z] \geq G_{U|Z}(\mathcal{S}|z) \geq \mathbb{P}[\mathcal{U}(Y,Z,m) \subseteq \mathcal{S} | Z = z] \tag{12}$$

hold for all $z \in \mathcal{R}_Z$ and all intervals, $\mathcal{S}$, on the real line.

There are moment *equalities* in the characterization of the identified set of structures when there are z and $\mathcal{S}$ such that the probabilities on the left and right hand sides above are equal. In a sufficiently restrictive model and for particular collections of distributions $\mathcal{F}_{Y|Z}$ and support $\mathcal{R}_Z$, there is the possibility that these moment inequalities deliver point identification.

In the IV Tobit model the upper bounding probability $\mathbb{P}[\mathcal{U}(Y,Z,m) \cap \mathcal{S} \neq \emptyset | Z = z]$ is as follows.

$$\begin{aligned}\mathbb{P}[\mathcal{U}(Y,Z,m) \cap [t_1,t_2] \neq \varnothing | Z = z] = \mathbb{P}[(Y_1 = 0) \wedge (0 \geq m(Y_2,Z,t_1)) | Z = z] \\ + \mathbb{P}[(Y_1 > 0) \wedge (m(Y_2,Z,t_1) \leq Y_1 \leq m(Y_2,Z,t_2)) | Z = z].\end{aligned} \tag{13}$$

Considering (6), (7), and (13), bounding probabilities in (12) are equal for semi-infinite intervals $(-\infty, t]$ when z and m are such that

$$P[Y_1 = 0 | Z = z] = P[Y_1 = 0 \wedge 0 \leq m(Y_2,Z,t) | Z = z]$$

and for finite intervals $[t_1, t]$ when z and m are such that

$$P[Y_1 = 0 \wedge 0 \geq m(Y_2,Z,t) | Z = z] = 0.$$

Both conditions are satisfied when $m(Y_2,Z,t) > 0$ almost surely conditional on $Z = z$. A leading case in which this can occur is when the endogenous explanatory variable Y_2 has bounded support and the function m is unbounded above as t becomes large.

Conditions such that the inequalities (12) reduce to equalities for some values of z, can

[8] Here $\emptyset$ denotes the empty set. The probability on the right hand side is known as a capacity functional or the hitting probability of the random set.

be the basis for establishing sufficient conditions for point identification. For example, in models for censored outcomes with a conditional median restriction (Hong and Tamer, 2003, p. 908) provide support conditions under which certain resulting moment equalities can establish point identification and a $\sqrt{n}$-consistent and asymptotically normal estimator of model parameters when m is linear in parameters. Under the restriction that $med\,(U|Z) = 0$ and $m(Y_2, Z, 0) > 0$ almost surely conditional on $Z = z$ with m linear in parameters and additive U there is in our notation

$$med\,(Y_1 - Y_2\alpha + Z\beta|Z = z) = 0,$$

which corresponds to the moment equality delivered by the inequalities (12) applied to the set $(-\infty, t]$. A condition requiring that the set of values of $z \in \mathcal{R}_Z$ such that $m(Y_2, Z, 0) > 0$ almost surely conditional on $Z = z$ has positive measure, in conjunction with a condition requiring sufficient variation in included endogenous variables conditional on instruments, is then used to establish sufficient conditions for point identification in Lemma 2 of Hong and Tamer (2003).

3.4 Restrictions on the influence of exogenous Z

So far no restrictions on the dependence between U and Z or on the influence of z on the values taken by the function $m(y_2, z, u)$ have been considered. So far as the latter is concerned consider the restriction requiring m to depend on z solely through the variation in a function, $w(z)$, that arises as z varies across the support of Z.

Restriction ZD: Restricted Z dependence

$$\exists w(\cdot) \quad s.t. \quad \forall (z, z') \in \mathcal{R}_Z \times \mathcal{R}_Z, \forall (y_2, u), \quad w(z) = w(z') \implies m(y_2, z, u) = m(y_2, z', u)$$

A case which commonly arises involves exclusion restrictions with $z = (z_1, z_2)$ and $w(z) = z_1$. Index restrictions in which m only varies with z through variation in some parametric linear functions of z are also commonly employed.

Define the set of values that $w(z)$ can take as z varies across its support

$$\mathcal{W}(\mathcal{R}_Z) \equiv \{w(z) : z \in \mathcal{R}_Z\}$$

and for each element, w, of this set define the set of values of z such that $w(z) = w$:

$$\mathcal{Z}(w, \mathcal{R}_Z) \equiv \{z \in \mathcal{R}_Z : w(z) = w\}.$$

When Z is excluded from the structural function define $w(z) = z$ in which case $\mathcal{W}(\mathcal{R}_Z) = \mathcal{R}_Z$ and $\mathcal{Z}(w, \mathcal{R}_Z) = \mathcal{R}_Z$.

In the case in which there is a *stochastic independence condition* so that $\mathcal{G}_{U|Z} = \{G_U\}$ the sets $\mathcal{I}_1(\mathcal{F}_{Y|Z}, \mathcal{R}_Z, A)$ and $\mathcal{I}_2(\mathcal{F}_{Y|Z}, \mathcal{R}_Z, A)$ are then as follows, where $\widetilde{G}_U(t)$ and $\widetilde{g}_U(t)$ are respectively the marginal distribution and density functions of U.

$$\mathcal{I}_1(\mathcal{F}_{Y|Z}, \mathcal{R}_Z, A) = \left\{ (m, \mathcal{G}_{U|Z}) : \forall w \in \mathcal{W}(\mathcal{R}_Z),\ t \in \mathbb{R},\ \widetilde{G}_U(t) \geq \sup_{z \in \mathcal{Z}(w, \mathcal{R}_Z)} C(z, t, m) \right\},$$

$$\mathcal{I}_2(\mathcal{F}_{Y|Z}, \mathcal{R}_Z, A) = \left\{ \left(m, \mathcal{G}_{U|Z}\right) : \forall w \in \mathcal{W}(\mathcal{R}_Z),\ t \in \mathbb{R},\ \widetilde{g}_U(t) \geq \sup_{z \in \mathcal{Z}(w, \mathcal{R}_Z)} b(z, t, m) \right\},$$

and the identified set of structures is

$$\mathcal{I}(\mathcal{F}_{Y|Z}, \mathcal{R}_Z, A) = \mathcal{I}_1(\mathcal{F}_{Y|Z}, \mathcal{R}_Z, A) \cap \mathcal{I}_2(\mathcal{F}_{Y|Z}, \mathcal{R}_Z, A).$$ [9]

4 The impact of restrictions on the dependence between U and Z

In this section we consider the identifying power of a conditional mean independence restriction, a conditional quantile independence restriction focussing on median independence, a stochastic independence restriction with no parametric specification of the distribution G_U and stochastic independence restriction with U restricted to be Gaussian.

We present characterizations of identified sets for models in which $Y_1 = \max(0, m(Y_2, Z, U))$ and illustrate for the case in which $m(Y_2, Z, U) = \beta + \alpha Y_2 + U$. In the illustrations probabilities delivered by two specific structures are employed, as follows.[10]

[9] In order to deal with possibilities of zero measure sets and conditions required to hold almost everywhere, here and througout the paper the sup and inf operators are to be understood to mean "essential supreumum" and "essential infimum" when applied to functions of realizations of random variables. So for instance $\sup_{z \in \mathcal{Z}} f(z)$ indicates the smallest value c such that $f(Z) \leq c$ with probability one given $Z \in \mathcal{Z}$.

[10] In order to fully determine the conditional distributions of $Y|Z = z$ for each $z \in \mathcal{R}_z$ it is necessary to specify complete structures in which the process delivering Y_2 is specified. However, the equation relating Y_2 to Z and a stochastic unobservable is not a restriction used by the single equation IV Tobit model, and

- Structure 1

$$\begin{aligned} Y_1 &= \max(0, b + aY_2 + U_1) \\ Y_2 &= d_0 + d_1 Z + U_2 \end{aligned}$$

- Structure 2

$$\begin{aligned} Y_1 &= \max(0, b + aY_2 + U_1) \\ Y_2 &= g(Z, U_2) \\ g(Z, U_2) &= k - e^{(-d_0 - d_1 Z - U_2)} \end{aligned}$$

In both cases, for all $Z \in \mathcal{R}_Z$

$$U = \begin{bmatrix} U_1 \\ U_2 \end{bmatrix} \sim N_2 \left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} s_{11} & s_{12} \\ s_{12} & s_{22} \end{bmatrix} \right)$$

and in all the examples the support of scalar Z is

$$\mathcal{R}_Z = \{-1, -0.9, -0.8, \ldots, 0, \ldots, 0.8, 0.9, 1\},$$

It is important to understand that these are specifications of complete structures. The incomplete models that we consider do not employ all the restrictions that are embodied in the structures. In particular none of the models specify a structural equation for Y_2 and in only one of the models is U_1 restricted to be Gaussian.

A crucial feature of Structure 2 is that Y_2 is bounded above. Containment and other probabilities are calculated to high accuracy using numerical integration procedures.

4.1 Mean independence

First consider models in which U is restricted to be *mean independent* of the instrumental variables. Absent censoring such a model would be point identifying under a suitable rank condition.

RESTRICTION MI - Mean Independence: Let $\mathcal{G}_{U|Z}$ comprise all collections of conditional distributions for U given Z, $G_{U|Z}$, satisfying $E[U|z] = 0$, a.e. $z \in \mathcal{R}_Z$.

is thus not brought to bear in the identification analysis.

Table 1: Parameter values employed in numerical illustrations using Structures 1 and 2

	Structure 1			Structure 2
Parameter	Case 1	Case 2	Case 3	Case 1
b	1.0	1.0	1.0	0.0
a	1.0	1.0	1.0	1.0
k	–	–	–	1.0
d_0	0.5	0.25	0.0	0.0
d_1	1.0	0.5	0.25	1.0
s_{11}	2.0	1.5	1.0	1.0
s_{12}	0.5	0.375	0.25	0.5
s_{22}	1.0	1.0	1.0	1.0

Manski and Tamer (2002) used mean independence restrictions conditional on included exogenous variables in regressions with censored outcomes or covariates. Here we impose an IV version of a conditional mean restriction, conditioning an included exogenous variable and instruments. This is also Restriction MI in CR17 except that here, to simplify, the value of the conditional expectation is restricted to be zero rather than a member of a specified set of values. Modifying Theorem 5 of CR17 delivers the result that the identified set for structural function m comprises those functions m such that zero is an element of the Aumann expectation of $\mathcal{U}(Y, Z, m)$ conditional on $Z = z$ a.e. $z \in \mathcal{R}_Z$.

Proposition 3 *Under Restriction MI the identified set for the function m is*

$$\{m : 0 \leq E[m^{-1}(Y_2, Z, Y_1)|Z = z] \ a.e. \ z \in \mathcal{R}_Z\}.$$

In the case in which $m(y_2, z, u)$ is linear and equal to $\alpha y_2 + \beta z + u$ there is

$$m^{-1}(y_2, z, y_1^*) = y_1^* - \alpha y_2 - \beta z.$$

Define $\mathcal{W}(\beta)$, the set of values that βz can take as z varies across its support[11]

$$\mathcal{W}(\beta) \equiv \{\beta z : z \in \mathcal{R}_Z\}$$

[11] The set $\mathcal{W}(\beta)$ depends upon $\mathcal{R}_Z$ but this is not shown in the notation.

and for each element, w, of this set define the set of values of z such that $\beta z = w$:

$$\mathcal{Z}(w, \beta) = \{z \in \mathcal{R}_Z : \beta z = w\}.$$

The identified set for (α, β) is as follows.

$$\left\{(\alpha, \beta) : \forall w \in \mathcal{W}(\beta), \quad 0 \leq \inf_{z \in \mathcal{Z}(w,\beta)} (E[Y_1|z] - \alpha E[Y_2|z] - \beta z)\right\}$$

This is an intersection of linear half spaces and is therefore a convex set.

In the case in which there are no included exogenous variables in the structural equation for the censored outcome, so βz is simply a scalar intercept term denoted β, the identified set is as follows.

$$\left\{(\alpha, \beta) : \quad 0 \leq \inf_{z \in \mathcal{R}_Z} (E[Y_1|z] - \alpha E[Y_2|z] - \beta)\right\}. \tag{14}$$

Figure 1 shows an example of this set for Structure 1 with the parameter values shown in the column headed Case 1 in Table 1.

The identified set comprises the region below all of the blue drawn straight lines. The projection of the set onto the space of α is the entire real line. The projection of the set onto the space of β is the entire real line unless there exist z and z' in $\mathcal{R}_Z$ such that $E[Y_2|z] \leq 0 \leq E[Y_2|z']$ (a condition which holds with strong inequalities in the case pictured) in which case the projection is a semi-infinite interval with finite upper limit. The value of α and β in the structure that generates the probabilities used in the calculations is the green plotted point.

The identified sets under mean independence are similar under the other cases of Structure 1 and in Structure 2 as will be shown shortly.

Clearly a conditional *expectation* restriction does not lead to particularly informative identified sets. We now turn to consider the identifying power of conditional *quantile* restrictions. These can be much more informative.

4.2 Quantile independence

We now study the power of the following quantile independence restriction.

RESTRICTION QI - Quantile Independence: Let $\mathcal{G}_{U|Z}$ comprise all collections of conditional distributions of U given Z satisfying $\mathbb{P}[U \leq q_j|z] = \lambda_j$, a.e. $z \in \mathcal{R}_Z$ for all $j \in \mathcal{J} \equiv \{1, ..., J\}$ where $\Lambda \equiv \{\lambda_1, ..., \lambda_J\}$ is a collection of specified known values, and some collection of values $\{q_1, ..., q_J\} \in \mathcal{Q}$ with λ_j and q_j both increasing in j and $\mathcal{Q}$ a specified set of possible values of $\{q_1, ..., q_J\}$.

Restriction QI restricts the J conditional quantiles of U given $Z = z$ specified by Λ to be invariant with respect to z. $\mathcal{Q}$ is a known set of possible values for these conditional quantiles. For example, a conditional median restriction corresponds to $\Lambda = \{0.5\}$ and the usual normalization that this conditional median is zero is then captured by setting $\mathcal{Q} = \{0\}$. In this case $J = 1$. However, Restriction QI allows one to restrict the conditional distributions of U to be invariant at multiple quantiles. For instance, specifying $\Lambda = \{0.25, 0.5, 0.75\}$ and $\mathcal{Q} = \{\{q_1, q_2, q_3\} \in \mathbb{R}^3 : q_2 = 0\}$ constitutes a conditional quantile restriction at $J = 3$ quantiles, with the 0.5 quantile set to zero a typical location normalization. In many cases all but one (normalized) value q_j will be unrestricted in which case they can either be added to the list of unknown model parameters or treated as nuisance parameters.

Sharp characterization of the identified set of structures will require consideration of all test sets comprising intervals of the form $(-\infty, q_j]$ for all $j = 1, ..., J$ and $[q_j, q_{j+1}]$ for all $j = 0, ..., J$ where $q_0 = -\infty$ and $q_{J+1} \equiv \infty$.

Proposition 4 *Let Restriction QI hold. Then the identified set of structral functions delivered by the IV Tobit Model is the set of functions* $m \in \mathcal{M}$ *such that for some* $\{q_1, ..., q_J\} \in \mathcal{Q}$*: (1)* $\forall j \in \mathcal{J} \quad C(z, q_j, m) \leq \lambda_j$*, and (2)* $\forall j \in \{0, 1, ..., J\} \quad \Delta(z, q_j, q_{j+1}, m) \leq \lambda_{j+1} - \lambda_j$.

Note that if $\tilde{G}(t|z; m)$ and $B(z, t, m)$ are differentiable with respect to t then the final condition here is equivalent to

$$\tilde{g}(t|z, m) \geq b(z, t, m)$$

holding for all t and z where

$$\begin{aligned} \tilde{g}(t|z; m) &\equiv \nabla_t \tilde{G}(t|z; m) \\ b(z, t, m) &\equiv \nabla_t B(z, t, m). \end{aligned}$$

In the application to tobacco expenditure shares in Section 5.2 multiple quantile restrictions are considered as a means of relaxing the restriction that U have a Gaussian distribution independent of Z.

In numerical illustrations considered now Restriction QI is imposed encompassing a single median independence restriction such that $J = 1$ and $\lambda_1 = 0.5$ with the normalization $q_1 = 0$. Thus $\mathcal{Q}$ =$\{0\}$ and

$$\mathsf{Q} = \{(-\infty, 0), (0, \infty)\},$$

and for all $z \in \mathcal{R}_Z$

$$\begin{aligned} G_{U|Z}((-\infty, 0]|z) &= 0.5, \\ G_{U|Z}([0, \infty)|z) &= 0.5, \end{aligned}$$

leading respectively to the inequality from (6)

$$\mathbb{P}[Y_1 = 0 \wedge 0 \leq m(Y_2, Z, 0)|Z = z] + \mathbb{P}[Y_1 > 0 \wedge Y_1 \leq m(Y_2, Z, 0)|Z = z] \leq 0.5,$$

and the inequality from (8)

$$\mathbb{P}[Y_1 = 0|z] + \mathbb{P}[0 < Y_1 \leq m(Y_2, Z, 0)|Z = z] \geq 0.5$$

which together deliver the identified set of structural functions, m, under the median independence restriction, as follows.

$$\begin{aligned} \{m : \mathbb{P}[Y_1 = 0 \wedge 0 \leq m(Y_2, Z, 0)|Z = z] + \mathbb{P}[Y_1 > 0 \wedge Y_1 \leq m(Y_2, Z, 0)|Z = z] \\ \leq 0.5 \leq \\ \mathbb{P}[Y_1 = 0|z] + \mathbb{P}[0 < Y_1 \leq m(Y_2, Z, 0)|Z = z] \text{ a.e. } z \in \mathcal{R}_Z\}. \end{aligned}$$

In the case in which $m(y_2, z, u)$ is linear and equal to $\alpha y_2 + \beta + u$ the identified set of values of α and β is as follows.

$$\begin{aligned} \{(\alpha, \beta) : \mathbb{P}[Y_1 = 0 \wedge 0 \leq \alpha Y_2 + \beta|Z = z] + \mathbb{P}[Y_1 > 0 \wedge Y_1 \leq \alpha Y_2 + \beta|Z = z] \\ \leq 0.5 \leq \\ \mathbb{P}[Y_1 = 0|z] + \mathbb{P}[0 < Y_1 \leq \alpha Y_2 + \beta|Z = z] \text{ a.e. } z \in \mathcal{R}_Z\}. \end{aligned}$$

Identified sets of values of α and β under a zero median independence restriction QI are calculated and displayed for each of the lists of parameter values shown in the columns of Table 1 relating to Structure 1.

In each case the identified set obtained under a zero conditional mean independence restriction is shown as the region below all the blue drawn lines and the value of α and β in the structure generating the probabilities is the green plotted point.

Figure 2 shows the results using the parameter values in the column Case 1. These are the parameter values that deliver the results on the mean independence restriction shown in

Figure 1. The identified set obtained in this Case under the median independence restriction is the extremely small pink filled region. At the parameter values of Case 1 the median independence restriction is very powerful.

Figure 3 shows the results using the parameter values in the column Case 2 of Table 1. In this case the instrumental variable has lower predictive power for endogenous Y_2 because d_1 is closer to zero. The pink filled region is the identified set under the median independence restriction, much larger than in Case 1 but still very informative compared with the set obtained under the conditional mean independence restriction. In Case 3, shown in Figure 4 the identified set under the median independence restriction is larger still, containing unboundedly large positive values of α and unboundedly large negative values of β, but it remains informative relative to the identified set obtained under the mean independence restriction.

Figure 5 shows identified sets under mean independence and quantile independence restrictions using probabilities delivered by Structure 2 with the parameter values shown in the final column of Table 1. Also shown in Figure 5 is the projection onto the space of (α, β) of the identified set for (α, β, σ) obtained under a parametric Gaussian specification of the distribution of U in which the variance parameter is σ^2. This is a bounded set unlike the sets obtained at these parameter values under mean or quantile independence restrictions.

The median independence restriction delivers substantially smaller identified sets than the mean independence restriction in the cases studied and the identified sets it delivers can be very small indeed.

4.3 Stochastic independence

Now consider a restriction requiring U and Z to be independently distributed but with no parametric specification of the distribution of U.

RESTRICTION NPSI - Nonparametric Stochastic Independence: The random variables U and Z are independently distributed and $\mathcal{G}_{U|Z}$ is the singleton set $\{G_U\}$.

Recall the characterization of identified sets using singly infinite collections of inequalities given in (3) and (10), repeated here for convenience with the restriction NPSI imposed. The identified set of structures, $\mathcal{I}(\mathcal{F}_{Y|Z}, \mathcal{R}_Z, A)$, is the intersection of these two sets.

$$\mathcal{I}(\mathcal{F}_{Y|Z}, \mathcal{R}_Z, A) = \mathcal{I}_1(\mathcal{F}_{Y|Z}, \mathcal{R}_Z, A) \cap \mathcal{I}_2(\mathcal{F}_{Y|Z}, \mathcal{R}_Z, A) \tag{15}$$

$$\mathcal{I}_1(\mathcal{F}_{Y|Z},\mathcal{R}_Z,A) = \left\{(h,\mathcal{G}_{U|Z}) \in \mathcal{M}_A : \forall t \in \Re, \widetilde{G}_U(t) \geq \sup_{z\in\mathcal{R}_Z} C(z,t,m)\right\} \tag{16}$$

$$\mathcal{I}_2(\mathcal{F}_{Y|Z},\mathcal{R}_Z,A) = \left\{(h,\mathcal{G}_{U|Z}) \in \mathcal{M}_A : \forall t \in \Re, \widetilde{g}_U(t) \geq \sup_{z\in\mathcal{R}_Z} b(z,t,m)\right\} \tag{17}$$

Here $\widetilde{g}_U(t)$ is the probability density function of U, the first derivative of the distribution function $\widetilde{G}_U(t) \equiv G_U([-\infty,t])$, and $b(z,t,m) = \nabla_t B(z,t,m)$ where

$$B(z,t,m) \equiv \mathbb{P}[Y_1 > 0 \wedge Y_1 \leq m(Y_2,Z,t)|z]$$

and

$$C(z,t,m) = \mathbb{P}[Y_1 = 0 \wedge 0 \leq m(Y_2,Z,t)|z] + \mathbb{P}[Y_1 > 0 \wedge Y_1 \leq m(Y_2,Z,t)|z].$$

In the linear case used in the illustrations in which there is no exogenous variable in the structural equation, these functions are as follows.

$$B(z,t,\theta) \equiv \mathbb{P}[Y_1 > 0 \wedge Y_1 \leq \alpha Y_2 + \beta + t|z]$$

$$C(z,t,\theta) \equiv \mathbb{P}[Y_1 = 0 \wedge 0 \leq \alpha Y_2 + \beta + t|z] + \mathbb{P}[Y_1 > 0 \wedge Y_1 \leq \alpha Y_2 + \beta + t|z].$$

Rather than m we use $\theta = (\alpha,\beta)$ as an argument of these functions in this linear case. Absent additional restrictions on the distribution of U the distribution of observable variables contains no information about the value of the constant term, β, so, in determining the identified set of values of α, β can be set equal to an arbitrary value. It is set to zero in the numerical illustrations.

We now develop a method for calculating an outer set for the identified set of structures onto the space of structural functions under the NPSI restriction. For this purpose partition the support of U into N intervals: $(-\infty,t_1]$, $(t_1,t_2]$, $\ldots,(t_{N-1},\infty)$ where N is large. For each $n = 1,...,N$ define

$$p_n \equiv G_U\left((t_{n-1},t_n]\right), \tag{18}$$

where it is understood that $(t_0,t_1]$ means $(-\infty,t_1]$ and $(t_{N-1},t_N]$ means (t_{N-1},∞).

It follows from Proposition 1 that for any structural function m for which there exists a distribution G_U such that (m,g_U) is in the identified set, there must exist probabilities

$p_1, \ldots, p_N$ each nonnegative and summing to one such that

$$\forall n = 1, ..., N: \quad \sum_{i=1}^{n} p_i \geq C(z, t_n, m), \tag{19}$$

$$\forall n = 1, ..., N: \quad p_n \geq B(z, t_n, m) - B(z, t_{n-1}, m), \tag{20}$$

for almost every $z \in \mathcal{R}_Z$. Let $\mathcal{I}_M\left(\mathcal{F}_{Y|Z}, \mathcal{R}_Z, A, \mathcal{T}\right)$ denote the set of admissible m that satisfy these inequalities applied to $\mathcal{T} \equiv \{t_0, t_1, ..., t_N\}$. Inequalities of the form (19) correspond to those defining $\mathcal{I}_1(\mathcal{F}_{Y|Z}, \mathcal{R}_Z, A)$ in the statement of Proposition 1 with $t = t_n$. Inequalities of the form (20) are those characterizing $\mathcal{I}_1(\mathcal{F}_{Y|Z}, \mathcal{R}_Z, A)$ with t_{n-1} and t_n in place of t_1 and t_2, respectively.[12] Indeed, Proposition 1 implies that the identified set for m are those admissible structural functions that satisfy these inequalities for all t_n and t_{n-1} for some possible distribution G_U, with $G_U\left((t_{n-1}, t_n]\right)$ replacing p_n, almost surely. Indeed, it is precisely the use of only a finite set of intervals $(t_{n-1}, t_n]$ that makes the resulting characterization nonsharp.

For any such sequence of N intervals the conditions determining when structural function m is in $\mathcal{I}_M\left(\mathcal{F}_{Y|Z}, \mathcal{R}_Z, A, \mathcal{T}\right)$ can be checked by solving a linear program. We set out how this can be done using the method developed in Theorem 4 of Chesher and Rosen (2020a) in the study of interdependent determination of discrete outcomes. Define the following quantities.[13]

$$\underset{1\times N}{A} \equiv \underbrace{[1, 1, \ldots, 1]}_{N \text{ times}} \qquad \underset{N\times 1}{x} \equiv [p_1, \ldots, p_N]' \qquad b = 1 \tag{21}$$

$$\underset{R(2N-1)\times N}{\mathbf{B}} = \begin{bmatrix} \mathbf{B}^* \\ \vdots \\ \mathbf{B}^* \end{bmatrix} \qquad \underset{R(2N-1)\times 1}{c(m)} = \begin{bmatrix} c(z_1, m) \\ \vdots \\ c(z_R, m) \end{bmatrix} \qquad R = \sharp(\mathcal{R}_Z) \tag{22}$$

[12] These are equivalently conditional containment probabilities applied to intervals of the form $(-\infty, t_n]$ and $(t_{n-1}, t_n]$, respectively.

[13] This is set out in the notation in Chesher and Rosen (2020a) which employs matrix B, here denoted $\mathbf{B}$, not to be confused with the function $B(z, t, m)$.

$$\underset{(2N-1)\times N}{\mathbf{B}^*} = -\begin{bmatrix} 1 & 0 & \cdots & & 0 & 0 \\ 0 & 1 & \cdots & & 0 & 0 \\ \vdots & & \ddots & & \vdots & \vdots \\ \cdots & & & \ddots & & \\ 0 & 0 & \cdots & \cdots & 1 & 0 \\ 0 & 0 & \cdots & \cdots & 0 & 1 \\ 1 & 0 & 0 & \cdots & 0 & 0 \\ 1 & 1 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & & & \vdots & \vdots \\ 1 & 1 & 1 & \cdots & 0 & 0 \\ 1 & 1 & 1 & \cdots & 1 & 0 \end{bmatrix} \qquad \underset{(2N-1)\times 1}{c(z,m)} = \begin{bmatrix} B(z,t_0,m) - B(z,t_1,m) \\ B(z,t_1,m) - B(z,t_2,m) \\ \vdots \\ \vdots \\ B(z,t_{N-2},m) - B(z,t_{N-1},m) \\ B(z,t_{N-1},m) - B(z,t_N,m) \\ -C(z,t_1,m) \\ -C(z,t_2,m) \\ \vdots \\ -C(z,t_{N-2},m) \\ -C(z,t_{N-1},m) \end{bmatrix} \tag{23}$$

Application of Theorem 4 of Chesher and Rosen (2020a) to the characterization of $\mathcal{I}_M\left(\mathcal{F}_{Y|Z}, \mathcal{R}_Z, \mathcal{T}\right)$ given by (19) and (20) then yields the following result.

Proposition 5 *Let $\mathcal{T} \equiv \{t_0, t_1, ..., t_N\}$ be an increasing sequence of scalars with $t_N = -t_0 = \infty$. In an IV Tobit Model A restricting structural function m to the set $\mathcal{M}_A^0$ and in which Restriction NPSI holds, the set*

$$\mathcal{I}_M\left(\mathcal{F}_{Y|Z}, \mathcal{R}_Z, A, \mathcal{T}\right) = \left\{m \in \mathcal{M}_A^0 : v^*(m) \geq 0\right\}$$

comprises bounds on m, where $v^*(m)$ is the value attained by the linear program

$$\min_{s,t,v} v \tag{24}$$

subject to the constraints

$$sA + t\mathbf{B} \geq 0, \tag{25}$$

$$t \geq 0 \tag{26}$$

and

$$s + t \cdot c(m) \leq v, \tag{27}$$

where $s \in \mathbb{R}^1$, $t \in \mathbb{R}^K$, $v \in \mathbb{R}$, and $K = R(2N-1)$.[14]

[14] The proof of this proposition given in Chesher and Rosen (2020a) makes use of a version of Farkas' Alternative. Again we stay with the notation in Chesher and Rosen (2020a) which employs a decision variable t in R2, not to be confused with t_n used to signify boundaries of intervals that partition the support

The linear program of Proposition 5 is easy to check for any given structural function m given a choice of points $\mathcal{T}$ defining a partition of the real line. The outer set approximation to the identified set of structural functions can be made closer to the identified set by finer choice of $\mathcal{T}$. A large number of points $t_1, ..., t_N$ can be used because linear programs can be solved very quickly on modern computers. The outer set further has the interpretation of the sharp identified set under an additional restriction that the distribution G has a piecewise constant density on each of the N intervals defined by successive points in $\mathcal{T}$.

This procedure has been implemented using probabilities generated by the Gaussian triangular structures 1 and 2 in which the structural function m is parametrically specified and indexed by parameter θ.[15] The results obtained using Structures 1 and 2 are shown in respectively Table 2 and Table 3. In all cases the value of β is normalized to zero[16] since the location of G_U is unrestricted. The tables show parameter values of structures used to generate probabilities in these calculations, N, the number of intervals on the support of U, and the projection of the identified set onto the space of α which is an interval $[\underline{\alpha}, \overline{\alpha}]$. The set is an outer region because we approximate the distribution of U using a finite number of intervals. In addition the minimum and maximum conditional censoring probabilities $p_0(z) \equiv \mathbb{P}[Y = 0|Z = z]$ with respect to $z \in \mathcal{R}_Z$ are provided, as well as the marginal probability $p_0 \equiv \mathbb{P}[Y = 0]$ when Z is distributed with equal mass on each point of $\mathcal{R}_Z$.

These projections are, to the accuracy obtained in the calculations, population values of $[\underline{\alpha}, \overline{\alpha}]$, not estimates. This is so because the calculations employ probabilities delivered by the structures, not estimates of the probabilities. When applying this method to data, not done here, it will be necessary to account for sampling variation to perform inference. This is the topic of research in progress.

4.3.1 Structure 1

First consider the results shown in Table 2 using probabilities generated by Structure 1 which is a triangular Gaussian structure with Y_2 a linear function of Z and U_2 and so unbounded below and above. The support of Z is bounded and the model employed imposes the restriction that U and Z are stochastically independent.

of U.

[15] Linear programs are solved using R's (R Core Team (2020)) package `lpSolveAPI (Konis and Schwendinger (2020))`. Projections onto the space of α are calculated using R's `uniroot` function. Some of the programs are large. In Case 6 (see Table 2) there are $21 \times (2 \times 1000 - 1) + 2 = 41977$ decision variables and 1001 constraints. No calculation takes longer than a few minutes to run on an iMac with a 4.2GHz Intel Core i7 processor and 32 GB memory.

[16] This is a normalization - any value of β can be used.

Table 2: Projections of the identified set onto the space of α under a stochastic independence restriction and a nonparametric specification of the distribution of U with the number of intervals $N = 100$. Structure 1. Censoring probabilities p_0^ℓ, $\overline{p_0}$, and p_0^u denote $\min_{z \in \mathcal{R}_Z} \mathbb{P}[Y=0|Z=z]$, $\mathbb{P}[Y=0]$, and $\max_{z \in \mathcal{R}_Z} \mathbb{P}[Y=0|Z=z]$, respectively.

	Structure 1 parameter values							Censoring Probabilities			Interval for α	
Case	b	a	d_0	d_1	s_{11}	s_{12}	s_{22}	p_0^ℓ	$\overline{p_0}$	p_0^u	$\underline{\alpha}$	$\overline{\alpha}$
4	0.0	1	0	1	1	0.5	1	0.28	0.50	0.72	0.93	1.01
5	2.0	1	0	1	1	0.5	1	0.04	0.14	0.28	0.99	1.00
6	−1.0	1	0	1	1	0.5	1	0.50	0.71	0.88	0.84	∞
7	−1.0	1	0	2	1	0.5	1	0.28	0.68	0.96	0.97	1.00
8	−1.0	1	0	1.5	1	0.5	1	0.39	0.69	0.93	0.91	1.02
9	−1.5	1	0	1.5	1	0.5	1	0.50	0.78	0.96	0.86	∞

Table 3: Projections of the identified set onto the space of α under a stochastic independence restriction and a nonparametric specification of the distribution of U with the number of intervals $N = 50$. Structure 2. Censoring probabilities p_0^ℓ, $\overline{p_0}$, and p_0^u denote $\min_{z \in \mathcal{R}_Z} \mathbb{P}[Y=0|Z=z]$, $\mathbb{P}[Y=0]$, and $\max_{z \in \mathcal{R}_Z} \mathbb{P}[Y=0|Z=z]$, respectively.

	Structure 2 parameter values								Censoring Probabilities			Interval for α	
Case	b	a	k	d_0	d_1	s_{11}	s_{12}	s_{22}	p_0^ℓ	$\overline{p_0}$	p_0^u	$\underline{\alpha}$	$\overline{\alpha}$
1	0	1	1	0	1	1	0.5	1	0.34	0.54	0.76	0.38	1.77
2	10	1	1	0	1	1	0.5	1	0.00	0.02	0.09	0.98	1.00
3	0	1	10	0	1	1	0.5	1	0.00	0.03	0.11	0.98	1.00
4	0	1	1	0	5	1	0.5	1	0.16	0.57	1.00	1.00	1.00
5	0	1	1	0	1	1	0.0	0.1	0.27	0.55	0.92	1.00	1.00

In Case 4 the interval for α is short. Increasing N to 1000 reduces the interval but only slightly to $[0.96, 1.00]$. Case 5 has a large intercept $(b = 2)$ in the structural equation for Y_1 so the marginal probability of censoring is small (0.14). For some values of Z the conditional censoring probability $p_0(z)$ is indeed quite small, with a minimum of 0.04. For values of Z for which the conditional censoring probability is very small, the distribution of Y given $Z = z$ will be close to that produced by a linear IV model which is point identifying under a rank condition. Thus, while the identified set for α depends on the entire collection of conditional distributions $\mathcal{F}_{Y|Z}$, values of Z that produce small conditional censoring probabilities should be helpful in achieving narrow intervals for α. We see that in Case 5 the identified interval for α is indeed extremely short. By contrast in Case 6 the low value of the intercept $(b = -1)$ causes the probability of censoring to be high (0.71). The high intercept similarly induces higher conditional censoring probabilities than for Case 5, ranging from 0.50 to 0.88, and we observe that the identified interval is unbounded above.

In Case 7 the intercept is held at the Case 6 value but the effect of the instrument on endogenous Y_2 is much larger $(d_1 = 2)$. In this case the conditional censoring probability $\mathbb{P}[Y_1 = 0|Z = z]$ ranges from 0.28 to 0.96. Even though the marginal probability of censoring remains high (0.68), the identified interval is very short. In Case 8 the coefficient on the instrument in the equation for Y_2 is reduced from 2.0 to 1.5 and the identified interval lengthens slightly, but when the intercept is further reduced to -1.5 (Case 9) causing the probability of censoring to rise to 0.78 the identified interval for α becomes unbounded above.

4.3.2 Structure 2

Now consider the results obtained using probabilities generated by Structure 2 which is a triangular Gaussian structure with Y_2 a *nonlinear* function of Z and U_2, unbounded below and *bounded above* at the value of the parameter k.

The parameter values in Case 1 are the values delivering the identified sets of values of (α, β) drawn in Figure 5 for mean and quantile independence restrictions and for stochastic independence with U restricted Gaussian. The identified interval for α is $[0.38, 1.77]$, which is bounded unlike the sets obtained under mean and median independence restrictions. Increasing the number of intervals to $N = 100$ has no effect on the interval.

The intercept is increased in Case 2, from $b = 0$ to $b = 10$, reducing the conditional censoring probabilities to all lie between 0.00 and 0.09. This causes the length of the identified interval to become close to zero. The result is similar on returning the intercept to $b = 0$ and increasing the value of k to 10 as in Case 3. In Cases 4 and 5 it remains that $b = 0$

and we observe the effects of (i) increasing the coefficient on Z in the equation for Y_2 to $d_1 = 5$, and (ii) reducing the variance of the unobserved variable in the equation for Y_2 from $s_{22} = 1$ to $s_{22} = 0.1$, respectively. In these last two cases the marginal censoring probability is relatively high 0.54, but the instrument has great predictive power for Y_2. In these cases the minimum censoring probabilities p_0^ℓ are small and indeed the identified interval is very short, nearly a point.

4.4 Parametric restrictions - the Gaussian IV Tobit model

Finally, before coming to our empirical application, we consider models in which a stochastic independence restriction is imposed and, additionally, the distribution of U is restricted to belong to a parametric family of distributions. In the examples here a Gaussian distribution is employed as in the classic Tobit model.

RESTRICTION GaussSI - Stochastic Independence - Gaussian U: The random variables U and Z are independently distributed and $\mathcal{G}_{U|Z}$ is the singleton set $\{G_U\}$ where G_U is the $N(0, \sigma^2)$ distribution.

In these illustrations we use the characterization given in (3), (10) and (15) with $\widetilde{G}_U(\cdot)$ and $\widetilde{g}_U(\cdot)$ respectively the distribution and density function of a $N(0, \sigma^2)$ random variable. For Structures 1 and 2 at selected values of the parameters $\theta = (\alpha, \beta, \sigma)$ we calculate the value of the derivative $b(z, t, \theta)$ and the probability $C(z, t, \theta)$ and examine the inequalities over a long sequence of values of t pronouncing the value of θ in the identified set if none of the inequalities are violated. The sequences of values of t are

$$\{\nu + \xi\Phi^{-1}(\varepsilon) : \varepsilon \in \{1/n, \ldots, (n-1)/n\}\} \tag{28}$$

where Φ^{-1} is the standard normal quantile function. In the illustrations we set $\nu = 0$, $\xi = 2$ and $n = 500$.[17]

For the parameter values in Case 4 of Structure 1 shown in Table 2 the stochastic independence model is not point identifying but the Gaussian model is. We determine this by randomly sampling[18] points uniformly distributed on a sphere with small radius r centered at the probability generating value of θ, namely $(0, 1, 1)$. When the radius of the sphere is as small as 0.001 we find no points on the sphere that lie inside the identified set, which we

[17] Smaller values of n are used to determine the rough location and extent of an identified set and then results are refined using $n = 500$.

[18] We use the function `runif_on_sphere` available in the R package `uniformly`, Laurant (2018).

take to indicate point identification bearing in mind the slight inaccuracies arising because we are conducting finite precision arithmetic and in particular calculating some probabilities using numerical integration routines. The situation is the same if the value of the intercept, b, in Structure 1 is reduced to -0.5. However when $b$is reduced further to -1 (Case 6 in Table 2) there is no longer point identification under the Gaussian restriction.

Figure 6 shows projections of the identified set for (α, β, σ) onto the space of each pair of parameters in turn for Structure 2 in which Y_2 is bounded above. The solid filled regions are convex hulls of points found lying in the projections.[19] Parameter values for this case are those in Case 1 of Table 3.[20]

Figure 5 shows identified sets for (α, β) under mean and median independence restrictions at these parameter values and additionally the projection onto the space of (α, β) of the identified set for (α, β, σ) obtained under the Gaussian stochastic independence restriction (filled in magenta).[21] The identified interval for α under stochastic independence absent the Gaussian restriction, maintaining the stochastic independence condition is $[0.38, 1.77]$ - see Row 1 of Table 3. On additionally imposing the Gaussian restriction this is reduced to $[0.47, 1.42]$ as shown in Figure 5.

Calculations with both structures show that the IV Tobit model can be effectively point identifying and that when it is not the model can still be highly informative about the values of structural parameters. We now turn to an application involving real data and consider issues of estimation and inference.

5 Implementation

This section finishes with an application to a Tobit model of tobacco expenditure using UK survey data from the period 2000-2009. First we set out the method employed to calculate confidence regions on projections of the identified sets and we explain how implementation is done when, as in the application, exogenous variables have rich support.

[19] There is no evidence to suggest the projections are not convex. The parameter values whose membership of the identified set was evaluated were obtained in a trail and error process by randomly sampling points in spheres of varying radii centred on points lying centrally in the identified set. The function `runif_in_sphere` in R's `uniformly` package was employed.

[20] Also shown in the final column of Table 1.

[21] This is the projection shown in the (β, α) and (α, β) panes of Figure 6

5.1 Inference

We employ a procedure proposed in Belloni et al. (2018) (BBC18) to calculate 95% confidence regions for individual parameter components of partially identified parameter vectors. We use a self-normalized critical value, shown to be asymptotically valid in Chernozhukov et al. (2019) (CCK19). Belloni et al. (2018) additionally provide theoretical justification for alternative critical values using a bootstrap procedure that can further refine these confidence sets. We employ the self-normalized critical value for its computational simplicity.

In applications there will typically be many exogenous variables, Z, and some of these may be continuous. In this circumstance it is hard to make progress using conditional moment inequalities in which conditioning is on Z taking singleton values.[22] Instead we conduct inference on outer regions obtained from moment inequalities in which conditioning is on Z taking a value in a *set* of values, $\mathcal{Z}$.

In the models that we employ the moment inequalities arising when an interval $[t_1, t_2]$ is considered are

$$w(t_1, t_2, \theta, z) \leq p(t_1, t_2, \theta, z). \tag{29}$$

Here, employing (9),

$$\begin{aligned} w(t_1, t_2, \theta, z) \equiv 1[t_1 = -\infty] \times \mathbb{P}[Y_1 = 0 \wedge 0 \leq \alpha Y_2 + \beta Z + t_2)|Z = z] + \\ \mathbb{P}[Y_1 > 0 \wedge \alpha Y_2 + \beta Z + t_1 \leq Y_1 \leq \alpha Y_2 + \beta Z + t_2 | Z = z] \end{aligned}$$

and

$$p(t_1, t_2, \theta, z) \equiv G_{U|Z}([t_1, t_2]|z]$$

is the probability mass placed on the interval $[t_1, t_2]$ by the distribution of U given $Z = z$ admitted by the model. This may depend on the value of the parameter θ.

We estimate using the GaussSI model for which

$$p(t_1, t_2, \theta, z) = \Phi\left(\frac{t_2}{\sigma}\right) - \Phi\left(\frac{t_1}{\sigma}\right)$$

and we also estimate using two models in which the Gaussian restriction is dropped and quantile independence is imposed at 3, 5 or 7 quantiles associated with selected probabili-

[22] When Z has rich support it will be difficult to obtain accurate estimates of conditional probabilities. Kernel or sieve estimation would lead to estimated moment functions that are not simple means of contributions which is required when the BBC18 procedure is used.

ties. In these two cases, as set out at the start of Section 4.2, t_1 and t_2 are unknown values of quantiles at the selected specified probabilities and $p(t_1, t_2, \theta, z)$ is the difference between those probabilities, independent of z. In these two cases the unknown values of the quantiles are elements of the parameter vector θ. In all the cases considered in the application $p(t_1, t_2, \theta, z)$ does not depend on z which we make explicit now by writing the probability as $p(t_1, t_2, \theta)$.

Let the distribution function of Z be $F_Z(z)$. If for some value of θ, t_1 and t_2 the moment conditions (29) hold for all z there is, for all sets $\mathcal{Z} \subseteq \mathcal{R}_Z$

$$\int_{\mathcal{Z}} w(t_1, t_2, \theta, z) dF_Z(z) \leq p(t_1, t_2, \theta) \int_{\mathcal{Z}} dF_Z(z)$$

and thus the moment conditions imply that, with

$$\begin{aligned}\tilde{w}(t_1, t_2, \theta, \mathcal{Z}) \equiv 1[t_1 = -\infty] \times \mathbb{P}[Y_1 = 0 \wedge 0 \leq \alpha Y_2 + \beta Z + t_2) \wedge Z \in \mathcal{Z}] + \\ \mathbb{P}[Y_1 > 0 \wedge \alpha Y_2 + \beta Z + t_1 \leq Y_1 \leq \alpha Y_2 + \beta Z + t_2 \wedge Z \in \mathcal{Z}]\end{aligned}$$

the inequality

$$\tilde{w}(t_1, t_2, \theta, \mathcal{Z}) \leq p(t_1, t_2, \theta) \mathbb{P}[Z \in \mathcal{Z}]$$

holds for that θ and and the interval $[t_1, t_2]$ and for all sets $\mathcal{Z} \subseteq \mathcal{R}_Z$.

Introducing the notation employed in BBC18, let the functions of moments required to be nonpositive at a parameter value θ in an identified set be denoted $m_j(\theta)$, $j \in \mathcal{J}$.

The test set employed in constructing $m_j(\theta)$ is the interval $[t_1^{k(j)}, t_2^{k(j)}]$ where $\{k(j) : j \in \mathcal{J}\}$ is a list of indexes identifying the endpoints of intervals and a set of values of Z, $\mathcal{Z}^{l(j)}$, is employed where $\{l(j) : j \in \mathcal{J}.\}$ is a list of indexes identifying sets $\mathcal{Z}^{l(j)} \subseteq \mathcal{R}_Z$.[23]

A moment function $m_j(\theta)$ therefore has the following form.

$$\begin{aligned}m_j(\theta) = E\Big[1[t_1^{k(j)} = -\infty] \times 1[Y_1 = 0 \wedge 0 \leq \alpha Y_2 + \beta Z + t_2^{k(j)}) \wedge Z \in \mathcal{Z}_{l(j)}] + \\ 1[Y_1 > 0 \wedge \alpha Y_2 + \beta Z + t_1^{k(j)} \leq Y_1 \leq \alpha Y_2 + \beta Z + t_2^{k(j)} \wedge Z \in \mathcal{Z}_{l(j)}] \\ - p(t_1, t_2, \theta) 1[Z \in \mathcal{Z}_{l(j)}]\Big]\end{aligned}$$

[23] When selected quantile independence is imposed the unique unknown elements in $\{t^{k(j)}\}_{j=1}^{J}$ are parameters, elements of θ.

With data $\{(Y_{1i}, Y_{2i}, Z_i) : i \in \{1, \ldots, N\}\}$ define the estimator

$$\hat{m}_j(\theta) \equiv N^{-1} \sum_{i=1}^{N} m_{ji}(\theta)$$

where the contributions, $m_{ji}(\theta)$, are obtained by replacing Y_1, Y_2, and Z in the expression whose expectation appears in $m_j(\theta)$ by realized values, respectively Y_{1i}, Y_{2i}, and Z_i Define an estimator of the variance of $N^{1/2}\hat{m}_j(\theta)$

$$\hat{\sigma}_j^2(\theta) \equiv N^{-1} \sum_{i=1}^{N} \left(m_{ij}(\theta) - \hat{m}_j(\theta)\right)^2 .$$

Using the self-normalization-based critical value given in BBC18 and CCK19 the $100(1-\gamma)\%$ confidence region for the projection of the identified set onto the space of an element θ_k of θ is

$$CI(\theta_k, \gamma) \equiv \{r : T_{n,k}(r) \leq c_N(J, \gamma)\} \tag{30}$$

where

$$T_{n,k}(r) \equiv \inf_{\{\theta:\theta_k = r\}} \max_{j \in \{1,\ldots J\}} \left(N^{1/2} \frac{\hat{m}_j(\theta)}{\hat{\sigma}_j(\theta)} \right), \tag{31}$$

and where $c_N(J, \gamma)$ is the critical value:

$$c_N(J, \gamma) \equiv \frac{\Phi^{-1}(1 - \gamma/J)}{\sqrt{1 - \Phi^{-1}(1 - \gamma/J)^2/N}}.$$

with Φ denoting the standard Gaussian cumulative distribution function.

In implementation we obtain sets $\mathcal{Z}$ by specifying intervals for each exogenous variable and the connected unions of those intervals. The sets $\mathcal{Z}$ that we employ are the across-variable intersections of these connected unions. An example is given in the next section.

In the models that restrict quantile independence at specified quantile probabilities the test sets comprise all intervals with endpoints selected from the set of (unknown) values of the quantiles augmented with plus and minus infinity.

In models that impose the GaussSI restriction a collection of test sets (intervals) is obtained by choosing a large positive integer n and defining a sequence of n values:

$$\mathcal{W}(n, \nu, \xi) \equiv \{\nu + \xi\Phi^{-1}(\varepsilon) : \varepsilon \in \{0, 1/(n-1), \ldots, (n-2)/(n-1), 1\}\}$$

where Φ^{-1} is the standard normal quantile function. The values of (n, ν, ξ) are chosen to deliver good coverage of the main part of the support of U.[24] The collection of test sets is

$$\{[t_1, t_2] : t_1 < t_2 \wedge t_1 \in \mathcal{W}(n, \nu, \xi) \wedge t_2 \in \mathcal{W}(n, \nu, \xi)\}. \tag{32}$$

5.2 Application: tobacco expenditure share

In this section, inspired by Adams et al. (2019) (ABBC), we present confidence regions and estimates of parameters of models for the share of household nondurable expenditure spent on tobacco.

ABBC estimate a Tobit model for tobacco expenditures as a share of total nondurable expenditure with explanatory variables log total expenditure on nondurables (potentially endogenous), an OECD equivalence scale[25], with log household disposable income as an excluded instrumental variable.

The data used in ABBC come from the UK Family Expenditure Survey (FES) 1980-2000 and are a sample of households with head of household aged 25-35 years in 1980. We do not aim to reproduce their analysis, and while we use the same explanatory variables and data source, aiming for reasonably large samples we focus on households in the FES and its successor surveys[26] in 2000-09 with head of household aged 25-60 at the time of observation.[27] We conduct separate analysis of the periods 2000-04 and 2005-09 in which respectively 68% and 74% of households record zero tobacco expenditures.

ABBC take what is described as a quantile control function approach which employs a model in which nondurable expenditure is exogenous when there is conditioning on a control function which depends on nondurable expenditures, log household disposable income and the equivalence scale.[28] This control function restriction can arise in a complete triangular model in which there is a structural equation for log nondurable expenditure with, as explanatory variables, log disposable household income and the equivalence scale, and an unobserved variable which along with the unobserved variable in the structural equation for the tobacco expenditure share is jointly independent of the exogenous variables.

[24] In the application U is specified $N(0, 1)$ so this is easily done.

[25] The OECD equivalence scale is 1+0.7×the number of adults in excess of one + 0.5*×the number of children.

[26] These are, from 2001, the Expenditure and Food Survey and from 2008, the Living Costs and Food Survey.

[27] We exclude households comprising more than one tax unit, all of which have one or two adult members, and households with disposable weekly income recorded as £20 or less.

[28] Expenditures and income are recorded in UK pounds per week.

Here we employ the incomplete, single equation, IV model introduced in this paper. The model has the structural equation, as in ABBC,

$$Y_1 = \max(0, \beta_0 + \alpha Y_2 + \beta_1 Z_1 + \sigma U_1)$$

where Y_1 denotes tobacco expenditure share, Y_2 denotes log nondurable expenditure and there are exogenous variables Z_1 (the OECD equivalence scale), and Z_2 (log household disposable income), the latter excluded from the structural equation for the tobacco expenditure share. The model places no restriction on the process delivering Y_2. The partial correlation between log nondurable expenditure and log disposable household income given the OECD equivalence scale, $r_{Y_2 Z_2 . Z_1}$, is 0.64 in 2000-04 and 0.62 in 2005-09.

We calculate confidence regions for the value of α using the method set out in Section 5.1.[29] We first consider a model requiring U_1 to have a Gaussian distribution independent of $Z = (Z_1, Z_2)$ and then drop the Gaussian restriction and turn to models imposing conditional quantile independence restrictions.

We compare with estimates of α obtained using a classical Tobit model making no allowance for endogeneity and with estimates of α obtained using a complete, point identifying, triangular model in which there is the additional structural equation

$$Y_2 = \gamma_0 + \gamma_1 Z_1 + \gamma_2 Z_2 + U_2$$

and the restriction that (U_1, U_2) have a Gaussian distribution independent of Z_1 and Z_2.

In the calculations using the IV model, sets of values of the exogenous variables are obtained as follows. Lists of intervals of values of exogenous variables are constructed: for Z_1, the OECD equivalence scale,

$$\begin{bmatrix} 1.0 & 1.0 \\ 1.5 & 2.7 \\ 3.0 & 5.7 \end{bmatrix}$$

[29] Notice that in the equation for Y_1 the parameter σ multiplies U_1 which is restricted to be $N(0,1)$ when the GaussSI restriction is imposed. By this device we are able to define intervals $[t_1, t_2]$ that span the effective range of U_1 at all values of the parameters.

and for Z_2, log disposable household income,

$$\begin{bmatrix} Q(0) & Q(0.125) \\ Q(0.125) & Q(0.250) \\ \vdots & \vdots \\ Q(0.750) & Q(0.875) \\ Q(0.875) & Q(1) \end{bmatrix}$$

where $Q(p)$ is the sample p-quantile in the data under consideration. For each of Z_1 and Z_2 in turn a list of sets (denoted respectively Z_1 and Z_2) containing all connected unions of the intervals is constructed. The sets employed in the calculations are the sets in the collection

$$\{\{Z : Z_1 \in \mathcal{Z}_1 \wedge Z_2 \in \mathcal{Z}_2\} : \mathcal{Z}_1 \in \mathsf{Z}_1, \mathcal{Z}_2 \in \mathsf{Z}_2\}.$$

In the calculations reported here there are 216 such sets.

In analysis using the Gaussian model the test sets (intervals) employed comprise all intervals with endpoints chosen from the list: $(-\infty, -1.35, 0, 1.35, \infty)$, excluding the interval $(-\infty, \infty)$. There are 9 such intervals and combining test sets and Z sets leads to $1,944$ moment inequalities.

For each of the samples from the periods 2000-04 and 2005-09, respectively, there is no value of α found that satisfies all 1944 sample moment inequalities. Equivalently, there is no value of α such that $T_n(\alpha) = 0$.[30] Given statistical uncertainty due to sampling variation and the large number of moment inequalities, this is unsurprising. We thus report both half-median-unbiased set estimates and 95% confidence intervals for α using the inference method described in Section 5.1. The half-median-unbiased set estimates employ a conservative median-bias correction such that the upper (lower) endpoint of the interval is lower than (greater than) the upper (lower) bound of the population bounds for α with probability no greater than 1/2 asymptotically. These correspond to 50% confidence intervals, as advocated by Chernozhukov et al. (2013).

Results using the GaussSI restriction are shown in Table 4. For comparison results using two point identifying models are also provided. These are (1) maximum likelihood (ML) estimates of α and estimated standard errors obtained using a simple Gaussian Tobit model,

[30] Relative to equation (31) the second subscript k indexing the parameter component in $T_{n,k}(\cdot)$ has been dropped in a slight abuse of notation. The parameter component under consideration is to be understood to be α throughout.

Table 4: Estimates and confidence intervals for the coefficient on log total expenditure α in determination of tobacco expenditure share. IV estimates are half-median-unbiased interval estimates. Triangular and Tobit results obtained by maximum likelihood.

Years	N	%0	-	IV model GaussSI	Triangular model	Tobit model
2000-04	18473	68	*estimate*	$[-0.199, -0.117]$	-0.130	-0.048
			std err	–	(0.0043)	(0.0027)
			95% CI	$[-0.211, -0.114]$	$[-0.139, -0.122]$	$[-0.053, -0.043]$
2005-09	15885	74	*estimate*	$[-0.299, -0.103]$	-0.121	-0.043
			std err	–	(0.0046)	(0.0028)
			95% CI	$[-0.339, -0.096]$	$[-0.130, -0.112]$	$[-0.048, -0.038]$

which makes no allowance for endogeneity of nondurable expenditure, and (2) ML estimates of α obtained using the complete two equation triangular model with estimated standard errors and 95% confidence intervals.[31]

The Tobit estimates of α obtained without accounting for endogeneity are slightly less than one third of the magnitude of the estimates obtained using the complete triangular model. The hypothesis of exogeneity of log nondurable expenditure is soundly rejected with tests delivering p-values less than 0.0001.

The incomplete IV model delivers somewhat wider confidence regions than the complete triangular model. This is to be expected as the IV model uses only a subset of the restrictions used in the triangular model. The IV model estimates are robust to failure of the additional restrictions embodied in the complete triangular model and the IV model delivers encouragingly informative confidence regions. In both periods the 95% confidence regions delivered by the complete triangular model are subsets of the 95% confidence regions produced using the IV model.

Dropping the Gaussian restriction we imposed conditional quantile independence restrictions, in three distinct cases restricting quantiles of U_1 at three quantile probabilities $(0.25, 0.5, 0.75)$, five quantile probabilities $(0.167, 0.333, 0.5, 0.666, 0.833)$, and at seven quantile probabilities $(0.125, 0.25, 0.375, 0.5, 0.635, 0.75, 0.875)$, to be independent of the values of the exogenous variables, Z, in all cases normalizing the median by setting it equal to zero. As explained in Section 4.2 this is done by including as unknown parameters the unknown values of the quantiles at the nominated probabilities.

Restricting quantile independence at J quantile probabilities, $p_1, \ldots, p_J$, denote the values

[31] These estimates are calculated using the `Stata` 16 commands `tobit` and `ivtobit`, StataCorp (2019). The triangular two equation model estimated using the `ivtobit` command restricts the unobservable variables in the two structural equations to be jointly Gaussian, independent of Z_1 and Z_2.

Table 5: Half-median-unbiased interval estimates and 95% confidence intervals for the coefficient on log total expenditure α in determination of tobacco expenditure share under the GaussSI restriction and with quantile independence restrictions at 3, 5 and 7 quantiles.

Years	-	GaussSI	Quantile independence restricting:		
			3 quantiles	5 quantiles	7 quantiles
2000-04	*estimate*	$[-0.199, -0.177]$	$[-0.448, -0.078]$	$[-0.227, -0.093]$	$[-0.198, -0.108]$
	95% CI	$[-0.211, -0.114]$	$[-0.641, -0.070]$	$[-0.258, -0.091]$	$[-0.229, -0.103]$
2005-09	*estimate*	$[-0.299, -0.103]$	$(-\infty, -0.052]$	$(-\infty, -0.075]$	$[-0.492, -0.088]$
	95% CI	$[-0.339, -0.096]$	$(-\infty, -0.045]$	$(-\infty, -0.072]$	$(-\infty, -0.085]$

of the quantiles at these probabilities by $q_1, \ldots, q_J$. As set out in Proposition 4, the core determining collection of test sets comprises the collection of intervals:

$$\{(-\infty, q_j], \quad j \in \{1, \ldots, J\}\}$$

together with the collection of intervals

$$\{[q_j, q_{j+1}] : j \in \{1, \ldots, J-1\}\}.$$

In the estimations reported here *all connected unions* of these intervals were employed as test sets. We include additional test sets because it is possible that the functions of moments appearing in the inequalities delivered by some of the unions of core determining sets are more accurately estimated than the moment functions arising if only core determining test sets are employed. There is no great computational cost unless J is large and there is likely improved finite sample performance.

Placing conditional independence restrictions on 3, 5 and 7 quantiles leads to identified sets defined by respectively 1944, 4320 and 7560 moment inequalities. The parameters in the model employing quantile independence restrictions at J quantile probabilities are $\left(\beta_0, \beta_1, \alpha, q_1, \ldots, q_{(J+1)/2-1}, q_{(J+1)/2+1}, \ldots, q_J\right)$ the value of $q_{(J+1)/2}$ associated with $p_{(J+1)/2+1} = 0.5$ being normalized, set equal to zero.

The half-median-unbiased interval estimates and 95% confidence intervals obtained under the conditional quantile independence restrictions are shown in Table 5 which, for comparison, also includes the confidence regions obtained under the Gaussian stochastic independence restriction.

For the period 2000-04 the 95% confidence regions and the median-bias-corrected set estimates under conditional independence restrictions on 7 quantiles are only slightly longer then the regions obtained under the GaussSI restriction. Reducing the strength of the conditional quantile independence restriction, restricting just 5 quantiles leads to only a moderate increase in the length of the confidence regions, and even when independence restrictions are placed on just 3 quantiles the confidence regions remain informative. It seems that, in the 2000-04 period, restricting the distribution of U_1 to be Gaussian does not buy much over and above independence restrictions.

For the period 2005-09 the Gaussian restriction contributes much more. Absent the Gaussian restriction all but one of the confidence regions is unbounded below, but the upper bounds are all negative, so the value of α is effectively signed. In this period there are more zero tobacco expenditures and the GaussSI confidence regions are nearly twice as long as in 2000-04. There is more censoring, and there seems to be more noise in the process in 2005-09 and possibly some drift in the values of parameters.

6 Concluding remarks

When working with censored data and endogenous explanatory variables the easy way to obtain estimates of structural parameters is to employ a complete triangular model like the Gaussian model underlying STATA's `ivtobit` command[32] or to assume directly that a valid identifiable control function exists. When there is no economic rationale for such restrictions the IV model developed here provides a route to robust estimation. Even when more restrictive models are thought to be appropriate the IV model can deliver useful information regarding the force of additional restrictions. The IV model can signal misspecification of more restrictive models. It can deliver results when the complete model attack is not available, for example when endogenous variables are discrete or are determined by multiple sources of heterogeneity.

The IV model can be partially or point identifying and it may not be possible to determine identification status in particular applications. So it is important to use methods for estimation and inference that deliver results regardless of whether there is point or partial identification, as has been done in the application presented here.

There is rarely a good economic argument for particular parametric restrictions on the distribution of the unobserved variable U in the structural equation for a censored outcome. Nevertheless such restrictions are frequently imposed. We have shown how estimation and inference can be done using an IV Tobit model, dispensing with the commonly used Gaussian

[32] StataCorp (2019).

restriction, using multiple quantile independence restrictions, requiring the p-quantiles of U given instrumental variables, Z, to be independent of Z at selected quantile probabilities. The values of quantiles at the selected probabilities become parameters with unknown values and we calculate confidence regions on projections of the identified set of values of the augmented parameter vector onto the spaces of particular parameters of interest. In the application the procedure reveals the varying power of the Gaussian restriction.

We have shown how to calculate outer regions for identified sets of structural parameter values and their projections onto the space of individual parameters under a restriction requiring U and Z to be stochastically independent with no further restriction on the distribution of U. A parameter value lies in the identified set if and only if the solution to a linear program is nonnegative. Although the program can involve a very large number of inequalities the solution is quick to calculate. Conducting inference using this method is a research challenge not addressed here, and there remain other challenges. For example, we have proposed and applied a procedure for conducting inference on a partially identified parameter capturing the marginal effect of an endogenous variable on an outcome of interest when instrumental variables are continuously distributed, in which one calculates joint probabilities of events defined by sets of values of endogenous variables and sets of values of instrumental variables. Finite sample performance will of course depend on the sets that are chosen and future research to help guide these choices is warranted.

References

Adams, A., R. Blundell, M. Browning, and I. Crawford (2019): "Prices versus Preferences: Taste Change and Revealed Preference," IFS Working Paper.

Amemiya, T. (1979): "The Estimation of a Simultaneous Equation Tobit Model," *International Economic Review*, 20, 169–181.

Belloni, A., F. Bugni, and V. Chernozhukov (2018): "Subvector Inference in PI Models with with Many Moment Inequalities," ArXiv:1806.11466.

Blundell, R. and J. L. Powell (2007): "Censored Regression Quantiles with Endogenous Regressors," *Journal of Econometrics*, 141, 65–83.

Blundell, R. and R. J. Smith (1989): "Estimation in a Class of Simultaneous Equation Limited Dependent Variable Models," *Review of Economic Studies*, 56, 37–58.

Border, K. C. (2020): "Alternative Linear Inequalities," CalTech, version 2020.10.15::09.50.

Chernozhukov, V., D. Chetverikov, and K. Kato (2019): "Inference on Causal and Structural Parameters Using Many Moment Inequalities," *Review of Economic Studies*, 86, 1867–1900.

Chernozhukov, V., I. Fernandez-Val, and A. Kowalski (2015): "Quantile Regression with Censoring and Endogeneity," *Journal of Econometrics*, 186, 201–221.

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

Chesher, A. and A. Rosen (2017): "Generalized Instrumental Variable Models," *Econometrica*, 85, 959–989.

Chesher, A. and A. M. Rosen (2020a): "Econometric Modeling of Interdependent Discrete Choice," CeMMAP working paper CWP25/20.

——— (2020b): "Generalized Instrumental Variable Models, Methods, and Applications," in *The Handbook of Econometrics*, ed. by S. Durlauf, L. P. Hansen, J. J. Heckman, and R. Matzkin, Elsevier, vol. 7a.

Das, M. (2002): "Estimation and Inference in a Censored Regression Model with Endogenous Covariates," Columbia Department of Economics Discussion Paper 0102-61.

Frandsen, B. R. (2015): "Treatment Effects with Censoring and Endogeneity," *Journal of the American Statistical Association*, 110, 1745–1752.

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

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

Khan, S., M. Ponomareva, and E. Tamer (2011): "Sharpness in Randomly Censored Linear Models," *Economics Letters*, 113, 23–25.

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

Konis, K. and F. Schwendinger (2020): "lpSolveAPI: R Interface to 'lp_solve' Version 5.5.2.0. R package version 5.5.2.0-17.7." `https://CRAN.R-project.org/package=lpSolveAPI`.

Lancaster, T. (1985): "Simultaneous Equations Models in Applied Search Theory," *Journal of Econometrics*, 28, 113–126.

Lancaster, T. and A. Chesher (1984): "Simultaneous Equations with Endogenous Hazards," in *Studies in Labor Market Dynamics, Studies in Contemporary Economics, vol 11.*, ed. by G. Neumann and N. Westergå rd Nielsen, Springer, Berlin, Heidelberg.

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

Molchanov, I. and F. Molinari (2018): *Random Sets in Econometrics*, Cambridge, U.K.: Cambridge University Press.

Nelson, F. and L. Olson (1978): "Specification and Estimation of a Simultaneous Equation Model with Limited Dependent Variables," *International Economic Review*, 19, 695–709.

Newey, W. K. (1987): "Efficient Estimation of Limited Dependent Variable Models with Endogenous Explanatory Variables," *Journal of Econometrics*, 36, 231–250.

Olsen, R. J. and G. Farkas (1989): "Endogenous Covariates in Duration Models and the Effect of Adolescent Childbirth on Schooling," *Journal of Human Resources*, 24, 39–53.

R Core Team (2020): *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria, ISBN 3-900051-07-0.

Smith, R. J. and R. Blundell (1986): "An Exogeneity Test for a Simultaneous Equation Tobit Model with an Application to Labor Supply," *Econometrica*, 54, 679–685.

StataCorp (2019): "Stata Statistical Software: Release 16," College Station, TX: StataCorp LLC.

Tobin, J. (1958): "Estimation of Relationships for Limited Dependent Variables," *Econometrica*, 26, 24–36.

Wrenn, D. H., H. A. Klaiber, and D. A. Newburn (2017): "Confronting Price Endogeneity in a Duration Model of Residential Subdivision Development," *Journal of Applied Econometrics*, 32, 661–682.

Appendix: Proofs and Figures

Proof of Proposition 1. Applying Theorem 1, Corollary 1 of CR17 gives the following characterization of the identified set of structures delivered by a model, A, a collection of conditional distributions of Y given Z and the support of Z:

$$\mathcal{I}(\mathcal{F}_{Y|Z},\mathcal{R}_Z,A)=\big\{(m,\mathcal{G}_{U|Z})\in\mathcal{M}_A : \\ \forall\mathcal{S}\in\mathsf{F}(\mathcal{R}_U),G_{U|Z}(\mathcal{S}|z)\geq\mathbb{P}[\mathcal{U}(Y,Z,m)\subseteq\mathcal{S}|Z=z]\text{ a.e. } z\in\mathcal{R}_Z\big\}, \quad (33)$$

where $\mathsf{F}(\mathcal{R}_U)$ denotes the collection of closed subsets of $\mathcal{R}_U$. By Lemma 1 of CR17 it follows that the requirement that the inequality holds for all closed sets $\mathcal{S}$ can be replaced by the requirement that it holds for all $\mathcal{S}$ that are unions of sets on the support of $\mathcal{U}(Y,Z,m)$ conditional on $Z=z$. Each such set can be be written as a union of sets of the form $(-\infty,t]$ and $[t_1,t_2]$, where if $t_1=t_2=t$, the set $[t_1,t_2]$ is simply the point $\{t\}$. All such unions are themselves either of the form $\mathcal{S}=(-\infty,t]$ or $\mathcal{S}=[t_1,t_2]$. The collections $\mathcal{I}_1(\mathcal{F}_{Y|Z},\mathcal{R}_Z,A)$ and $\mathcal{I}_2(\mathcal{F}_{Y|Z},\mathcal{R}_Z,A)$ comprise those structures satisfying $G_{U|Z}(\mathcal{S}|z)\geq\mathbb{P}[\mathcal{U}(Y,Z,m)$ for each of these two types of sets $\mathcal{S}$, respectively, completing the proof. ■

Proof of Proposition 2. Existence of $b(z,t,m)\equiv\nabla_t B(z,t,m)$ follows from differentiability of $m(y_2,z,t)$ with respect to t and the existence of the density $\widetilde{g}_{U|Z}(\cdot|z)$. The inequality

$$\widetilde{G}_{U|Z}(t_2|z)-\widetilde{G}_{U|Z}(t_1|z)\geq B(z,t_2,m)-B(z,t_1,m).$$

can then be expressed as

$$\int_{t_1}^{t_2}(\widetilde{g}_{U|Z}(t|z)-b(z,t,m))dt\geq 0,$$

which for any z holds for all $[t_1,t_2]\subseteq\mathbb{R}$ if and only if for all t, $\widetilde{g}_{U|Z}(t|z)\geq b(z,t,m)$. ■

Proof of Proposition 3. Theorem 5 of CR17 implies that identified set for structural function m comprises those functions m such that zero is an element of the Aumann expectation of $\mathcal{U}(Y, Z, m)$ conditional on $Z = z$ a.e. $z \in \mathcal{R}_Z$. Recall that the residual set in the model under study is

$$\mathcal{U}(y, z, m) = \begin{cases} (-\infty, m^{-1}(y_2, z, 0)] & , \quad y_1 = 0 \\ \{m^{-1}(y_2, z, y_1)\} & , \quad y_1 > 0 \end{cases}$$

and let $\mathbb{E}[\mathcal{A}|z]$ denote the Aumann expectation of random set $\mathcal{A}$ conditional on $Z = z$.[33] There is

$$\mathbb{E}[\mathcal{U}(Y, Z, m)|z] = \mathbb{E}[\mathcal{U}(Y, Z, m)|z, Y_1 = 0]P[Y_1 = 0|z] + \mathbb{E}[\mathcal{U}(Y, Z, m)|z, Y_1 > 0]P[Y_1 > 0|z]$$

where the sum is a Minkowski sum.[34] Considering each term in turn there is[35]

$$\mathbb{E}[\mathcal{U}(Y, Z, m)|z, Y_1 = 0] = (-\infty, E[m^{-1}(Y_2, Z, 0)|z, Y_1 = 0]]$$

which is a semi-infinite interval and there is[36]

$$\mathbb{E}[\mathcal{U}(Y, Z, m)|z, Y_1 > 0] = \{E[m^{-1}(Y_2, Z, Y_1)|z, Y_1 > 0]\}$$

which is a singleton. The Minkowski sum of a semi-infinite interval and a singleton set is a semi-infinite interval. The result is that the conditional (on Z) Aumann expectation of the residual set is the semi-infinite interval

$$\mathbb{E}[\mathcal{U}(Y, Z, m)|z] = (-\infty, E[m^{-1}(Y_2, Z, Y_1)|z]]$$

which leads directly to the result of the Proposition. ■

[33]The Aumann expectation of a random set $\mathcal{A}$ is the set of expected values of all random variables A with finite expected values having the property that with probability one $A \in \mathcal{A}$.

[34]The Minkowski sum of sets A and B is the set of values obtained by adding each element of A to each element of B.

$$A + B = \{a + b : a \in A, b \in B\}$$

[35]See Example 3.14 in Molchanov and Molinari (2018).

[36]See Example 3.12 in Molchanov and Molinari (2018).

Proof of Proposition 4. From Proposition 1 the identified set for $(m, \mathcal{G}_{U|Z})$ is

$$\mathcal{I}(\mathcal{F}_{Y|Z}, \mathcal{R}_Z, A) = \mathcal{I}_1(\mathcal{F}_{Y|Z}, \mathcal{R}_Z, A) \cap \mathcal{I}_2(\mathcal{F}_{Y|Z}, \mathcal{R}_Z, A).$$

It will be shown that under Restriction QI the identified set for structural function m, which is the projection of $\mathcal{I}(\mathcal{F}_{Y|Z}, \mathcal{R}_Z, A)$ onto $\mathcal{M}$, here denoted $\mathcal{M}^*$, is equivalent to the set of functions $m \in \mathcal{M}$ that satisfy conditions (1) and (2) in the statement of the Proposition.

Suppose first that $m \in \mathcal{M}^*$ such that for some $\mathcal{G}_{U|Z}$ satisfying Restriction QI $(m, \mathcal{G}_{U|Z}) \in \mathcal{I}(\mathcal{F}_{Y|Z}, \mathcal{R}_Z, A)$. Conditions (1) and (2) then hold because these are implied by the inequalities that define the sets $\mathcal{I}_1(\mathcal{F}_{Y|Z}, \mathcal{R}_Z, A)$ and $\mathcal{I}_2(\mathcal{F}_{Y|Z}, \mathcal{R}_Z, A)$, respectively.

Now suppose that m satisifies conditions (1) and (2) for some $\{q_1, ..., q_J\} \in \mathcal{Q}$. To show that $m \in \mathcal{M}^*$ it will be shown by construction that there exists a collection of conditional distributions $\mathcal{G}_{U|Z}(\cdot|z;m)$ with cumulative distribution functions $\tilde{G}(\cdot|z;m) = \tilde{G}_{U|Z}(\cdot|z;m)$ for each $z \in \mathcal{R}_Z$ satisfying Restriction QI such that $(m, \mathcal{G}_{U|Z}(\cdot|z;m)) \in \mathcal{I}(\mathcal{F}_{Y|Z}, \mathcal{R}_Z, A)$. The inclusion of m in the notation $\mathcal{G}_{U|Z}(\cdot|z;m)$ signifies that the associated collection of conditional distributions $\mathcal{G}_{U|Z}(\cdot|z) = \mathcal{G}_{U|Z}(\cdot|z;m)$ in the construction will in general vary with m.

Specifically, it needs to be shown that for almost every $z \in \mathcal{R}_Z$ there exists a cumulative distribution function $\tilde{G}(\cdot|z;m)$ such that the following three conditions (34)–(36) hold. For all $j \in \mathcal{J}$:

$$\tilde{G}(q_j|z;m) = \lambda_j. \tag{34}$$

For all $t \in \mathbb{R}$:

$$\tilde{G}(t|z;m) \geq C(z,t,m). \tag{35}$$

For all $s \leq t$, each in $\mathbb{R}$:

$$\tilde{G}(t|z;m) - \tilde{G}(s|z;m) \geq \Delta(z,s,t,m). \tag{36}$$

Condition (34) ensures that Restriction QI holds and conditions (35) and (36) are the conditions defining the identified set for $(m, \mathcal{G}_{U|Z})$, as shown in Proposition 1. Note that for (36) it is equivalent to show that

$$\tilde{G}(t|z;m) - B(z,t,m) \tag{37}$$

is weakly increasing in t.

Construction of $\tilde{G}(t|z;m) : \mathbb{R} \to [0,1]$ for each z, m is as follows, divided into separate cases according to where argument t lies relative to $q_1, ..., q_J$.

1. **Case 1**: $t \in (-\infty, q_1]$. Define

$$\tilde{G}(t|z;m) \equiv C(z,t,m) + (\lambda_1 - C(z,q_1,m)) \exp(\eta(t-q_1)),$$

where $\eta > 0$ is arbitrary.[37] Since $\lim_{t\to-\infty} C(z,t,m) = 0$ and $\lim_{t\to-\infty} \exp(\eta(t-q_1)) = 0$ it follows that $\lim_{t\to-\infty} \tilde{G}(t|z;m) = 0$. There is also $\tilde{G}(q_1|z;m) = \lambda_1$. $\tilde{G}(t|z;m)$ is an increasing function of t because it is the sum of two increasing functions of t, $\tilde{G}(t|z;m) \geq C(z,t,m)$ by definition and

$$\tilde{G}(t|z;m) - B(z,t,m) = D(z,t,m) + (\lambda_2 - C(z,q_2,m)) \exp(\eta(t-q_2))$$

is an increasing function of t because it is the sum of two increasing functions of t.

2. **Case 2**: $t \in [q_j, q_{j+1}]$, each $j = 1, ..., J-1$. Define

$$L_j(z,t,m) \equiv B(z,t,m) + \lambda_j - B(z,q_j,m)$$

which is an increasing function of t with $L_j(z,t,m) - B(z,t,m)$ constant and $L_j(z,q_j,m) = \lambda_j$. Condition (2) ensures

$$\lambda_{j+1} - B(z,q_{j+1},m) \geq \lambda_j - B(z,q_j,m),$$

from which it follows that $L_j(z,q_{j+1},m) \leq \lambda_{j+1}$. Define

$$M_j(z,t,m) \equiv C(z,t,m) + (\lambda_{j+1} - C(z,q_{j+1},m)) \frac{(t-q_j)}{(q_{j+1}-q_j)}.$$

This is an increasing function of t with $M_j(z,t,m) \geq C(z,t,m)$, $M_j(z,q_j,m) = C(z,q_j,m)$, and $M_j(z,q_{j+1},m) = \lambda_{j+1}$. Define

$$\tilde{G}(t|z;m) = \max(L_j(z,t,m), M_j(z,t,m)).$$

There is

$$\tilde{G}(q_j|z;m) = \max(\lambda_j, C(z,q_j,m)) = \lambda_j,$$

[37] Construction of $\tilde{G}(t|z;m)$ employing functions other than the exponential function could also be used.

and

$$\tilde{G}(q_{j+1}|z;m) = \max(L_j(z, q_{j+1}, m), \lambda_{j+1}) = \lambda_{j+1},$$

because as just shown, $L_j(z, q_{j+1}, m) \leq \lambda_{j+1}$. This is an increasing function of t in the interval $[q_j, q_{j+1}]$ because it is the maximum of two increasing functions of t, and $\tilde{G}(t|z;m) \geq C(z,t,m)$ because $\tilde{G}(t|z;m)$ is the maximum of two functions one of which is at least equal to $C(z,t,m)$ in the interval under consideration. Moreover,

$$\tilde{G}(t|z;m) - B(z,t,m) = \max\left(\lambda_j - B(z, q_j, m), D(z,t,m) + (\lambda_{j+1} - C(z, q_{j+1}, m))\frac{(t-q_j)}{(q_{j+1}-q_j)}\right)$$

which is increasing in t because it is the maximum of two increasing functions of t, verifying condition (37).

- **Case 3**: $t \in [q_J, \infty)$. Define

$$\tilde{G}(t|z;m) \equiv \max\left(C(z,t,m), B(z,t,m) - B(z,q_J,m) + \lambda_J\right).$$

There is

$$\tilde{G}(q_J|z;m) = \max(C(z,q_J,m), \lambda_J) = \lambda_J,$$

and then

$$\lim_{t\to\infty} \tilde{G}(t|z;m) = \max(1, B(z,\infty,m) - B(z,q_J,m) + \lambda_J) = 1$$

because from condition (2)

$$\lambda_J - B(z,q_J,m) \leq \lambda_{J+1} - B(z,q_{J+1},m)$$

which implies

$$B(z,q_{J+1},m) - B(z,q_J,m) + \lambda_J \leq \lambda_{J+1} = 1.$$

$\tilde{G}(t|z;m)$ is an increasing function of t because it is the maximum of two increasing functions of t. Finally

$$\tilde{G}(t|z;m) - B(z,t,m) = \max\left(D(z,t,m), -B(z,q_J,m) + \lambda_J\right)$$

which is an increasing function of t.

We have shown that for any $z \in \mathcal{R}_Z$ and any function m satisfying the conditions of the Proposition the piecewise function $\tilde{G}(t|z,m)$ defined above can be constructed and we have shown that it satisfies conditions (34), (35) and (36) Therefore any function m satisfying the conditions of the Proposition is contained in the identified set of structural functions delivered by the model. ■

Proof of Proposition 5. Suppose that m is in the projection of the identified set of structures $\left(m, \mathcal{G}_{U|Z}\right)$ deliverd by the IV Tobit Model A under consideration. Then under Restriction NPSI there exists a distribution G such that $(m, \{G\}) \in \mathcal{I}(\mathcal{F}_{Y|Z}, \mathcal{R}_Z, A)$ and as explained in the text it follows from Proposition 1 that (19) and (20) hold with $p_1, ..., p_N$ as defined in (18) as a function of that distribution G. The existence of a vector of proper probabilities $\mathbf{p} = (p_1, ..., p_N)$ such that (19) and (20) hold almost surely is equivalent to the existence of $\mathbf{p} \in \mathbb{R}_N$ satisfying

$$\begin{aligned} A\mathbf{p} &= b, \\ \mathbf{B}p &\leq c\,(m)\,, \\ \mathbf{p} &\geq 0, \end{aligned}$$

a linear program in $\mathbf{p}$. Then, applying the same steps as in Chesher and Rosen (2020a), and in particular using the version of Farkas's Alternative provided in Border (2020) – see paragraph 12, Section 1.4 – such probabilities exist *if and only if* there is *no* solution for (s,t), $s \in \mathbb{R}$, $t \in \mathbb{R}^K$, $v \in \mathbb{R}$, to the system

$$sA + t\mathbf{B} \geq 0,$$

$$t \geq 0,$$

$$s + t \cdot c(m) < 0,$$

which is equivalent there being a nonnegative value of the linear program (24) subject to (25), (26), and (27). ■

Figure 1: Structure 1, Y_2 linear in Z and U_2. The identified set of values of α and β in a linear IV Tobit model under a zero conditional expectation restriction, $E[U|z] = 0$ for all z. The identified set comprises the region below all the drawn lines whose formula is $\beta = E[Y_1|z] - \alpha E[Y_2|z]$ for $z \in \{-1, -0.9, \ldots, 0.9, 1\}$. The value of a and b generating the probabilities used to calculate the set is $a = 1$, $b = 1$, plotted in green. The parameter value used to calculate the conditional expectations is Case 1 in Table 1.

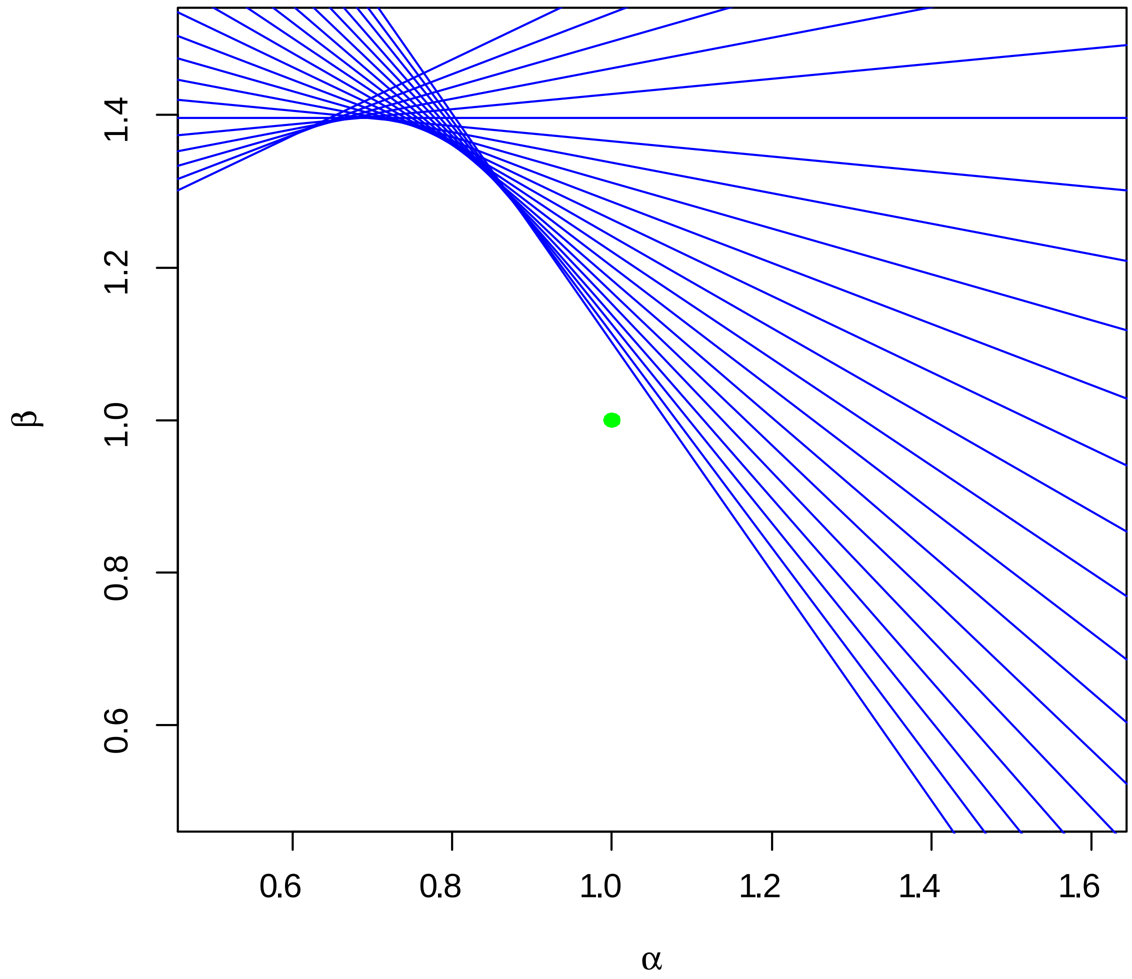

Figure 2: Structure 1, Y_2 linear in Z and U_2. The pink filled region is the identified set of values of α and β in a linear IV Tobit model under a zero conditional median restriction. The region below the blue drawn lines is the identified set under a zero conditional mean restriction. The parameter values in the triangular Gaussian structure generating the probabilities used in this calculation are shown in the column Case 1 in Table 1. The value of a and b generating the probabilities used to calculate the set is $a = 1$, $b = 1$, plotted in green.

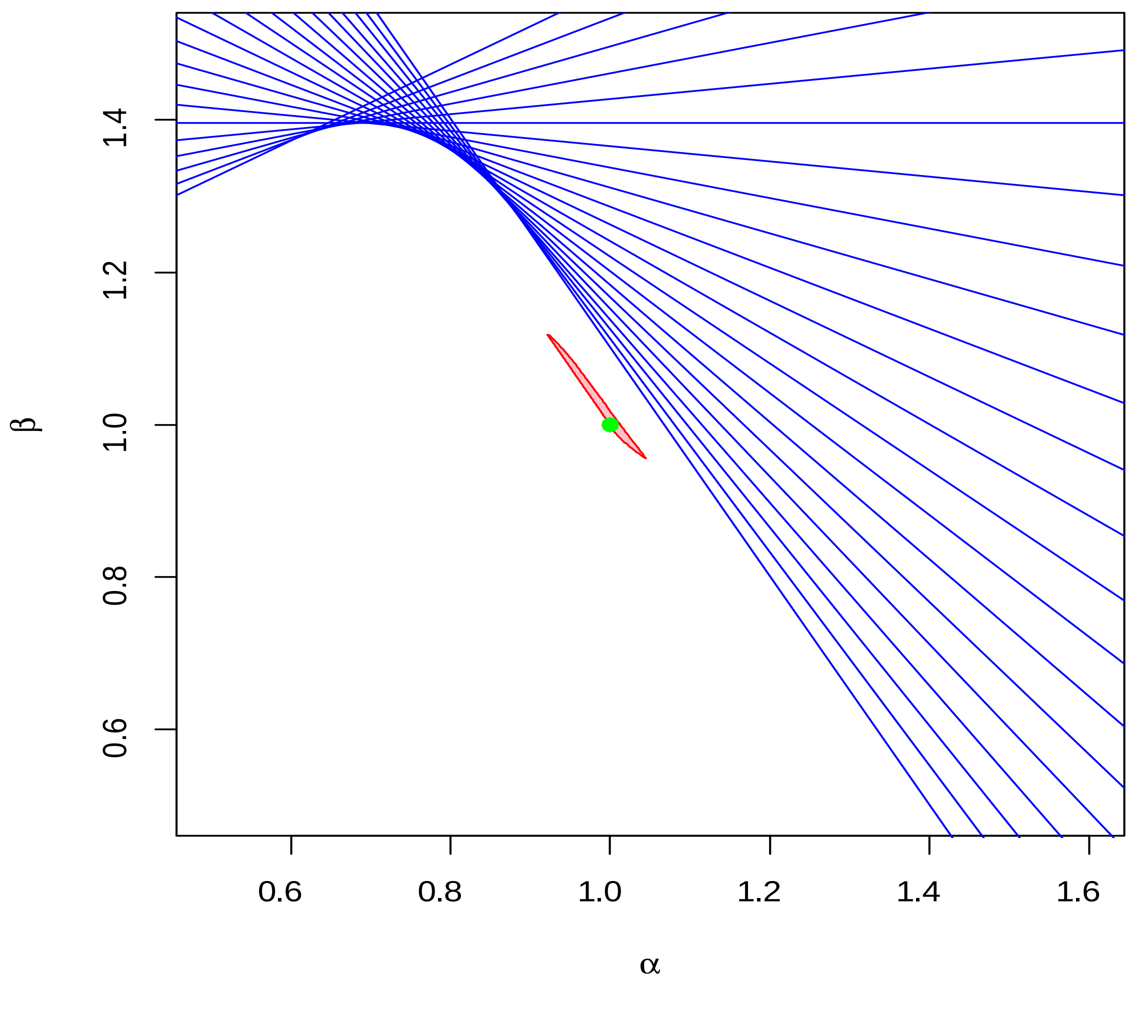

Figure 3: Structure 1, Y_2 linear in Z and U_2. The pink filled region is the identified set of values of α and β in a linear IV Tobit model under a zero conditional median restriction. The region below the blue drawn lines is the identified set under a zero conditional mean restriction. The parameter values in the triangular Gaussian structure generating the probabilities used in this calculation are shown in the column Case 2 in Table 1. The value of a and b generating the probabilities used to calculate the set is $a = 1$, $b = 1$, plotted in green.

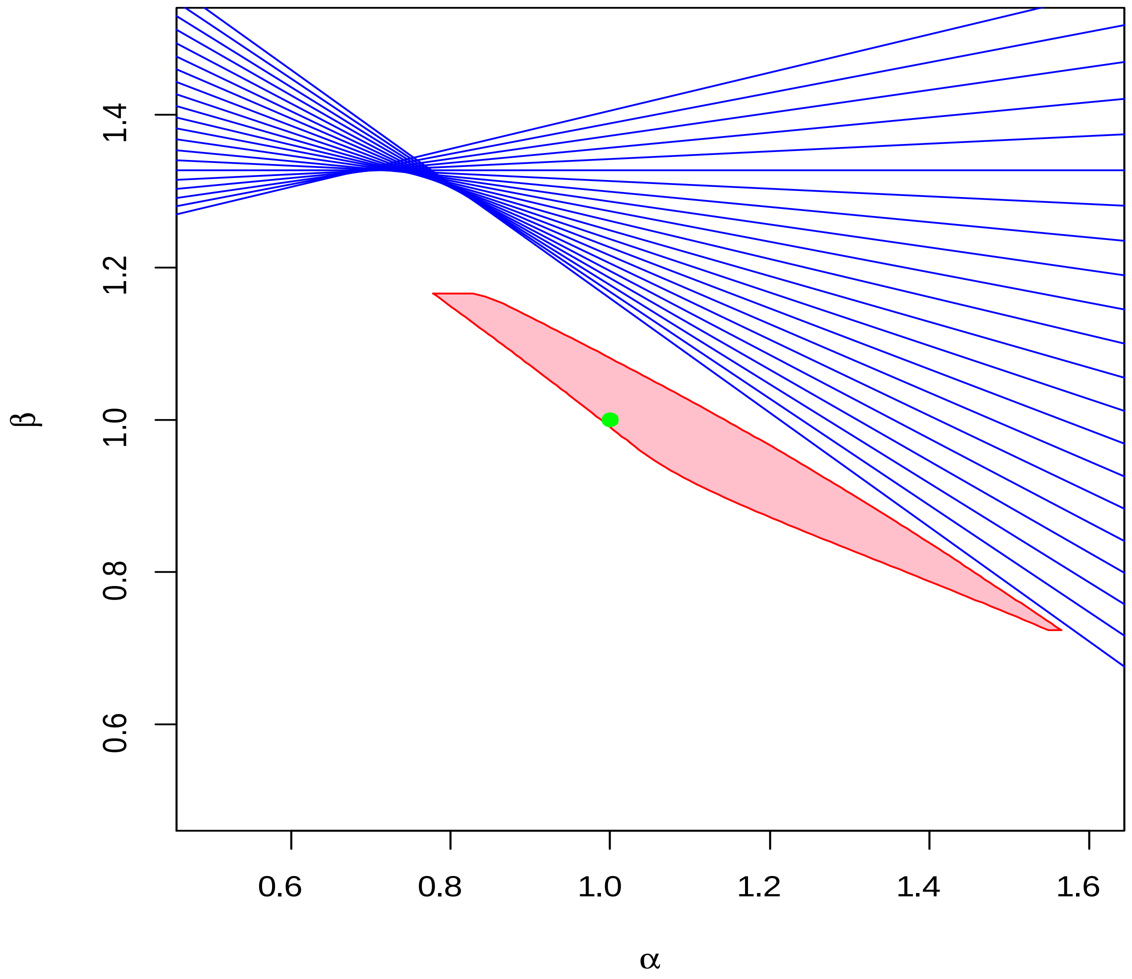

Figure 4: Structure 1, Y_2 linear in Z and U_2. The pink filled region is the identified set of values of α and β in a linear IV Tobit model under a zero conditional median restriction. The region below the blue drawn lines is the identified set under a zero conditional mean restriction. The parameter values in the triangular Gaussian structure generating the probabilities used in this calculation are shown in the column Case 3 in Table 1. The value of a and b generating the probabilities used to calculate the set is $a = 1$, $b = 1$, plotted in green.

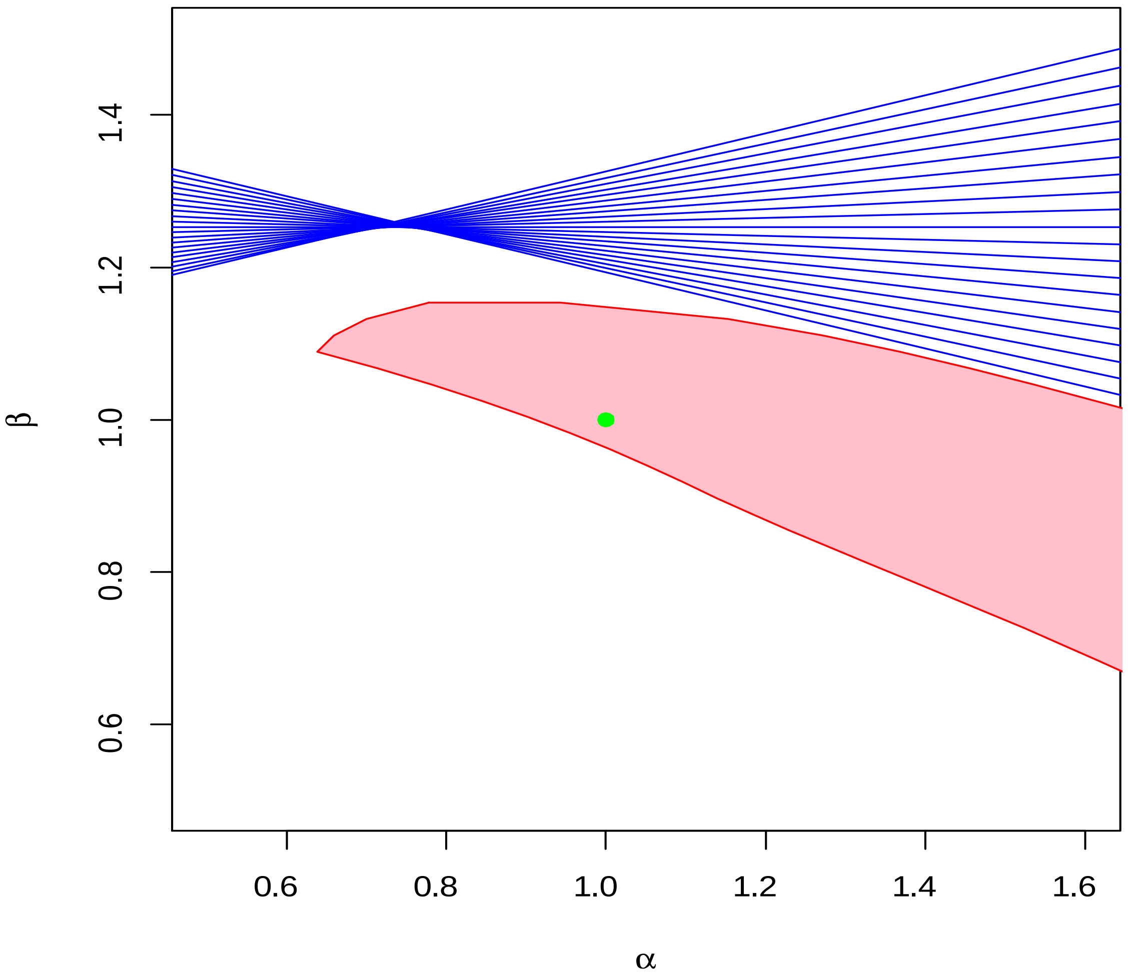

Figure 5: Structure 2, Y_2 bounded above. The pink filled region is the identified set of values of α and β in a linear IV Tobit model under a zero conditional median restriction. The magenta filled region is the projection of the identified set onto the space of (α, β) when U is restricted to be Gaussian, independent of Z with mean 0 and unknown variance. The region below the blue drawn lines is the identified set under a zero conditional mean restriction. The parameter values in the triangular Gaussian structure generating the probabilities used in this calculation are shown in the column Structure 2 Case 1 in Table 1. The value of a and b generating the probabilities used to calculate the set is $a = 1$, $b = 0$, plotted in green.

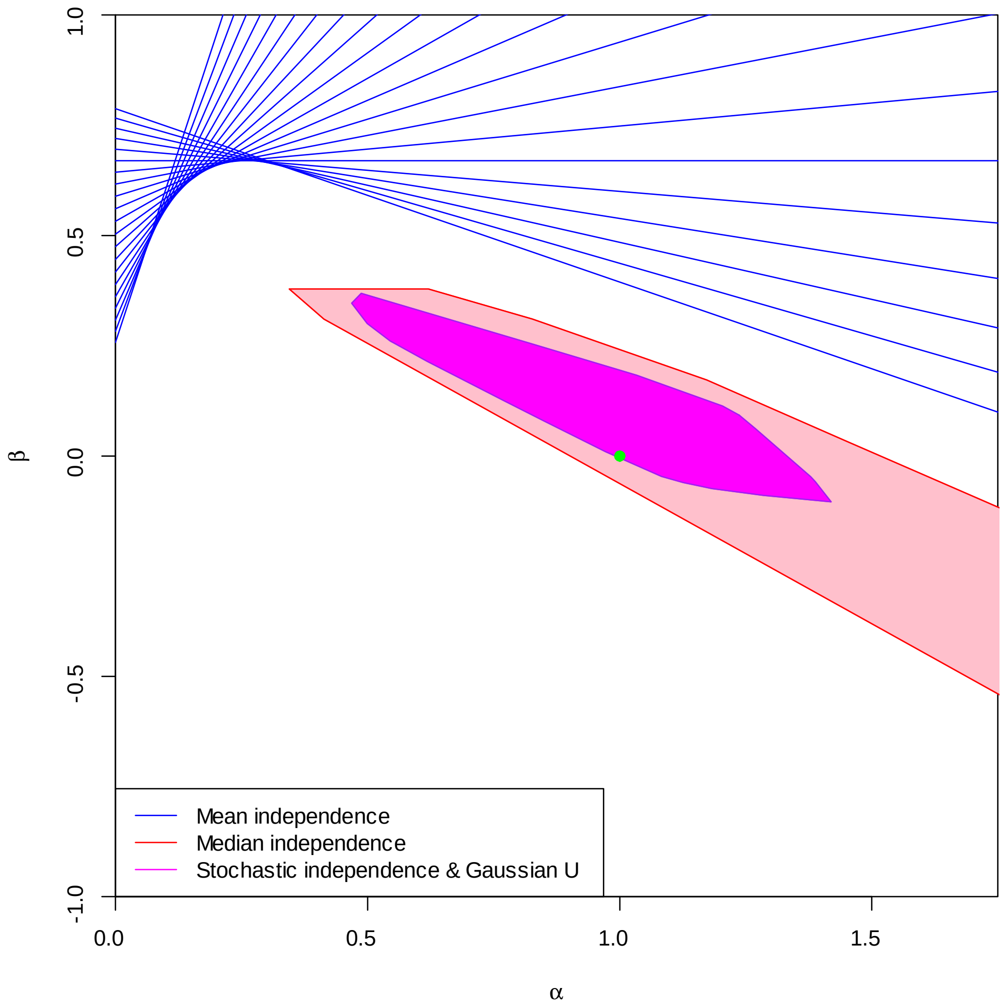

Figure 6: IV tobit model with U restricted Gaussian independent of Z, Structure 2 Case 1: projections of the identified set onto the space of each pair of parameters in turn. Values of the parameters generating probabilities are plotted in green. One dimensional projections are drawn in pale blue. Red filled regions are convex hulls of points found to lie in the projections.