Discussion Papers in Economics

STRUCTURAL EQUATIONS AND INVARIANCE

By

Giovanni Forchini
(University of Surrey)

DP 03/12

Department of Economics
University of Surrey
Guildford
Surrey GU2 7XH, UK
Telephone +44 (0)1483 689380
Facsimile +44 (0)1483 689548
Web www.econ.surrey.ac.uk
ISSN: 1749-5075

Structural Equations and Invariance

G. Forchini[1][2]
University of Surrey
18 January 2012

Abstract
The paper approaches structural econometric models using an algebraic approach. It shows that the invariance properties of the reduced form and the decision to exclude some of the exogenous variables from the structural equations fundamentally affect the functional form of the structural equation itself. A local approach based on Lie group theory shows that the functional form of the structural equation can be partially recovered from the invariance properties of the reduced form equations.

1. Introduction

A series of recent papers have investigated the identification of features of nonparametric structural equations with non-additive errors. The feature of interest is usually a quantification of the effect of an economic policy. Often it takes the form of *average treatment effect*, *average treatment effect on the treated* (cf. Heckman and Robb (1984) and Heckman, Ichimura and Todd (1997)), or the *quantile treatment effect* for the whole population or just for the treated (cf. Abadie, Angrist and Imbens (2002), Chernozhukov and Hansen (2005)). Sometimes, the feature of interest is a derivative (cf. Brown (1983), Roehrig (1988), Chesher (2003), Matzkin (2008)).

Seminal contributions to the literature on non-parametric identification of such quantities include work by Brown (1983) and Roehrig (1988) which has been recently critically re-examined by Benkard and Berry (2006) and revisited by Matzkin (2008). Triangular simultaneous equations models have been the focus of much recent research including Newey, Powell and Vella (1999), Chesher (2003) and Imbens and Newey (2009). Other notable recent contributions are Chernozhukov

[1] Address for correspondence: Giovanni Forchini, School of Economics, Ground Floor AD Building, University of Surrey, Guildford, Surrey, GU2 7XH, United Kingdom. E-mail: G.Forchini@surrey.ac.uk.

[2] I thank Grant Hillier for helpful discussions, Andrew Leung and Chirok Han for comments. This research was partially supported by Australian Research Council grant DP0985432.

and Hansen (2005), Chernozhukov, Imbens and Newey (2007), Hoderlein and Mammen (2007) and Matzkin (2008). Matzkin (2007) provides an excellent survey of existing results.

This paper studies how the properties of the joint density of the endogenous variables and the decision to exclude some exogenous variables from the structural equation affect the functional form of the structural equation itself. For example, suppose that the n-dimensional vector y of endogenous variables - the "reduced form" - is $y-\pi' z=u$, $u \sim N(0,\Omega)$) where π is a $k\times n$ vector with $k\geq n$, z is fixed and Ω is a positive definite covariance matrix Ω. Suppose we state that the structural equation does not contain z, so that it can be written as $t(y)=v$ for a function $t:\mathbb{R}^n\rightarrow\mathbb{R}$ where v is the structural error. Suppose that the exogenous variables could be transformed to new ones in a smooth way. For example, we could change $z\rightarrow z^*=z+g$ where $g\in\mathbb{R}^k$. This implies that, instead of focussing on an individual characterized by z, we focus on an individual characterized by $z^*=z+g$. Since $y-\pi' z=u=(y+\pi' g)-\pi' z^*$, the transformation of the exogenous variables induces a transformation of the endogenous variable $y\rightarrow y^*=y+\pi' g$ which produces a reduced form that is observationally equivalent to $y-\pi' z=u$. Therefore, the structural equation must satisfy $t(y)=v=t(y^*)=t(y+\pi' g)$. Thus, t cannot be arbitrary, and it must be, at least partially, determined by the reduced form as well as by the statement that the variable z does not appear in the structural equation. This paper extends this "invariance" approach to very general (possibly nonparametric) models, and by so doing investigates (i) to what extent the structural equation is determined by the reduced form and (ii) what features of the structural equation are uniquely defined (and are thus identified).

Our approach is algebraic in nature and is particularly appropriate to exploit the symmetries which are implicitly imposed on the model by for example stating that the structural equation does not depend on (some of) the exogenous variables. We make use of the group invariance methods described for example by Lehmann (1997) and, in more detail, by Eaton (1989) and Wijsman (1990), and Lie group theory as described for example by Olver (1993) and Olver (1996). Our contribution recognizes that group invariance methods provide a tool for a structural approach to over-identification and identification of econometric models. In particular, we show that the invariance properties of the reduced form fundamentally affect the functional form of the structural equation.

The set-up considered in this paper is very close to the one used by Brown (1983) and Matzkin (2008). However, while Brown (1983) and Matzkin (2008) are interested in system identification, we focus on an individual structural equation. We also explicitly adopt the specific conceptual framework whereby the specification of the joint distribution of the endogenous variables - the reduced form - logically precedes the formulation of the structural equation (see Poskitt and Skeels (2008) for a discussion in the context of linear structural equations). Our line of enquiry involves the compatibility of the structural equation with the reduced form and as such it is primarily concerned with the specification of the model. We show that one often needs to impose restrictions on both the reduced form and the structural equation for them to be compatible. By re-examining some examples by Matzkin (2008), we find that some of her conditions for observational equivalence can be interpreted as statements of compatibility of structural and reduced form equations. Moreover, since the characteristics of the structural equation depend on the characteristics of the reduced form, our analysis also helps to understand which structural properties are identified and which ones are not. The algebraic approach allows us to relax some of the assumptions in the work of Brown (1983) and Matzkin (2008) as well as to highlight different aspects of the problem. Finally, the use of Lie group theory allows us to express the structural equation as the solution to a system of partial differential equations so that, in some cases, the functional form of the structural equations can be partly recovered in closed form from the invariance properties of the reduced form.

The rest of the paper is organised as follows. Section 2 proposes a very general global framework to determine the compatibility between the structural equation and the reduced form. Section 3 suggests a practical implementation of the ideas developed in Section 2. This is based on local considerations and Lie group actions. The conclusions follow.

2. A general framework

This section introduces a general framework for defining over-identification of a general structural model. First, we study the reduced form and establish conditions for observational equivalence of two reduced forms. Then, we discuss the "structural equation" and its *compatibility* with the reduced form. Some examples follow.

2.1 The reduced form

Our starting point is the joint distribution of the endogenous variables y having the form

(1) $$\eta(z,y)=u\ ,$$

where $\eta:\mathbb{R}^k\times\mathbb{R}^n\to\mathbb{R}^n$ and u has probability distribution P_u. Extending the terminology from the linear structural equations model, we refer to (1) as the *reduced form*. We define the *structure* of the model for y as the pair $S=(\eta,P_u)$.

Consider another model for y written as

$$\eta^*(z,y)=u^*,$$

where u^* has probability P_{u^*} and $\eta^*:\mathbb{R}^n\times\mathbb{R}^k\to\mathbb{R}^n$. The structure for this model is denoted by $S^*=(\eta^*,P_{u^*})$.

The probability measures P_u and P_{u^*} do not depend on z or z^* (c.f. Matzkin (2008)). This measure is defined on the Borel sigma-algebra B of $\mathbb{R}^n$. We denote the probability measure for y given z and the structure S by $P_{y|z,S}(Y) = \mathrm{P}_u(\eta(z,Y))$ for any set Y for which $\eta(z,Y)$ is B-measurable (so that $Y \in \eta^{-1}(z,\mathrm{B})$).

Definition 1. Two models for y are observationally equivalent if $P_{y|z,S} = P_{y|z,S^*}$ for all $z \in \mathbb{R}^k$.

Notice that it is necessary that the two probability measures $P_{y|z,S}$ and $P_{y|z,S^*}$ are defined on the same sigma-algebra for the two models for y to be observationally equivalent. For simplicity we will assume that this is the Borel sigma-algebra B of $\mathbb{R}^n$.

We now investigate how a model for y can be transformed into an observationally equivalent one. To do this we set some restrictions on the set of structures, and assume that it is possible to transform P_u into P_{u^*} by transforming the random variable $u \in U$ to $u^* \in U$, where U is a subset of all n dimensional random variables. Our first assumption states that this can be done in a smooth way.

Assumption 1. Any random variable $u^* \in U$ can be written as $u^* = f^{-1} \cdot u = f^{-1}(u)$ for $u \in U$ and for some $f \in F$, where F is a group of measurable transformations acting transitively on $\mathbb{R}^n$.

If $n = 1$, Assumption 1 is satisfied provided the cumulative distribution functions (CDFs) of u and u^* are strictly increasing. This is the case if both u and u^* are continuous random variables with densities having the whole real line as support. If $n > 1$, one can assume that u and u^* are one-to-one functions of random vectors $\underset{\sim}{u}$ and $\underset{\sim}{u}^*$ having components that are independent with CDFs that are strictly increasing (cf. assumption AA.5' of Matzkin (2003)). Rosenblatt (1952) describes a simple transformation that allows one to transform an n-dimensional random vector having an absolutely continuous distribution to an n dimensional hypercube. Rosenblatt's transformation can be used to change uniformly continuous distributions into uniformly continuous distributions.

The fact that F is a group makes sure that one can go from u^* to u as well as from u to u^*. Thus, we can think of all random variables in U as generated from a fixed random variable u, as $f \cdot u$ for each $f \in F$.

We now investigate the implications of Assumption 1. Since $\eta^*(z,y)=u^*=f^{-1}(u)$ we can write $f(\eta^*(z,y))=u$ so that

(2) $$S^*=(\eta^*,P_{u^*})=(\eta^*,P_{f^{-1}\cdot u})=(f\circ\eta^*,P_u),$$

where $(f\circ\eta^*)(z,y)=f(\eta^*(z,y))$ is the composition of two functions. The following result investigates how the properties of $P_{u|z}$ affect the relationship between two observationally equivalent structures of the form $S=(\eta,P_u)$ and $S^*=(f\circ\eta^*,P_u)$ for given η and η^*.

Lemma 1. Two structures $S=(\eta,P_u)$ and $S^*=(f\circ\eta^*,P_u)$ are observationally equivalent if and only if $\eta_i=g_i\circ f\circ\eta^*$ where $g_i:\mathbb{R}^n\to\mathbb{R}^n$ is a B measurable function and η_i denotes the i-th component of η.

Proof. Let z be fixed. Observational equivalence means that a given set of values for y is $(f\circ\eta)^{-1}(z,\mathrm{B})$ measurable (and also $\eta^{-1}(z,\mathrm{B})$ measurable). This means that η is $(f\circ\eta)^{-1}(z,\mathrm{B})$ measurable. Now apply Theorem 4.2.8 of Dudley (2002): (a) $f\circ\eta^*$ is a (measurable) function from $Y=\mathbb{R}^n$ into (U,B) where $U=\mathbb{R}^n$; (b) the i-th component of η, η_i, is a real-valued B measurable function on Y; (c) the function η_i is $(f\circ\eta^*)^{-1}(\mathrm{B})$ measurable on $Y=\mathbb{R}^n$ if and only if $\eta=g\circ f\circ\eta^*$ for some B measurable function $g:\mathbb{R}^n\to\mathbb{R}^n$.■

Notice that Lemma 1 implies the existence of a B-measurable function $g:\mathbb{R}^n\to\mathbb{R}^n$, but this may not be uniquely defined. It follows from Lemma 1 that, in our setup, observational equivalence involves η only.

Theorem 1. Given Assumption 1, two structures $S=(\eta,P_u)$ and $S^*=(\eta^*,P_{u^*})$ are observationally equivalent if and only if $\eta_i=f_i^*\circ\eta^*$ B-a.e. for measurable function $f^*:\mathbb{R}^n\to\mathbb{R}^n$ for every component η_i of η.

Proof. The corollary follows from Lemma 1 by setting $f_i^*=g_i\circ f$ for $f\in F$ and by using (2).■

Observational equivalence is the condition $\eta=f^*\circ\eta^*$. Notice that f^* takes the range of any function η^* and projects it to the range of η. The function f^* could be standardized as in Matzkin (2003). If one is interested in derivatives, and makes sufficient assumptions to guarantee differentiability of η and η^*, then some of their ratios will be identified.

After dealing with observational equivalence connected to the way P_u can be changed into P_{u^*}, we now focus on η. Precisely, we show that, under some conditions, the function η can be modified in such way that the resulting structure is observationally equivalent to the original one. To do this we make two further assumptions.

Assumption 2. There exists a group G that acts on the exogenous variables by $z \to g \cdot z$.

Assumption 3. For each $g \in G$ and each y there exists a unique y' such that $\eta(g \cdot z, y') = \eta(z, y)$ for all z.

Assumption 2 implies that we can always transform the exogenous variables into new ones. Assumption 3 implies the function has $\eta(z, y)$ some symmetry. Precisely, it entails that G induces a group of transformations on the endogenous variables via the group action $g \cdot y = y'$, and that with this group action $\eta(g \cdot z, g \cdot y) = \eta(z, y)$ so that $\eta(z, y)$ is invariant (e.g. Theorem 2.6 of Eaton (1989)). Notice that in this notation G also denotes the group of transformations acting on y by $y \to g \cdot y$. Therefore,

(3) $$P_{y|z,S}(B) = P_u(\eta(z, B)) = P_u(\eta(g \cdot z, g \cdot B)),$$

for all $B \in \mathrm{B}$. Since $S = (\eta, P_u)$ we can define $S^* = (\eta^*, P_u)$ where $\eta^*(z, y) = \eta(g \cdot z, g \cdot y)$ and $z^* = g \cdot z$. Thus,

$$P_{y|z,S^*} = P_u(\eta^*(z, B)) = P_u(\eta(g \cdot z, g \cdot B)).$$

This is summarized in the following result.

Lemma 2. Given Assumptions 2 and 3, there is a group of transformations G such that $\eta(g \cdot z, g \cdot y) = \eta(z, y)$ for all $g \in G$. Moreover, the structures $S = (\eta, P_u)$ and $S^* = (\eta^*, P_u)$ where $\eta^*(z, y) = \eta(g \cdot z, g \cdot y)$ are observationally equivalent.

We will regard G as largest group leaving the reduced form $\eta(z,y)$ invariant. Notice that Assumption 3 (and also Lemma 2) is compatible with the case where G consists of the identity transformation only. However, this case is trivial since it implies that a structure is observationally equivalent to itself.

Notice also that if the two structures $S=(\eta,P_u)$ and $S^*=(\eta^*,P_{u^*})$ are observationally equivalent, and η^* is invariant under the action of a group G (i.e. $\eta^*(z,y)=\eta^*(g\cdot z,g\cdot y)$) then Theorem 1 implies that $\eta_i(z,y)=(f_i^*\circ\eta^*)(z,y)=f_i(\eta^*(z,y))$ and

$$\begin{aligned}\eta_i(g\cdot z,g\cdot y)&=f_i(\eta^*(g\cdot z,g\cdot y))\\&=f_i(\eta^*(z,y))\\&=\eta_i(z,y).\end{aligned}$$

Hence, η is also invariant under the action of G. Thus,

Lemma 3. If $S=(\eta,P_u)$ and $S^*=(\eta^*,P_{u^*})$ are observationally equivalent, and η^* is invariant under the action of a group G then η is also invariant under the action of G.

This result has an important consequence. Suppose that G is the largest group for which η is invariant and that η^* is such that the structures $S=(\eta,P_u)$ and $S^*=(\eta^*,P_{u^*})$ are observationally equivalent. Then any other group G^* for which η^* is invariant must coincide with G or must be a proper subgroup of G. If η^* is invariant under G^*, then η would also be invariant under G^* because of Lemma 3, so G^* cannot be larger than G or one would have a contradiction. In the rest of the paper we take G to be the largest group for which η is invariant.

The orbits $O_{(z,y)}=\{g(z,y):g\in G\}$ form a partition of the space of the exogenous and the endogenous variables and are uniquely defined. The function η is invariant only if it is constant on each orbit. A function ι that is constant and takes on a different value on each orbit is called a *maximal invariant* under the action of the group G (e.g. Lehmann (1997)). Thus, we have

Theorem 2. Let $S^*=(\eta^*,P_{u^*})$ be a structure for which $\eta^*(z,y)$ is invariant under the action of the group G defined in Lemma 2. Let $S=(\eta,P_u)$ be a structure observationally equivalent to $S^*=(\eta^*,P_{u^*})$. Then, given Assumptions 1, 2 and 3, $\eta(z,y)=\mu(\iota(z,y))$ B-a.e where μ is a suitable function to $\mathbb{R}^n$ and ι is the maximal invariant under the action of G.

Proof. It follows from Theorem 1 that $\eta_i = f_i^* \circ \eta^*$ B-a.e. for measurable function $f_i^* : \mathbb{R}^n \to \mathbb{R}^n$. Proposition 7.7 of Eaton (1982) says every invariant function is a function of the maximal invariant. Therefore, $\eta_i(z,y) = (f_i^* \circ \eta^*)((z,y)) = f_i^*(\mu^*(\iota(z,y))) = \mu_i(\iota(z,y))$.■

It is worth considering a familiar example. Let the model be $\eta(z,y) = u$ where $\eta(z,y) = y - \pi' z$, π is a $k \times n$ matrix with $k \geq n$ and z is a $k \times 1$ vector. The group G acts on the exogenous variables by translation $g \cdot z = z + g$ for $g \in \mathbb{R}^k$. The induced group of transformations of the endogenous variables is $g \cdot y = y + \pi' g$. Thus G acts on (z,y) by $g(z,y) = (z,y) + (g, \pi' g)$. To find the maximal invariant set $g = -z$ to obtain $g(z,y) = (z,y) + (-z, -\pi' z) = (0, y - \pi' z)$. Thus, $\eta(z,y) = y - \pi' z$ is the maximal invariant. Notice that all models of the form $u^* = \mu(y - \pi' z)$ where $\mu : \mathbb{R}^n \to \mathbb{R}^n$ and $\mu \in F$ are observationally equivalent to the model $u = y - \pi' z$. For more complicated examples one needs to use local arguments and the tools developed in Section 3.

Theorem 2 identifies the maximal invariant ι as the component of any reduced form observationally equivalent to $\eta(z,y)$. Notice that ι may have dimension smaller than $\eta(z,y)$ and that $\iota(z,y) = u_*$ could be regarded as a reduced form itself .

2.2 The structural equation

After having defined the reduced form and established that under certain conditions its identification is only a property of η, we add a *structural equation* to the model. We define a *structural equation* as a measurable function of (z,y) following the notation of equation (2.1) of Matzkin (2008). Let $t : Z \times Y \subset \mathbb{R}^{k+n} \to \mathbb{R}$ be a measurable mapping taking $(z,y) \in Z \times Y \subset \mathbb{R}^{k+n}$ to the set of real numbers satisfying the functional restrictions $R[t] = 0$. These could be for example exclusion restrictions for some of the exogenous variables z_2 (where $z = (z_1, z_2)$) from the structural equation: for example if t is differentiable, the partial derivatives of $t(z,y) = t(z_1, y)$ with respect to the components of z_2 would vanish for all (z,y).

The distribution function of the structural equation $t(z,y)$ given the reduced form structure $S = (\eta, P_u)$ is

$$F_{t(z,y)|z,S}(v) = \int_{\{y \in \mathbb{R}^n : t(z,y) \leq v\}} P_{y|z,S}(dy).$$

Notice that no further assumption is needed apart from the measurability of t.

Theorem 2 shows that there are observationally equivalent ways of writing the reduced form. One could use $\eta(z,y)$ or $\eta^*(z,y)=\eta(g\cdot z, g\cdot y)$ for any $g\in G$. Therefore, a structural equation t, in order to be compatible with the reduced form, must be the same irrespective of the way the reduced form is written. Therefore the structural equation must be invariant under the action of G. Proposition 7.7 of Eaton (1982) says every invariant function is a function of the maximal invariant. Thus, the structural must be of the form $t(z,y)=r(h(z,y))$, where r is a function and h is the maximal invariant under the action of the group G. Notice that t satisfies the restrictions $R[t]=0$ so that it must also be true that $R\left[r(h(z,y))\right]=0$.

Theorem 3. Given Assumptions 2 and 3, the structural equation $t(z,y)$ (satisfying $R[t]=0$) compatible with the reduced form $u=\eta(z,y)$ must be of the form $t(z,y)=r(h(z,y))$, where r is a function and h is the maximal invariant under the action of the group G, moreover $R\left[r(h(z,y))\right]=0$.

Therefore, the function $t(z,y)$ in the structural equation cannot be arbitrary. It needs have the form specified in Theorem 3 which depends on the reduced form equation through the maximal invariant. The family of structural equations compatible with the reduced form $\eta(z,y)=u$ is the set of all functions $r(h(z,y))$ obtained by varying r in such a way that $R\left[r(h(z,y))\right]=0$. The features of the structural equation that are identified depend on both h and the restrictions R. Moreover, if two structures are equivalent, a structural equation which is compatible with one reduced form is also compatible with a reduced form having an observationally equivalent structure.

Notice that in the absence of functional restrictions on the reduce form, $t(z,y)=r(h(z,y))$ would not contain any more information than the reduced form itself.

2.3 Examples

Some examples are now considered to illustrate the theory developed above. First, we continue the example in the introduction.

Example 1. (*Linear structural equations*) Let the reduced form be $\eta(z,y)=u$ where $\eta(z,y)=y-\pi' z$ and π is a $k\times n$ vector with $k\geq n$ and z is fixed. The group G acts on the exogenous variables by translation $g \cdot z = z + g$ for $g \in \mathbb{R}^k$. The induced group of transformations of the endogenous variables is $g \cdot y = y + \pi' g$. Thus G is the group of transformations of the form $g(z, y) = (z + g, y + \pi' g)$ for which the maximal invariant is, as we saw before, $y - \pi' z$. Suppose that the structural equation does not depend on the exogenous variable and can be written as $t(y)$ so that invariance implies that $t(y + \pi' g) = t(y)$. Since from Theorem 3 we must have that $t(y)$ is a function of the maximal invariant then we can write $t(y) = r(y - \pi' z)$. The function r must have a specific form that we now derive. In fact, $t(y)$ does not depend on z so that $r(y - \pi' z)$ cannot depend on z either. Notice that π has at most n rows which are linearly independent and span a linear subspace V, say, of $\mathbb{R}^n$. Let P_V be the orthogonal projection onto V and $M_V = I_n - P_V$ be the orthogonal projection onto the space orthogonal to V. Let $CC' = M_V$ and $C'C = I_{n-\dim(V)}$ then t must depend on y only through $C'y$, i.e. $t(y) = r(C'y)$. If V has dimension n then $C'y = 0$. If V has dimension $n-1$, then the space orthogonal to V is one dimensional so that it can be generated by multiples of the vector m such that $m'm = 1$ and $\pi m = 0$. The structural equation must be of the form $t(y) = r(m'y)$. In this case, m is uniquely determined, after normalization, by π. Therefore, any known function of m is identified. For example, even if we do not know r we can still identify $\partial y_1 / \partial y_2$. To do this notice that

$$\frac{\partial(u - r(m'y))}{\partial y_2} = 0 = \left.\frac{dr}{dx}\right|_{x=m'y}\left(m_1 \frac{\partial y_1}{\partial y_2} + m_2\right)$$

so that $\partial y_1 / \partial y_2 = -m_2 / m_1$, provided $\left. dr/dx \right|_{x=m'y} \neq 0$ and $m_1 \neq 0$.

The analysis can be easily generalized to the case where $t(y, z_1)$ and $z = (z_1, z_2)$. This is illustrated in Example 2.

Example 2. A slightly more complex version of Example 1 is obtained by taking $y \sim N(\pi_1' z_1 + \pi_2' z_2, \Omega)$ where π_1 is a $k_1 \times n$ vector, π_2 is a $k_2 \times n$ vector with $k_2 \geq n$ and z_1 and z_2 are fixed. Consider a function $t(z, y) = t(z_1, y)$ so that this depends on z_1 but not on z_2. The group G of translations acts on $z_1 \to g \cdot z_1 = z_1 + \zeta_1$ and $z_2 \to g \cdot z_2 = z_2 + \zeta_2$. The induced group of transformations on the endogenous variables is $g \cdot y = y + \pi_1' \zeta_1 + \pi_2' \zeta_2$ where $\zeta_1 \in \mathbb{R}^{k_1}$ and $\zeta_2 \in \mathbb{R}^{k_2}$. The maximal invariant under the action of G is $y - \pi_1' z_1 - \pi_2' z_2$ so that

$$t(z_1, y) = r(y - \pi_1' z_1 - \pi_2' z_2).$$

Once again the function r cannot be arbitrary because the right hand side in the last equation cannot depend on z_2. Suppose that π_2 has at most n rows which are linearly independent and span a linear subspace V of $\mathbb{R}^n$. Let P_V be the orthogonal projection onto V and $M_V = I_n - P_V$ be the orthogonal projection onto the space orthogonal to V. Therefore t does not depend on z_2 if $t(z_1, y) = r(C'(y - \pi_1' z_1))$ where C is defined as in Example 1. If V has dimension $n-1$, then the space orthogonal to V is one-dimensional so that it can be generated by multiples of the vector m such that $m'm = 1$ and $\pi m = 0$. So the functions which are compatible with the reduced form are of the form $r(m'(y - \pi_1' z_1))$ with $m'm = 1$ and $\pi_2 m = 0$. This can be written in a more standard form using the same strategy as in Example 1.

Example 3. Suppose we allow the structural equation to depend on the whole vector z, i.e. $t(z, y)$. The action on the exogenous variables is $g \cdot z = z + g$ for $g \in G = \mathbb{R}^k$, and the induced action on the endogenous variables is $g \cdot y = y + \pi' g$ and the maximal invariant is $y - \pi' z$. In this case $g(z, y) = r(y - \pi' z) = r(u)$ where $u \sim N(0, \Omega)$. No restrictions are available to restrict the functional form of t because in this case r is totally arbitrary.

3. Lie groups

Assumptions 2 and 3 imply the existence of a group G that acts on $\mathbb{R}^{k+n}$ and leaves $\eta(z, y)$ unchanged. Finding such a group G may be difficult. Moreover, sometimes the group action of G on $\mathbb{R}^{k+n}$ is not defined for all elements of G nor for all points in $\mathbb{R}^{k+n}$. This section deals with this situation. We will use the theory of Lie groups to justify our procedure. However, once this is done application of invariance arguments are easily done.

The group $G = \mathbb{R}^k$ acts on $\mathbb{R}^k$ by $g \cdot z = z + g$. This induces a group of transformations of the endogenous variables through the invariance of the function $\eta(z, y)$. So we make enough assumptions for this action to be well defined locally. We investigate the local properties however these results can be made into global results by using partition-of-unity type arguments.

Assumption 4. (a) The function $\eta:\mathbb{R}^{k+n}\to\mathbb{R}^n$ is smooth and the ($n\times n$) matrix of partial derivatives $\frac{\partial\eta_s}{\partial y_i}$, denoted by $D_y\eta$, has rank n for all $(z,y)\in\mathbb{R}^k\times\mathbb{R}^n$. (b) The ($n\times k$) matrix of partial derivatives $\frac{\partial\eta_s}{\partial z_j}$, denoted by $D_z\eta$, has rank r, $0\le r\le\min\{n,k\}$, for all $(z,y)\in\mathbb{R}^{k+n}$.

Assumption 4 implies the validity of the implicit function theorem. Therefore, there is, at least locally, a smooth function $y:\mathbb{R}^k\to\mathbb{R}^n$ such that $\eta(z,y(z))=u$. The fact that $D_z\eta$ has rank r implies that only r components of y are locally functionally independent. Notice that the group $G=\mathbb{R}^k$ induces an action on the space of the endogenous variables $\mathbb{R}^n$ by $g\cdot(z,y(z))=(z+g,y(z+g))$, so that $g\cdot y=y(z+g)$.

Denote by $M_u=\{(z,y)\in\mathbb{R}^{k+n}:\eta(z,y)=u\}$ the level sets of the function $\eta:\mathbb{R}^{k+n}\to\mathbb{R}^n$. Assumption 4 implies that such function is a submersion. The submersion theorem (Abraham, Marsden and Ratiu (1988), p.197) implies that M_u is a k-dimensional closed sub-manifold of $\mathbb{R}^{k+n}$ and its tangent space is $\ker T_{(z,y)}\eta=\{v\in\mathbb{R}^{k+n}:D\eta\cdot v=0\}$. Therefore the tangent space of M_u is given by all vectors $v=\begin{pmatrix}v_1\\v_2\end{pmatrix}\in\mathbb{R}^{k+n}$ solving $[D_z\eta]v_1+[D_y\eta]v_2=0$. Assumption 4 allows us to write $v_2=-[D_y\eta]^{-1}[D_z\eta]v_1$, and $v=\begin{pmatrix}I_k\\\Xi(z,y)\end{pmatrix}v_1$ where $\Xi(z,y)=-[D_y\eta]^{-1}[D_z\eta]$.

The tangent space of M_u is generated by linear combinations of the k vectors

(4) $$\zeta_j=\frac{\partial}{\partial z_j}+\xi_{1j}(z,y)\frac{\partial}{\partial y_1}+\ldots+\xi_{nj}(z,y)\frac{\partial}{\partial y_n},$$

$j=1,2,\ldots,k$ where $\xi_{ji}(z,y)$ are the components of $\Xi(z,y)$. These vectors ζ_j vary smoothly with (z,y) and are thus vector fields on M_u. We now need to make sure that the differential operator (4) captures our group action.

Given two vector fields ζ_j and ζ_i, their Lie bracket is defined by $[\zeta_j,\zeta_i]f=\zeta_j(\zeta_i f)-\zeta_i(\zeta_j f)$ for any smooth function $f:\mathbb{R}^{k+n}\to\mathbb{R}$. If we take $f=\eta_s$, $s=1,2,\ldots,n$, we find $[\zeta_j,\zeta_i]\eta_s=\zeta_j(\zeta_i\eta_s)-\zeta_i(\zeta_j\eta_s)=0$ since $\zeta_j\eta_s=0$ by definition. Therefore, $[\zeta_j,\zeta_i]$ belongs to the tangent space of M_u, and can be written as $[\zeta_j,\zeta_i]=C_{ji}^1\zeta_1+\ldots+C_{ji}^k\zeta_k$.

Each operator ζ_j generates a flow $\exp\{t\zeta_j\}(z,y)$ on M_u. Let $\zeta_j\big|_{(z,y)} = \frac{d}{dt}\exp\{t\zeta_j\}(z,y)\Big|_{t=0}$ so that $\zeta_j\big|_{(z,y)} \in T_{(z,y)}M_u$ is the tangent vector to M_u. Then, $\exp\{t\zeta_j\}(z,y) = (z,y) + t\,\zeta_j\big|_{(z,y)} + O(t^2)$. Since $[\zeta_j, \zeta_i]$ belongs to the tangent space to M_u, all the flows generated by the vector fields ζ_j fit together to form a family of sub-manifolds. Therefore, every group element $g \in G$ can be written as $g = \exp\{\varepsilon_1\zeta_{i_1} + \varepsilon_2\zeta_{i_2} + ... + \varepsilon_k\zeta_{i_k}\}$ for suitable $\varepsilon_j \in \mathbb{R}$ and $1 \le i_j \le k$, $j = 1,2,...,k$ and the local group action is defined by

(5) $$g\cdot(z,y) = \exp\{\varepsilon_1\zeta_{i_1} + \varepsilon_2\zeta_{i_2} + ... + \varepsilon_k\zeta_{i_k}\}(z,y).$$

The vector fields ζ_j are called the infinitesimal generators of the flows $\exp\{t\zeta_j\}(z,y)$, and the vector space spanned by these infinitesimal generators ζ_j is denoted by $\mathfrak{g}$. Since this is closed under Lie-brackets operations, it is a Lie algebra.

Suppose that the function $t(z,y)$ is defined on an open subset $Z\times Y$ of $\mathbb{R}^{k+n}$. The function $t(z,y)$ is locally invariant if $t(g\cdot(z,y)) = t(z,y)$ for all $(z,y)\in Z\times Y$, and for all $g \in B_{(z,y)}$ in some neighbourhood $B_{(z,y)} \subset G$ of the identity element of G. Notice that $B_{(z,y)}$ may depend on (z,y). Analogously, $t(z,y)$ is globally invariant if $t(g\cdot(z,y)) = t(z,y)$ for all $(z,y)\in Z\times Y$, and for all $g\in G$. Notice that the fact that the function is locally invariant depends on the allowed transformations being local to the identity element. It does not depend on the set on which the (z,y) are defined.

Theorem 4. Let G be a Lie group acting on $\mathbb{R}^{k+n}$ according to (5). Then, there exists n functionally independent global invariants $I_1,...,I_n$ defined on a neighbourhood $Z\times Y$ of $\mathbb{R}^{k+n}$, with the property that any other local invariant I defined on $Z\times Y$ can be written as a function of the fundamental invariants $I = H(I_1,...,I_n)$. Moreover, in the regular case, two points in $(z,y),(z',y')\in Z\times Y$ lie in the same orbit of G if and only if all the invariants have the same value, $I_i(z,y) = I_i(z',y')$, $i = 1,...,n$.

Proof. The result follows from Theorem 2.34 in Olver (1996), by noting that the orbits of G are k dimensional and are pathwise connected (e.g. the discussion preceding equation (5)).■

In our case the invariants are the n components of η in the reduced form. In order to obtain nontrivial results one has to impose restrictions on the structural equation beyond invariance. These usually take the form of exclusion restrictions.

Let t be a function defined on a subset of $\mathbb{R}^k \times \mathbb{R}^n$. The following theorem gives an infinitesimal criterion to verify that t is invariant, and thus provides the main tool to establish the compatibility of structural equation and reduced form. This is a fairly natural thing to do since most of restrictions being imposed on a structural equation are exclusion restrictions.

Theorem 5. Let $t:\mathbb{R}^{k+n} \to \mathbb{R}$. Then t is invariant under G if and only if $\zeta_j\left[t\right]=0$ and $R\left[t\right]=0$ for all $\left(z,y\right)\in \mathbb{R}^k \times \mathbb{R}^n$, for every infinitesimal generator ζ_j $j=1,2,...,k$.

Proof. It follows from Theorem 2.74 of Olver (1996) noticing that G is connected and the ζ_j, $j=1,2,...,k$, form a Lie algebra. ■

In order to determine the compatibility of a structural equation and a given reduced form one proceeds as follows:

1. find the vector fields ζ_j by noticing that the reduced form equations are invariant under the action of ζ_j (Theorem 4) and that by Theorem 5 one must have $\zeta_j\left[\eta_i\right]=0$. These kn equations determine the components of $\Xi\left(z,y\right)$.
2. Determine whether the structural equation t satisfying $R\left[t\right]=0$ is also invariant and satisfies $\zeta_j\left[t\right]=0$.

The matrix of coefficients Ξ is entirely determined by the invariance properties of the reduced form. Given Assumption 4 it equals $\Xi\left(z,y\right)=-\left[D_y\eta\right]^{-1}\left[D_z\eta\right]$. Theorem 5 links the reduced form and the structural equation. For a given infinitesimal generator, the condition $\zeta_j\left[t\right]=0$ can be written as

(6) $$\left(D_y t\right)\Xi\left(z,y\right)=D_z t,$$

for all $\left(z,y\right)\in \mathbb{R}^k \times \mathbb{R}^n$. Equation (6) defines a system of k partial differential equations. Locally (6) can be regarded a system of k linear equations in $n+k$ unknowns, $D_y t$ and $D_z t$. Clearly, such system cannot have a unique (local) solution unless we impose further restrictions on $D_y t$ and $D_z t$.

We can now impose the functional restrictions $R\left[r\left(h\left(z,y\right)\right)\right]=0$ in the form of at least n local restrictions of the form:

(7) $$\left(D_y t\right)\Xi_{12}\left(z,y\right)+\left(D_z t\right)\Xi_{22}\left(z,y\right)=0\,.$$

These can be exclusion restrictions for the exogenous variables (e.g. $\Xi_{12}\left(z,y\right)=0$ and $\Xi_{22}\left(z,y\right)=I_k$) but other restrictions could also be employed. Notice that given $D_y t$, $D_z t$ is determined by (6). Replacing (6) in (7) we obtain

(8) $$\left(D_y t\right)\left[\Xi_{12}\left(z,y\right)+\Xi\left(z,y\right)\Xi_{22}\left(z,y\right)\right]=0\,.$$

This is a homogenous linear equation, and the number of solutions depends on the rank of the matrix $\Xi_{12}\left(z,y\right)+\Xi\left(z,y\right)\Xi_{22}\left(z,y\right)$. If it has rank n, there will be only one solution $D_y t=0$. If it has rank $n-1$, there will be only one solution subject to a normalization constraint, for example, restricting $D_y t$ to be on the unit circle. This means that ratios of any two components of $D_y t$ are uniquely defined.

The nature of the solution to (6) and (7) determines the structural equation t. This may depend on the reduced form through $\Xi\left(z,y\right)$, but, clearly, one could potentially have $D_y t$ uniquely defined by the restrictions (8) alone (i.e. one could set $\Xi_{22}\left(z,y\right)=0$ and choose $\Xi_{12}\left(z,y\right)$ to be any matrix of rank $n-1$).

3.1 Examples

We first consider the linear case again and then study some nonlinear models.

Example 4. The results in Example 1 can be put in the local context described above. In such a case we have $\eta\left(z,y\right)=y-\pi' z$ and

$$\zeta_j\eta_i=-\pi_{ij}+\xi_{ij}\left(z,y\right)=0$$

so that $\xi_{ji}\left(z,y\right)=\pi_{ji}$, and in matrix form $\Xi=\left\{\xi_{ji}\right\}=\pi$. That is, in this case,

$$\zeta_j=\frac{\partial}{\partial z_j}+\pi_{1j}\frac{\partial}{\partial y_1}+\pi_{2j}\frac{\partial}{\partial y_2}+\ldots+\pi_{nj}\frac{\partial}{\partial y_n}\,.$$

Now consider the structural equation $t:\mathbb{R}^n\to\mathbb{R}$. This is invariant under the action of the vector field ζ_j if $\zeta_j t=0$. We impose restrictions on t by assuming that it depends only on y, i.e. we let $t\left(z,y\right)=t\left(y\right)$, then

$$\zeta_j t=\left(\frac{\partial}{\partial z_j}+\sum_{i=1}^{n}\pi_{ij}\frac{\partial}{\partial y_i}\right)t\left(y\right)=\sum_{i=1}^{n}\pi_{ij}\frac{\partial t\left(y\right)}{\partial y_i}=0$$

for $j=1,2,...,k$, that can be written in matrix forms as $\left[D_y t\right]\pi'=0$. The classical results follow:

1. if π has full rank n then $D_y t=0$ that is $t\left(y\right)$ equals a constant;
2. if π has rank $n-1$ then the rows of π span a space of dimension $n-1$, so that $D_y t$ must be orthogonal to π for all y, so that after normalization $t\left(y\right)=t\left(m'y\right)$ for $m'm=1$ and $\pi m=0$.
3. if π has rank $0\le r\le n-1$, then one goes back to the classical case of partial identification (e.g. Phillips (1989) and Choi and Phillips (1992)), and $t\left(y\right)=t\left(m'y\right)$ for $m'm=I_r$ and $\pi m=0$.

Example 5. Now assume that $t\left(z_1,y\right)$ is as in Example 2. Then

$$\zeta_j t=\frac{\partial t\left(z_1,y\right)}{\partial z_j}+\sum_{i=1}^{n}\pi_{ij}\frac{\partial t\left(z,y\right)}{\partial y_i}=0 \text{ for } j=1,...,k_1$$
$$\zeta_j t=\sum_{i=1}^{n}\pi_{ij}\frac{\partial t\left(z,y\right)}{\partial y_i}=0 \text{ for } j=k_1+1,...,k$$

Partition π as $\pi=\begin{pmatrix}\pi_1\\ \pi_2\end{pmatrix}$, where π_1 contains k_1 rows. Then the two equations above can be written as

$$D_{z_1}t+\left(D_y t\right)\pi_1'=0$$
$$\left(D_y t\right)\pi_2'=0$$

so that if the rank of π_2 is $n-1$ we need $D_y t=m$, where $m'm=1$ and $\pi_2 m=0$. Then we must have $D_{z_1}t=\gamma=-m\pi_1'$. So $t\left(z_1,y\right)=t\left(m'\left(y-\pi_1'z_1\right)\right)$ as before.

Example 6. (Single equation model of Matzkin (2008)). Assume $v=\eta\left(z,y\right)$ where all variables are univariate and let $\zeta=\frac{\partial}{\partial z}+\xi_1\left(z,y\right)\frac{\partial}{\partial y}$ be a vector field on $\mathbb{R}^2$. Such vector field leaves the function η invariant if

$$\zeta\eta\left(z,y\right)=\frac{\partial\eta\left(z,y\right)}{\partial z}+\xi_1\left(z,y\right)\frac{\partial\eta\left(z,y\right)}{\partial y}=0.$$

If $\frac{\partial\eta\left(z,y\right)}{\partial y}\neq 0$, this equation can be solved to yield $\xi_1\left(z,y\right)=-\frac{\partial\eta\left(z,y\right)}{\partial z}/\frac{\partial\eta\left(z,y\right)}{\partial y}$.

Now, we consider a function $t(z,y)$ which is also invariant under the action of the operator ζ and investigate when this is compatible with $v=\eta(z,y)$. We must have

(9) $$\zeta t(z,y)=\frac{\partial t(z,y)}{\partial z}-\xi_1(z,y)\frac{\partial t(z,y)}{\partial y}=0$$

which can be rewritten to give

(10) $$\frac{\dfrac{\partial t(z,y)}{\partial z}}{\dfrac{\partial t(z,y)}{\partial y}}=\xi_1(z,y)=\frac{\dfrac{\partial \eta(z,y)}{\partial z}}{\dfrac{\partial \eta(z,y)}{\partial y}}.$$

This compatibility condition is the same as the condition for observational equivalence given in equation (5.1) of Matzkin (2008). From a geometric point of view the partial differential equation (9) implies that the function $t(z,y)$ is constant on a manifold parameterized by $\eta(z,y)=c$, where c is generic constant, having a tangent vector at (z,y) which has the same direction as the vector $(1,-\xi_1(z,y))$. Notice that $\frac{\partial \eta(z,y)}{\partial z}/\frac{\partial \eta(z,y)}{\partial y}=\frac{dy}{dz}$ using the implicit function theorem, so that

$$\frac{dy}{dz}=\xi_1(z,y).$$

This is a differential equation which under some regularity conditions (i.e. continuity of $\xi_1(z,y)$ and $\frac{d\xi_1(z,y)}{dy}$) has the general solution $y=f(z,c)$ which also satisfy $\eta(z,y)=c$. Therefore a general solution to (9) is $t(z,y)=f^*(\eta(z,y))$ where f^* is an arbitrary differentiable function.

Example 7. (A triangular model of Matzkin (2008)). Assume that the reduced form consists of two equations $v_1=\eta_1(y_1,y_2)$ and $v_2=\eta_2(y_2,z)$ (notice y_2 does not appear in the second equation) and investigate the structural equations $t(y_1,y_2)$ that depend only on (y_1,y_2). Let

$$\zeta=\frac{\partial}{\partial z}+\xi_1(z,y)\frac{\partial}{\partial y_1}+\xi_2(z,y)\frac{\partial}{\partial y_2},$$

be a vector field on $\mathbb{R}^3$. Then, this vector field leaves the structural equations η_1 and η_2 unchanged if

$$\zeta v_1 = \left(\frac{\partial}{\partial z} + \xi_1(z,y)\frac{\partial}{\partial y_1} + \xi_2(z,y)\frac{\partial}{\partial y_2} \right)\eta_1(y_1,y_2) = 0$$
$$\zeta v_2 = \left(\frac{\partial}{\partial z} + \xi_1(z,y)\frac{\partial}{\partial y_1} + \xi_2(z,y)\frac{\partial}{\partial y_2} \right)\eta_2(y_2,z) = 0.$$

Therefore, ξ_1 and ξ_2 satisfy the system of equations

$$\xi_1(z,y)\frac{\partial \eta_1(y_1,y_2)}{\partial y_1} + \xi_2(z,y)\frac{\partial \eta_1(y_1,y_2)}{\partial y_2} = 0$$
$$\frac{\partial \eta_2(y_2,z)}{\partial z} + \xi_2(z,y)\frac{\partial \eta_2(y_2,z)}{\partial y_2} = 0.$$

This can be solved to give

$$\xi_2(z,y) = -\frac{\dfrac{\partial \eta_2(y_2,z)}{\partial z}}{\dfrac{\partial \eta_2(y_2,z)}{\partial y_2}}$$
$$\xi_1(z,y) = -\xi_2(z,y)\frac{\dfrac{\partial \eta_1(y_1,y_2)}{\partial y_2}}{\dfrac{\partial \eta_1(y_1,y_2)}{\partial y_1}}.$$

Now consider a structural equation $t(y_1,y_2)$. This is invariant under ζ and satisfies $\zeta t(y_1,y_2) = 0$ so that

$$\zeta t(y_1,y_2) = \left(\frac{\partial}{\partial z} - \xi_2(z,y)\frac{\dfrac{\partial \eta_1(y_1,y_2)}{\partial y_2}}{\dfrac{\partial \eta_1(y_1,y_2)}{\partial y_1}}\frac{\partial}{\partial y_1} + \xi_2(z,y)\frac{\partial}{\partial y_2} \right) t(y_1,y_2)$$
$$= \xi_2(z,y)\left(-\frac{\dfrac{\partial \eta_1(y_1,y_2)}{\partial y_2}}{\dfrac{\partial \eta_1(y_1,y_2)}{\partial y_1}}\frac{\partial t(y_1,y_2)}{\partial y_1} + \frac{\partial t(y_1,y_2)}{\partial y_2} \right) = 0.$$

The function $t(y_1,y_2)$ that is compatible with the reduced form $v_1 = \eta_1(y_1,y_2)$ and $v_2 = \eta_2(y_2,z)$ must satisfy

(11)
$$\frac{\dfrac{\partial t(y_1,y_2)}{\partial y_1}}{\dfrac{\partial t(y_1,y_2)}{\partial y_2}} = \frac{\dfrac{\partial \eta_1(y_1,y_2)}{\partial y_1}}{\dfrac{\partial \eta_1(y_1,y_2)}{\partial y_2}}.$$

Once again, this is the same condition for observational equivalence given by Matzkin (2008).

If $\xi_2(z,y)\neq 0$, $t(y_1,y_2)$ satisfies the partial differential equation

$$-\frac{\frac{\partial\eta_1(y_1,y_2)}{\partial y_2}}{\frac{\partial\eta_1(y_1,y_2)}{\partial y_1}}\frac{\partial t(y_1,y_2)}{\partial y_1}+\frac{\partial t(y_1,y_2)}{\partial y_2}=0$$

which is formally the same as (9) with obvious changes of variables, so that the general solution is $t(y_1,y_2)=f^*(\eta_1(y_1,y_2))$ where f^* is an arbitrary differentiable function.

If $\eta_1(y_1,y_2)$ is specified as $v_1=\eta_1(y_1,y_2)=y_1-m(y_2)$ we have a nonparametric regression model with an endogenous regressors. Suppose also we choose $t(y_1,y_2)=y_1-\tilde{m}(y_2)$. The latter is compatible with the reduced form $v_1=y_1-m(y_2)$ and $v_2=\eta_2(y_2,z)$ if

$$\frac{\partial\tilde{m}(y_2)}{\partial y_2}=\frac{\partial m(y_2)}{\partial y_2}.$$

In this case one must have $\tilde{m}(y_2)=m(y_2)+C$ where C is constant.

Example 8. (Nonlinear simultaneous equations). Consider a reduced form $u=\eta(z,y)$ and suppose that the first r rows of $\Xi(z,y)$ have rank $0\leq r\leq\min\{n,k\}$ and that the structural equation $t(y)$ does not depend on the exogenous variables. Then from Theorem 5, $t(y)$ solves the system of equations:

(12) $$(D_y t)\Xi(z,y)=0$$

where $D_y t$ be the $(1\times n)$ vector of derivatives $\frac{\partial t(y)}{\partial y_i}$. If $r=n$ the only solution to (12) requires $\frac{\partial t(y)}{\partial y_i}=0$, $i=1,...,n$, so that $t(y)$ must be a constant function. If $r=0$, $t(y)$ can be any function. To obtain a nontrivial solution we need $0<r<n$. Partition $D_y t$ conformably to $\Xi(z,y)$ as $D_y t=(D_{y_1}t,D_{y_2}t)$. Then, a solution to (12) is a solution to $(D_{y_1}t)\Xi_{11}(z,y)+(D_{y_2}t)\Xi_{21}(z,y)=0$ so that compatibility requires

(13) $$(D_{y_1}t)=-(D_{y_2}t)\Xi_{21}(z,y)\left[\Xi_{11}(z,y)\right]^{-1}.$$

Notice that the left hand side of (13) does not depend on the exogenous variables so that the right hand side cannot depend on the exogenous variables either. This imposes restrictions on the reduced form.

Example 9. (Nonlinear reduced form). We now assume the reduced form is $v_1 = y_1 - \eta_1(z_1, z_2)$ and $v_2 = y_2 - \eta_2(z_1, z_2)$. The operators to consider are

$$\begin{aligned}
\zeta_1 &= \frac{\partial}{\partial z_1} + \xi_{11}(z, y)\frac{\partial}{\partial y_1} + \xi_{21}(z, y)\frac{\partial}{\partial y_2} \\
\zeta_2 &= \frac{\partial}{\partial z_2} + \xi_{12}(z, y)\frac{\partial}{\partial y_1} + \xi_{22}(z, y)\frac{\partial}{\partial y_2}
\end{aligned}.$$

The coefficients are determined by the invariance of the reduced form:

$$\begin{aligned}
\xi_{11}(z, y) &= -\frac{\partial \eta_1}{\partial z_1} \\
\xi_{12}(z, y) &= -\frac{\partial \eta_1}{\partial z_2} \\
\xi_{21}(z, y) &= -\frac{\partial \eta_2}{\partial z_1} \\
\xi_{22}(z, y) &= -\frac{\partial \eta_2}{\partial z_2}.
\end{aligned}$$

Suppose that the structural equation has the form $t(y_1, y_2) = u$. Then, compatibility with the reduced form implies that

$$\begin{aligned}
\zeta_1 t(y_1, y_2) &= -\frac{\partial \eta_1}{\partial z_1}\frac{\partial t}{\partial y_1} - \frac{\partial \eta_2}{\partial z_1}\frac{\partial t}{\partial y_2} = 0 \\
\zeta_2 t(y_1, y_2) &= -\frac{\partial \eta_1}{\partial z_2}\frac{\partial t}{\partial y_1} - \frac{\partial \eta_2}{\partial z_2}\frac{\partial t}{\partial y_2} = 0.
\end{aligned}$$

Re-writing this we have

$$\frac{\frac{\partial t}{\partial y_1}}{\frac{\partial t}{\partial y_2}} = -\frac{\frac{\partial \eta_2}{\partial z_2}}{\frac{\partial \eta_1}{\partial z_2}} = -\frac{\frac{\partial \eta_2}{\partial z_1}}{\frac{\partial \eta_1}{\partial z_1}} \tag{14}$$

The equality on the right hand side implies a restriction on the reduced form given by the partial differential equation for η_2

$$\frac{\partial \eta_2}{\partial z_2}\frac{\partial \eta_1}{\partial z_1} = \frac{\partial \eta_2}{\partial z_1}\frac{\partial \eta_1}{\partial z_2}.$$

This can be solved by using the method of characteristics:

$$\frac{dz_1}{\frac{\partial \eta_1}{\partial z_2}} = \frac{dz_2}{\frac{\partial \eta_1}{\partial z_1}},$$

so the solution is $\eta_1(z_1,z_2)=c_1$ where c_1 is an arbitrary constant, and η_2 can be written as $\eta_2=F(\eta_1(z_1,z_2))$, where F is an arbitrary function. The first part of (14) yields

(15)
$$\frac{\frac{\partial t}{\partial y_1}}{\frac{\partial t}{\partial y_2}}=-\frac{\frac{\partial \eta_2}{\partial z_2}}{\frac{\partial \eta_1}{\partial z_2}}=-\frac{\partial F}{\partial \eta_1}\frac{\frac{\partial \eta_1}{\partial z_2}}{\frac{\partial \eta_1}{\partial z_2}}=-\frac{\partial F}{\partial \eta_1}$$

So that t must satisfy the partial differential equation

$$\frac{\partial t}{\partial y_1}+\frac{\partial F}{\partial \eta_1}\frac{\partial t}{\partial y_2}=0$$

Notice that $\eta_2=F(\eta_1(z_1,z_2))$ depends only on (z_1,z_2), however $\frac{\partial F}{\partial \eta_1}$ does not because the left-hand-side of (15) is only a function of (y_1,y_2). Thus $\frac{\partial F}{\partial \eta_1}$ must be a constant. The general solution has the form

$$t(y_1,y_2)=G(y_1-\beta y_2),$$

where $\beta=1/\left(\frac{\partial F(\eta_1(z_1,z_2))}{\partial \eta_1(z_1,z_2)}\right)$ is constant.

Example 10. (Control function separability of Blundel and Matzkin (2010)) Blundel and Matzkin (2010) show that, under some conditions, the model $v_1=\eta_1(y_1,y_2)$ and $v_2=\eta_2(y_1,y_2,z)$ is observationally equivalent to $v_1=\eta_1(y_1,y_2)$ and $v_2=s(y_2,z)$ if and only if $\eta_2(y_1,y_2,z)=v(q(y_2,z),\eta_1(y_1,y_2))$ (the latter condition is labeled as control function separability). Here we look at this problem under the conditions of this paper.

Since there is one exogenous and two endogenous variables we let

$$\zeta=\frac{\partial}{\partial z}+\xi_1(z,y)\frac{\partial}{\partial y_1}+\xi_2(z,y)\frac{\partial}{\partial y_2}.$$

We now apply invariance to the reduced form $v_1=\eta_1(y_1,y_2)$ and $v_2=s(y_2,z)$

$$\zeta[\eta_1]=0$$
$$\zeta[s]=0.$$

Thus, $\xi_1(z,y)$ and $\xi_2(z,y)$ satisfy

$$\xi_1(z,y)\frac{\partial\eta_1(y_1,y_2)}{\partial y_1}+\xi_2(z,y)\frac{\partial\eta_1(y_1,y_2)}{\partial y_2}=0$$
$$\xi_2(z,y)\frac{\partial s(y_2,z)}{\partial y_2}=-\frac{\partial s(y_2,z)}{\partial z}.$$

If $\frac{\partial s(y_2,z)}{\partial y_2}\neq 0$ we also have

$$\xi_2(z,y)=-\frac{\frac{\partial s(y_2,z)}{\partial z}}{\frac{\partial s(y_2,z)}{\partial y_2}}.$$

In this equation the right hand side does not depend on y_1 so that we also have $\xi_2(z,y)=\xi_2(z,y_2)$. If $\frac{\partial\eta_1(y_1,y_2)}{\partial y_2}\neq 0$ and $\xi_1(z,y)\neq 0$ the first equation can be written as

$$\frac{\frac{\partial\eta_1(y_1,y_2)}{\partial y_1}}{\frac{\partial\eta_1(y_1,y_2)}{\partial y_2}}=-\frac{\xi_2(z,y)}{\xi_1(z,y)}=\frac{1}{\xi_1(z,y)}\frac{\frac{\partial s(y_2,z)}{\partial z}}{\frac{\partial s(y_2,z)}{\partial y_2}}=\varphi(y_1,y_2)$$

and the left hand side does not depend on z so that the right hand side does not depend on z either. We now consider $v_1=\eta_1(y_1,y_2)$ and $v_2=\eta_2(y_1,y_2,z)$. The first equation is invariant to the transformations above. The second one is invariant if

$$\left(\frac{\partial}{\partial z}-\frac{\xi_2(z,y_2)}{\varphi(y_1,y_2)}\frac{\partial}{\partial y_1}+\xi_2(z,y_2)\frac{\partial}{\partial y_2}\right)\eta_2(y_1,y_2,z)=0,$$

or equivalently η_2 satisfies the partial differential equation

(16) $$\frac{1}{\xi_2(z,y_2)}\frac{\partial\eta_2(y_1,y_2,z)}{\partial z}-\frac{1}{\varphi(y_1,y_2)}\frac{\partial\eta_2(y_1,y_2,z)}{\partial y_1}+\frac{\partial\eta_2(y_1,y_2,z)}{\partial y_2}=0.$$

Using the method of characteristics we integrate the equations

$$\frac{dz}{\frac{1}{\xi_2(z,y_2)}}=\frac{dy_1}{-\frac{1}{\varphi(y_1,y_2)}}=dy_2$$

This gives two ordinary differential equations

$$\frac{\partial s(y_2,z)}{\partial y_2}\frac{dy_2}{dz}+\frac{\partial s(y_2,z)}{\partial z}=0$$

and

$$\frac{\partial \eta_1(y_1, y_2)}{\partial y_2}\frac{dy_2}{dy_1} + \frac{\partial \eta_1(y_1, y_2)}{\partial y_1} = 0.$$

These are exact differential equations which can be integrated to $s(y_2, z) = c_1$ and $\eta_1(y_1, y_2) = c_2$ respectively, where c_1 and c_2 are arbitrary constants. Thus the general solution to (16) is

(17) $$\eta_2(y_1, y_2, z) = \nu\left(s(y_2, z), \eta_1(y_1, y_2)\right).$$

Thus, for $v_1 = \eta_1(y_1, y_2)$ and $v_2 = \eta_2(y_1, y_2, z)$ to be observationally equivalent to $v_1 = \eta_1(y_1, y_2)$ and $v_2 = s(y_2, z)$ the function $\eta_2(y_1, y_2, z)$ has the form specified in equation (17).

4. Conclusions

The algebraic approach employed in this paper allows us to investigate the problem of overidentification and identification of econometric models from an alternative point of view from those currently used in the literature and to obtain results offering some new insights. In particular, it is shown that the functional form of the structural equation is partly determined by the invariance properties of the reduced form.

The main result of the paper is that for the structural equation and the reduced form to be compatible, the former must be a functional of the maximal invariant under the action of a group leaving the reduced form invariant. In this paper, we assume that the endogenous variables as well as all the errors are continuous and thus rule out some cases of interest (e.g. Chesher (2005) and Vytlacil and Yildiz (2006)). We are currently working on this case and hope to be able to report some results soon.

An easy to implement local approach based on Lie group theory is developed. This is based on an Abelian group action on the exogenous variables. There are other possible group actions on the exogenous variables and it may be worth investigating them. Again, this is currently under investigation. This approach also informs about identification. In fact, some of the features of the structural equation are uniquely defined by the reduced form and are thus identified. Other features, on the other hand, are not uniquely defined as so are unidentified.

6. References

Abadie, A., J. D. Angrist and G. W. Imbens (2002). "Instrumental Variables Estimates and the Effect of Subsidized Training on the Quantiles of Trainee Earnings." *Econometrica* 70: 91-117.

Abraham, R., J. E. Marsden and T. S. Ratiu (1988). *Manifolds, Tensor Analysis, and Applications*, Springer.

Benkard, C. L. and S. Berry (2006). "On the Nonparametric Identification of Nonlinear Simultaneous Equations Models: Comment on Brown (1983) and Roehrig (1988)." *Econometrica* 74: 1429-1440.

Blundel, R. and R. L. Matzkin (2010). "Conditions of the Existence of Control Functions in Nonseparable Simultaneous Equations Models." *cemmap working paper CWP28/10*.

Brown, B. W. (1983). "The Identification Problem in System Nonlinear in the Variables." *Econometrica* 51(1): 175-196.

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

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

Chesher, A. (2003). "Identification in Nonseparable Models." *Econometrica* 71: 1405-1441.

Chesher, A. (2005). "Nonparametric Identification under Discrete Variation." *Econometrica* 73: 1525-2550.

Choi, I. and P. C. B. Phillips (1992). "Asymptotic and Finite Sample Distribution Theory for *Iv* Estimators and Tests in Partially Identified Structural Equations." *Journal of Econometrics* 51: 113-150.

Dudley, R. M. (2002). *Real Analysis and Probability*. Cambridge, Cambridge University Press.

Eaton, M. L. (1982). *Multivariate Statistics: A Vector Space Approach*. Beachwood, Ohio, USA, Institute of Mathematical Statistics, 2007. .

Eaton, M. L. (1989). *Group Invariance Applications in Statistics*. Alexandria, VA, American Statistical Association.

Heckman, J., H. Ichimura and P. Todd (1997). "Matching as an Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Program." *The Review of Economic Studies* 64: 605-654.

Heckman, J. and R. Robb (1984). Alternative Methods of Evaluating the Impact of Interventions. *Longitudinal Analysis of Labor Market Data*. J. Heckman and B. Singer. Cambridge, Cambridge University Press.

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

Imbens, G. W. and W. K. Newey (2009). "Identification and Estimation of Triangular Simultaneous Equations Models without Additivity." *Econometrica* 77(5): 1481-1512.

Lehmann, E. L. (1997). *Testing Statistical Hypotheses*. New York, Springer-Verlag.

Matzkin, R. L. (2003). "Nonparametric Estimation of Nonadditive Random Functions." *Econometrica* 71: 1339-1376.

Matzkin, R. L. (2007). Nonparametric Identification. *The Handbook of Econometrics*. J. J. Heckman and E. E. Leamer. Amsterdam, North-Holland.

Matzkin, R. L. (2008). "Identification in Nonparametric Simultaneous Equations Models." *Econometrica* 76(5): 945-978.

Newey, W. K., J. L. Powell and F. Vella (1999). "Nonparametric Estimation of Triangular Simultaneous Equations Models." *Econometrica* 67: 565-603.

Olver, P. J. (1993). *Applications of Lie Groups to Differential Equations*, Springer-Verlag.

Olver, P. J. (1996). *Equivalence, Invariants and Symmetry*, Cambridge University Press.

Phillips, P. C. B. (1989). "Partially Identified Econometric Models." *Econometric Theory* 5: 181-240.

Poskitt, D. S. and C. L. Skeels (2008). "Conceptual Frameworks and Experimental Design in Simultaneous Equations." *Economics Letters* 100: 138-142.

Roehrig, C. S. (1988). "Conditions for Identification in Nonparametric and Parametric Models." *Econometrica* 56: 433-447.

Rosenblatt, M. (1952). "Remark on a Multivariate Transformation." *Annals of Mathematical Statistics* 23: 470-472.

Vytlacil, E. and N. Yildiz (2006). "Dummy Endogenous Variables in Weakly Separable Models." *Econometrica*.

Wijsman, R. A. (1990). *Invariant Measures on Group and Their Use in Statistic*. Hayward, California, Institute of Mathematical Statistics.