



# QUANTILE DRIVEN IDENTIFICATION OF STRUCTURAL DERIVATES

---

*Andrew Chesher*

THE INSTITUTE FOR FISCAL STUDIES  
DEPARTMENT OF ECONOMICS, UCL  
**cemmap** working paper CWP08/01

# Quantile Driven Identification of Structural Derivatives

ANDREW CHESHER\*

CENTRE FOR MICRODATA METHODS AND PRACTICE  
INSTITUTE FOR FISCAL STUDIES AND UNIVERSITY COLLEGE LONDON

December 1st 2001

ABSTRACT. Conditions are derived under which there is local nonparametric identification of derivatives of structural equations in nonlinear triangular simultaneous equations systems. The attack on this problem is via conditional quantile functions and exploits local quantile independence conditions. The identification conditions include local analogues of the order and rank conditions familiar in the analysis of linear simultaneous equations models. The objects whose identification is sought are derivatives of structural equations at a point defined by values of covariates and quantiles of the distributions of the stochastic drivers of the system. These objects convey information about the distribution of the exogenous impact of variables potentially endogenous in the data generating process. The identification conditions point directly to analogue estimators of derivatives of structural functions which are functionals of quantile regression function estimators.

## 1. IDENTIFICATION OF DERIVATIVES OF NONLINEAR EQUATION SYSTEMS

**1.1. Introduction.** This paper derives conditions under which there is local nonparametric identification of partial derivatives of equations in triangular nonlinear structural systems under weak local quantile independence conditions. Points at which identification of derivatives is sought are defined by values of covariates upon which inference is conditioned, and values of probabilities defining quantiles of the distributions of the unobservable random variables that are the stochastic drivers of the system.

The identification conditions include local analogues of the order and rank conditions (Koopmans, Rubin and Leipnik (1950)) that arise in the analysis of identification in linear simultaneous equations models. These familiar conditions appear in this context because the partial derivatives of the structural equations, whose identification at a point of interest is sought, are, under the assumptions maintained, the coefficients of a linear approximation to the structural equations at the point of interest.

The derivatives of the structural equations studied here are the Exogenous Impact Functions (EIF) studied in Chesher (2001a). They convey information about the impact of *exogenous changes* in variables that are *endogenous* in a data generating process. They are of interest in the analysis of policy interventions because they convey distributional information, in the sense that they measure exogenous impacts at chosen quantiles of the distributions of the stochastic drivers of the model. They are also of interest when the

---

\*I started working on this topic while preparing a paper for Tom Rothenberg's Festschrift Conference held in the summer of 2001 at the University of California Berkeley on the occasion of Tom's retirement. My interest in the problem of identification, a problem at the core of econometrics, was stimulated many years ago by Tom's insightful approach to the problem, the deep results he obtained, and his elegant expression of them. This paper is dedicated to him. I am grateful to Art Goldberger, Tony Lancaster, Whitney Newey, Richard Smith, Richard Spady, Tieman Woutersen and participants at the Malinvaud Seminar at CREST and at seminars at University College London and the European University Institute for comments on earlier versions of this and related papers.

structural equations emanate from economic theory and structural derivatives measure the presence, direction and strength of behavioural responses.

As an example, consider a model explaining the joint determination of a person's wage and potentially endogenous schooling, in which schooling is determined by ability, and the wage given schooling is determined by good fortune and other labour demand side heterogeneity, and also by ability. The exogenous impact function for the wage relative to schooling is the derivative of the wage with respect to schooling, at a *fixed* level of ability and demand side heterogeneity, these levels determined by probabilities identifying quantiles of the distributions of ability and heterogeneity<sup>1</sup>.

The analysis of identification of nonlinear systems in this paper is, in essence, very simple. The key to this simplicity is the focus on *local* identification via conditional *quantile* functions.

Focusing on local identification of derivatives of structural functions reduces the problem to one of identification of coefficients - the derivatives of interest - in a linear model which is a local approximation to the structural system around the point of interest.

An exhaustive set of iterated conditional quantile functions<sup>2</sup> carries all information contained in the joint distribution function of continuously distributed observable responses so it is a suitable basis from which to develop identification conditions. An analysis via quantiles is well suited to the study of nonlinear systems because of the *equivariance* property of quantiles, namely that the  $\tau$ -quantile of  $h(Y)$  for strictly monotonic increasing (decreasing)  $h(\cdot)$ , is the result of applying the function  $h(\cdot)$  to the  $\tau$ -quantile ( $(1 - \tau)$ -quantile) of  $Y$ .

The identification conditions developed here point directly to analogue estimators<sup>3</sup> of the derivatives of structural equations. The estimators are functionals of quantile regression function estimators. In parametric linear location shift models they provide an alternative to the family of Two Stage Least Absolute Deviations estimators introduced by Amemiya (1982).

The next part of this Section sets out the structure of the models considered here, the nature of the identification conditions and of the identification result. An example of the sort of models treated in this paper, drawn from demand analysis, is provided. The Section concludes with a brief review of related literature and a plan of the rest of the paper.

**1.2. Triangular structural models.** The structural models considered in this paper have the following, triangular, form.

$$\begin{aligned} Y_1 &= \tilde{h}_1(Y_2, Y_3, \dots, Y_M, Z, \varepsilon_1, \varepsilon_2, \dots, \varepsilon_M) \\ Y_2 &= \tilde{h}_2(Y_3, \dots, Y_M, Z, \varepsilon_2, \dots, \varepsilon_M) \\ &\vdots \\ Y_M &= \tilde{h}_M(Z, \varepsilon_M) \end{aligned}$$

These models recursively determine values of  $M$  scalar variables,  $Y = \{Y_i\}_{i=1}^M$ , given values of covariates,  $Z = \{Z_i\}_{i=1}^K$ , and of continuously distributed stochastic unobservable variables,  $\varepsilon = \{\varepsilon_i\}_{i=1}^M$ .

This paper is concerned with the identifiability of the values of the *derivatives* of the structural equations at a point  $\mathcal{X}$  defined by: a value of  $Z$ ,  $z^*$ , and  $M$  probabilities,  $\tau^* = \{\tau_i^*\}_{i=1}^M$ .

<sup>1</sup>This example is considered in more detail in Chesher (2001a).

<sup>2</sup>An exhaustive set of iterated conditional quantile functions for  $M$  variates  $Y_1, \dots, Y_M$  give the relationships between the conditional quantiles of each variate  $Y_i$ ,  $i \in \{1, \dots, M\}$ , and the values of conditioning variables,  $Y_j$ ,  $j > i$ , and values of any other covariates. There are  $M!$  distinct exhaustive sets of iterated conditional quantile functions.

<sup>3</sup>See Manski (1988).

At the point  $\mathcal{X}$  the value of  $\varepsilon_i$  is  $\varepsilon_i^*$ , its conditional  $\tau_i^*$ -quantile conditional on  $\varepsilon_j = \varepsilon_j^*$ ,  $j > i$  and on  $Z = z^*$ . The value of  $Y$  at  $\mathcal{X}$  is the value implied by the system at  $\varepsilon = \varepsilon^* = \{\varepsilon_i^*\}_{i=1}^M$  when  $Z = z^*$ .

Formally, at the point  $\mathcal{X}$ ,  $Z = z^*$ , and

$$\begin{aligned} \varepsilon_M &= \varepsilon_M^* = Q_{\varepsilon_M|Z}(\tau_M, z^*) \\ \varepsilon_i &= \varepsilon_i^* = Q_{\varepsilon_i|\varepsilon_{i+1}\dots\varepsilon_M Z}(\tau_i^*, \varepsilon_{i+1}^*, \dots, \varepsilon_M^*, z^*), \quad i \in \{1, \dots, M-1\} \\ Y_M &= y_M^* = \tilde{h}_M(z^*, \varepsilon_M^*) \\ Y_i &= y_i^* = \tilde{h}_i(y_{i+1}^*, \dots, y_M^*, z^*, \varepsilon_{i+1}^*, \dots, \varepsilon_M^*), \quad i \in \{1, \dots, M-1\}. \end{aligned}$$

Here  $Q_{A|B_1\dots B_K}(\tau, b_1, \dots, b_K)$  denotes the conditional  $\tau$ -quantile of random variable  $A$  given  $B_1 = b_1, \dots, B_K = b_K$ . Henceforth special forms of expressions for the case  $i = M$  will not usually be stated explicitly.

The derivatives whose identification is sought are as follows.

$$\left. \begin{aligned} \nabla_{y_j} \tilde{h}_i(\mathcal{X}) &= \left. \frac{\partial}{\partial y_j} \tilde{h}_i(y_{i+1}, \dots, y_M, z, e_1, \dots, e_M) \right|_{\mathcal{X}} \\ \nabla_{z_k} \tilde{h}_i(\mathcal{X}) &= \left. \frac{\partial}{\partial z_k} \tilde{h}_i(y_{i+1}, \dots, y_M, z, e_1, \dots, e_M) \right|_{\mathcal{X}} \end{aligned} \right\} \quad (1)$$

Here  $|_{\mathcal{X}}$  indicates evaluation of arguments at the values defined by the point  $\mathcal{X}$ .

The additional conditions under which identification is demonstrated are of four types<sup>4</sup>.

1. *Local Smoothness*: At  $\mathcal{X}$  each structural function,  $\tilde{h}_i$ , is differentiable<sup>5</sup> with respect to its arguments.
2. *Monotonicity*: Each structural function,  $\tilde{h}_i$ , is a *strictly monotonic* function of its associated stochastic unobservable,  $\varepsilon_i$ , when variables other than  $\varepsilon_i$  are set equal to their values at  $\mathcal{X}$ . Throughout the functions are normalised to be strictly increasing.
3. *Local quantile independence*: At  $\mathcal{X}$  each  $\varepsilon_i$  is *conditional  $\tau_i$ -quantile independent* of  $Z$  given  $\varepsilon_j$ ,  $j > i$ , that is

$$Q_{\varepsilon_i|\varepsilon_{i+1}\dots\varepsilon_M Z}(\tau_i^*, \varepsilon_{i+1}^*, \dots, \varepsilon_M^*, z^*) = Q_{\varepsilon_i|\varepsilon_{i+1}\dots\varepsilon_M}(\tau_i^*, \varepsilon_{i+1}^*, \dots, \varepsilon_M^*).$$

4. *Local order and rank conditions*: At  $\mathcal{X}$  the first partial derivatives of the structural equations with respect to elements of  $Z$  satisfy certain *a priori* restrictions. In a leading case of interest there exist elements of  $Z$  (local instrumental variables) for which, at  $\mathcal{X}$ , certain first partial derivatives of the structural functions are zero and certain others are nonzero.

The nature of the identification result provided in this paper is as follows. It is shown that all structural systems satisfying conditions of these four types (conditions made precise in Section 2) with a common set of values,  $D_{\mathcal{X}}$ , of the structural derivatives (1) at  $\mathcal{X}$  generate distributions of  $Y$  given  $Z$ , and associated iterated conditional quantile functions such that certain well-defined functionals of the derivatives of the iterated conditional quantile functions take the values  $D_{\mathcal{X}}$  at  $\mathcal{X}$ .

<sup>4</sup>Precise conditions are given for the three equation model studied in Section 2.

<sup>5</sup>Differentiability with respect to an element,  $Z_i$ , of  $Z$  is not required if (a)  $Z_i$  is not a local instrumental variable (to be defined) and (b) identification of structural derivatives with respect to  $Z_i$  is not required.

**1.3. Example: Food Engel curves.** Let  $Y_1$  be household expenditure on a good, say food, in some period, let  $Y_2$  be the logarithm of total household expenditure in the period, and let  $Z$  be a list of household characteristics and other covariates. The Working-Leser form of the food Engel curve, which finds frequent application in the Almost Ideal Demand System (Deaton and Muellbauer (1980)), is, for some function  $h_1(Z)$  as follows.

$$\frac{Y_1}{\exp(Y_2)} = \theta^+ Y_2 + h_1(Z) + \varepsilon_1^+$$

In the two stage budgeting model implicit in this representation, total expenditure and food expenditure are jointly determined given household income and other household characteristics,  $Z$ , and then  $Y_2$  may be endogenous. The following, semiparametric specification is one possible expression of this.

$$\begin{aligned} Y_1 &= \exp(Y_2) \left( \theta^{g(\varepsilon_2)} Y_2 + h_1(Z) + \varepsilon_1 + \kappa_1 \varepsilon_2 \right) \\ Y_2 &= h_2(Z) + \varepsilon_2 \end{aligned}$$

Note that  $\varepsilon_1^+$  has been rewritten in terms of two unobserved continuously distributed stochastic drivers,  $\varepsilon_1$  and  $\varepsilon_2$  which it is convenient in this example to suppose are independently distributed. Further,  $\theta^+$  has been allowed to be a nonparametric function of  $\varepsilon_2$ . Endogeneity arises unless  $\kappa_1 = 0$  and  $g(\varepsilon_2)$  is invariant with respect to  $\varepsilon_2$ .

The results of this paper are relevant to, for example, the identification of the first partial derivative of the first structural equation with respect to  $Y_2$  evaluated at a particular value,  $z^*$ , of  $Z$  and at  $\tau_1^*$  and  $\tau_2^*$  quantiles of continuously distributed  $\varepsilon_1$  and  $\varepsilon_2$ . Let  $y_1^*$ ,  $y_2^*$  and  $e_2^*$  be the values taken at this point by respectively  $Y_1$ ,  $Y_2$  and  $\varepsilon_2$ . Results derived later in the paper imply that  $y_1^*$  is the value of the conditional  $\tau_1^*$ -quantile of  $Y_1$  given  $Y_2 = y_2^*$  and  $Z = z^*$  and that  $y_2^*$  is the conditional  $\tau_2^*$ -quantile of  $Y_2$  given  $Z = z^*$ . Both  $y_1^*$  and  $y_2^*$  are clearly identifiable.

In the context of this problem the value of the structural derivative whose identification is studied is  $\pi_{12}(\tau_1^*, \tau_2^*, z^*)$  where

$$\begin{aligned} \pi_{12}(\tau_1^*, \tau_2^*, z^*) &= \left. \frac{\partial Y_1}{\partial Y_2} \right|_{\varepsilon_1 = \varepsilon_1^*, \varepsilon_2 = e_2^*, z = z^*} \\ &= y_1^* + \theta^{g(e_2^*)} \exp(y_2^*). \end{aligned}$$

Suppose that  $\varepsilon_1$  and  $\varepsilon_2$  are quantile independent of  $Z$  at the probability levels considered. The results of this paper imply that  $\pi_{12}(\tau_1^*, \tau_2^*, z^*)$  is then identifiable as a well defined functional of the conditional quantile functions for  $Y_1$  given  $Y_2$  and  $Z$  and for  $Y_2$  given  $Z$  if there exists an element of  $Z$  ( $Z_1$ , say, perhaps household income) such that  $\nabla_{z_1} h_1(z^*) = 0$  and  $\nabla_{z_1} h_2(z^*) \neq 0$ . It then follows that, for example, the “random coefficient”:

$$\theta^{g(e_2^*)} = \exp(-y_2^*) (\pi_{12}(\tau_1^*, \tau_2^*, z^*) - y_1^*)$$

and the total expenditure elasticity of demand for food,

$$\eta(\tau_1^*, \tau_2^*, z^*) = \pi_{12}(\tau_1^*, \tau_2^*, z^*) / y_1^*$$

are identifiable under such local rank and order conditions.

If the identification conditions apply across a range of values  $\tau_1^*$  and  $\tau_2^*$  then it will be possible to identify how elasticities and “random coefficients” such as  $\theta^+ = \theta^{g(\varepsilon_2)}$  vary across quantiles of the distributions of the unobservable  $\varepsilon_1$  and  $\varepsilon_2$ . Clearly, the attack on identification via conditional quantile functions, which is extremely convenient in the analysis of nonlinear equation systems, has the considerable added benefit of bringing attention to bear on distributional aspects of the sensitivity of the relationships amongst responses.

**1.4. Related literature.** The study of parametric identification has a long history in econometrics starting with Working (1925, 1927) and Frisch (1934) and with notable contributions by, among others, Koopmans, Rubin and Leipnik (1950), Wald (1950), Fisher (1959, 1961, 1966), and Rothenberg (1971). One product of this research was the order and rank conditions in linear models, local versions of which feature in the results of this paper.

There has been considerable attention paid to the problem of determining conditions under which there is *nonparametric* identification of structural models. Roehrig (1988), extending the work of Brown (1983), considered global nonparametric identification of structural equations under the assumption that the stochastic drivers of the system (disturbances) are distributed independently of covariates. Much of Roehrig’s development is for the case in which the stochastic drivers appear additively in the equations of the model. Newey and Powell (1988), Newey, Powell and Vella (1999), Pinkse (2000), Darolles, Florens and Renault (2000) study models with additive disturbances which satisfy mean independence conditions of various types. Blundell and Powell (2000) provide a survey of mean independence based work in this area.

There is a large recent literature concerning identification and estimation in models involving responses, treatments and instrumental variables, some or all of which are discrete, usually binary<sup>6</sup>. The methods of this paper are not applicable in these problems because they require endogenous variables and local instruments to exhibit continuous variation near the points at which identification is sought.

Recently there has been interest in determining when global identification can be achieved in models with nonseparable disturbances. Brown and Matzkin (1996) study the identification of nonparametric primitive functions (e.g., production or utility functions) associated with simultaneous equations systems under the assumption that disturbances and covariates are independently distributed. Altonji and Matzkin, (2001) study panel data models with endogeneity under conditional exchangeability assumptions. Their procedure for problems with continuously distributed responses exploits strict monotonicity assumptions to secure global identification, slightly stronger than the monotonicity assumptions needed to achieve local identification using the quantile based attack of this paper.

The aim of this paper is in some respects less ambitious than those just cited. Firstly the aim is to identify first partial derivatives of structural functions rather than the structural functions themselves. This is a worthwhile aim because if it can be achieved then it is possible to take a view of the sensitivity of responses to exogenous changes in conditions.

Secondly, the focus is on local identification, rather than global identification of structural equations or “global functionals” of them, for example the average structural functions treated in much of the literature. Local identification is relatively easy to demonstrate as will be seen in Section 3. When conditions under which it is achievable have been obtained, and one is dealing with a problem in which there is sufficient smoothness, one can address global identification by asking whether the local conditions hold everywhere. In parametric models the object that is identified locally may be a “global” parameter in which case one automatically has global identification of that parameter. Chesher (2001b) explores identification of average derivatives of structural equations using quantile based methods.

Imbens and Newey (2001) study a triangular, two equation, structural model of the sort addressed in this paper. Global identification is demonstrated without assuming additive disturbances but under full independence of disturbances and covariances. The approach taken by Imbens and Newey is different, but complementary, to that taken here. They attack the identification problem from the standpoint of the distribution functions of the observable variates whereas here inverse distribution functions, that is

<sup>6</sup>See for example Heckman (1990), Imbens and Angrist (1994), Das (2000) and Vytlačil (2001).

quantile functions, are the starting point. The Imbens and Newey estimation procedure entails nonparametric distribution and mean regression function estimation in contrast to the estimation procedure suggested by the results of this paper, which entails (perhaps non-parametric) conditional quantile function estimation.

The objects whose identification are considered in this paper, derivatives of structural functions evaluated at quantiles of the distributions of stochastic drivers of the system, are the Exogenous Impact Functions (EIF) introduced in Chesher (2001a). Estimates of EIFs can give valuable information about the distribution of policy impacts across a population. There is a recent literature aimed at developing estimators of such distributions. For example Heckman, Smith and Clements (1997) explore non-quantile based approaches in a programme evaluation setting. Abadie, Angrist and Imbens (2001) propose a Quantile Treatment Effect estimator in a study of the impact of subsidised training on the distribution of earnings.

Identification is considered from a conditional quantile perspective in Matzkin (1999). Matzkin considers a model  $Y = m(X, \varepsilon)$  in which  $\varepsilon$  is distributed independently of  $X$  (so  $X$  is exogenous) and  $m(\cdot, \cdot)$  is strictly monotonic in  $\varepsilon$ . Conditions under which the function  $m(\cdot, \cdot)$  and the distribution function of  $\varepsilon$  are identifiable are obtained. The value of  $m(\cdot, \cdot)$  at a point  $(x, e)$  is shown, under suitable conditions, to be identifiable as the value of the conditional  $\tau$ -quantile of  $Y$  given  $X = x$  where  $\tau$  is such that  $e$  is the  $\tau$ -quantile of the marginal distribution of  $\varepsilon$ . An analogous result for *derivatives* of structural functions with respect to exogenous covariates is obtained in this paper, but our main interest here is in identification when there is endogeneity, that is, in Matzkin's problem, when  $\varepsilon$  is not independent of  $X$ .

The quantile based approach of this paper produces results relevant to the literature on quantile regression function estimation in the presence of endogeneity. Amemiya (1982) develops<sup>7</sup> Two Stage Least Absolute Deviations (2SLAD) estimators for parametric models in which a conditional median is linear in endogenous and exogenous variables and independent of instrumental variables. The estimation procedure suggested in this paper can be applied to the parametric problem studied by Amemiya and Powell, producing a simple alternative to the family of 2SLAD estimators.

The identification conditions of this paper include local quantile independence conditions. There are recent uses of quantile independence conditions as the basis for developing estimators in Newey and Powell (1990), Chaudhuri, Doksum and Samarov (1997) and Kahn (2001).

**1.5. Plan of the remainder of the paper.** Section 2 provides a detailed analysis of identification in a triangular three equation system. At the end of Section 2 the result is illustrated in the context of a linear simultaneous equations location shift model. This analysis of the linear model suggests a general attack which is outlined in Section 3. There,  $M$  equation triangular structural systems, as set out at the start of this Section, are studied, and the case in which there are restrictions on derivatives of the structural functions with respect to elements of  $Y$ ,  $Z$  and  $\varepsilon$  is considered. Section 4 considers identification of average structural derivatives and more generally expected values of functions of structural derivatives. Section 5 concludes.

---

<sup>7</sup>See also Powell (1983).

## 2. IDENTIFICATION IN A THREE EQUATION MODEL

This Section is concerned with a three equation triangular structure, as follows.

$$Y_1 = \tilde{h}_1(Y_2, Y_3, Z, \varepsilon_1, \varepsilon_2, \varepsilon_3) \quad (2a)$$

$$Y_2 = \tilde{h}_2(Y_3, Z, \varepsilon_2, \varepsilon_3) \quad (2b)$$

$$Y_3 = \tilde{h}_3(Z, \varepsilon_3) \quad (2c)$$

The random variables  $\varepsilon_1$ ,  $\varepsilon_2$  and  $\varepsilon_3$  are continuously distributed and unobservable. The symbol  $Z$  stands for a list of covariates.

An identification theorem for this three equation system is stated. The Theorem relates to the case in which, there are local exclusion restrictions pertaining to the covariates  $Z$ . Section 3 outlines the extension to the  $M$  equation case and considers more general restrictions. After the statement of the Theorem there are brief remarks, followed by a proof.

**Theorem**

Consider the three equation model (2a) - (2c),  $\tau_i^* \in (0, 1)$ ,  $i \in \{1, 2, 3\}$ , and a value of  $K$ -element  $Z$ ,  $z^*$ . Let  $\tau^* = \{\tau_i^*\}_{i=1}^3$ . Define

$$\varepsilon_3^* = Q_{\varepsilon_3|Z}(\tau_3^*, z^*) \quad \varepsilon_2^* = Q_{\varepsilon_2|\varepsilon_3 Z}(\tau_2^*, \varepsilon_3^*, z^*) \quad \varepsilon_1^* = Q_{\varepsilon_1|\varepsilon_2, \varepsilon_3 Z}(\tau_1^*, \varepsilon_2^*, \varepsilon_3^*, z^*)$$

and

$$y_3^* = Q_{Y_3|Z}(\tau_3^*, z^*) \quad y_2^* = Q_{Y_2|Y_3 Z}(\tau_2^*, y_3^*, z^*)$$

which, together with  $z^*$ , define a point<sup>8</sup> referred to as  $\mathcal{X}$ . Consider distinct elements of  $Z$  which are denoted  $Z_{(2)}$  and  $Z_{(3)}$ . Assume that the following conditions hold.

1. At  $Z = z^*$ ,  $\varepsilon = \{\varepsilon_i\}_{i=1}^3$  is an absolutely continuous random vector with positive density everywhere over compact support.
2. For each  $i \in \{1, 2, 3\}$ , the function  $\tilde{h}_i$ , with arguments other than  $\varepsilon_i$  set equal to their values at  $\mathcal{X}$ , is an increasing function<sup>9</sup> of  $\varepsilon_i$  for all  $\varepsilon_i$  and a strictly increasing function of  $\varepsilon_i$  at  $\mathcal{X}$ .
3. For each  $i \in \{1, 2, 3\}$ , at  $\mathcal{X}$ ,  $\tilde{h}_i$  is a differentiable function of its arguments, its first partial derivatives being finite at  $\mathcal{X}$ .
4. At  $\mathcal{X}$  the following local quantile independence conditions hold<sup>10</sup>.

$$Q_{\varepsilon_1|\varepsilon_2, \varepsilon_3 Z}(\tau_1, \varepsilon_2, \varepsilon_3, z) = Q_{\varepsilon_1|\varepsilon_2, \varepsilon_3}(\tau_1, \varepsilon_2, \varepsilon_3) \quad (3a)$$

$$Q_{\varepsilon_2|\varepsilon_3 Z}(\tau_2, \varepsilon_3, z) = Q_{\varepsilon_2|\varepsilon_3}(\tau_2, \varepsilon_3) \quad (3b)$$

$$Q_{\varepsilon_3|Z}(\tau_3, z) = Q_{\varepsilon_3}(\tau_3) \quad (3c)$$

<sup>8</sup>Note that under the monotonicity assumption 2 below,  $y_2^*$  and  $y_3^*$  can be written as follows.

$$\begin{aligned} y_3^* &= h_3(z^*, Q_{\varepsilon_3|Z}(\tau_3^*, z^*)) \\ y_2^* &= h_2(y_3^*, z^*, Q_{\varepsilon_2|\varepsilon_3 Z}(\tau_2^*, Q_{\varepsilon_3|Z}(\tau_3^*, z^*), z^*), Q_{\varepsilon_3|Z}(\tau_3^*, z^*)) \\ &= h_2(h_3(z^*, Q_{\varepsilon_3|Z}(\tau_3^*, z^*)), z^*, Q_{\varepsilon_2|\varepsilon_3 Z}(\tau_2^*, Q_{\varepsilon_3|Z}(\tau_3^*, z^*), z^*), Q_{\varepsilon_3|Z}(\tau_3^*, z^*)) \end{aligned}$$

They are the values taken by  $Y_2$  and  $Y_3$  when  $Z = z^*$  and the stochastic drivers  $\varepsilon_2$  and  $\varepsilon_3$  are set equal to the  $\tau_2^*$ - and  $\tau_3^*$ -quantiles of their conditional distributions.

<sup>9</sup>The functions are required to be strictly monotonic and normalised to be *increasing*. This monotonicity assumption is slightly stronger than is required. The minimal condition is that at  $\mathcal{X}$  the function  $\tilde{h}_i$  does give the relationship between (a) the conditional  $\tau_i^*$ -quantile of  $Y_i$  given  $Y_j$ ,  $j > i$ , and  $z^*$  and  $y_j^*$ ,  $j > i$ , and (b) the conditional  $\tau_i^*$ -quantile of  $\varepsilon_i$ . For example, considering  $\tilde{h}_3$ , all that is needed is that the  $\tau_3^*$ -quantiles of  $\varepsilon_3$  and of  $Y_3$  given  $Z$  satisfy  $Q_{Y_3|Z}(\tau_3^*, z^*) = \tilde{h}_3(z^*, Q_{\varepsilon_3|Z}(\tau_3^*, z^*))$ .

<sup>10</sup>These conditions are stronger than is required. As shown in the Annex, it is only required that at  $\mathcal{X}$  the derivatives of these conditional quantiles functions with respect to  $z$  exist, and are zero. Only derivatives with respect to local instrumental variables are relevant.



5. At  $\mathcal{X}$  the following local *order* conditions hold.

$$\nabla_{z(2)} \tilde{h}_1(y_2, y_3, z, \varepsilon_1, \varepsilon_2, \varepsilon_3) = 0 \quad (4a)$$

$$\nabla_{z(3)} \tilde{h}_1(y_2, y_3, z, \varepsilon_1, \varepsilon_2, \varepsilon_3) = 0 \quad (4b)$$

$$\nabla_{z(3)} \tilde{h}_2(y_3, z, \varepsilon_2, \varepsilon_3) = 0 \quad (4c)$$

6. At  $\mathcal{X}$  the following local *rank* conditions hold.

$$\nabla_{z(2)} \tilde{h}_2(y_3, z, \varepsilon_2, \varepsilon_3) \neq 0 \quad (5a)$$

$$\nabla_{z(3)} \tilde{h}_3(z, \varepsilon_3) \neq 0 \quad (5b)$$

The Theorem is concerned with the identification of the following derivatives of the structural functions when  $\tau = \tau^*$  at the point  $\mathcal{X}$ . Here  $Z_k$  is an arbitrary element of  $Z$  (possibly  $Z_{(2)}$  or  $Z_{(3)}$ ) for which the differentiability condition of Assumption 3 holds<sup>11</sup>.

$$\pi_{12}(\tau^*, z^*) = \nabla_{y_2} \tilde{h}_1(y_2^*, y_3^*, z^*, \varepsilon_1^*, \varepsilon_2^*, \varepsilon_3^*)$$

$$\pi_{13}(\tau^*, z^*) = \nabla_{y_3} \tilde{h}_1(y_2^*, y_3^*, z^*, \varepsilon_1^*, \varepsilon_2^*, \varepsilon_3^*)$$

$$\pi_{1z_k}(\tau^*, z^*) = \nabla_{z_k} \tilde{h}_1(y_2^*, y_3^*, z^*, \varepsilon_1^*, \varepsilon_2^*, \varepsilon_3^*)$$

$$\pi_{23}(\tau^*, z^*) = \nabla_{y_3} \tilde{h}_2(y_3^*, z^*, \varepsilon_2^*, \varepsilon_3^*)$$

$$\pi_{2z_k}(\tau^*, z^*) = \nabla_{z_k} \tilde{h}_2(y_3^*, z^*, \varepsilon_2^*, \varepsilon_3^*)$$

$$\pi_{3z_k}(\tau^*, z^*) = \nabla_{z_k} \tilde{h}_3(z^*, \varepsilon_3^*)$$

These derivatives are identifiable as functionals of conditional quantile functions as shown below. In these equations arguments of functions are suppressed. The  $\tau_i$ -quantiles for  $Y_i$  conditional on other variables are evaluated at  $\tau_i = \tau_i^*$  and other arguments are set equal to their values at the point  $\mathcal{X}$ .

$$\pi_{12} = \nabla_{y_2} Q_{Y_1|Y_2Y_3Z} + \frac{\nabla_{z(2)} Q_{Y_1|Y_2Y_3Z} - \frac{\nabla_{z(3)} Q_{Y_1|Y_2Y_3Z}}{\nabla_{z(3)} Q_{Y_3|Z}} \nabla_{z(2)} Q_{Y_3|Z}}{\nabla_{z(2)} Q_{Y_2|Y_3Z} - \frac{\nabla_{z(3)} Q_{Y_2|Y_3Z}}{\nabla_{z(3)} Q_{Y_3|Z}} \nabla_{z(2)} Q_{Y_3|Z}}$$

$$\begin{aligned} \pi_{13} &= \nabla_{y_3} Q_{Y_1|Y_2Y_3Z} + \frac{\nabla_{z(3)} Q_{Y_1|Y_2Y_3Z}}{\nabla_{z(3)} Q_{Y_3|Z}} \\ &\quad - \pi_{23}(\tau^*, z^*) \times \frac{\nabla_{z(2)} Q_{Y_1|Y_2Y_3Z} - \frac{\nabla_{z(3)} Q_{Y_1|Y_2Y_3Z}}{\nabla_{z(3)} Q_{Y_3|Z}} \nabla_{z(2)} Q_{Y_3|Z}}{\nabla_{z(2)} Q_{Y_2|Y_3Z} - \frac{\nabla_{z(3)} Q_{Y_2|Y_3Z}}{\nabla_{z(3)} Q_{Y_3|Z}} \nabla_{z(2)} Q_{Y_3|Z}} \end{aligned}$$

$$\begin{aligned} \pi_{1z_k} &= \nabla_{z_k} Q_{Y_1|Y_2Y_3Z} - \nabla_{z_k} Q_{Y_3|Z} \times \frac{\nabla_{z(3)} Q_{Y_1|Y_2Y_3Z}}{\nabla_{z(3)} Q_{Y_3|Z}} \\ &\quad - \left( \nabla_{z(2)} Q_{Y_1|Y_2Y_3Z} - \frac{\nabla_{z(3)} Q_{Y_1|Y_2Y_3Z}}{\nabla_{z(3)} Q_{Y_3|Z}} \nabla_{z(2)} Q_{Y_3|Z} \right) \times \frac{\pi_{2z_k}}{\pi_{2z(2)}} \end{aligned}$$

$$\pi_{23} = \nabla_{y_3} Q_{Y_2|Y_3Z} + \frac{\nabla_{z(3)} Q_{Y_2|Y_3Z}}{\nabla_{z(3)} Q_{Y_3|Z}}$$

<sup>11</sup>The notation for derivatives here is as follows. For a function  $f(a, b)$ ,  $\nabla_a f(a^*, b^*)$  indicates  $\frac{\partial}{\partial a} f(a, b) \Big|_{a=a^*, b=b^*}$ .

$$\pi_{2z_k} = \nabla_{z_k} Q_{Y_2|Y_3Z} - \nabla_{z_k} Q_{Y_3|Z} \times \frac{\nabla_{z(3)} Q_{Y_2|Y_3Z}}{\nabla_{z(3)} Q_{Y_3|Z}}$$

$$\pi_{3z_k} = \nabla_{z_k} Q_{Y_3|Z}$$

### Remarks

Arguments of functions ( $\tau^*$  and the values of  $Y_1$ ,  $Y_2$ ,  $Y_3$  and  $Z$  at the point  $\mathcal{X}$ ) are suppressed in these comments.

1. The Theorem locally identifies structural derivatives as functionals of iterated conditional quantile functions in the following sense. All structural models (that is, all systems of equations  $\tilde{h}_1, \dots, \tilde{h}_M$ ) that satisfy the assumptions of the Theorem, with derivatives taking values  $D_{\mathcal{X}}$  at the point  $\mathcal{X}$ , generate iterated conditional quantile functions with the property that the functionals at the end of the statement of the Theorem take the value  $D_{\mathcal{X}}$  at  $\mathcal{X}$ .
2. In general the structural derivatives evaluated at  $\mathcal{X}$  vary with  $\tau^*$ , that is the sensitivity of responses  $Y$  to exogenous shifts in  $Y$  and  $Z$  varies across quantiles of the distributions of the unobserved stochastic drivers. It is in this sense that the objects identified bear on the distributional impacts of policy interventions.
3. In the expression<sup>12</sup> for  $\pi_{1z_k}$  the term  $\pi_{2z(2)}$  is  $\pi_{2z_k}$  with  $z(2)$  replacing  $z_k$ .
4. In the expression for  $\pi_{12}$  the denominator of the second right hand side term is  $\pi_{2z(2)}$  which is nonzero by virtue of the rank condition (5a).
5. Setting  $Z_k = Z_{(3)}$  yields  $\pi_{2z(3)} = 0$  and  $\pi_{1z(3)} = 0$  which are in accord with the order conditions (4c) and (4b).
6. Setting  $Z_k = Z_{(2)}$  yields  $\pi_{1z(2)} = 0$  which is in accord with the order condition (4a).
7. The derivative  $\pi_{12}$  may be identified when  $\pi_{23}$  is not, but unless  $\pi_{23}$  is identified  $\pi_{13}$  is not identified.
8. If there is more than one element of  $Z$  that satisfies the assumptions of the Theorem regarding  $Z_{(3)}$  ( $Z_{(2)}$ ) then  $\pi_{23}$  ( $\pi_{12}$  and  $\pi_{13}$ ) is (are) overidentified.
9. In order to identify the separate impact of elements of  $Z$  on the derivatives of the structural equations, it must be possible to identify their separate impacts on the conditional quantile functions that appear above. This rules out exact functional dependencies amongst elements of  $Z$ .

<sup>12</sup>Note that  $\pi_{1z_k}$  can be expressed as follows

$$\pi_{1z_k} = \nabla_{z_k} Q_{Y_1|Y_2Y_3Z} - \nabla_{z_k} Q_{Y_3|Z} \times \frac{\nabla_{z(3)} Q_{Y_1|Y_2Y_3Z}}{\nabla_{z(3)} Q_{Y_3|Z}}$$

$$- \pi_{2z_k} \times \frac{\nabla_{z(2)} Q_{Y_1|Y_2Y_3Z} - \frac{\nabla_{z(3)} Q_{Y_1|Y_2Y_3Z}}{\nabla_{z(3)} Q_{Y_3|Z}} \nabla_{z(2)} Q_{Y_3|Z}}{\nabla_{z(2)} Q_{Y_2|Y_3Z} - \frac{\nabla_{z(3)} Q_{Y_2|Y_3Z}}{\nabla_{z(3)} Q_{Y_3|Z}} \nabla_{z(2)} Q_{Y_3|Z}}$$

**Proof**

First the equation system is rewritten in terms of mutually independently distributed<sup>13</sup>  $\nu = \{\nu_i\}_{i=1}^3$ . Define

$$\nu_3 = F_{\varepsilon_3|Z}(\varepsilon_3|z) \quad \nu_2 = F_{\varepsilon_2|\varepsilon_3Z}(\varepsilon_2|\varepsilon_3, z) \quad \nu_1 = F_{\varepsilon_1|\varepsilon_2\varepsilon_3Z}(\varepsilon_1|\varepsilon_2, \varepsilon_3, z)$$

which have independent uniform distributions, each on  $[0, 1]$ . Each  $\varepsilon_i$  is then related to the elements of  $\nu$  by the following equations

$$\begin{aligned} \varepsilon_3 &= Q_{\varepsilon_3|Z}(\nu_3|z) \\ \varepsilon_2 &= Q_{\varepsilon_2|\varepsilon_3Z}(\nu_2|Q_{\varepsilon_3|Z}(\nu_3|z), z) \\ \varepsilon_1 &= Q_{\varepsilon_1|\varepsilon_2\varepsilon_3Z}(\nu_1|Q_{\varepsilon_2|\varepsilon_3Z}(\nu_2|Q_{\varepsilon_3|Z}(\nu_3|z), z), Q_{\varepsilon_3|Z}(\nu_3|z), z) \end{aligned}$$

and at  $Z = z^*$  and with  $\nu = \tau^*$  in which case there is local quantile independence relative to  $Z$  (Assumption 4),

$$\begin{aligned} \varepsilon_3 &= Q_{\varepsilon_3}(\nu_3) \\ \varepsilon_2 &= Q_{\varepsilon_2|\varepsilon_3}(\nu_2|Q_{\varepsilon_3}(\nu_3)) \\ \varepsilon_1 &= Q_{\varepsilon_1|\varepsilon_2\varepsilon_3}(\nu_1|Q_{\varepsilon_2|\varepsilon_3}(\nu_2|Q_{\varepsilon_3}(\nu_3)), Q_{\varepsilon_3}(\nu_3)). \end{aligned}$$

Note that each  $\varepsilon_i$  is a strictly increasing function of its associated  $\nu_i$ . Let the equation system with each  $\varepsilon_i$  replaced by its associated  $\nu_i$  be written as follows<sup>14</sup>.

$$Y_1 = h_1(Y_2, Y_3, Z, \nu_1, \nu_2, \nu_3) \quad (6a)$$

$$Y_2 = h_2(Y_3, Z, \nu_2, \nu_3) \quad (6b)$$

$$Y_3 = h_3(Z, \nu_3) \quad (6c)$$

At  $\mathcal{X}$  this substitution ensures that  $\nu = \tau^*$  and leaves the dependence of each  $Y_i$  on the  $Y_j$ ,  $j > i$ , and on  $Z$  unchanged. In particular the derivatives of the functions  $h_i$  with respect to the observable variables, evaluated at  $\mathcal{X}$  are identically equal to the corresponding derivatives of the functions  $\tilde{h}_i$  evaluated at  $\mathcal{X}$ . Note that without the local quantile independence assumption this would not necessarily be true<sup>15</sup>.

Assumptions 1 - 6 apply to the system (6a)-(6c) involving independent  $\nu$  and the proof proceeds employing this version of the system.

First substitute for  $Y_3$  in (6b) using (6c) and then, using the result, substitute for  $Y_2$  in (6a). This gives the following.

$$Y_1 = h_1(h_2(h_3(z, \nu_3), z, \nu_2, \nu_3), h_3(z, \nu_3), z, \nu_1, \nu_2, \nu_3)$$

Monotonicity of  $h_1$  with respect to  $\nu_1$  ensures that, (suppressing arguments  $\tau_1, \nu_2, \nu_3, z$ , of the two conditional quantile functions):

$$Q_{Y_1|\nu_2\nu_3Z} = h_1(h_2(h_3(z, \nu_3), z, \nu_2, \nu_3), h_3(z, \nu_3), z, Q_{\nu_1|\nu_2\nu_3Z}, \nu_2, \nu_3)$$

and at  $\mathcal{X}$  the quantile independence condition (3a) and the mutual independence of the elements of  $\nu$  simplifies this as follows.

$$Q_{Y_1|\nu_2\nu_3Z} = h_1(h_2(h_3(z, \nu_3), z, \nu_2, \nu_3), h_3(z, \nu_3), z, Q_{\nu_1}, \nu_2, \nu_3)$$

<sup>13</sup>This step is not essential. It reduces the notational complexity later in the proof. The Annex indicates the stages in a proof which does not take this step.

<sup>14</sup>For example

$$h_3(Z, \nu_3) = \tilde{h}_3(Z, Q_{\varepsilon_3}(\nu_3)).$$

<sup>15</sup>In the absence of local quantile independence the transformation from  $\varepsilon$  to  $\nu$  would not be free of  $z$ . As shown in the Annex, local identification could still be achieved if first derivatives of conditional quantile functions with respect to local instrumental variables were zero at  $\mathcal{X}$ .

The monotonicity conditions of Assumption 2 and the continuity conditions of Assumption 3 ensure the existence of well behaved inverse functions as follows<sup>16</sup>.

$$\begin{aligned}\nu_3 &= g_3(Z, Y_3) \\ \nu_2 &= g_2(Y_3, Z, Y_2, g_3(Z, Y_3))\end{aligned}$$

Conditioning on  $\nu_2$ ,  $\nu_3$  and  $Z$  is equivalent to conditioning on  $Y_2$ ,  $Y_3$  and  $Z$  and so

$$Q_{Y_1|Y_2Y_3Z}(\tau_1, y_2, y_3, z) = h_1(y_2, y_3, z, Q_{\nu_1}(\tau_1), g_2(y_3, z, y_2, g_3(z, y_3)), g_3(z, y_3)). \quad (7)$$

Similarly, exploiting (3b) and (3c),

$$Q_{Y_2|Y_3Z}(\tau_2, y_3, z) = h_2(y_3, z, Q_{\nu_2}(\tau_2), g_3(z, y_3)) \quad (8)$$

$$Q_{Y_3|Z}(\tau_3, z) = h_3(z, Q_{\nu_3}(\tau_3)). \quad (9)$$

The Theorem is proved as follows. The conditional quantile functions (7) - (9) are differentiated, as appropriate, with respect to  $y_2$ ,  $y_3$  and elements of  $Z$ . The derivatives are evaluated at the point  $\mathcal{X}$  producing expressions involving the derivatives of the  $h_i$ 's at  $\mathcal{X}$  and derivatives of their inverse functions at  $\mathcal{X}$ . The latter are expressed in terms of derivatives of conditional quantile functions at  $\mathcal{X}$  and the results combined to obtain the required expressions for the structural function derivatives. Some of the manipulations are lengthy and just three of the simpler cases are worked through in detail here.

(a) *Identification of  $\pi_{3z_k}(\tau^*, z^*)$*

Note that  $Z_k$  is an arbitrary element of  $Z$  and could be  $Z_{(2)}$  or  $Z_{(3)}$ . The conditional  $\tau_3^*$ -quantile of  $Y_3$  given  $Z = z^*$  is<sup>17</sup>

$$Q_{Y_3|Z}(\tau_3^*, z^*) = h_3(z^*, \tau_3^*)$$

and on differentiating with respect to  $z_k$ ,

$$\begin{aligned}\nabla_{z_k} Q_{Y_3|Z}(\tau_3^*, z^*) &= \nabla_{z_k} h_3(z^*, \tau_3^*) \\ &= \pi_{3z_k}(\tau^*, z^*).\end{aligned}$$

This is the analogue for structural derivatives of the result of Matzkin (1999).

(a) *Identification of  $\pi_{23}(\tau^*, z^*)$*

Henceforth arguments of functions are suppressed, it being understood that all derivatives of functions are evaluated at the point  $\mathcal{X}$  defined in the assumptions, where existence is assured, and derivatives appearing in denominators of expressions are guaranteed non-zero.

Differentiating (8) with respect to  $y_3$  and  $z_{(3)}$  gives the following.

$$\nabla_{y_3} Q_{Y_2|Y_3Z} = \nabla_{y_3} h_2 + \nabla_{\nu_3} h_2 \nabla_{y_3} g_3 \quad (10)$$

$$\nabla_{z_{(3)}} Q_{Y_2|Y_3Z} = \nabla_{\nu_3} h_2 \nabla_{z_{(3)}} g_3 \quad (11)$$

There is only one term in (11) because of the local order condition (4c). Note that the first term on the right hand side of (10) is  $\pi_{23}(\tau^*, z^*)$  whose identification is sought.

Since

$$Q_{\nu_3} = g_3(z, Q_{Y_3|Z}) \quad (12)$$

<sup>16</sup>For example,  $g_3(Z, Y_3)$  satisfies

$$Y_3 = h_3(Z, g_3(Z, Y_3)).$$

<sup>17</sup>Note, we use  $Q_{\nu_3}(\tau_3^*) = \tau_3^*$ , which follows because  $\nu_3 \sim U[0, 1]$ .

there is, on differentiating with respect to  $z_{(3)}$ ,

$$\nabla_{z_{(3)}} g_3 + \nabla_{y_3} g_3 \nabla_{z_{(3)}} Q_{Y_3|Z} = 0$$

and therefore

$$\nabla_{z_{(3)}} Q_{Y_2|Y_3Z} = -\nabla_{\nu_3} h_2 \nabla_{y_3} g_3 \nabla_{z_{(3)}} Q_{Y_3|Z}$$

and so,

$$\nabla_{\nu_3} h_2 \nabla_{y_3} g_3 = -\frac{\nabla_{z_{(3)}} Q_{Y_2|Y_3Z}}{\nabla_{z_{(3)}} Q_{Y_3|Z}}, \quad (13)$$

the denominator on the right hand side being nonzero at  $\mathcal{X}$  by virtue of the local rank condition (5b).

Substituting (13) in (10) gives

$$\nabla_{y_3} Q_{Y_2|Y_3Z} = \nabla_{y_3} h_2 - \frac{\nabla_{z_{(3)}} Q_{Y_2|Y_3Z}}{\nabla_{z_{(3)}} Q_{Y_3|Z}}$$

and finally

$$\pi_{23}(\tau^*, z^*) = \nabla_{y_3} Q_{Y_2|Y_3Z} + \frac{\nabla_{z_{(3)}} Q_{Y_2|Y_3Z}}{\nabla_{z_{(3)}} Q_{Y_3|Z}}.$$

(c) *Identification of  $\pi_{2z_k}(\tau^*, z^*)$*

Differentiating (8) with respect to  $z_k$  and  $z_{(3)}$  gives the following.

$$\nabla_{z_k} Q_{Y_2|Y_3Z} = \nabla_{z_k} h_2 + \nabla_{\nu_3} h_2 \nabla_{z_k} g_3 \quad (14)$$

$$\nabla_{z_{(3)}} Q_{Y_2|Y_3Z} = \nabla_{\nu_3} h_2 \nabla_{z_{(3)}} g_3 \quad (15)$$

Note that if  $Z_k = Z_{(3)}$  then the first term on the right hand side of (14) is absent. Differentiating (12),

$$\nabla_{z_k} g_3 + \nabla_{y_3} g_3 \nabla_{z_k} Q_{Y_3|Z} = 0 \quad (16)$$

$$\nabla_{z_{(3)}} g_3 + \nabla_{y_3} g_3 \nabla_{z_{(3)}} Q_{Y_3|Z} = 0 \quad (17)$$

and multiplying (17) by  $\nabla_{\nu_3} h_2$ ,

$$\begin{aligned} \nabla_{\nu_3} h_2 \nabla_{z_{(3)}} g_3 &= -\nabla_{\nu_3} h_2 \nabla_{y_3} g_3 \nabla_{z_{(3)}} Q_{Y_3|Z} \\ &= \nabla_{z_{(3)}} Q_{Y_2|Y_3Z} \end{aligned}$$

the second line following from (15). Therefore

$$\nabla_{\nu_3} h_2 \nabla_{y_3} g_3 = -\frac{\nabla_{z_{(3)}} Q_{Y_2|Y_3Z}}{\nabla_{z_{(3)}} Q_{Y_3|Z}}. \quad (18)$$

The denominator on the right hand side of (18) is nonzero by virtue of the rank condition (5b). From (16)

$$\begin{aligned} \nabla_{\nu_3} h_2 \nabla_{z_k} g_3 &= -\nabla_{\nu_3} h_2 \nabla_{y_3} g_3 \nabla_{z_k} Q_{Y_3|Z} \\ &= \frac{\nabla_{z_{(3)}} Q_{Y_2|Y_3Z}}{\nabla_{z_{(3)}} Q_{Y_3|Z}} \nabla_{z_k} Q_{Y_3|Z} \end{aligned}$$

the second line following from (18). Finally, substituting in (14)

$$\nabla_{z_k} Q_{Y_2|Y_3Z} = \nabla_{z_k} h_2 + \frac{\nabla_{z_{(3)}} Q_{Y_2|Y_3Z}}{\nabla_{z_{(3)}} Q_{Y_3|Z}} \nabla_{z_k} Q_{Y_3|Z}$$

and on rearranging

$$\pi_{2z_k}(\tau^*, z^*) = \nabla_{z_k} Q_{Y_2|Y_3Z} - \nabla_{z_k} Q_{Y_3|Z} \times \frac{\nabla_{z_{(3)}} Q_{Y_2|Y_3Z}}{\nabla_{z_{(3)}} Q_{Y_3|Z}}$$

The manipulations to obtain the remaining results of the Theorem are lengthy and are not reported here. The general approach set out in Section 3 shows how such results can be obtained in an elementary, rule driven fashion. ■

This Section concludes with an illustration of the use of the Theorem in a parametric linear simultaneous equations location shift model.

**2.1. Example: A linear simultaneous equations location shift model.** Consider the following linear simultaneous equations location shift model.

$$\begin{aligned} Y_1 &= \theta_{12}Y_2 + \theta_{13}Y_3 + \beta'_1 Z_1 + \nu_1 + \kappa_{12}\nu_2 + \kappa_{13}\nu_3 \\ Y_2 &= \theta_{23}Y_3 + \beta'_2 Z_1 + \beta_{2(2)}Z_{(2)} + \nu_2 + \kappa_{23}\nu_3 \\ Y_3 &= \beta'_3 Z_1 + \beta_{3(2)}Z_{(2)} + \beta_{3(3)}Z_{(3)} + \nu_3 \end{aligned}$$

Here  $Z_1$  contains covariates which may appear in any of the equations,  $Z_{(2)}$  is a covariate which does not appear in the first equation and  $Z_{(3)}$  is a covariate which appears in neither the first or the second equation. The covariates  $Z_{(2)}$  and  $Z_{(3)}$  satisfy the local order conditions<sup>18</sup> (4a) - (4c) and do not appear in  $Z_1$ .

Assume that the unobservable random variables  $\nu_1$ ,  $\nu_2$  and  $\nu_3$  are continuously distributed and that the quantile independence assumptions of the Theorem hold at some set of probabilities,  $\tau^*$ , for example  $\tau^* = \{0.5, 0.5, 0.5\}$ , requiring median independence. The quantile independence assumptions need not apply at other values of  $\tau^*$ , allowing the conditional distributions of the  $\nu$ 's to exhibit dispersion and shape dependence on  $Z_1$ ,  $Z_{(2)}$  and  $Z_{(3)}$ . There is no requirement that the  $\nu$ 's have moments of any order. The monotonicity, continuity and differentiability assumptions of the Theorem are obviously satisfied.

There are the following derivatives of the structural functions.

$$\left. \begin{aligned} \pi_{12}(\tau^*, z^*) &= \theta_{12} & \pi_{13}(\tau^*, z^*) &= \theta_{13} & \pi_{1z_{ik}}(\tau^*, z^*) &= \beta_{1k} \\ \pi_{23}(\tau^*, z^*) &= \theta_{23} & \pi_{2z_{1k}}(\tau^*, z^*) &= \beta_{2k} & \pi_{3z_{1k}}(\tau^*, z^*) &= \beta_{3k} \end{aligned} \right\} \quad (19)$$

Here  $k$  may identify any of the  $Z$  variables. Note that in this linear location shift model these derivatives are invariant with respect to  $\tau^*$  and  $z^*$ .

Suppose that the elements of  $\nu$  are mutually independently distributed<sup>19</sup>. Then the conditional quantile derivatives (at  $\tau^*$ ) appearing in the equations identifying the structural function derivatives are as follows.

$$\begin{aligned} \nabla_{y_2} Q_{Y_1|Y_2Y_3Z} &= \theta_{12} + \kappa_{12} & \nabla_{y_3} Q_{Y_1|Y_2Y_3Z} &= \theta_{13} - \kappa_{12}\theta_{23} - (\kappa_{12}\kappa_{23} - \kappa_{13}) \\ \nabla_{z_{1k}} Q_{Y_1|Y_2Y_3Z} &= \beta_{1k} - \kappa_{12}\beta_{2k} + (\kappa_{12}\kappa_{23} - \kappa_{13})\beta_{3k} \\ \nabla_{z_{(2)}} Q_{Y_1|Y_2Y_3Z} &= -\kappa_{12}\beta_{2(2)} + (\kappa_{12}\kappa_{23} - \kappa_{13})\beta_{3(2)} & \nabla_{z_{(3)}} Q_{Y_1|Y_2Y_3Z} &= (\kappa_{12}\kappa_{23} - \kappa_{13})\beta_{3(3)} \\ \nabla_{y_3} Q_{Y_2|Y_3Z} &= \theta_{23} + \kappa_{23} & \nabla_{z_{1k}} Q_{Y_2|Y_3Z} &= \beta_{2k} - \kappa_{23}\beta_{3k} & \nabla_{z_{(2)}} Q_{Y_2|Y_3Z} &= \beta_{2(2)} - \kappa_{23}\beta_{3(2)} \end{aligned}$$

<sup>18</sup>Of course in this linear model when the local order conditions hold, they hold globally.

<sup>19</sup>This independence assumption is made for convenience, as in the proof in Section 2. If the  $\nu_i$ 's were dependent then they could be transformed to independence, as in the proof, but this would in general introduce nonlinear functions of the transformed variables into the system. This would be of no consequence for the identification result but would complicate this demonstration.

$$\nabla_{z(3)} Q_{Y_2|Y_3Z} = -\kappa_{23}\beta_{3(3)} \quad \nabla_{z_{1k}} Q_{Y_3|Z} = \beta_{3k} \quad \nabla_{z(2)} Q_{Y_3|Z} = \beta_{3(2)} \quad \nabla_{z(3)} Q_{Y_3|Z} = \beta_{3(3)}$$

If the rank conditions (5a) and (5b) are satisfied, that is:  $\beta_{2(2)} \neq 0$  and  $\beta_{3(3)} \neq 0$ , then it is clear, on working in reverse order through the equations above, that all the derivatives of the structural functions are just identified, and that the resulting expressions are those given in the Theorem. Along the way the  $\kappa_{ij}$ 's are identified, for example

$$\kappa_{23} = -\frac{\nabla_{z(3)} Q_{Y_2|Y_3Z}}{\nabla_{z(3)} Q_{Y_3|Z}}.$$

If the quantile independence assumption were replaced by a mean independence assumption then the structural equation derivatives (19) would be derivatives of mean regressions, that is of  $E_{Y_1|Y_2Y_3Z}[Y_1|Y_2 = y_2, Y_3 = y_3, Z = z]$ , and so forth. This suggests an estimation strategy in this case, namely: calculate OLS estimates of the coefficients in these regression functions and combine the results as indicated in the Theorem. In fact this is just a rather unorthodox rendition of the Indirect Least Squares estimator.

Under median independence the analogous procedure is to calculate estimates of coefficients of the median regression functions, for example using the Least Absolute Deviations method, and combine the results as above. This provides an alternative to the Two Stage Least Absolute Deviations procedures proposed by Amemiya (1982).

More complex versions of this parametric model are easily handled. For example the first equation could be replaced by

$$Y_1 = \theta_{12}^{\phi_{11} + \phi_{12}\varepsilon_2} Y_2 + \theta_{13}^{\phi_{13} + \phi_{14}\varepsilon_3} Y_3 + \beta'_1 Z_1 + \varepsilon_1 + \kappa_{12}\varepsilon_2 + \kappa_{13}\varepsilon_3$$

and the assumptions of the Theorem would remain satisfied. With this amendment some of the structural derivatives depend upon the stochastic drivers of the system and the EIFs (Chesher (2001a)) will vary with the chosen probability levels,  $\tau^*$ . Much more complex structures also fall within the scope of the Theorem, there being essentially no limit on the way in which  $\varepsilon_j$ ,  $j > i$ , appear in the equation for  $Y_i$  as long as the resulting equation is differentiable (at the point  $\mathcal{X}$ ) with respect to the  $\varepsilon_j$ 's and strictly monotonic in  $\varepsilon_i$ .

### 3. IDENTIFICATION IN M EQUATION SYSTEMS

The manipulations in the Proof of the previous Section are quite tedious. One would not wish to proceed in this fashion in a much larger problem.

In fact the analysis of the linear model in Section 2.1 gives the clue to obtaining an easily applicable result in triangular systems of any size, because the linear model can be regarded as an approximation at the point  $\mathcal{X}$  to the nonlinear structure of interest, the coefficients in the linear model being the derivatives of the structural equations at the point  $\mathcal{X}$ . These derivatives are the functionals of the structural model whose identification is sought.

Consider a triangular structural system determining  $M$  responses as set out in Section 1. It is convenient to work in terms of independently distributed stochastic drivers,  $\nu = \{\nu_i\}_{i=1}^M$  which, as shown in the proof of the Theorem of Section 2, constitutes an innocuous normalisation<sup>20</sup>. It is assumed that obvious  $M$ -variate extensions of the conditions of the Theorem apply at  $\mathcal{X}$ .

Variations in  $Y$ ,  $Z$  and  $\nu$  at the point  $\mathcal{X}$ , satisfy

$$dY = AdY + BdZ + Cdv \tag{20}$$

---

<sup>20</sup>Under this normalisation the values of the  $\nu_i$ 's are the probabilities defining iterated conditional quantiles of the stochastic drivers.

where  $A$ ,  $B$  and  $C$  are matrices of structural derivatives evaluated at  $\mathcal{X}$  and the vectors of differentials are defined as follows.

$$dY = \begin{bmatrix} dY_1 \\ dY_2 \\ \vdots \\ dY_M \end{bmatrix} \quad dZ = \begin{bmatrix} dZ_1 \\ dZ_2 \\ \vdots \\ dZ_K \end{bmatrix} \quad d\nu = \begin{bmatrix} d\nu_1 \\ d\nu_2 \\ \vdots \\ d\nu_M \end{bmatrix}$$

The matrices of derivatives,  $A$ ,  $B$  and  $C$ , are as shown below. Note that these derivatives are evaluated at the point  $\mathcal{X}$  and that the assumed triangular structure has been imposed on  $A$  and  $C$ . The leading diagonal elements of  $C$ ,  $\nabla_{\nu_i} h_i$ , have been set equal to 1. This is an innocuous normalisation given the assumption that each function  $h_i$  is, at the point  $\mathcal{X}$ , a differentiable and strictly monotonic function of  $\nu_i$ .

$$A = \begin{bmatrix} 0 & \nabla_{y_2} h_1 & \nabla_{y_3} h_1 & \dots & \nabla_{y_M} h_1 \\ 0 & 0 & \nabla_{y_3} h_2 & \dots & \nabla_{y_M} h_2 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & & \nabla_{y_M} h_{M-1} \\ 0 & 0 & 0 & \dots & 0 \end{bmatrix}$$

$$B = \begin{bmatrix} \nabla_{z_1} h_1 & \nabla_{z_2} h_1 & \dots & \nabla_{z_K} h_1 \\ \nabla_{z_1} h_2 & \nabla_{z_2} h_2 & \dots & \nabla_{z_K} h_2 \\ \vdots & \vdots & & \vdots \\ \nabla_{z_1} h_M & \nabla_{z_2} h_M & \dots & \nabla_{z_K} h_M \end{bmatrix}$$

$$C = \begin{bmatrix} 1 & \nabla_{\nu_2} h_1 & \nabla_{\nu_3} h_1 & \dots & \nabla_{\nu_M} h_1 \\ 0 & 1 & \nabla_{\nu_3} h_2 & \dots & \nabla_{\nu_M} h_2 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \dots & \nabla_{\nu_M} h_{M-1} \\ 0 & 0 & 0 & \dots & 1 \end{bmatrix}$$

The task now is to forge a link between the matrices of derivatives,  $A$ ,  $B$  and  $C$ , above, and the derivatives of the iterated conditional quantile functions at the point  $\mathcal{X}$ .

The iterated conditional quantile functions are determined by a transformed structural system in which, in each equation  $i$ , the *single* unobservable,  $\nu_i$ , appears, the remaining  $\nu_j$ 's ( $j > i$ ) having been substituted away, using the inverse functions associated with the equations for  $Y_j$ ,  $j > i$ . Equations (7), (8) and (9) are the equations of this transformed structural form in the three equation case.

The derivatives of the transformed structural equations, and hence of the iterated conditional quantile functions, are obtained as follows.

First rewrite (20) as follows.

$$dY = AdY + BdZ + d\nu + (C - I_M) d\nu \quad (21)$$

The first appearance of  $d\nu$  in (21) isolates  $\nu_i$  in the  $i$ th structural equation. Now note that the matrix  $C$  is upper triangular with nonzero leading diagonal elements and so is nonsingular. It follows that, from (20),

$$d\nu = C^{-1} (dY - AdY - BdZ) \quad (22)$$

and on substituting (22) in (21) there is

$$dY = AdY + BdZ + d\nu + (C - I_M) C^{-1} (dY - AdY - BdZ)$$



and, after rearrangement, the following.

$$dY = (I_M - C^{-1}(I_M - A)) dY + C^{-1}BdZ + d\nu \quad (23)$$

Equation (23) shows how each differential  $dY_i$  is related to the differentials  $dZ$ ,  $d\nu_i$  and  $dY_j$ ,  $j > i$ , all the differentials  $d\nu_j$ ,  $j > i$ , having been expressed in terms of differentials  $dY_j$ ,  $j > i$  and  $dZ$ .

Define the vector of differentials of the iterated conditional quantile functions,  $dQ$ , as follows.

$$dQ = \begin{bmatrix} dQ_{Y_1|Y_2\dots Y_M Z} \\ dQ_{Y_2|Y_3\dots Y_M Z} \\ \vdots \\ dQ_{Y_{M-1}|Y_M Z} \\ dQ_{Y_M|Z} \end{bmatrix}$$

At a point  $\mathcal{X}$  at which there is monotonicity and quantile independence as assumed in the Theorem, equation (23) with  $dQ$  replacing  $dY$  on its left hand side gives the relationship between the differentials of the iterated conditional quantile functions and differentials,  $dZ$  and  $dY$ , of the variates upon which conditioning is done, that is:

$$dQ = (I_M - C^{-1}(I_M - A)) dY + C^{-1}BdZ.$$

It follows that the derivatives with respect to  $Y$  and  $Z$  of the conditional quantile functions at the point  $\mathcal{X}$  are  $D_Y$  and  $D_Z$  defined below.

$$D_Y = \frac{dQ}{dY} = I_M - C^{-1}(I_M - A)$$

$$D_Z = \frac{dQ}{dZ} = C^{-1}B$$

Knowledge of the joint distribution of  $Y$  given  $X$  implies knowledge of the iterated conditional quantile functions, and *vice versa*, and under differentiability assumptions of the sort imposed in the Theorem it implies knowledge of the derivatives of the iterated conditional quantile functions at  $\mathcal{X}$ , and therefore of the following matrices.

$$D_Y = I_M - C^{-1}(I_M - A) \quad (24)$$

$$D_Z = C^{-1}B \quad (25)$$

Elements of  $A$ ,  $B$  and  $C$  are therefore identifiable if *a priori* restrictions placed on  $A$ ,  $B$  and  $C$  are sufficient to allow those elements to be deduced uniquely from  $D_Y$  and  $D_Z$ .

**3.1. Identification via restrictions on  $B$ .** The upper triangular structures of  $A$  (whose leading diagonal contains zeros) and of  $C$  imply that  $D_Y$  is upper triangular with leading diagonal elements containing zeros. But this is an intrinsic property of the matrix of derivatives of the iterated conditional quantile functions. Therefore the restrictions on  $A$  and  $C$  imposed by triangularity contain no additional information relevant to identification of the superdiagonal elements of  $A$  and  $C$ .

The relationship between  $A$ ,  $C$  and  $D_Y$  implied by (24) can be written as follows.

$$A = I_M - C(I_M - D_Y)$$

It is clear that if the superdiagonal elements of  $C$  can be deduced from (25) then the superdiagonal elements of  $A$  can be deduced from  $D_Y$ . Now consider the identification of  $C$  and  $B$  from (25) using just restrictions on the elements of  $B$ . Identification when there are additionally restrictions on  $A$  and  $C$  is considered in Section 3.2.

Rewrite the relationship (25) as

$$(D'_Z \otimes I_M) \text{vec } C = \text{vec } B$$

and consider linear restrictions on  $B$  which are expressed as

$$W_B \text{vec } B = w$$

where  $W$  is a  $N \times (MK)$  matrix of known constants and  $w$  is a  $N \times 1$  vector of known constants. If the restrictions are all exclusion restrictions then  $w = 0$ . Note that in this local (at the point  $\mathcal{X}$ ) analysis, linear restrictions on  $B$  can be regarded as local approximations at the point  $\mathcal{X}$  to *nonlinear* restrictions subject to satisfaction of appropriate assumptions concerning differentiability of the functions involved.

It is convenient now to proceed with  $C$  completely unrestricted and, when considering identification of  $C$  and  $B$  from (25), to explicitly impose the constraints on  $C$  implied by its triangular structure with the leading diagonal elements of  $C$  normalised to be one.

These restrictions on  $C$  are written as

$$R \text{vec } C = r$$

where  $R$  is a  $M(M+1)/2 \times M^2$  matrix and  $r$  is a  $M(M+1)/2 \times 1$  vector, as shown below. Here  $O_j$  denotes a  $(M-j) \times j$  matrix of zeros,  $O^j$  denotes a  $(M-j) \times M$  matrix of zeros and  $o_j$  denotes a  $(M-j) \times 1$  vector of zeros.

$$R = \begin{bmatrix} [I_M] & O^0 & O^0 & \dots & O^0 & O^0 \\ O^1 & [O_1 : I_{M-1}] & O^1 & \dots & O^1 & O^1 \\ O^2 & O^2 & [O_2 : I_{M-2}] & \dots & O^2 & O^2 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ O^{M-2} & O^{M-2} & O^{M-2} & \dots & [O_{M-2} : I_2] & O^{M-2} \\ O^{M-1} & O^{M-1} & O^{M-1} & \dots & O^{M-1} & [O_{M-1} : I_1] \end{bmatrix}$$

$$r' = \begin{bmatrix} 1 & o'_1 & \vdots & 1 & o'_2 & \vdots & 1 & o'_3 & \vdots & \dots & \vdots & 1 & o'_{M-1} & \vdots & 1 \end{bmatrix} \quad (26)$$

There is then the following equation.

$$\begin{bmatrix} -I_{KM} & (D'_Z \otimes I_M) \\ W_B & O_{N \times M^2} \\ O_{M(M+1)/2 \times (KM)} & R \end{bmatrix} \begin{bmatrix} \text{vec } B \\ \text{vec } C \end{bmatrix} = \begin{bmatrix} O_{M^2 \times 1} \\ O_{N \times 1} \\ r \end{bmatrix} \quad (27)$$

where  $O_{(a \times b)}$  indicates an  $a \times b$  matrix of zeros. This equation is written as follows.

$$\underbrace{\hspace{15em}}_F \times \underbrace{\hspace{5em}}_\Psi = \underbrace{\hspace{15em}}_f$$

$$(KM + N + M(M+1)/2) \times (KM + M^2) \quad (KM + M^2) \times 1 \quad (KM + N + M(M+1)/2) \times 1$$

Clearly it is impossible to solve for  $\Psi$  if the column order of  $F$  is less than its row order. There is therefore the following necessary *order condition* for identification.

$$\text{Order condition: } N \geq \frac{M(M-1)}{2}$$

The following rank condition ensures identifiability of  $B$  and  $C$  and therefore, by the argument given above, of  $A$ .

$$\text{Rank condition: } \text{rank}(F) \geq M^2 + KM.$$

It is clear that  $R$  has full column rank. Suppose  $D_Z$  also has full column rank, which rules out exact functional relationships amongst the elements of  $Z$ . Then the rank condition is satisfied if and only if  $\text{rank}(W_B) \geq M(M-1)/2$ .

**3.2. Identification using restrictions on  $A$ ,  $B$  and  $C$ .** It is easy enough to consider more general (possibly nonlinear) restrictions involving  $A$ ,  $B$ , and  $C$  which are written (possibly linearised at the point  $\mathcal{X}$ ) as follows.

$$W_A \text{vec } A + W_B \text{vec } B + W_C \text{vec } C = w \tag{28}$$

The restrictions on  $A$  imposed by the assumption of triangularity (with zero leading diagonal elements) are expressed by the following equation

$$R \text{vec } A = O_{M(M+1)/2 \times 1}$$

in which  $R$  is as defined earlier. As before the triangularity of  $C$  together with the normalisation restrictions on its leading diagonal elements are embodied in equation (26).

With the restrictions now all in place there is the following equation, from which the order and rank conditions can be deduced.

$$\begin{bmatrix} R & O_{M(M+1)/2 \times (KM)} & O_{M(M+1)/2 \times M^2} \\ I_{M^2} & O_{M^2 \times KM} & I_{M^2} - (D'_Y \otimes I_M) \\ O_{KM \times M^2} & -I_{KM} & (D'_Z \otimes I_M) \\ W_A & W_B & W_C \\ O_{M(M+1)/2 \times M^2} & O_{M(M+1)/2 \times (KM)} & R \end{bmatrix} \begin{bmatrix} \text{vec } A \\ \text{vec } B \\ \text{vec } C \end{bmatrix} = \begin{bmatrix} O_{M(M+1)/2 \times 1} \\ \text{vec } I_M \\ O_{M^2 \times 1} \\ w \\ r \end{bmatrix}$$

**3.3. Estimation.** Estimation of  $A$ ,  $B$  and  $C$ , the derivatives of the structural equations at the point  $\mathcal{X}$ , can proceed using the analogue principle<sup>21</sup>. There are the following steps.

1. Probability levels,  $\tau^*$ , are selected and the required iterated conditional quantile functions are estimated using a parametric, semi- or non-parametric method, as desired<sup>22</sup>.
2. A value of  $Z$ ,  $z^*$ , is selected and the values of  $Y_1, \dots, Y_M$  at the point  $\mathcal{X}$  are estimated using the estimated iterated conditional quantile functions.
3. The estimated iterated conditional quantile functions are used to produce estimates of derivatives of the iterated conditional quantile functions which are evaluated at the point  $\mathcal{X}$ , yielding estimates,  $\hat{D}_Y$  and  $\hat{D}_Z$ , of the matrices  $D_Y$  and  $D_Z$ .
4. The restrictions on  $A$ ,  $B$  and  $C$  are assembled as in (28), the final equation in the preceding sub-section is constructed using  $\hat{D}_Y$  and  $\hat{D}_Z$ , and subject to satisfaction of the order condition, and subject to satisfaction of the rank condition at the estimated values of  $D_Y$  and  $D_Z$ , the equation is solved for estimates of  $A$ ,  $B$  and  $C$ .

<sup>21</sup>See Manski (1988).

<sup>22</sup>For parametric estimation, see Koenker and Bassett (1978), Koenker and d'Orey (1987); for semi-parametric estimation see Chaudhuri, Doksum and Samarov (1997) and Kahn (2001); for nonparametric estimation, see Chaudhuri (1991).

At step 4, if there are abundant restrictions, for example when  $N > M(M - 1)/2$ , then there is overidentification and there is unlikely to be a solution. Solutions can be obtained by eliminating restrictions so that the order condition is exactly satisfied, but there will be many ways of doing this, each leading in general to an inefficient estimate. A solution is to retain all the restrictions and employ a minimum distance estimation procedure in place of step 4. In some parametric models one could contemplate imposing the identifying restrictions at step 1.

#### 4. AVERAGE STRUCTURAL DERIVATIVES

The results so far have concerned local nonparametric identification of structural derivatives evaluated at points defined by a value,  $z^*$ , of a vector of covariates  $Z$  and by  $M$  probabilities,  $\tau^*$ , defining iterated conditional quantiles of unobservable continuously distributed stochastic drivers,  $\varepsilon$ . If such identification can be achieved at  $z^*$  at all combinations of the  $M$  probability levels each varying in  $[0, 1]$  then it is a simple matter to identify conditional (on  $Z = z^*$ ) expected values of structural derivatives and of functions of them.

Consider a function  $g(\varepsilon, Z)$  and its conditional expectation (assumed to exist) given  $Z = z$

$$E_{\varepsilon|Z}[g(\varepsilon, Z)|Z = z] = \int \dots \int g(e, z) f_{\varepsilon|Z}(e|z) de_1 \dots de_M.$$

Let  $\tilde{g}(\tau, Z) = g(\varepsilon(\tau), Z)$  where  $\tau = \{\tau_i\}_{i=1}^M$ , and  $\varepsilon(\tau) = \{\varepsilon_i(\tau)\}_{i=1}^M$  whose elements are defined recursively in terms of conditional quantile functions as follows.

$$\begin{aligned} \varepsilon_i(\tau) &= Q_{\varepsilon_i|\varepsilon_{i+1}\dots\varepsilon_M Z}(\tau_i|\varepsilon_{i+1}, \dots, \varepsilon_M, z), \quad i \in \{1, \dots, M - 1\} \\ \varepsilon_M(\tau) &= Q_{\varepsilon_M|Z}(\tau_M|z) \end{aligned}$$

Since the elements of  $\tau$  are mutually independently uniformly distributed, each on  $[0, 1]$ , and distributed independently of  $Z$ , it follows directly that

$$\begin{aligned} E_{\varepsilon|Z}[g(\varepsilon, Z)|Z = z] &= E_{\tau}[g(\varepsilon(\tau), z)] \\ &= \int_0^1 \dots \int_0^1 g(e(t), z) dt_1 \dots dt_M. \end{aligned}$$

The structural derivatives evaluated at the point  $\mathcal{X}$  considered in the Theorem of Section 2 are precisely structural derivatives expressed, not in terms of the unobservable stochastic drivers, but instead in terms of probabilities ( $\tau^*$ ) defining iterated conditional quantiles of the conditional distribution of  $\varepsilon$  given  $Z = z^*$ . Therefore expected values of functions of structural derivatives (for example average derivatives, variances) can be identified, assuming the local identification conditions hold at  $Z = z^*$  for all combinations of quantile probability levels - the conditional expected value of a function  $r(\cdot)$  of a structural derivative  $\pi(\tau, z)$  is simply  $\int_0^1 \dots \int_0^1 r(\pi(t, z)) dt_1 \dots dt_M$ .

#### 5. CONCLUDING REMARKS

This paper has shown that in nonlinear triangular structural systems, conditions under which there is nonparametric identification of certain interesting features of structures are quite easily obtained if the question of identifiability is couched in terms of the feasibility of uniquely deducing structural features from knowledge of an exhaustive set of iterated conditional quantile functions. An analysis via conditional quantile functions is well suited to nonlinear structural systems because of the equivariance property of quantiles.

Focusing on local identification of the derivatives of structural functions at a point of interest reduces the problem to a search for identification conditions in a linear approximation to the nonlinear structure, as long as there is sufficient smoothness to allow linear approximation at the point of interest. As a result the familiar order and rank conditions for identification in the linear simultaneous equations system appear in an unfamiliar setting.

A quantile based analysis naturally leads to identification of structural features which convey information about the distribution of exogenous impacts of variables perhaps endogenous in the data generating process. This will be of value in the analysis of policy interventions.

The estimation procedures that flow from the identification conditions involve various sorts of quantile regression function estimation. Quantile regression methods have frequently been advocated because they may be less sensitive to data contamination and because they can provide a wealth of distributional information. The results of this paper suggest a further, and very important, virtue of quantile regression methods, namely that, because of their equivariance property, exploited here, they are the natural tool to employ in identification, estimation and inference in nonlinear structural models.

### Annex: Weakening the local quantile independence conditions

This Annex examines a two equation version of the model studied in this paper and examines the identification conditions when the stochastic drivers of the system are *not* transformed to be independently distributed. This allows easy appreciation of one way in which the identification conditions of the Theorem of Section 2 can be weakened while retaining the identification result.

Consider the following two equation model for scalar  $Y_1$  and  $Y_2$ .

$$\begin{aligned} Y_1 &= h_1(Y_2, Z, \varepsilon_1, \varepsilon_2) \\ Y_2 &= h_2(Z, \varepsilon_2) \end{aligned}$$

The conditions of the Theorem of Section 2 are assumed to apply.

Substituting for  $Y_2$  in  $h_1$  and exploiting the monotonicity of  $h_1$  with respect to  $\varepsilon_1$  gives, at  $\mathcal{X}$ ,

$$Q_{Y_1|Y_2Z}(\tau_1^*, y_2^*, z^*) = h_1(y_2^*, z^*, Q_{\varepsilon_1|\varepsilon_2Z}(\tau_1^*, g_2(z^*, y_2^*), z^*), g_2(z^*, y_2^*))$$

where  $g_2$  is the inverse function associated with  $h_2$  satisfying

$$Y_2 = h_2(Z, g_2(Z, Y_2)).$$

Differentiating with respect to  $y_2$  and an element,  $z_i$ , of  $z$  satisfying the local order and rank conditions of the Theorem, and evaluating at  $\mathcal{X}$  gives the following (arguments suppressed).

$$\nabla_{y_2} Q_{Y_1|Y_2Z} = \nabla_{y_2} h_1 + (\nabla_{\varepsilon_1} h_1 \nabla_{\varepsilon_2} Q_{\varepsilon_1|\varepsilon_2Z} + \nabla_{\varepsilon_2} h_1) \nabla_{y_2} g_2$$

$$\nabla_{z_i} Q_{Y_1|Y_2Z} = \nabla_{z_i} h_1 + \nabla_{\varepsilon_1} h_1 \nabla_{z_i} Q_{\varepsilon_1|\varepsilon_2Z} + (\nabla_{\varepsilon_1} h_1 \nabla_{\varepsilon_2} Q_{\varepsilon_1|\varepsilon_2Z} + \nabla_{\varepsilon_2} h_1) \nabla_{z_i} g_2 \quad (\text{A1})$$

The ratio  $\nabla_{z_i} g_2 / \nabla_{y_2} g_2$  is identifiable as the  $z_i$  derivative of  $Q_{Y_2|Z}$  at  $\mathcal{X}$ .

To achieve identifiability of  $\nabla_{y_2} h_1$  using the approach adopted in this paper the *first two* terms on the right-hand side of (A1) must be zero<sup>23</sup> at  $\mathcal{X}$ . The first term is zero by virtue of the local order condition.

---

<sup>23</sup>At least restricted to known values.

Equation (A1) reveals that all that is required so far as “quantile independence” is concerned is that the first partial derivative with respect to the local instrument,  $z_i$ , of the conditional  $\tau_1^*$ -quantile of  $\varepsilon_1$  given  $\varepsilon_2$  and  $Z$  be zero at  $\mathcal{X}$ . Full local quantile independence is not required - for example higher order derivatives of the conditional  $\tau_i^*$ -quantile of  $\varepsilon_1$  with respect to  $z_i$  could be nonzero at  $\mathcal{X}$  and local identification of the first<sup>24</sup> derivative of  $h_1$  with respect to  $Y_2$  at  $\mathcal{X}$  could be achieved. Nor is there any need to assume zero values at  $\mathcal{X}$  for derivatives of the conditional  $\tau_1^*$ -quantile of  $\varepsilon_1$  given  $\varepsilon_2$  and  $Z$  with respect to elements of  $z$  which are not candidate local instrumental variables.

## REFERENCES

- ABADIE, A., ANGRIST, J., AND G. IMBENS (2001): “Instrumental variables estimates of the effect of subsidized training on the quantiles of trainee earnings,” *Econometrica*, forthcoming.
- ALTONJI, J.G., AND R.L. MATZKIN, (2001): “Panel data estimators for nonseparable models with endogenous regressors,” NBER Technical Working Paper 267.
- AMEMIYA, T., (1982): “Two stage least absolute deviations estimators,” *Econometrica*, 50, 689-711.
- BLUNDELL, R.W., AND J.L. POWELL (2000): “Endogeneity in Nonparametric and Semiparametric Regression Models,” invited lecture, 2000 World Congress of the Econometric Society, Seattle.
- BROWN, B.W., (1983): “The identification problem in systems nonlinear in the variables,” *Econometrica*, 51,175-196.
- BROWN, B.W., AND R.L. MATZKIN (1996): “Estimation of nonparametric functions in simultaneous equations models, with an application to consumer demand,” mimeo, Department of Economics, Northwestern University.
- CHAUDHURI, P., (1991): “Nonparametric estimation of regression quantiles and their local Bahadur representation,” *Annals of Statistics*, 19, 760-777.
- CHAUDHURI, P., K. DOKSUM AND A. SAMAROV (1997): “On average derivative quantile regression,” *Annals of Statistics*, 25, 715-744.
- CHESHER, A.D., (2001a): “Exogenous impact and conditional quantile functions,” Centre for Microdata Methods and Practice Working Paper 01/01.
- CHESHER, A.D., (2001b): “Weak endogeneity,” mimeo.
- DAROLLES, S., J-P FLORENS AND E. RENAULT, (2000): “Nonparametric instrumental regression,” CREST Documents de Travail 2000-17.
- DAS, M., (2000): “Nonparametric instrumental variable estimation with discrete endogenous regressors,” presented at the 2000 World Congress of the Econometric Society, Seattle.
- DEATON, A., AND J. MUELLBAUER, (1980): “An almost ideal demand system,” *American Economic Review*, 70, 312-326.
- FISHER, F.M., (1959): “Generalization of the rank and order conditions for identifiability,” *Econometrica*, 27, 431-447.
- FISHER, F.M., (1961): “Identifiability criteria in nonlinear systems,” *Econometrica*, 29, 574-590.
- FISHER, F.M., (1966): *The identification problem in econometrics*, New York: McGraw Hill.
- FRISCH, R., (1934): *Statistical confluence analysis by means of complete regression systems*, Publication No. 5, Oslo: Universitets Økonomiske Institutt.
- HECKMAN, J.J., (1990): “Varieties of selection bias,” *American Economic Review, Papers and Proceedings*, 80, 313-318.

<sup>24</sup>But not, e.g., the second derivative.

- HECKMAN, J.J., J. SMITH AND N. CLEMENTS (1997): "Making the most out of programme evaluations and social experiments: accounting for heterogeneity in programme impacts," *The Review of Economic Studies*, 64, 487-535.
- IMBENS, G.W., AND J. ANGRIST (1994): "Identification and estimation of average treatment effects," *Econometrica*, 62, 467-476.
- IMBENS, G.W., AND W.K. NEWEY (2001): "Identification and estimation of triangular simultaneous equations models without additivity," mimeo.
- KAHN, S., (2001): "Two-stage rank estimation of quantile index models," *Journal of Econometrics*, 100, 319-355.
- KOENKER, R., AND G. BASSETT JR. (1978): "Regression quantiles," *Econometrica*, 46, 33-50.
- KOENKER, R.W. AND V. D'OREY (1987): "Computing regression quantiles," *Journal of the Royal Statistical Society, Series C*, 36, 383-393.
- KOOPMANS, T.C., H. RUBIN AND R.B. LEIPNIK (1950): "Measuring the equation systems of dynamic economics," in *Statistical inference in dynamic economic models*. Cowles Commission Monograph 10, New York, John Wiley.
- MANSKI, C.F., (1988): *Analog estimation methods in econometrics*, New York: Chapman and Hall.
- MATZKIN, R.L., (1999): "Nonparametric estimation of nonadditive random functions," Invited lecture on New Development in the Estimation of Preferences and Production Functions, 1999 Latin American Meeting of the Econometric Society.
- NEWEY, W.K., AND J.L. POWELL (1988): "Nonparametric Instrumental Variables Estimation," mimeo, Department of Economics, MIT.
- NEWEY, W.K., AND J.L. POWELL (1990): "Efficient estimation of linear and Type-1 censored regression models under conditional quantile restrictions," *Econometric Theory*, 6, 295-317.
- NEWEY, W.K., J.L. POWELL, AND F. VELLA (1999): "Nonparametric Estimation of Triangular Simultaneous Equations Models," *Econometrica* 67, 565-603.
- PINKSE, J., (2000): "Nonparametric two-step regression functions when regressors and errors are dependent," *Canadian Journal of Statistics*, 28, 289-300.
- POWELL, J.L., (1983): "The asymptotic normality of two-stage least absolute deviations estimators," *Econometrica*, 51, 1569-1576.
- ROEHRIG, C.S., (1988): "Conditions for identification in nonparametric and parametric models", *Econometrica* 56, 433-447.
- ROTHENBERG, T.J., (1971): "Identification in parametric models," *Econometrica*, 39, 577-591.
- VYTLACIL, E., (2001): "Independence, monotonicity and latent index models: an equivalence result," *Econometrica*, forthcoming.
- WALD, A., (1950): "Note on identification of economic relations," in *Statistical inference in dynamic economic models*. Cowles Commission Monograph 10, New York, John Wiley.
- WORKING, E.J., (1925): "The statistical determination of demand curves," *Quarterly Journal of Economics*, 39, 503-543.
- WORKING, E.J., (1927): "What do statistical 'demand curves' show?" *Quarterly Journal of Economics*, 41, 212-235.