

IDENTIFICATION IN NONPARAMETRIC SIMULTANEOUS EQUATIONS

Rosa L. Matzkin*
Department of Economics
Northwestern University

October 2005

Abstract

This paper considers identification in parametric and nonparametric models, with additive or nonadditive unobservables, and with or without simultaneity among the endogenous variables. Several characterizations of observational equivalence are presented and conditions for identification are developed based on these. It is shown that the results can be extended to situations where the dependent variables are latent. We also demonstrate how the results may be used to derive constructive ways to calculate the unknown functions and distributions in simultaneous equations models, directly from the probability density of the observable variables. Estimators based on this do not suffer from the ill-posed inverse problem that other methods encounter.

* The support of NSF is gratefully acknowledged. I have benefited from the comments and suggestions of seminar participants at Northwestern University, Universidad Torcuato Di Tella, University College London, University of Chicago, Yale University, the Microeconomic Conference at NYU, the Berkeley Conference in Honor of Daniel McFadden, and the BIRS workshop in the series "Mathematical Structures in Economic Theory and Econometrics", and from conversations with Richard Blundell, Donald Brown, Steven Berry, Andrew Chesher, Jerry Hausman, James Heckman, Joel Horowitz, Guido Imbens, Arthur Lewbel, Daniel McFadden, Lars Nesheim, Whitney Newey, Elie Tamer, and Chris Sims.

1. Introduction

Many economic models, specially those concerning optimization and equilibrium conditions, involve, either explicitly or implicitly, several equations. In such situations, it is often the case that some of the observable explanatory variables are not distributed independently of some or all of the unobservable variables in the system. Identification becomes then of concern.

In this paper, we develop several conditions to determine identification in parametric and non-parametric models, with additive or nonadditive unobservables, and with or without simultaneity among the endogenous variables. We show how our results can be used to determine identification in simultaneous equations with discrete endogenous variables. We also indicate how to derive constructive ways to calculate the unknown functions and distributions in simultaneous equation models, directly from the probability density of the observable variables.

We analyze identification using two approaches. Both are based on the methodology developed in seminal work, by B. Brown (1983). B. Brown (1983) analyzed identification in systems nonlinear in variables and linear in parameters. He developed rank conditions by first defining a mapping from the observable and unobservable explanatory variables of the true model to the unobservables of an alternative model. Then, he asked what restrictions of the alternative model guarantee that the unobservable of the alternative model is independent of the observable explanatory variables. The resulting conditions were extensions of the familiar rank conditions developed by Koopmans, Rubin, and Leipnik (1950) for the identification of systems linear in variables and parameters. Earlier fundamental works in the study of identification of systems nonlinear in variables were Fisher (1961, 1965, 1966), Rothenberg (1971), and Bowden (1973). (See Hausman (1983) and Hsiao (1983) for an analysis of these and other works on identification.) Roehrig (1988) extended the ideas in B. Brown (1983) to systems of nonparametric equations, obtaining similar rank conditions. Recently, Benkard and Berry (2004) established that in some situations B. Brown's conditions are not necessary for observational equivalence.

In part of this paper, we revisit the question analyzed by B. Brown and Roehrig, and derive different sets of conditions. As in B. Brown (1983), our conditions are derived from relating the explanatory variables of the true model to the explanatory variables of the alternative model, but, unlike B. Brown (1983), our conditions relate these through a mapping connecting the density of the unobservable variables of the true model to the density of the unobservable variables under the alternative model. We derive new rank conditions, as well as alternative conditions. We also present several restrictions under which our conditions for identification are satisfied when those developed by B. Brown (1983) and Roehrig (1988) are satisfied.

An approach that has been more recently taken when analyzing identification of a nonparametric function within a set of nonparametric functions proceeds by first determining what type of transformations, or perturbations, of a function are observationally equivalent to the function. For example, in binary response, generalized models, and, more recently in single-equations with nonadditive errors, Matzkin (1992, 1994, 2003) establishes identification of nonparametric functions up to monotone transformations. Determining the set of transformations of the true function that are observationally equivalent to the true function allows one to determine a set of functions under which the true function is identified, by restricting the set to include, for each function, only one observationally equivalent transformation. In Section 3, we analyze identification by providing conditions for observational equivalence between a true function and a transformation of it. We exemplify our results by describing a simple transformation that satisfies the "rotation" property described in Benkard and Berry (2004), and showing that it is observationally equivalent to the original function. As described in Benkard and Berry (2004), rotations can generate situations where B. Brown (1983) and Roehrig (1988) rank conditions imply non-observational equivalence.

Under some conditions, our analysis of identification can be used also in models where the

dependent variables are unobservable, but one can observe, instead, at least a threshold crossing indicator of the dependent variable. For example, one may consider a situation where the profits of firms in an industry are determined simultaneously. Each firm knows the profits of the other firms, but the econometrician knows only whether those profits are above or below a threshold.

Our identification conditions can be used to develop consistent estimators by using, as in D. Brown and Matzkin (1998), a nonparametric version of the semiparametric “minimum distance from independence” criterion, developed by Manski (1983). D. Brown and Wegkamp (2002) showed that the distribution-free estimator is consistent and asymptotically normal, but, thus far, the nonparametric version is only known to be consistent. In Section 6, we indicate a new method to estimate simultaneous equations, which can produce consistent and asymptotically normal estimators. The method uses normalizations and/or restrictions, derived from our identification results, to recover the unknown nonparametric functions and distribution of the unobservable variables, directly from the distribution of the observable variables. Estimation can then proceed by substituting the true distribution of the observable variables by a nonparametric estimator of it. The derived estimator does not suffer from the ill-posed inverse problem that typically appears in other estimation methods.

Our results add to an increasingly large literature that has developed in recent years, dealing with identification and estimation of nonparametric models with endogenous explanatory variables. Newey, Powell and Vella (1999) and Pinske (2000) considered nonparametric triangular systems with additive unobservable random terms. Altonji and Ichimura (2000) considered models with latent variables. Altonji and Matzkin (2001) provided estimators for average derivatives in nonseparable models, using conditional independence, and for nonparametric nonseparable functions, using exchangeability. Altonji and Matzkin (2003) extended their 2001 results to discrete endogenous regressors. Chesher (2003) considered local identification of derivatives in triangular systems with nonadditive random terms. Imbens and Newey (2003) studied global identification and estimation of triangular systems with nonadditive random terms. Matzkin (2003) considered estimation under conditional independence, using either normalizations or restrictions on nonadditive functions. Vytlacil and Yildiz (2004) studied estimation of average effects in models with weak separability and dummy endogenous variables. Chesher (2004) considered nonparametric index models. Chesher (2005) analyzed identification in additive error models with discrete endogenous variables. For nontriangular systems, Newey and Powell (1989, 2003), Darolles, Florens, and Renault (2003), Ai and Chen (2003), Hall and Horowitz (2003), and Chernozhukov, Imbens, and Newey (2004) considered estimation using conditional moment conditions between unobservables and instruments. Das (2004) analyzed identification with endogenous discrete variables. Matzkin (2004) considered identification imposing independence among the unobservable variables. Chernozhukov and Hansen (2005) used a rank invariance assumption on a nonadditive random term, conditional on an instrument and another variable. Blundell and Powell (2003) analyzed several nonparametric and semiparametric models and provided many other references.

The outline of the paper is as follows. In the next section, we characterize observational equivalence between the true function in a simultaneous equations model and an alternative function. In Section 3, we characterize observational equivalence of transformations of the true function. Section 4 connects our results to the previous results of B. Brown and Roehrig, and presents several conditions under which determining identification using those previous results, imply identification using our new results. The extension of our results to determine identification in a model with latent endogenous variables is presented in Section 5. In Section 6, we demonstrate how the results may be used to derive constructive ways to calculate the unknown functions and distributions in simultaneous equation models, directly from the probability density of the observable variables. Section 7 concludes and discusses extensions. The Appendix contains the proofs of the theorems and lemmas presented in the main sections.

2. Observational equivalence of two structural functions

We consider a system of equations, described as

$$(2.1) \quad U = r(Y, X)$$

where $r : R^{G+K} \rightarrow R^G$ is an unknown, twice continuously differentiable function, Y is a vector of G observable endogenous variables, X is vector of K observable exogenous variables, and U is a vector of G unobservable variables, which is assumed to be distributed independently of X . A typical example of such a system is a demand and supply model,

$$\begin{aligned} Q &= D(P, I, U_1) \\ P &= S(Q, W, U_2) \end{aligned}$$

where Q and P denote the quantity and price of a commodity, I denotes consumers' income, W denotes producers' input prices, U_1 denotes an unobservable demand shock, and U_2 denotes an unobservable supply shock. If the demand function, D , is strictly increasing in U_1 and the supply function, S , is strictly increasing in U_2 , one can invert these functions and write this system as in (2.1), with $Y = (P, Q)$, $X = (I, W)$, $U = (U_1, U_2)$, r_1 denoting the inverse of D with respect to U_1 , and r_2 denoting the inverse of S with respect to U_2 .

Given the distribution, $f_{Y,X}$, of the observable variables (Y, X) , we want to determine under what restrictions on model (2.1) we can recover the function r and the distribution of U from $f_{Y,X}$. An equivalent question asks what restrictions on the set of pairs $(\tilde{r}, f_{\tilde{U}})$, of functions \tilde{r} and densities $f_{\tilde{U}}$, to which (r, f_U) belongs, guarantee that any pair $(\tilde{r}, f_{\tilde{U}})$ in this set that is different from (r, f_U) generates a distribution of (Y, X) that is different from the true one.

In seminal work, B. Brown (1983) proposed to answer this question, in a model where U is independent of X , by asking the restrictions that independence between the unobservable exogenous variables and the observable exogenous variables imposes on any function \tilde{r} . Given any function \tilde{r} , one can define a random variable, \tilde{U} , by

$$(2.2) \quad \tilde{U} = \tilde{r}(Y, X),$$

and given the true distribution of (Y, X) , one can then use this distribution, together with (2.2), to generate the implied distribution of \tilde{U} , conditional on X . This conditional distributions will, in general, vary with X . B. Brown (1983) asked under what restrictions will these conditional distributions be the same for all values of X . Since \tilde{U} depends on \tilde{r} and $f_{Y,X}$, and the latter depends on r and f_U , the condition that the distribution of \tilde{U} be the same for all X imposes restrictions on \tilde{r} , r , and f_U . Suppose that such restrictions were found, and that they were established to be necessary and sufficient for the independence between \tilde{U} and X . Then, any function \tilde{r} that does not satisfy those restrictions can not generate the true distribution of the observable variables, $f_{Y,X}$, with a distribution, $f_{\tilde{U}}$, that is independent of X . Consider then a set of pairs $(\tilde{r}, f_{\tilde{U}})$ where \tilde{U} is independent of X and \tilde{r} does not satisfy the restrictions for \tilde{U} to be independent of X . Then, within this set of pairs $(\tilde{r}, f_{\tilde{U}})$, it must be the case that (r, f_U) is the only pair that generates the true distribution of (Y, X) . If any other pair $(\tilde{r}, f_{\tilde{U}})$ would generate such distribution, then letting $\tilde{U} = \tilde{U}$, we would get that \tilde{U} is independent of X , which is impossible by the way \tilde{r} was selected into the set. Hence, as long as the functions \tilde{r} are so selected, (r, f_U) will be identified within any set of pairs formed by those \tilde{r} functions with distributions that are independent of X . To derive conditions on \tilde{r} , for \tilde{U} , defined by (2.2), to be independent of X , B. Brown proposed to use the reduced form, $Y = h(X, U)$, to define a mapping from (X, U) to \tilde{U} , by

$$\tilde{U} = \tilde{r}(h(X, U), X)$$

He then argued that a necessary and sufficient conditions for \tilde{U} to be independent of X is that for all y, x ,

$$(2.3) \quad \left[\frac{\partial r(y, x)}{\partial x} - \frac{\partial r(y, x)}{\partial y} \left(\frac{\partial \tilde{r}(y, x)}{\partial y} \right)^{-1} \frac{\partial \tilde{r}(y, x)}{\partial x} \right] = 0$$

Under some assumptions, this condition is equivalent to requiring that the rank of the matrix

$$\begin{pmatrix} \frac{\partial \tilde{r}(y, x)}{\partial y} & \frac{\partial \tilde{r}(y, x)}{\partial x} \\ \frac{\partial r(y, x)}{\partial y} & \frac{\partial r(y, x)}{\partial x} \end{pmatrix}$$

is G .¹ Benkard and Berry (2004) recently established that, in some situations, (2.3) is not a necessary condition for independence between \tilde{U} and X .

In this section, we derive new conditions for \tilde{U} , defined by (2.2), to be independent of X . Hence, after replacing the old conditions with our new ones, one can continue using B. Brown's (1983) approach to determining sets where the true function and distribution are identified.

To derive our first set of results, we will make almost the same assumptions that B. Brown and Roehrig made. We will assume that r is such that for every values u and x of, respectively, U and X , there exists a unique value, y , of Y satisfying (2.1). Denote the function that maps X and U into Y by $h : R^{K+G} \rightarrow R^G$, so that

$$(2.2) \quad Y = h(X, U)$$

We will assume that for all y, x , $|\partial r(y, x)/\partial y| > 0$. Let f_U denote the density of U , assumed to be differentiable and everywhere positive on R^G . Let $f_{Y,X}$ denote the density of (Y, X) , assumed to be differentiable and everywhere positive on R^{G+K} . Since $f_{Y,X}$ is generated from r and f_U , for all y, x

$$(2.3) \quad f_{Y|X=x}(y) = f_U(r(y, x)) \left| \frac{\partial r(y, x)}{\partial y} \right|$$

Consider an alternative function \tilde{r} that satisfies the same regularity conditions that r satisfies. That is, letting

$$(2.4) \quad \tilde{U} = \tilde{r}(Y, X)$$

we will assume that \tilde{r} is such that for every values, \tilde{u} and x , of \tilde{U} and X , there exists a unique value, y , of Y satisfying (2.3). We will denote the function that maps X and \tilde{U} into Y by $\tilde{h} : R^{K+G} \rightarrow R^G$, so that

$$(2.5) \quad Y = \tilde{h}(X, \tilde{U})$$

We will also assume that for all y, x , $|\partial \tilde{r}(y, x)/\partial y| > 0$.

Let $f_{\tilde{U},X}$ be a density of (\tilde{U}, X) such that $\tilde{U} = \tilde{r}(Y, X)$ & $f_{\tilde{U},X}$ generate $f_{Y,X}$. Then, for all y, x

$$(2.6) \quad f_{Y|X=x}(y) = f_{\tilde{U}|X=x}(\tilde{r}(y, x)) \left| \frac{\partial \tilde{r}(y, x)}{\partial y} \right|$$

¹The expressions are taken from Roehrig (1988), who presented them in terms of nonparametric functions, rather than from B. Brown (1983), who presented them in terms of parametric functions.

or, using (2.3), for all y, x

$$(2.7) \quad f_U(r(y, x)) \left| \frac{\partial r(y, x)}{\partial y} \right| = f_{\tilde{U}|X=x}(\tilde{r}(y, x)) \left| \frac{\partial \tilde{r}(y, x)}{\partial y} \right|$$

We can now ask, in analogy to the reasoning in B. Brown (1983), what are the restrictions on \tilde{r} guaranteeing that \tilde{U} is independent of X , when (2.7) is satisfied for all y, x . Under our assumptions, this is equivalent to asking the restrictions on \tilde{r} guaranteeing that for all \tilde{u} and all x

$$(2.8) \quad \frac{\partial f_{\tilde{U}|X=x}(\tilde{u})}{\partial x} = 0$$

when (2.7) is satisfied.

To establish these conditions, we use the invertibility, given X , between U and Y and between \tilde{U} and Y . These invertibility conditions allow us to establish an invertible relationship between U and \tilde{U} , given X . In particular, for any values \tilde{u}, x of \tilde{U} and X , let y be the unique value satisfying

$$y = \tilde{h}(x, \tilde{u})$$

Then, the mapping between \tilde{u} and u is given by

$$u = r(\tilde{h}(x, \tilde{u}), x)$$

Substituting into (2.7), we get that

$$(2.9) \quad f_{\tilde{U}|X=x}(\tilde{u}) = f_U(r(\tilde{h}(x, \tilde{u}), x)) \left| \frac{\partial r(\tilde{h}(x, \tilde{u}), x)}{\partial y} \right| \left| \frac{\partial \tilde{r}(\tilde{h}(x, \tilde{u}), x)}{\partial y} \right|^{-1}$$

Theorem 1, which is proved in the Appendix, establishes the restrictions that (2.8) imposes on $f_{\tilde{U}|X=x}$. It uses the following assumption:

ASSUMPTION 2.1: *The models (2.1) and (2.2) are such that*

(i) *the densities f_U, f_X , and for all $x, f_{\tilde{U}|X=x}$ are everywhere positive and continuously differentiable,*

(ii) *the functions r and \tilde{r} are continuously differentiable,*

(iii) *for all y, x $|\partial r(y, x) / \partial y| > 0$ and $|\partial \tilde{r}(y, x) / \partial y| > 0$,*

(iv) *for all x, u there exists a unique y satisfying $u = r(y, x)$, and*

(v) *for all x, \tilde{u} there exists a unique \tilde{y} satisfying $\tilde{u} = \tilde{r}(\tilde{y}, x)$*

Assumption 2.1(i) allows us to characterize the distributions of the random variables by their densities. It guarantees that the support of the densities is the whole Euclidean space, and that

these densities converge to 0 as the value of one of their arguments tends to infinity. Relaxing the full support condition would involve dealing with supports for \tilde{U} conditional on X that depend on the value of X . Note that the condition that the support of $f_{\tilde{U}|X=x}$ is R^G requires that $\tilde{r}(\cdot, x)$ is unbounded. Assumption 2.1(ii) allow us to express our conditions in terms of derivatives of r and \tilde{r} . The continuity of the derivatives, together with Assumption 2.1(i), guarantees that some conditions are satisfied on a neighborhood of (y, x) that possesses positive probability, whenever these conditions are satisfied at (y, x) . Assumption 2.1(iii) is analogous to the monotonicity condition that was used in Matzkin (1999) to identify single-equation models with nonadditive unobservable random terms. Assumptions 2.1(iv) and 2.1(v) guarantee that r and \tilde{r} possess reduced form systems, $y = h(x, u)$ and $y = \tilde{h}(x, \tilde{u})$, respectively. In equilibrium price models, where y denotes a vector of equilibrium prices, this assumption requires that for each vector of the exogenous variables X , and the unobservable random variables, equilibrium is unique. The result in Roehrig (1988) used Assumption 2.1(i), 2.1(ii), 2.1(iv) and 2.1(v).

Theorem 2.1: *Suppose that Assumption 2.1 is satisfied, and that $f_{\tilde{U},X}$ satisfies (2.7). Then, $f_{\tilde{U},X}$ satisfies (2.8) if and only if for all y, x*

$$\begin{aligned}
(T2.1) \quad & \frac{\partial \log f_U(r(y, x))}{\partial u} \left[\frac{\partial r(y, x)}{\partial x} - \frac{\partial r(y, x)}{\partial y} \left(\frac{\partial \tilde{r}(y, x)}{\partial y} \right)^{-1} \frac{\partial \tilde{r}(y, x)}{\partial x} \right] \\
& + \left[\left(\frac{\partial}{\partial x} \log \left| \frac{\partial r(y, x)}{\partial y} \right| - \frac{\partial}{\partial x} \log \left| \frac{\partial \tilde{r}(y, x)}{\partial y} \right| \right) \right] \\
& - \left[\left(\frac{\partial}{\partial y} \log \left| \frac{\partial r(y, x)}{\partial y} \right| - \frac{\partial}{\partial y} \log \left| \frac{\partial \tilde{r}(y, x)}{\partial y} \right| \right) \left(\frac{\partial \tilde{r}(y, x)}{\partial y} \right)^{-1} \frac{\partial \tilde{r}(y, x)}{\partial x} \right] \\
& = 0
\end{aligned}$$

Condition (T2.1) expresses a necessary and sufficient condition for independence between \tilde{U} and X in terms of the true density f_U , the true function r , and the alternative function, \tilde{r} . It can be shown that

$$\left(\frac{\partial \tilde{r}(y, x)}{\partial y} \right)^{-1} \frac{\partial \tilde{r}(y, x)}{\partial x} = - \frac{\partial \tilde{h}(x, \tilde{u})}{\partial x}$$

where $\tilde{h}(x, \tilde{u})$ is the reduced form function of the alternative model evaluated at $\tilde{u} = \tilde{r}(y, x)$. Hence, a different way of writing condition (T2.1), in terms of the reduced form function of the alternative model is

$$\begin{aligned}
(T2.2) \quad & \frac{\partial \log f_U(r(y, x))}{\partial u} \left[\frac{\partial r(y, x)}{\partial x} + \frac{\partial r(y, x)}{\partial y} \frac{\partial \tilde{h}(x, \tilde{u})}{\partial x} \right] \\
& + \frac{\partial}{\partial x} \left(\log \left| \frac{\partial r(y, x)}{\partial y} \right| \right) + \frac{\partial}{\partial y} \left(\log \left| \frac{\partial r(y, x)}{\partial y} \right| \right) \frac{\partial \tilde{h}(x, \tilde{u})}{\partial x} \\
& = \frac{\partial}{\partial x} \log \left| \frac{\partial \tilde{r}(y, x)}{\partial y} \right| + \frac{\partial}{\partial y} \left(\log \left| \frac{\partial \tilde{r}(y, x)}{\partial y} \right| \right) \frac{\partial \tilde{h}(x, \tilde{u})}{\partial x}
\end{aligned}$$

This can be interpreted as stating that the proportional change in the conditional density of Y given X , when the value of X changes and Y responds to that change according to \tilde{h} has to equal the proportional change in the value of the determinant determined by \tilde{r} when X changes and Y responds to that change according to \tilde{h} . Or, in other words, the effects of the alternative model on the density and the determinant have to cancel out.

The result of Theorem 2.1 can be used to determine observational equivalence between two functions, according to the following definition.

Definition 2.1: A function \tilde{r} will be said to be **observationally equivalent** to r , when f_U is the true density function of U , if there exists a random vector \bar{U} that is independent of X and is such that (r, f_U) and $(\tilde{r}, f_{\bar{U}})$ generate the same distribution of (Y, X) .

The following result connects the definition of observational equivalence to condition (T2.1).

Theorem 2.2: Suppose that r and \tilde{r} satisfy Assumption 2.1. Then, \tilde{r} is observationally equivalent to r if and only if for all y, x , (T2.1) is satisfied.

To express conditions (T2.1) in a more succinct way, we define, for any y, x , the $G \times K$ matrix $A(y, x; \partial r, \partial \tilde{r})$, the $1 \times K$ vector $b(y, x; \partial r, \partial \tilde{r}, \partial^2 r, \partial^2 \tilde{r})$ and the $1 \times G$ vector $\gamma(y, x; f_U, r)$ by

$$A(y, x; \partial r, \partial \tilde{r}) = \left[\frac{\partial r(y, x)}{\partial x} - \frac{\partial r(y, x)}{\partial y} \left(\frac{\partial \tilde{r}(y, x)}{\partial y} \right)^{-1} \frac{\partial \tilde{r}(y, x)}{\partial x} \right]$$

$$\begin{aligned}
& b(y, x; \partial r, \partial \tilde{r}, \partial^2 r, \partial^2 \tilde{r}) \\
&= - \left[\left(\frac{\partial}{\partial x} \log \left(\left| \frac{\partial r(y, x)}{\partial y} \right| \right) - \frac{\partial}{\partial x} \log \left(\left| \frac{\partial \tilde{r}(y, x)}{\partial y} \right| \right) \right) \right] \\
&+ \left[\left(\frac{\partial}{\partial y} \log \left(\left| \frac{\partial r(y, x)}{\partial y} \right| \right) - \frac{\partial}{\partial y} \log \left(\left| \frac{\partial \tilde{r}(y, x)}{\partial y} \right| \right) \right) \left(\frac{\partial \tilde{r}(y, x)}{\partial y} \right)^{-1} \frac{\partial \tilde{r}(y, x)}{\partial x} \right]
\end{aligned}$$

and

$$\gamma(y, x; f_U, r) = \frac{\partial \log(f_U(r(y, x)))}{\partial u}$$

We index $A(y, x)$ by $(\partial r, \partial \tilde{r})$, $b(y, x)$ by $(\partial r, \partial \tilde{r}, \partial^2 r, \partial^2 \tilde{r})$, and $\gamma(y, x)$ by (f_U, r) to emphasize that the value of A depends on the first order derivatives of the functions r and \tilde{r} , the value of b depends on the first and second derivatives of the functions r and \tilde{r} , and the value of γ depends of the function f_U and the value of the function r . Condition (T2.1) can be expressed as stating that for all y, x

$$\gamma(y, x; f_U, r) A(y, x; \partial r, \partial \tilde{r}) = b(y, x; \partial r, \partial \tilde{r}, \partial^2 r, \partial^2 \tilde{r})$$

Let A_j denote the j -th column of A , b_j denote the j -th coordinate of b , and a_{ij} denote the ij -th element of A . To show that a function \tilde{r} satisfying Assumption 2.1 is not observationally equivalent to r it suffices to show that for some j and some y, x ,

$$\gamma(y, x; f_U, r) A_j(y, x; \partial r, \partial \tilde{r}) \neq b_j(y, x; \partial r, \partial \tilde{r}, \partial^2 r, \partial^2 \tilde{r})$$

Example 2.1: Consider the model described in B. Brown (1983, page 179), where

$$\begin{aligned}
U_1 &= Y_1 + \alpha_1 \\
U_2 &= \alpha_2 Y_1^2 + Y_2 + \alpha_3 X + \alpha_4
\end{aligned}$$

and

$$\begin{aligned}
\tilde{U}_1 &= Y_1 + \tilde{\alpha}_1 \\
\tilde{U}_2 &= \tilde{\alpha}_2 Y_1^2 + Y_2 + \tilde{\alpha}_3 X + \tilde{\alpha}_4
\end{aligned}$$

In this example, for all y, x , $b(y, x) = 0$, since

$$\frac{\partial \left| \frac{\partial r(y,x)}{\partial y} \right|}{\partial x} = \frac{\partial \left| \frac{\partial \tilde{r}(y,x)}{\partial y} \right|}{\partial x} = \frac{\partial \left| \frac{\partial r(y,x)}{\partial y} \right|}{\partial y} = \frac{\partial \left| \frac{\partial \tilde{r}(y,x)}{\partial y} \right|}{\partial y} = 0.$$

Moreover, for all y, x

$$A(y, x) = \begin{bmatrix} 0 \\ \alpha_3 - \tilde{\alpha}_3 \end{bmatrix}.$$

Then, for all y, x ,

$$\gamma(y, x; f_U, r) A(y, x; \partial r, \partial \tilde{r}) = \left(\frac{\partial f_U(y_1 + \alpha_1, \alpha_2 y_1^2 + y_2 + \alpha_3 x + \alpha_4) / \partial u_2}{f_U(u_1, u_2)} \right) (\alpha_3 - \tilde{\alpha}_3)$$

and

$$b(y, x; \partial r, \partial \tilde{r}, \partial^2 r, \partial^2 \tilde{r}) = 0$$

Let (u_1, u_2) be such that $f_U(u_1, u_2) \neq 0$ and $\partial f_U(u_1, u_2) / \partial u_2 \neq 0$. Let $y_1 = u_1 - \alpha_1$ and $y_2 = u_2 - \alpha_2 y_1^2 - \alpha_3 x - \alpha_4$. Then, if $\alpha_3 \neq \tilde{\alpha}_3$,

$$\gamma(y, x; f_U, r) A(y, x; \partial r, \partial \tilde{r}) \neq b_j(y, x; \partial r, \partial \tilde{r}, \partial^2 r, \partial^2 \tilde{r})$$

Hence, α_3 is identified.

2.1. Interpretation and Rank Conditions

To interpret the result of the theorem, we recall (2.7). Given r and f_U , consider an alternative function \tilde{r} that satisfies Assumption 2.1 and generates with \tilde{U} the same distribution of (Y, X) as r and f_U generate. Then, for all y, x

$$(2.7) \quad f_{\tilde{U}|X=x}(\tilde{r}(y, x)) \left| \frac{\partial \tilde{r}(y, x)}{\partial y} \right| = f_U(r(y, x)) \left| \frac{\partial r(y, x)}{\partial y} \right|$$

Taking log's on both sides and differentiating the expression first with respect to y and then with respect to x , one gets that

$$(2.10) \quad \frac{\partial \log f_{\tilde{U}|X=x}(\tilde{r}(y, x))}{\partial \tilde{u}} \frac{\partial \tilde{r}(y, x)}{\partial y} + \frac{\partial}{\partial y} \log \left| \frac{\partial \tilde{r}(y, x)}{\partial y} \right|$$

$$= \frac{\partial \log (f_U(r(y, x)))}{\partial u} \frac{\partial r(y, x)}{\partial y} + \frac{\partial}{\partial y} \log \left| \frac{\partial r(y, x)}{\partial y} \right|$$

and

$$\begin{aligned}
(2.11) \quad & \frac{\partial \log f_{\tilde{U}|X=x}(\tilde{r}(y, x))}{\partial \tilde{u}} \frac{\partial \tilde{r}(y, x)}{\partial x} + \frac{\partial}{\partial x} \log \left| \frac{\partial \tilde{r}(y, x)}{\partial y} \right| + \frac{\partial \log f_{\tilde{U}|X=x}(t)}{\partial x} \Big|_{t=\tilde{r}(y, x)} \\
& = \frac{\partial \log (f_U(r(y, x)))}{\partial u} \frac{\partial r(y, x)}{\partial x} + \frac{\partial}{\partial x} \log \left| \frac{\partial r(y, x)}{\partial y} \right|
\end{aligned}$$

The critical term in these expressions, whose value determines the dependence between \tilde{U} and X is $\partial \log f_{\tilde{U}|X=x}(t) / \partial x$. Given r , f_U , and \tilde{r} , one can view (2.10) and (2.11) as a system of equations with unknown vectors

$$\frac{\partial \log f_{\tilde{U}|X=x}(\tilde{r}(y, x))}{\partial \tilde{u}} \quad \text{and} \quad \frac{\partial \log f_{\tilde{U}|X=x}(t)}{\partial x} \Big|_{t=\tilde{r}(y, x)}$$

We may then ask under what conditions a solution exists and satisfies

$$\frac{\partial \log f_{\tilde{U}|X=x}(t)}{\partial x} \Big|_{t=\tilde{r}(y, x)} = 0$$

Using (2.10) to solve for $\partial \log f_{\tilde{U}|X=x}(\tilde{r}(y, x)) / \partial \tilde{u}$ and substituting the resulting expression into (2.11) to solve for $\partial \log f_{\tilde{U}|X=x}(t) / \partial x$ when $t = \tilde{r}(y, x)$, one gets that

$$\begin{aligned}
\frac{\partial \log f_{\tilde{U}|X=x}(t)}{\partial x} \Big|_{t=\tilde{r}(y, x)} & = \frac{\partial \log f_U(r(y, x))}{\partial u} \left[\frac{\partial r(y, x)}{\partial x} - \frac{\partial r(y, x)}{\partial y} \left(\frac{\partial \tilde{r}(y, x)}{\partial y} \right)^{-1} \frac{\partial \tilde{r}(y, x)}{\partial x} \right] \\
& + \left[\left(\frac{\partial}{\partial x} \log \left| \frac{\partial r(y, x)}{\partial y} \right| - \frac{\partial}{\partial x} \log \left| \frac{\partial \tilde{r}(y, x)}{\partial y} \right| \right) \right] \\
& - \left[\left(\frac{\partial}{\partial y} \log \left| \frac{\partial r(y, x)}{\partial y} \right| - \frac{\partial}{\partial y} \log \left| \frac{\partial \tilde{r}(y, x)}{\partial y} \right| \right) \left(\frac{\partial \tilde{r}(y, x)}{\partial y} \right)^{-1} \frac{\partial \tilde{r}(y, x)}{\partial x} \right]
\end{aligned}$$

Condition (T2.1) states that the right hand side of this expression equals zero. Using, instead, Cramer's rule, we can obtain an equivalent condition, in terms of the rank of a matrix. Define

$$\tilde{\gamma}_{\tilde{u}}(\tilde{r}(y, x), x) = \frac{\partial \log f_{\tilde{U}|X=x}(\tilde{r}(y, x))}{\partial \tilde{u}},$$

$$\tilde{\gamma}_x(\tilde{r}(y, x), x) = \frac{\partial \log f_{\tilde{U}|X=x}(t)}{\partial x} \Big|_{t=\tilde{r}(y, x)}$$

$$\gamma(r(y, x)) = \frac{\partial \log f_U(r(y, x))}{\partial u},$$

$$\Delta_y(y, x; \partial r, \partial^2 r, \partial \tilde{r}, \partial^2 \tilde{r}) = \frac{\partial}{\partial y} \log \left| \frac{\partial r(y, x)}{\partial y} \right| - \frac{\partial}{\partial y} \log \left| \frac{\partial \tilde{r}(y, x)}{\partial y} \right|$$

$$\Delta_x(y, x; \partial r, \partial^2 r, \partial \tilde{r}, \partial^2 \tilde{r}) = \frac{\partial}{\partial x} \log \left| \frac{\partial r(y, x)}{\partial y} \right| - \frac{\partial}{\partial x} \log \left| \frac{\partial \tilde{r}(y, x)}{\partial y} \right|$$

We can state an equivalent result to Theorem 2.1:

Theorem 2.3: *Suppose that Assumption 2.1 is satisfied. Then, \tilde{r} is observationally equivalent to r if and only if for all y, x the rank of the matrix*

$$\begin{pmatrix} \left(\frac{\partial \tilde{r}(y, x)}{\partial y} \right)' & \Delta_y(y, x; \partial r, \partial^2 r, \partial \tilde{r}, \partial^2 \tilde{r}) - \frac{\partial \log(f_U(r(y, x)))}{\partial u} \frac{\partial r(y, x)}{\partial y} \\ \left(\frac{\partial \tilde{r}(y, x)}{\partial x} \right)' & \Delta_x(y, x; \partial r, \partial^2 r, \partial \tilde{r}, \partial^2 \tilde{r}) - \frac{\partial \log(f_U(r(y, x)))}{\partial u} \frac{\partial r(y, x)}{\partial x} \end{pmatrix}$$

is G .

Example 2.2: Consider a demand and supply model specified as

$$\begin{aligned} U_1 &= D(p, q) + m(I) \\ U_2 &= S(p, q) + v(W) \end{aligned}$$

and an alternative model specified as

$$\begin{aligned} \tilde{U}_1 &= \tilde{D}(p, q) + \tilde{m}(I) \\ \tilde{U}_2 &= \tilde{S}(p, q) + \tilde{v}(W) \end{aligned}$$

Suppose that $m(I)$ and $v(W)$ have support equal to R , and that at some point, u^* , $\partial_U f(u^*)/\partial u = 0$. Then, for all p, q there exist I^* and W^* such that

$$\frac{\partial f_U(D(p, q) + m(I^*), S(p, q) + v(W^*))}{\partial u} = 0$$

It then follows from Theorem 2.3 that the alternative model is not observationally equivalent to the true model if at this particular p, q, I^*, W^* the rank of the matrix

$$\begin{pmatrix} \frac{\partial \tilde{D}(p, q)}{\partial p} & \frac{\partial \tilde{S}(p, q)}{\partial p} & \Delta_p(\partial r, \partial^2 r, \partial \tilde{r}, \partial^2 \tilde{r}) \\ \frac{\partial \tilde{D}(p, q)}{\partial q} & \frac{\partial \tilde{S}(p, q)}{\partial q} & \Delta_q(\partial r, \partial^2 r, \partial \tilde{r}, \partial^2 \tilde{r}) \\ \frac{\partial \tilde{m}(I^*)}{\partial I} & 0 & 0 \\ 0 & \frac{\partial \tilde{v}(W^*)}{\partial W} & 0 \end{pmatrix}$$

is strictly larger than 2. This would be satisfied, for example, if $\partial \tilde{m}(I^*)/\partial I \neq 0$, $\partial \tilde{v}(W^*)/\partial W \neq 0$ and either $\Delta_p(\partial r, \partial^2 r, \partial \tilde{r}, \partial^2 \tilde{r}) \neq 0$ or $\Delta_q(\partial r, \partial^2 r, \partial \tilde{r}, \partial^2 \tilde{r}) \neq 0$. It would also be satisfied if $\partial \tilde{m}(I^*)/\partial I \neq 0$ and

$$\frac{\frac{\partial \tilde{S}(p, q)}{\partial p}}{\frac{\partial \tilde{S}(p, q)}{\partial q}} \neq \frac{\Delta_p(\partial r, \partial^2 r, \partial \tilde{r}, \partial^2 \tilde{r})}{\Delta_q(\partial r, \partial^2 r, \partial \tilde{r}, \partial^2 \tilde{r})}$$

or if $\partial \tilde{v}(W^*)/\partial W \neq 0$ and

$$\frac{\frac{\partial \tilde{S}(p, q)}{\partial p}}{\frac{\partial \tilde{S}(p, q)}{\partial q}} \neq \frac{\Delta_p(\partial r, \partial^2 r, \partial \tilde{r}, \partial^2 \tilde{r})}{\Delta_q(\partial r, \partial^2 r, \partial \tilde{r}, \partial^2 \tilde{r})}$$

Note that $\Delta_p(\partial r, \partial^2 r, \partial \tilde{r}, \partial^2 \tilde{r})$ and $\Delta_q(\partial r, \partial^2 r, \partial \tilde{r}, \partial^2 \tilde{r})$ are not functions of (I, W) .

3. Observational equivalence of transformations of structural functions

An approach that has often been taken when analyzing identification of a nonparametric function within a set of nonparametric functions proceeds by first determining what type of transformations, or perturbations, of a function are observationally equivalent to the function. For example, in binary response, generalized models, and, more recently in single-equations with nonadditive errors, Matzkin (1992, 1994, 2003) establishes identification of nonparametric functions up to monotone transformations. Determining the set of transformations of the true function that are observationally equivalent to the true function allows one to determine a set of functions under which the true function is identified, by restricting the set to include only one of the observationally equivalent transformations. Moreover, the conditions for observational equivalence become less cumbersome

than the ones presented in Section 2. In this section, we develop analogous results to those in Section 2, in terms of transformations.

Given the function r and an alternative function \tilde{r} , satisfying the model and assumptions in Section 2, we can express \tilde{r} as a transformation g of u by

$$(3.1) \quad g(u, x) = \tilde{r}(h(x, u), x)$$

By our assumptions, it follows that

$$(3.2) \quad \left| \frac{\partial g(u, x)}{\partial u} \right| = \left| \frac{\partial \tilde{r}(h(x, u), x)}{\partial y} \right| \left| \frac{\partial h(x, u)}{\partial u} \right| \neq 0$$

Hence, g is invertible in u . The representation of \tilde{r} in terms of the transformation g implies that for all y, x

$$(3.3) \quad \tilde{r}(y, x) = g(r(y, x), x)$$

We can then proceed to analyze observational equivalence between r and \tilde{r} in terms of observational equivalence between $r(y, x)$ and $g(r(y, x), x)$. Recall that, for any given x , \tilde{r} generates the same distribution of Y given $X = x$ iff for all y

$$(3.4) \quad f_{\tilde{U}|X=x}(\tilde{r}(y, x)) \left| \frac{\partial \tilde{r}(y, x)}{\partial y} \right| = f_U(r(y, x)) \left| \frac{\partial r(y, x)}{\partial y} \right|$$

Hence, using (3.1)-(3.3), we can state that the transformation $g(U, x)$ generates the same distribution of Y given $X = x$ as U generates iff for all u

$$(3.5) \quad f_{\tilde{U}|X=x}(g(u, x)) \left| \frac{\partial g(u, x)}{\partial u} \right| = f_U(u)$$

The analogous result to Theorem 2.1 in this setup is the following

Theorem 3.1: *Suppose that Assumption 2.1 is satisfied for r and for \tilde{r} defined by (3.3). Suppose further that $f_{\tilde{U}|X=x}$ satisfies (3.5). Then $\partial f_{\tilde{U}|X=x}(\tilde{u})/\partial x = 0$ for all x and \tilde{u} iff for all u and x*

$$(T3.1) \quad \left[-\frac{\partial \log(f_U(u))}{\partial u} + \frac{\partial}{\partial u} \log \left(\left| \frac{\partial g(u, x)}{\partial u} \right| \right) \right] \left[\left(\frac{\partial g(u, x)}{\partial u} \right)^{-1} \frac{\partial g(u, x)}{\partial x} \right] \\ = \frac{\partial}{\partial x} \log \left(\left| \frac{\partial g(u, x)}{\partial u} \right| \right)$$

Example 3.1: Consider the model

$$U = r(Y, X)$$

Suppose that we wanted to study under what conditions on a function $v : R^K \rightarrow R^G$ and on f_U , it is the case that (T3.1) holds when

$$\tilde{U} = r(Y, X) + v(X)$$

Define

$$g(U, X) = U + v(X)$$

Then, for all u, x

$$\frac{\partial g(u, x)}{\partial u} = I, \quad \left| \frac{\partial g(u, x)}{\partial u} \right| = 1, \quad \text{and} \quad \frac{\partial}{\partial u} \left| \frac{\partial g(u, x)}{\partial u} \right| = \frac{\partial}{\partial x} \left| \frac{\partial g(u, x)}{\partial u} \right| = 0$$

Condition (T3.1) is then satisfied iff for all u, x

$$\left(\frac{\partial \log(f_U(u))}{\partial u} \right) \left[\frac{\partial v(x)}{\partial x} \right] = 0$$

or, equivalently, if

$$\left(\frac{\partial (f_U(r(y, x)))}{\partial u} \right) \left[\frac{\partial v(x)}{\partial x} \right] = 0$$

Suppose, for example, that $G = 2$, $K = 2$,

$$f_U(u_1, u_2) = \frac{e^{-(u_1+u_2)^2}}{\int e^{-(t_1+t_2)^2} dt_1 dt_2}$$

and

$$v(x_1, x_2) = \begin{pmatrix} x_1 + x_2 \\ -x_1 - x_2 \end{pmatrix}$$

Then, since in this case

$$\begin{aligned} & \left(\frac{\partial (f_U(r(y, x)))}{\partial u} \right) \frac{\partial v(x)}{\partial x} \\ &= \left(\frac{\partial (f_U(u_1, u_2))}{\partial u_1}, \frac{\partial (f_U(u_1, u_2))}{\partial u_2} \right) \begin{pmatrix} \frac{\partial v_1(x)}{\partial x_1} & \frac{\partial v_1(x)}{\partial x_2} \\ \frac{\partial v_2(x)}{\partial x_1} & \frac{\partial v_2(x)}{\partial x_2} \end{pmatrix} \\ &= \left[\frac{[-2+2] (u_1+u_2) e^{-(u_1+u_2)^2}}{\int e^{-(t_1+t_2)^2} dt_1 dt_2}, \frac{[-2+2] (u_1+u_2) e^{-(u_1+u_2)^2}}{\int e^{-(t_1+t_2)^2} dt_1 dt_2} \right] \\ &= (0, 0) \end{aligned}$$

the function

$$\tilde{r}(y, x) = r(y, x) + v(x)$$

is observationally equivalent to

$$r(y, x)$$

To interpret better the result in Theorem 3.1 and to express an equivalent rank condition, we perform an analysis analogous to that in Section 2.1. Consider the random variable generated by a transformation g that is invertible with respect to U :

$$\tilde{U} = g(U, X)$$

Recall from above that the transformation $\tilde{g}(U, x)$ generates the same distribution of Y given $X = x$ as U generates iff for all u

$$f_{\tilde{U}|X=x}(g(u, x)) \left| \frac{\partial g(u, x)}{\partial u} \right| = f_U(u)$$

Taking log's in both sides of (3.5) and differentiating the resulting expression first with respect to u and then with respect to x , we get

$$(3.6) \quad \frac{\partial \log f_{\tilde{U}|X=x}(g(u, x))}{\partial \tilde{u}} \frac{\partial g(u, x)}{\partial u} + \frac{\partial \log \left| \frac{\partial g(u, x)}{\partial u} \right|}{\partial u} = \frac{\partial \log f_U(u)}{\partial u}$$

and

$$(3.7) \quad \frac{\partial \log f_{\tilde{U}|X=x}(g(u, x))}{\partial \tilde{u}} \frac{\partial g(u, x)}{\partial x} + \frac{\partial \log f_{\tilde{U}|X=x}(\tilde{u})}{\partial x} \Big|_{\tilde{u}=g(u, x)} + \frac{\partial \log \left| \frac{\partial g(u, x)}{\partial u} \right|}{\partial x} = 0$$

Given f_U and g , this is a system of equations in $\partial \log f_{\tilde{U}|X=x}(g(u, x)) / \partial \tilde{u}$ and $\partial \log f_{\tilde{U}|X=x}(\tilde{u}) / \partial x \Big|_{\tilde{u}=g(u, x)}$, which can be written in matrix form as

$$\begin{bmatrix} \left(\frac{\partial g(u, x)}{\partial u} \right)' & 0 \\ \left(\frac{\partial g(u, x)}{\partial x} \right)' & I_K \end{bmatrix} \begin{bmatrix} \tilde{\lambda}_{\tilde{u}}(g(u, x), x) \\ \tilde{\lambda}_x(g(u, x), x) \end{bmatrix} = \begin{bmatrix} \frac{\partial \log f_U(u)}{\partial u} - \frac{\partial \log \left| \frac{\partial g(u, x)}{\partial u} \right|}{\partial u} \\ - \frac{\partial \log \left| \frac{\partial g(u, x)}{\partial u} \right|}{\partial x} \end{bmatrix}$$

where the ij -th element in the $G \times G$ matrix $(\partial g(u, x) / \partial u)'$ is $\partial g_j(u, x) / \partial u_i$, the ij -th element in the $K \times G$ matrix $(\partial g(u, x) / \partial x)'$ is $\partial g_j(u, x) / \partial x_i$,

$$\tilde{\lambda}_{\tilde{u}}(g(u, x), x) = \left(\frac{\partial \log f_{\tilde{U}|X=x}(g(u, x))}{\partial \tilde{u}} \right)'$$

and

$$\tilde{\lambda}_x(g(u, x), x) = \left(\frac{\partial \log f_{\tilde{U}|X=x}(\tilde{u})}{\partial x} \Big|_{\tilde{u}=g(u, x)} \right)'$$

Using (3.6) to solve for $\tilde{\lambda}_{\tilde{u}}(g(u, x), x)$ and substituting into (3.7), we get that

$$\begin{aligned} & \tilde{\lambda}_x(g(u, x), x) \\ = & \left[-\frac{\partial \log(f_U(u))}{\partial u} + \frac{\partial}{\partial u} \log \left(\left| \frac{\partial \tilde{g}(u, x)}{\partial u} \right| \right) \right] \left[\left(\frac{\partial \tilde{g}(u, x)}{\partial u} \right)^{-1} \frac{\partial \tilde{g}(u, x)}{\partial x} \right] \\ & - \frac{\partial}{\partial x} \log \left(\left| \frac{\partial \tilde{g}(u, x)}{\partial u} \right| \right) \end{aligned}$$

Condition (T3.1) is satisfied when this expression equal zero. An equivalent condition, in terms of the rank of a matrix is given in the next theorem:

Theorem 3.2: *Suppose that Assumption 2.1 is satisfied for r and for \tilde{r} defined by (3.3). Define the transformation g by (3.1). Then, $g(r(y, x), x)$ is observationally equivalent to $r(y, x)$ if and only if for all u, x , the rank of the matrix*

$$\begin{pmatrix} \left(\frac{\partial g(u, x)}{\partial u} \right)' & \frac{\partial \log f_U(u)}{\partial u} - \frac{\partial \log \left| \frac{\partial g(u, x)}{\partial u} \right|}{\partial u} \\ \left(\frac{\partial g(u, x)}{\partial x} \right)' & - \frac{\partial \log \left| \frac{\partial g(u, x)}{\partial u} \right|}{\partial x} \end{pmatrix}$$

is G .

Example 3.2: Suppose that

$$\tilde{U} = g(U)$$

where g is invertible. Then, $g(u)$ is observationally equivalent to U .

To see this, note that $\partial g(u, x)/\partial x = 0$ and $\partial \log |\partial g(u, x)/\partial u|/\partial x = 0$. Since in this case the determinant in the statement of Theorem 3.2 is 0, $g(u)$ is observationally equivalent to u .

Example 3.3: (Observational equivalence of monotone transformations in single equation models)

Suppose that $G = 1$ and g is strictly increasing in U . Then, if Assumption 2.1 is satisfied, $g(U, X)$ is observationally equivalent to U if and only if for all u, x

$$\frac{\partial g(u, x)}{\partial x} = 0$$

That $\partial g(u, x)/\partial x = 0$ implies observational equivalence was already argued in Example 3.2. To show the converse, note that by Theorem 3.2, $\tilde{U} = g(U, X)$ is observationally equivalent to $U = r(Y, X)$ iff for all u, x , the determinant of

$$\begin{pmatrix} \left(\frac{\partial g(u, x)}{\partial u} \right) & \frac{\partial \log f_U(u)}{\partial u} - \frac{\partial \log \left(\frac{\partial g(u, x)}{\partial u} \right)}{\partial u} \\ \left(\frac{\partial g(u, x)}{\partial x} \right) & - \frac{\partial \log \left(\frac{\partial g(u, x)}{\partial u} \right)}{\partial x} \end{pmatrix}$$

is zero, or

$$\begin{aligned} & - \frac{\partial g(u, x)}{\partial u} \frac{\partial \log \left(\frac{\partial g(u, x)}{\partial u} \right)}{\partial x} - \left(\frac{\partial g(u, x)}{\partial x} \right) \frac{\partial \log f_U(u)}{\partial u} + \left(\frac{\partial g(u, x)}{\partial x} \right) \frac{\partial \log \left(\frac{\partial g(u, x)}{\partial u} \right)}{\partial u} \\ &= - \frac{\partial^2 g(u, x)}{\partial u \partial x} - \frac{\partial f_U(u)}{\partial u} \left(\frac{\partial g(u, x)}{\partial x} \right) + \frac{\partial^2 g(u, x)}{\partial u^2} \left(\frac{\partial g(u, x)}{\partial x} \right) \\ &= 0 \end{aligned}$$

Since for all u, x , $f_U(u) > 0$ and $\partial g(u, x)/\partial u > 0$, this is equivalent to requiring that for all u, x

$$\frac{\partial}{\partial u} \left(f_U(u) \frac{\left(\frac{\partial g(u, x)}{\partial x} \right)}{\left(\frac{\partial g(u, x)}{\partial u} \right)} \right) = 0$$

or that the function

$$m(u, x) \equiv \left(f_U(u) \frac{\left(\frac{\partial g(u, x)}{\partial x} \right)}{\left(\frac{\partial g(u, x)}{\partial u} \right)} \right)$$

is constant in u . Since $\partial g(u, x)/\partial u > 0$ and $\lim_{u \rightarrow \infty} f_U(u) = 0$, it must then be that for all u, x , $m(u, x) = 0$. Hence, since $\partial g(u, x)/\partial u > 0$, it must be that either $f_U(u) = 0$ or $\partial g(u, x)/\partial x = 0$. Since $f_U(u) = 0$ violates Assumption 2.1, it must be that $\partial g(x, u)/\partial x = 0$.

4. Connection with previous results

The main difference between our results and those derived in B. Brown (1983) and Roehrig (1988) is that according to our results \tilde{r} is observationally equivalent to r iff for all y, x ,

$$(4.1) \quad \gamma(y, x; f_U, r) \quad A(y, x; \partial r, \partial \tilde{r}) = b(y, x; \partial r, \partial \tilde{r}, \partial^2 r, \partial^2 \tilde{r})$$

where

$$\begin{aligned} A(y, x; \partial r, \partial \tilde{r}) &= \left[\frac{\partial r(y, x)}{\partial x} - \frac{\partial r(y, x)}{\partial y} \left(\frac{\partial \tilde{r}(y, x)}{\partial y} \right)^{-1} \frac{\partial \tilde{r}(y, x)}{\partial x} \right] \\ b(y, x; \partial r, \partial \tilde{r}, \partial^2 r, \partial^2 \tilde{r}) &= - \left[\left(\frac{\partial}{\partial x} \log \left(\left| \frac{\partial r(y, x)}{\partial y} \right| \right) \right) - \frac{\partial}{\partial x} \log \left(\left| \frac{\partial \tilde{r}(y, x)}{\partial y} \right| \right) \right] \\ &+ \left[\left(\frac{\partial}{\partial y} \log \left(\left| \frac{\partial r(y, x)}{\partial y} \right| \right) \right) - \frac{\partial}{\partial y} \log \left(\left| \frac{\partial \tilde{r}(y, x)}{\partial y} \right| \right) \right] \left(\frac{\partial \tilde{r}(y, x)}{\partial y} \right)^{-1} \frac{\partial \tilde{r}(y, x)}{\partial x} \right] \end{aligned}$$

and

$$\gamma(y, x; f_U, r) = \frac{\partial \log(f_U(r(y, x)))}{\partial u}$$

and according to the results derived by Lemma 2 in B. Brown (1983), which were extended by Roehrig (1988), \tilde{r} is observationally equivalent to r iff for all y, x

$$(4.2) \quad A(y, x; \partial r, \partial \tilde{r}) = \left[\frac{\partial r(y, x)}{\partial x} - \frac{\partial r(y, x)}{\partial y} \left(\frac{\partial \tilde{r}(y, x)}{\partial y} \right)^{-1} \frac{\partial \tilde{r}(y, x)}{\partial x} \right] = 0$$

The arguments in Benkard and Berry (2004) demonstrated that (4.2) is not a necessary condition for \tilde{U} to be independent of X . The condition is sufficient, as they noted, because it implies that \tilde{U} is a function of only U and not X , and U is independent of X . Theorem 2.1 provides further evidence for Benkard and Berry's arguments. Since $A(y, x; \partial r, \partial \tilde{r}) = 0$ implies that $b(y, x; \partial r, \partial \tilde{r}, \partial^2 r, \partial^2 \tilde{r}) = 0$, (4.2) implies (4.1). However, it is not always true that if for some (y, x)

$$(4.3) \quad \left[\frac{\partial r(y, x)}{\partial x} - \frac{\partial r(y, x)}{\partial y} \left(\frac{\partial \tilde{r}(y, x)}{\partial y} \right)^{-1} \frac{\partial \tilde{r}(y, x)}{\partial x} \right] \neq 0$$

one has that (4.1) is not satisfied, as the following example demonstrates.

Example 4.1 (the "rotation" problem described in Benkard and Berry (2004)): Consider again Example 3.1 with

$$\begin{aligned} U_1 &= r_1(Y_1, Y_2, X_1) \\ U_2 &= r_2(Y_1, Y_2, X_2) \end{aligned}$$

$$\begin{aligned} \tilde{U}_1 &= r_1(Y_1, Y_2, X_1) + X_1 + X_2 \\ \tilde{U}_2 &= r_2(Y_1, Y_2, X_1) - X_1 - X_2 \end{aligned}$$

and

$$f_U(u_1, u_2) = \frac{e^{-(u_1+u_2)^2}}{\int e^{-(t_1+t_2)^2} dt_1 dt_2}$$

Then, since for all y, x

$$\left(\frac{\partial \tilde{r}(y, x)}{\partial y} \right) = \left(\frac{\partial r(y, x)}{\partial y} \right), \text{ and}$$

$$\left(\frac{\partial \tilde{r}(y, x)}{\partial x} \right) = \left(\frac{\partial r(y, x)}{\partial x} \right) + \begin{pmatrix} 1 & 1 \\ -1 & -1 \end{pmatrix}$$

it follows that

$$\begin{aligned} A(y, x; \partial r, \partial \tilde{r}) &= \left[\frac{\partial r(y, x)}{\partial x} - \frac{\partial \tilde{r}(y, x)}{\partial x} \right] \\ &= \begin{pmatrix} 1 & 1 \\ -1 & -1 \end{pmatrix} \\ &\neq 0 \end{aligned}$$

However, as it was shown in Example 3.1, \tilde{r} is observationally equivalent to r .

Since the rank conditions of Brown and Roehrig may be easier to verify than our new conditions, it would be useful to know under what situations (4.3) implies that at some \tilde{y}, \tilde{x}

$$(4.4) \quad \gamma(\tilde{y}, \tilde{x}; f_U, r) \quad A(\tilde{y}, \tilde{x}; \partial r, \partial \tilde{r}) \neq b(\tilde{y}, \tilde{x}; \partial r, \partial \tilde{r}, \partial^2 r, \partial^2 \tilde{r})$$

For this, we first note how (4.2) is related to the rank conditions developed by B. Brown and Roehrig. Recall that a_{ij} denotes the element in the i -th row and j -th column of $A(y, x; \partial r, \partial \tilde{r})$. The next Lemma establishes a rank condition that is equivalent to $a_{ij} \neq 0$.

Lemma 4.1: The condition $a_{ij}(y, x) \neq 0$ is equivalent to the condition that the rank of the matrix

$$\begin{pmatrix} \frac{\partial r_i(y, x)}{\partial y} & \frac{\partial r_i(y, x)}{\partial x_j} \\ \frac{\partial \tilde{r}(y, x)}{\partial y} & \frac{\partial \tilde{r}(y, x)}{\partial x_j} \end{pmatrix}$$

is $G + 1$

Consider the following condition:

Condition 4.1: *There exists (y, x) and j such that the rank of the matrix*

$$\begin{pmatrix} \frac{\partial r_i(y, x)}{\partial y} & \frac{\partial r_i(y, x)}{\partial x} \\ \frac{\partial \tilde{r}(y, x)}{\partial y} & \frac{\partial \tilde{r}(y, x)}{\partial x} \end{pmatrix}$$

is $G + 1$, $f_U(r(y, x)) > 0$, $\partial f_U(r(y, x))/\partial u_i \neq 0$, $\partial f_U(r(y, x))/\partial u_k = 0$ for all $k \neq i$, and $b_j(y, x; \partial r, \partial \tilde{r}, \partial^2 r, \partial^2 \tilde{r}) = 0$.

Under this condition,

$$\begin{aligned} \gamma(y, x; f_U, r) \quad A_j(y, x; \partial r, \partial \tilde{r}) &= \left(\frac{\partial f_U(r(y, x))/\partial u_i}{f_U(r(y, x))} \right) a_{ij}(y, x; \partial r, \partial \tilde{r}) \\ &\neq 0 \\ &= b_j(y, x; \partial r, \partial \tilde{r}, \partial^2 r, \partial^2 \tilde{r}) \end{aligned}$$

Hence, (4.4) is satisfied.

Alternatively, we may consider the following condition:

Condition 4.2: *There exist (y, x) , (\tilde{y}, \tilde{x}) , and j such that either (1.i)*

$$\gamma(y, x; f_U, r) A_j(y, x; \partial r, \partial \tilde{r}) \neq b_j(y, x; \partial r, \partial \tilde{r}, \partial^2 r, \partial^2 \tilde{r})$$

or (1.ii)

$$A_j(y, x; \partial r, \partial \tilde{r}) = A_j(\tilde{y}, \tilde{x}; \partial r, \partial \tilde{r}) \neq 0,$$

$$b_j(y, x; \partial r, \partial \tilde{r}, \partial^2 r, \partial^2 \tilde{r}) = b_j(\tilde{y}, \tilde{x}; \partial r, \partial \tilde{r}, \partial^2 r, \partial^2 \tilde{r})$$

$$\gamma_g(y, x; f_U, r) < \gamma_g(\tilde{y}, \tilde{x}; f_U, r) \quad \text{for all } g \text{ such that } a_{gj}(y, x; \partial r, \partial \tilde{r}) \geq 0,$$

$$\gamma_g(y, x; f_U, r) > \gamma_g(\tilde{y}, \tilde{x}; f_U, r) \quad \text{for all } g \text{ such that } a_{gj}(y, x; \partial r, \partial \tilde{r}) < 0,$$

This condition on two points, (y, x) and (\tilde{y}, \tilde{x}) , requires that the j -th column of A and the j -th coordinate of b attain the same values at (y, x) and (\tilde{y}, \tilde{x}) , and that the values of γ be different at these two points. Since A and b depend on the first and second order derivatives of r and \tilde{r} , while γ depends on the value and derivatives of the density f_U and on the value of r , this condition may be easy to establish when the values of γ have enough variation.

When Condition 4.2 is satisfied, \tilde{r} is not observationally equivalent to r . Since, if (1.i) is satisfied, then (T2.1) is not satisfied at (y, x) . If, on the other hand, (1.ii) is not satisfied, so that

$$\gamma(y, x; f_U, r) A_j(y, x; \partial r, \partial \tilde{r}) = b_j(y, x; \partial r, \partial \tilde{r}, \partial^2 r, \partial^2 \tilde{r})$$

then, (1.ii) implies that

$$\begin{aligned} \gamma(\tilde{y}, \tilde{x}; f_U, r) A_j(\tilde{y}, \tilde{x}; \partial r, \partial \tilde{r}) &= \sum_{g=1}^G \gamma_g(\tilde{y}, \tilde{x}; f_U, r) A_{gj}(\tilde{y}, \tilde{x}; \partial r, \partial \tilde{r}) \\ &= \sum_{g=1}^G \gamma_g(\tilde{y}, \tilde{x}; f_U, r) A_{gj}(y, x; \partial r, \partial \tilde{r}) \\ &> \sum_{g=1}^G \gamma_g(y, x; f_U, r) A_{gj}(y, x; \partial r, \partial \tilde{r}) \\ &= b_j(y, x; \partial r, \partial \tilde{r}, \partial^2 r, \partial^2 \tilde{r}) \\ &= b_j(\tilde{y}, \tilde{x}; \partial r, \partial \tilde{r}, \partial^2 r, \partial^2 \tilde{r}) \end{aligned}$$

Hence, (4.4) is satisfied at (\tilde{y}, \tilde{x}) .

A condition that is stronger than Condition 4.2, but which may be easier to verify is the following:

Condition 4.3: For some (y, x) , and some j ($j = 1, \dots, K$), there exist a set $\Theta \times \Xi$ such that

(i) for all $(\tilde{y}, \tilde{x}) \in \Theta \times \Xi$,

$$A_j(y, x; \partial r, \partial \tilde{r}) = A_j(\tilde{y}, \tilde{x}; \partial r, \partial \tilde{r}) \neq 0 \quad \&$$

$$b_j(y, x; \partial r, \partial \tilde{r}, \partial^2 r, \partial^2 \tilde{r}) = b_j(\tilde{y}, \tilde{x}; \partial r, \partial \tilde{r}, \partial^2 r, \partial^2 \tilde{r})$$

and

(ii) for $t^* = \partial \log(f_U(r(y, x)))/\partial u$ and some $\delta > 0$,

$$N(t^*; \delta) \subset \{w \in R^G \mid \text{for some } (\tilde{y}, \tilde{x}) \in \Theta \times \Xi, \partial \log(f_U(r(\tilde{y}, \tilde{x}))/\partial u = w)\}.$$

For requirement (i) to be satisfied, we need to find a set of vectors $\Theta \times \Xi$ on which the values of A_j and b_j are constant, and $A_j \neq 0$. For requirement (ii) to be satisfied, we need to establish that there is sufficient variation on the value of the function r across those vectors, and that, in turn, such variation in r generates sufficient variation in $\partial \log(f_U(r(y, x)))/\partial u$. Together, these imply that there exists (\tilde{y}, \tilde{x}) satisfying Condition 4.2. When either Condition 4.2 or Condition 4.3 is satisfied, \tilde{r} is not observationally equivalent to r .

The following Lemma connects Conditions 4.1, 4.2 and 4.3 to Theorem 2.1.

Lemma 4.2: Suppose that r and \tilde{r} satisfy Assumption 2.1, and that either Condition 4.1, 4.2 or 4.3 is satisfied, then \tilde{r} is not observationally equivalent to r .

Example 4.2: Suppose that r and \tilde{r} are specified to be linear:

$$U = CY + BX \quad \text{and} \quad \tilde{U} = \tilde{C}Y + \tilde{B}X$$

For each i , Let c_i and b_i denote the i -th row of C and B , respectively. Then, \tilde{r} is not observationally equivalent to r if for some i , the rank of the matrix

$$\begin{pmatrix} c_i & b_i \\ \tilde{C} & \tilde{B} \end{pmatrix}$$

is $G + 1$ and for some y, x and all $k \neq i$, $f_U(Cy + Bx) \neq 0$, $\partial f_U(Cy + Bx)/\partial u_k = 0$ and $\partial f_U(Cy + Bx)/\partial u_i \neq 0$.

To verify this, note that by the linearity of r and \tilde{r} , for all y, x , $b(y, x) = 0$, since

$$\frac{\partial \left| \frac{\partial r(y, x)}{\partial y} \right|}{\partial x} = \frac{\partial \left| \frac{\partial \tilde{r}(y, x)}{\partial y} \right|}{\partial x} = \frac{\partial \left| \frac{\partial r(y, x)}{\partial y} \right|}{\partial y} = \frac{\partial \left| \frac{\partial \tilde{r}(y, x)}{\partial y} \right|}{\partial y} = 0.$$

Hence, Condition 4.1 is satisfied. By Lemma 4.2, \tilde{r} is not observationally equivalent to r .

Example 4.4: Suppose that r and \tilde{r} are specified as:

$$U = m(Y, Z) + BX \quad \text{and} \quad \tilde{U} = \tilde{m}(Y, Z) + \tilde{B}X$$

where $Z \in R^{K_1}$ and $X \in R^{K_2}$ are exogenous. (K_1 may be 0.) Suppose further that f_U , m , and B are such that for some y, z , the range of the function $\partial \log(f_U(r(y, z, \cdot)))/\partial u : R^K \rightarrow R^G$ contains an open neighborhood. Let m_i and b_i denote, respectively, the i -th coordinate of the function m and the i -th row of B . Then, \tilde{r} is not observationally equivalent to r if for some i and y, z , the rank of the matrix

$$\begin{pmatrix} \frac{\partial m_i(y, z)}{\partial y} & \frac{\partial m_i(y, z)}{\partial z} & b_i \\ \frac{\partial \tilde{m}(y, z)}{\partial y} & \frac{\partial \tilde{m}(y, z)}{\partial z} & \tilde{b} \end{pmatrix}$$

is $G + 1$.

To verify the result, note that the structures of r and \tilde{r} imply that for all y, z, x

$$A(y, z, x) = \left[\left(\frac{\partial m(y, z)}{\partial z} \quad B \right) - \frac{\partial m(y, z)}{\partial y} \left(\frac{\partial \tilde{m}(y, z)}{\partial y} \right)^{-1} \left(\frac{\partial \tilde{m}(y, z)}{\partial z} \quad \tilde{B} \right) \right]$$

and

$$\begin{aligned} b(y, z, x) = & - \left[\left(\frac{\partial}{\partial z} \log \left| \frac{\partial m(y, z)}{\partial y} \right| - \frac{\partial}{\partial z} \log \left(\left| \frac{\partial \tilde{m}(y, z)}{\partial y} \right| \right) \right) \right] \\ & + \left[\left(\frac{\partial}{\partial y} \log \left| \frac{\partial m(y, z)}{\partial y} \right| - \frac{\partial}{\partial y} \log \left(\left| \frac{\partial \tilde{m}(y, z)}{\partial y} \right| \right) \right) \left(\frac{\partial \tilde{m}(y, z)}{\partial y} \right)^{-1} \left(\frac{\partial \tilde{m}(y, z)}{\partial z} \quad \tilde{B} \right) \right] \end{aligned}$$

Let y^*, z^* be given. Since for all x, x' , $A(y^*, z^*, x) = A(y^*, z^*, x')$ and $b(y^*, z^*, x) = b(y^*, z^*, x')$, we can let the set $\Xi \times \Theta$ be $\{(y^*, z^*, x) \mid x \in R^K\}$. This and the assumption on the density imply that Condition 4.3 is satisfied. Hence, the result follows by Lemma 4.2.

Example 4.4: Suppose that r and \tilde{r} are homogenous of degree one functions. Let

$$U = r(Y, X) \quad \text{and} \quad \tilde{U} = \tilde{r}(Y, X)$$

Suppose that for some (y, x) , some g and some j , the rank of the matrix

$$\begin{pmatrix} \frac{\partial r_g(y, x)}{\partial y} & \frac{\partial r_g(y, x)}{\partial x_j} \\ \frac{\partial \tilde{r}(y, x)}{\partial y} & \frac{\partial \tilde{r}(y, x)}{\partial x_j} \end{pmatrix}$$

is $G + 1$, and for some λ

$$\frac{\frac{\partial f_U(r(y, x))}{\partial u_g}}{f_U(r(y, x))} < \lambda \frac{\frac{\partial f_U(r(\lambda y, \lambda x))}{\partial u_g}}{f_U(r(\lambda y, \lambda x))} \quad \text{for all } g \text{ such that } a_{gj}(y, x; \partial r, \partial \tilde{r}) \geq 0, \text{ and}$$

$$\frac{\frac{\partial f_U(r(y, x))}{\partial u_g}}{f_U(r(y, x))} > \lambda \frac{\frac{\partial f_U(r(\lambda y, \lambda x))}{\partial u_g}}{f_U(r(\lambda y, \lambda x))} \quad \text{for all } g \text{ such that } a_{gj}(y, x; \partial r, \partial \tilde{r}) < 0.$$

Then, \tilde{r} is not observationally equivalent to r .

To see this, note that the rank condition implies, by Lemma 4.1, that

$$A_j(y, x; \partial r, \partial \tilde{r}) \neq 0$$

If

$$\gamma(y, x; f_U, r) A_j(y, x; \partial r, \partial \tilde{r}) \neq b_j(y, x; \partial r, \partial \tilde{r}, \partial^2 r, \partial^2 \tilde{r})$$

then, by Theorem 2.1, \tilde{r} is not observationally equivalent to r . Otherwise, consider the set $\Theta \times \Xi = \{(\tilde{y}, \tilde{x}) \mid \text{for some } \lambda > 0, (\tilde{y}, \tilde{x}) = \lambda(y, x)\}$. Since the functions are homogenous of degree one, for any such (\tilde{y}, \tilde{x}) , we have that

$$r(\tilde{y}, \tilde{x}) = \lambda r(y, x), \quad \frac{\partial r(\tilde{y}, \tilde{x})}{\partial y} = \frac{\partial r(y, x)}{\partial y}, \quad \frac{\partial r(\tilde{y}, \tilde{x})}{\partial x} = \frac{\partial r(y, x)}{\partial x},$$

$$\frac{\partial^2 r(\tilde{y}, \tilde{x})}{\partial y^2} = \left(\frac{1}{\lambda}\right) \frac{\partial^2 r(y, x)}{\partial y^2}, \text{ and } \frac{\partial^2 r(\tilde{y}, \tilde{x})}{\partial y \partial x} = \left(\frac{1}{\lambda}\right) \frac{\partial^2 r(y, x)}{\partial y \partial x}.$$

Hence,

$$A_j(\tilde{y}, \tilde{x}; \partial r, \partial \tilde{r}) = A_j(y, x; \partial r, \partial \tilde{r}) \neq 0 \quad \&$$

$$b_j(\tilde{y}, \tilde{x}; \partial r, \partial \tilde{r}, \partial^2 r, \partial^2 \tilde{r}) = \left(\frac{1}{\lambda}\right) b_j(y, x; \partial r, \partial \tilde{r}, \partial^2 r, \partial^2 \tilde{r})$$

It follows that for some $(\tilde{y}, \tilde{x}) \in \Theta \times \Xi$ to violate (T2.1) it suffices to show that

$$\lambda \gamma(\tilde{y}, \tilde{x}; f_U, r) A_j(y, x; \partial r, \partial \tilde{r}) \neq b_j(y, x; \partial r, \partial \tilde{r}, \partial^2 r, \partial^2 \tilde{r})$$

when

$$\gamma(y, x; f_U, r) A_j(y, x; \partial r, \partial \tilde{r}) = b_j(y, x; \partial r, \partial \tilde{r}, \partial^2 r, \partial^2 \tilde{r})$$

Note that

$$\gamma(y, x; f_U, r) = \left(\frac{\frac{\partial f_U(r(y, x))}{\partial u_1}}{f_U(r(y, x))}, \dots, \frac{\frac{\partial f_U(r(y, x))}{\partial u_G}}{f_U(r(y, x))} \right)$$

and for $(\tilde{y}, \tilde{x}) = (\lambda y, \lambda x)$

$$\lambda \gamma(\tilde{y}, \tilde{x}; f_U, r) = \lambda \left(\frac{\frac{\partial f_U(r(\lambda y, \lambda x))}{\partial u_1}}{f_U(r(\lambda y, \lambda x))}, \dots, \frac{\frac{\partial f_U(r(\lambda y, \lambda x))}{\partial u_G}}{f_U(r(\lambda y, \lambda x))} \right)$$

Hence, by the arguments in Lemma 4.2, (4.4) will be satisfied if for some g and some j

$$a_{gj}(y, x; \partial r, \partial \tilde{r}) \neq 0$$

and for some λ

$$\frac{\frac{\partial f_U(r(y, x))}{\partial u_g}}{f_U(r(y, x))} < \lambda \frac{\frac{\partial f_U(r(\lambda y, \lambda x))}{\partial u_g}}{f_U(r(\lambda y, \lambda x))} \quad \text{for all } g \text{ such that } a_{gj}(y, x; \partial r, \partial \tilde{r}) \geq 0, \text{ and}$$

$$\frac{\frac{\partial f_U(r(y, x))}{\partial u_g}}{f_U(r(y, x))} > \lambda \frac{\frac{\partial f_U(r(\lambda y, \lambda x))}{\partial u_g}}{f_U(r(\lambda y, \lambda x))} \quad \text{for all } g \text{ such that } a_{gj}(y, x; \partial r, \partial \tilde{r}) < 0.$$

In such case, \tilde{r} is not observationally equivalent to r .

Instead of using a condition where γ varies while A and b stay constant, we could also consider conditions where b varies while γ and A stay constant, or conditions where A varies, while γ and b stay constant. For example, we may consider the following condition

Condition 4.4: *There exist (y, x) , (\tilde{y}, \tilde{x}) and j such that either (i)*

$$\gamma(y, x; f_U, r) A_j(y, x; \partial r, \partial \tilde{r}) \neq b_j(y, x; \partial r, \partial \tilde{r}, \partial^2 r, \partial^2 \tilde{r})$$

or (ii)

$$A_j(y, x; \partial r, \partial \tilde{r}) = A_j(\tilde{y}, \tilde{x}; \partial r, \partial \tilde{r}),$$

$$\gamma(y, x; f_U, r) = \gamma(\tilde{y}, \tilde{x}; f_U, r), \text{ and}$$

$$b_j(y, x; \partial r, \partial \tilde{r}, \partial^2 r, \partial^2 \tilde{r}) \neq b_j(\tilde{y}, \tilde{x}; \partial r, \partial \tilde{r}, \partial^2 r, \partial^2 \tilde{r})$$

Clearly, if Condition 3 is satisfied, then (T2.1) is violated, for either (y, x) or for (\tilde{y}, \tilde{x}) , and hence, by Theorem 1, r and \tilde{r} can not be observationally equivalent. Since b_j depends on the second order derivatives of r and \tilde{r} , while A_j and γ depend on the derivatives of r and \tilde{r} , the density f_U , and the magnitude of r , one may consider points (y, x) and (\tilde{y}, \tilde{x}) at which only the curvatures of r and \tilde{r} are different, while the derivatives r and \tilde{r} as well as the magnitude of r are the same. If the restrictions on the functions are such that the different curvatures imply that the corresponding values of b_j are different, then (ii), and then Condition 4.4 will be satisfied.

Conditions similar to 4.1-4.4 can be developed to determine the non-observational equivalence of transformations. Specifically, for any u, x , define the $G \times K$ matrix $D(u, x; \partial \tilde{g})$, the $1 \times K$ vector $d(u, x; \partial \tilde{g}, \partial^2 \tilde{g})$ and the $1 \times G$ vector $\delta(u; f_U)$ by

$$D(u, x; \partial \tilde{g}) = \left[\left(\frac{\partial \tilde{g}(u, x)}{\partial u} \right)^{-1} \frac{\partial \tilde{g}(r(y, x), x)}{\partial x} \right]$$

$$\begin{aligned} d(u, x; \partial \tilde{g}, \partial^2 \tilde{g}) &= \frac{\partial}{\partial x} \log \left(\left| \frac{\partial \tilde{g}(u, x)}{\partial u} \right| \right) \\ &\quad - \frac{\partial}{\partial u} \log \left(\left| \frac{\partial \tilde{g}(u, x)}{\partial u} \right| \right) \left[\left(\frac{\partial \tilde{g}(r(y, x), x)}{\partial u} \right)^{-1} \frac{\partial \tilde{g}(r(y, x), x)}{\partial x} \right] \end{aligned}$$

and

$$\delta(u, x; f_U) = \frac{\partial \log(f_U(u))}{\partial u}$$

Condition (T3.1) can then be stated as

$$(4.5) \quad \delta(u; f_U) D(u, x; \partial \tilde{g}) = d(u, x; \partial \tilde{g}, \partial^2 \tilde{g})$$

Conditions analogous to the ones above can be used to determine that (4.5) is not satisfied when an element of the matrix $D(u, x; \partial \tilde{g})$ is different from zero.

5. Discrete Endogenous Variables

The results obtained in the previous section can be extended to the case where the endogenous variables are generated from limited dependent variables, with some additional structure. Suppose, for example, that the model is

$$(5.1) \quad \begin{aligned} U &= r(Y^* - Z, X) \\ &= r(Y_1^* - Z_1, \dots, Y_G^* - Z_G, X) \end{aligned}$$

where U is distributed independently of (X, Z) . Suppose that conditional on (X, Z) , we only observe whether each Y_g^* is above or below 0. Letting $I(\cdot)$ denote the vector of G indicators of the values of Y^* , we define the observable variable Y by

$$Y = I(Y^* \geq 0)$$

Assuming that r is invertible, the model can be written as

$$(5.2) \quad Y_g^* - Z_g = h_g(X, U)$$

for $g = 1, \dots, G$. Hence,

$$\begin{aligned} \Pr(Y|X, Z) &= \Pr(I(Y^* \geq 0)|X, Z) \\ &= \Pr(I(h(X, U) + Z \geq 0)|X, Z) \end{aligned}$$

Define the unobservable random vector $W \in R^G$ by

$$W = h(X, U)$$

Since U is independent of (X, Z) , W is independent of Z given X . It follows that

$$\Pr(Y|X, Z) = \Pr(I(W \geq -Z)|X, Z)$$

Letting X fixed and varying Z , this relationship identifies the distribution of W given X , since, for any z, x

$$\begin{aligned} \frac{\partial \Pr(Y|X = x, Z = z)}{\partial z} &= \frac{\partial}{\partial z} \int_{-z}^{\infty} f_{W|X=x, Z=z}(w) dw \\ &= \frac{\partial}{\partial z} \int_{-z}^{\infty} f_{W|X=x}(w) dw \\ &= -f_{W|X=x}(-z) \end{aligned}$$

The variable Z acts as a “special regressor” (Lewbel (2000).) Note that in semiparametric multinomial models with no endogeneity, where the distribution of the unobservable random terms is not specified parametrically, identification requires one to normalize the values of the coefficients of some regressors. Hence, the specification (5.2) is not too restrictive. (See Matzkin (1992, 1994) and Briesch, Chintagunta, and Matzkin (1997) for the use of such regressors in nonparametric multinomial models.)

Let $f_{W|X}$ denote the density of W given X and let f_X denote the density of X . Since X is observable, the established identification of W given X implies that the joint density of W and X is identified. Since, by assumption, h is invertible in U , and

$$W = h(X, U)$$

it follows that for some invertible function s

$$U = s(W, X)$$

This is the same model as that analyzed in the previous sections with W denoting now the vector of endogenous variables. Hence, the identification of s and the distribution of U can be analyzed by the same methods as in the previous sections

6. Constructive Identification

In this section, we demonstrate how our results in the previous sections may be used to calculate the unknown functions and distributions, directly from the probability density of the observable variables. When the probability density is substituted by a nonparametric estimator of it, one may then use this to derive consistent and asymptotically normal estimators. (See Matzkin (2005) for such developments.) The methodology uses the results in the previous sections to determine normalizations and/or restrictions on r and f_U . To describe the method, we will consider a simultaneous equations model described by

$$U_g = r_g(Y, Z) + X_g$$

for $g = 1, \dots, G$, where U is assumed to be independent of $(Z, X) \in R^{L+G}$. (We allow for $L = 0$.) We impose the normalizations that at some given (y^*, z^*) and some given $\alpha \in R^G$ and $\Delta^* \in R$

$$r_g(y^*, z^*) = \alpha_g$$

and

$$\left| \frac{\partial r(y^*, z^*)}{\partial y} \right| = \Delta^*$$

To show how f_U can be recovered directly from the density of the observable variables, we note that since for all y, z, x

$$f_{Y|(X,Z)=(x,z)}(y) = f_U(r(y, z) + x) \left| \frac{\partial r(y, z)}{\partial y} \right|$$

we have that for all (t_1, \dots, t_G)

$$\begin{aligned}
f_U(t_1, \dots, t_G) &= f_U(t_1 - \alpha_1 + r_1(y^*, z^*), \dots, t_G - \alpha_G + r_G(y^*, z^*)) \\
&= f_{Y|(X,Z)=(t-\alpha, z^*)}(y^*) \left| \frac{\partial r(y^*, z^*)}{\partial y} \right|^{-1} \\
&= f_{Y|(X,Z)=(t-\alpha, z^*)}(y^*) (\Delta^*)^{-1}
\end{aligned}$$

Hence, under our normalizations, the value of the density of U at $(U_1, \dots, U_G) = (t_1, \dots, t_G)$ can be calculated from the value of the conditional density of Y given (X, Z) when $X_g = t_g - \alpha_g$ ($g = 1, \dots, G$), $Z = z^*$, and $Y = y^*$ by

$$(6.1) \quad f_U(t_1, \dots, t_G) = f_{Y|(X,Z)=(t-\alpha, z^*)}(y^*) (\Delta^*)^{-1}$$

To calculate the values of $r(y, z)$ at any given y, z , we use again our specification to note that at that particular y, z

$$(6.3) \quad f_{Y|(X,Z)=(x,z)}(y) = f_U(r(y, z) + x) \left| \frac{\partial r(y, z)}{\partial y} \right|$$

Denote the values of $r_1(y, z), \dots, r_G(y, z)$ by, respectively, $\tilde{r}_1, \dots, \tilde{r}_G$. Denote $|\partial r(y, z)/\partial y|$ by $\tilde{\Delta}$. Then, since the value of $r_1(y, z), \dots, r_G(y, z)$ and of $|\partial r(y, z)/\partial y|$ are fixed for all values of the vector (x_1, \dots, x_G) , it follows that for all vectors (x_1, \dots, x_G) the following must be satisfied by $\tilde{r}_1, \dots, \tilde{r}_G$ and $\tilde{\Delta}$:

$$(6.4) \quad f_{Y|(X,Z)=(x,z)}(y) = f_U(\tilde{r}_1 + x_1, \dots, \tilde{r}_G + x_G) \tilde{\Delta}$$

In particular, given arbitrary vectors $x^{(0)}, \dots, x^{(G)} \in R^G$, we get that for all $j = 1, \dots, G$

$$(6.5) \quad \frac{f_{Y|(X,Z)=(x^{(j)}, z)}(y)}{f_{Y|(X,Z)=(x^{(0)}, z)}(y)} = \frac{f_U(\tilde{r}_1 + x_1^{(j)}, \dots, \tilde{r}_G + x_G^{(j)})}{f_U(\tilde{r}_1 + x_1^{(0)}, \dots, \tilde{r}_G + x_G^{(0)})}$$

Replacing f_U by its expression in terms of the conditional density of Y given (X, Z) , it follows that $\tilde{r}_1, \dots, \tilde{r}_G$ and $\tilde{\Delta}$ must satisfy for $j = 0, \dots, G$

$$f_{Y|(X,Z)=(x^{(j)},z)}(y) = f_{Y|(X,Z)=(\tilde{r}_1+x_1^{(j)}-\alpha_1,\dots,\tilde{r}_G+x_G^{(j)}-\alpha_G,z^*)}(y^*) \frac{\tilde{\Delta}}{\Delta^*}$$

Hence, taking any arbitrary values for $x^{(0)}, \dots, x^{(G)}$, the values, $\tilde{r}_1, \dots, \tilde{r}_G$ for $r_1(y, z), \dots, r_G(y, z)$ solve

$$(6.6) \quad \frac{f_{Y|(X,Z)=(x^{(j)},z)}(y)}{f_{Y|(X,Z)=(x^{(0)},z)}(y)} = \frac{f_{Y|(X,Z)=(\tilde{r}_1+x_1^{(j)}-\alpha_1,\dots,\tilde{r}_G+x_G^{(j)}-\alpha_G,z^*)}(y^*)}{f_{Y|(X,Z)=(\tilde{r}_1+x_1^{(0)}-\alpha_1,\dots,\tilde{r}_G+x_G^{(0)}-\alpha_G,z^*)}(y^*)}$$

Equation (6.6) is a system of G equations in G unknowns, which can be solved using standard algorithms for solutions of equations. Matzkin (2005) shows that when the conditional densities of Y given X , in (6.1) and (6.6), are substituted by kernel estimators, and additional restrictions for identification are satisfied, the values of $\tilde{r}_1, \dots, \tilde{r}_G$ that solve (6.6) subject to those restrictions can be consistent and asymptotically normally distributed estimators for the true values of $\tilde{r}_1, \dots, \tilde{r}_G$.

7. Conclusions and extensions

We have revisited the question studied by B. Brown (1983), Roehrig (1988), and Benkard and Berry (2004) regarding the conditions that independence imposes on otherwise observable equivalent functions. We have developed new sets of conditions, and have analyzed their relationship to the rank conditions studied by Brown and Roehrig. We have demonstrated how the conditions can be used to derive set of functions that are not observationally equivalent to the true function, and have extended the analysis to models with discrete endogenous variables. We have shown how estimation of the nonparametric structural function and distribution may proceed.

The results can be extended in several directions, using very similar analyses to the one presented in the previous sections. First, it is easy to develop analogous conditions for observational equivalence when some or all of the coordinates of X are discrete. The main difference is that instead of imposing the restriction that for all \tilde{u} and x

$$\frac{\partial}{\partial x} f_{U|X=x}(\tilde{u}) = 0$$

we would impose the restriction that for all \tilde{u}, x, x'

$$f_{U|X=x}(\tilde{u}) = f_{U|X=x'}(\tilde{u})$$

Second, the independence assumption between U and X can be easily relaxed. The analysis can be easily extended to the case where U and X are independent conditional on some other variable Z . Third, a similar analysis can be performed to determine the restrictions imposed by either independence, or conditional independence across the G unobservable variables, along the lines of Matzkin (2004). Fourth, the results about discrete endogenous regressors can be extended to other type of limited dependent variable models.

Appendix

Proof of Theorem 2.1: Condition (2.7) establishes that for all y, x

$$(A.1) \quad f_{\tilde{U}|X=x}(\tilde{r}(y, x)) = f_U(r(y, x)) \left| \frac{\partial r(y, x)}{\partial y} \right| \left| \frac{\partial \tilde{r}(y, x)}{\partial y} \right|^{-1}$$

Let \tilde{u} and x be given. We want to determine conditions guaranteeing that

$$(A.2) \quad \frac{df_{\tilde{U}|X=x}(\tilde{u})}{dx} = 0$$

when (A.1) is satisfied. Let y be the unique value for which

$$\tilde{u} = \tilde{r}(y, x)$$

Then,

$$y = \tilde{h}(x, \tilde{u})$$

Substituting in (A.1), we get

$$f_{\tilde{U}|X=x}(\tilde{u}) = f_U\left(r\left(\tilde{h}(x, \tilde{u}), x\right)\right) \left| \frac{\partial r(\tilde{h}(x, \tilde{u}), x)}{\partial y} \right| \left| \frac{\partial \tilde{r}(\tilde{h}(x, \tilde{u}), x)}{\partial y} \right|^{-1}$$

Taking derivatives, we get that $df_{\tilde{U}|X=x}(\tilde{u})/dx = 0$ iff

$$(A.3) \quad \begin{aligned} & \frac{\partial f_U\left(r\left(\tilde{h}(x, \tilde{u}), x\right)\right)}{\partial u} \left[\frac{\partial r\left(\tilde{h}(x, \tilde{u}), x\right)}{\partial y} \frac{\partial \tilde{h}(x, \tilde{u})}{\partial x} + \frac{\partial r\left(\tilde{h}(x, \tilde{u}), x\right)}{\partial x} \right] \\ & \cdot \left| \frac{\partial r(\tilde{h}(x, \tilde{u}), x)}{\partial y} \right| \left| \frac{\partial \tilde{r}(\tilde{h}(x, \tilde{u}), x)}{\partial y} \right|^{-1} \\ & + f_U\left(r\left(\tilde{h}(x, \tilde{u}), x\right)\right) \frac{d \left| \frac{\partial r(\tilde{h}(x, \tilde{u}), x)}{\partial y} \right|}{dx} \left| \frac{\partial \tilde{r}(\tilde{h}(x, \tilde{u}), x)}{\partial y} \right|^{-1} \\ & - f_U\left(r\left(\tilde{h}(x, \tilde{u}), x\right)\right) \left| \frac{\partial r(\tilde{h}(x, \tilde{u}), x)}{\partial y} \right| \frac{d \left| \frac{\partial \tilde{r}(\tilde{h}(x, \tilde{u}), x)}{\partial y} \right|}{dx} \left| \frac{\partial \tilde{r}(\tilde{h}(x, \tilde{u}), x)}{\partial y} \right|^{-2} \\ & = 0 \end{aligned}$$

Differentiating with respect to x the relationship

$$\tilde{u} = \tilde{r}\left(\tilde{h}(x, \tilde{u}), x\right)$$

we get

$$0 = \frac{\partial \tilde{r}\left(\tilde{h}(x, \tilde{u}), x\right)}{\partial y} \frac{\partial \tilde{h}(x, \tilde{u})}{\partial x} + \frac{\partial \tilde{r}\left(\tilde{h}(x, \tilde{u}), x\right)}{\partial x}$$

Hence

$$\frac{\partial \tilde{h}(x, \tilde{u})}{\partial x} = - \left(\frac{\partial \tilde{r}(\tilde{h}(x, \tilde{u}), x)}{\partial y} \right)^{-1} \frac{\partial \tilde{r}(\tilde{h}(x, \tilde{u}), x)}{\partial x}$$

For each g, j , ($g = 1, \dots, G$; $j = 1, \dots, K$) let $D_{g,j}$ denotes the matrix $\left[\frac{\partial r(y,x)}{\partial y} \right]$, evaluated at $y = \tilde{h}(x, \tilde{u})$, and with the g -th row replaced by

$$\left(\frac{\partial r_g(y,x)}{\partial y_1 \partial x_j} \quad \frac{\partial r_g(y,x)}{\partial y_2 \partial x_j} \quad \dots \quad \frac{\partial r_g(y,x)}{\partial y_G \partial x_j} \right)$$

For each g, i ($g = 1, \dots, G$; $i = 1, \dots, G$) let $B_{g,i}$ denotes the matrix $\left[\frac{\partial r(y,x)}{\partial y} \right]$, evaluated at $y = \tilde{h}(x, \tilde{u})$, and with the g -th row replaced by

$$\left(\frac{\partial r_g(y,x)}{\partial y_1 \partial y_i} \quad \frac{\partial r_g(y,x)}{\partial y_2 \partial y_i} \quad \dots \quad \frac{\partial r_g(y,x)}{\partial y_G \partial y_i} \right)$$

It can be shown that:

$$\frac{d \left| \frac{\partial r(\tilde{h}(x, \tilde{u}), x)}{\partial y} \right|}{dx} = \frac{\partial}{\partial x} \left| \frac{\partial r(\tilde{h}(x, \tilde{u}), x)}{\partial y} \right| + \frac{\partial}{\partial y} \left| \frac{\partial r(\tilde{h}(x, \tilde{u}), x)}{\partial y} \right| \frac{\partial \tilde{h}(x, \tilde{u})}{\partial x}$$

and similarly

$$\frac{d \left| \frac{\partial \tilde{r}(\tilde{h}(x, \tilde{u}), x)}{\partial y} \right|}{dx} = \frac{\partial}{\partial x} \left| \frac{\partial \tilde{r}(\tilde{h}(x, \tilde{u}), x)}{\partial y} \right| + \frac{\partial}{\partial y} \left| \frac{\partial \tilde{r}(\tilde{h}(x, \tilde{u}), x)}{\partial y} \right| \frac{\partial \tilde{h}(x, \tilde{u})}{\partial x}$$

where, for each j

$$\frac{\partial}{\partial x_j} \left| \frac{\partial r(\tilde{h}(x, \tilde{u}), x)}{\partial y} \right| = \sum_{g=1}^G |D_{g,j}|$$

and where the i -th element of the $1 \times G$ vector $\frac{\partial}{\partial y} \left| \frac{\partial r(\tilde{h}(x, \tilde{u}), x)}{\partial y} \right|$ is

$$\frac{\partial}{\partial y_i} \left| \frac{\partial r(\tilde{h}(x, \tilde{u}), x)}{\partial y} \right| = \sum_{g=1}^G |B_{g,i}|$$

Substituting in (A.3), we get that \tilde{U} is independent of X iff

$$\begin{aligned}
& \frac{\partial f_U \left(r \left(\tilde{h}(x, \tilde{u}), x \right) \right)}{\partial u} C(x, \tilde{u}) \cdot \left| \frac{\partial r(\tilde{h}(x, \tilde{u}), x)}{\partial y} \right| \left| \frac{\partial \tilde{r}(\tilde{h}(x, \tilde{u}), x)}{\partial y} \right|^{-1} \\
& + f_U \left(r \left(\tilde{h}(x, \tilde{u}), x \right) \right) \left[\frac{\partial}{\partial x} \left| \frac{\partial r(\tilde{h}(x, \tilde{u}), x)}{\partial y} \right| \right] \left| \frac{\partial \tilde{r}(\tilde{h}(x, \tilde{u}), x)}{\partial y} \right|^{-1} \\
& - f_U \left(r \left(\tilde{h}(x, \tilde{u}), x \right) \right) \frac{\partial}{\partial y} \left| \frac{\partial r(\tilde{h}(x, \tilde{u}), x)}{\partial y} \right| \left(\frac{\partial \tilde{r}(\tilde{h}(x, \tilde{u}), x)}{\partial y} \right)^{-1} \frac{\partial \tilde{r}(\tilde{h}(x, \tilde{u}), x)}{\partial x} \\
& \cdot \left| \frac{\partial \tilde{r}(\tilde{h}(x, \tilde{u}), x)}{\partial y} \right|^{-1} \\
& - f_U \left(r \left(\tilde{h}(x, \tilde{u}), x \right) \right) \left[\frac{\partial}{\partial x} \left| \frac{\partial \tilde{r}(\tilde{h}(x, \tilde{u}), x)}{\partial y} \right| \right] \left| \frac{\partial r(\tilde{h}(x, \tilde{u}), x)}{\partial y} \right| \left| \frac{\partial \tilde{r}(\tilde{h}(x, \tilde{u}), x)}{\partial y} \right|^{-2} \\
& + f_U \left(r \left(\tilde{h}(x, \tilde{u}), x \right) \right) \frac{\partial}{\partial y} \left| \frac{\partial \tilde{r}(\tilde{h}(x, \tilde{u}), x)}{\partial y} \right| \left(\frac{\partial \tilde{r}(\tilde{h}(x, \tilde{u}), x)}{\partial y} \right)^{-1} \frac{\partial \tilde{r}(\tilde{h}(x, \tilde{u}), x)}{\partial x} \\
& \cdot \left| \frac{\partial r(\tilde{h}(x, \tilde{u}), x)}{\partial y} \right| \left| \frac{\partial \tilde{r}(\tilde{h}(x, \tilde{u}), x)}{\partial y} \right|^{-2} \\
& = 0
\end{aligned}$$

where

$$C(x, \tilde{u}) = \left[\frac{\partial r(\tilde{h}(x, \tilde{u}), x)}{\partial x} - \frac{\partial r(\tilde{h}(x, \tilde{u}), x)}{\partial y} \left(\frac{\partial \tilde{r}(\tilde{h}(x, \tilde{u}), x)}{\partial y} \right)^{-1} \frac{\partial \tilde{r}(\tilde{h}(x, \tilde{u}), x)}{\partial x} \right]$$

Multiplying by $\left| \frac{\partial \tilde{r}(\tilde{h}(x, \tilde{u}), x)}{\partial y} \right|$ and dividing by $f_U \left(r \left(\tilde{h}(x, \tilde{u}), x \right) \right) \left| \frac{\partial r(\tilde{h}(x, \tilde{u}), x)}{\partial y} \right|$ we get

$$\begin{aligned}
& \frac{\partial \log f_U \left(r \left(\tilde{h}(x, \tilde{u}), x \right) \right)}{\partial u} C(x, \tilde{u}) \\
& + \left[\left(\frac{\frac{\partial}{\partial x} \left| \frac{\partial r(\tilde{h}(x, \tilde{u}), x)}{\partial y} \right|}{\left| \frac{\partial r(\tilde{h}(x, \tilde{u}), x)}{\partial y} \right|} - \frac{\frac{\partial}{\partial x} \left| \frac{\partial \tilde{r}(\tilde{h}(x, \tilde{u}), x)}{\partial y} \right|}{\left| \frac{\partial \tilde{r}(\tilde{h}(x, \tilde{u}), x)}{\partial y} \right|} \right) \right] \\
& - \left[\left(\frac{\frac{\partial}{\partial y} \left| \frac{\partial r(\tilde{h}(x, \tilde{u}), x)}{\partial y} \right|}{\left| \frac{\partial r(\tilde{h}(x, \tilde{u}), x)}{\partial y} \right|} - \frac{\frac{\partial}{\partial y} \left| \frac{\partial \tilde{r}(\tilde{h}(x, \tilde{u}), x)}{\partial y} \right|}{\left| \frac{\partial \tilde{r}(\tilde{h}(x, \tilde{u}), x)}{\partial y} \right|} \right) \left(\frac{\partial \tilde{r}(\tilde{h}(x, \tilde{u}), x)}{\partial y} \right)^{-1} \frac{\partial \tilde{r}(\tilde{h}(x, \tilde{u}), x)}{\partial x} \right] \\
& = 0
\end{aligned}$$

Substituting $y = \tilde{h}(x, \tilde{u})$, we obtain that

$$\begin{aligned}
(A.4) \quad & \frac{\partial \log f_U(r(y, x))}{\partial u} \widehat{C}(y, x) \\
& + \left[\left(\frac{\frac{\partial}{\partial x} \left| \frac{\partial r(y, x)}{\partial y} \right|}{\frac{\partial r(y, x)}{\partial y}} - \frac{\frac{\partial}{\partial x} \left| \frac{\partial \tilde{r}(y, x)}{\partial y} \right|}{\left| \frac{\partial \tilde{r}(y, x)}{\partial y} \right|} \right) \right] \\
& - \left[\left(\frac{\frac{\partial}{\partial y} \left| \frac{\partial r(y, x)}{\partial y} \right|}{\left| \frac{\partial r(y, x)}{\partial y} \right|} - \frac{\frac{\partial}{\partial y} \left| \frac{\partial \tilde{r}(y, x)}{\partial y} \right|}{\left| \frac{\partial \tilde{r}(y, x)}{\partial y} \right|} \right) \left(\frac{\partial \tilde{r}(y, x)}{\partial y} \right)^{-1} \frac{\partial \tilde{r}(y, x)}{\partial x} \right] \\
& = 0
\end{aligned}$$

where

$$\widehat{C}(y, x) = \left[\frac{\partial r(y, x)}{\partial x} - \frac{\partial r(y, x)}{\partial y} \left(\frac{\partial \tilde{r}(y, x)}{\partial y} \right)^{-1} \frac{\partial \tilde{r}(y, x)}{\partial x} \right]$$

Proof of Theorem 2.2: Suppose that (T2.1) is satisfied. Define

$$\begin{aligned}
(A.5) \quad & \frac{\partial \log f_{\tilde{U}|X=x}(\tilde{r}(y, x))}{\partial \tilde{u}} \\
= & \frac{\partial \log (f_U(r(y, x)))}{\partial u} \frac{\partial r(y, x)}{\partial y} \left(\frac{\partial \tilde{r}(y, x)}{\partial y} \right)^{-1} \\
& + \left(\frac{\partial}{\partial y} \log \left| \frac{\partial r(y, x)}{\partial y} \right| \right) \left(\frac{\partial \tilde{r}(y, x)}{\partial y} \right)^{-1} \\
& - \left(\frac{\partial}{\partial y} \log \left| \frac{\partial \tilde{r}(y, x)}{\partial y} \right| \right) \left(\frac{\partial \tilde{r}(y, x)}{\partial y} \right)^{-1}
\end{aligned}$$

Then,

$$\frac{\partial}{\partial y} \log \left[\frac{f_{\tilde{U}|X=x}(\tilde{r}(y, x)) \left| \frac{\partial \tilde{r}(y, x)}{\partial y} \right|}{f_U(r(y, x)) \left| \frac{\partial r(y, x)}{\partial y} \right|} \right] = 0$$

This implies that the function

$$s(y, x) = \log \left[\frac{f_{\tilde{U}|X=x}(\tilde{r}(y, x)) \left| \frac{\partial \tilde{r}(y, x)}{\partial y} \right|}{f_U(r(y, x)) \left| \frac{\partial r(y, x)}{\partial y} \right|} \right]$$

is constant in y . Since log is 1-1, its argument must then be constant in y . Hence, for some c ,

$$f_{\tilde{U}|X=x}(\tilde{r}(y, x)) \left| \frac{\partial \tilde{r}(y, x)}{\partial y} \right| = c f_U(r(y, x)) \left| \frac{\partial r(y, x)}{\partial y} \right|$$

Let $\tilde{u}_x = \tilde{r}(y, x)$ and $u_x = r(y, x)$. Integrating with respect to \tilde{u}_x and u_x , respectively, after the corresponding change of variables, we get that

$$\int f_{\tilde{U}|X=x}(\tilde{u}_x) d\tilde{u}_x = c \int f_U(u_x) du_x$$

Hence, for any solution $f_{\tilde{U}|X=x}$ of (A.5) that is a density, $c = 1$, and (2.7) is satisfied. Substituting (A.5) into (T2.1), we get that

$$\begin{aligned} (A.6) \quad & \frac{\partial \log f_{\tilde{U}|X=x}(\tilde{r}(y, x))}{\partial \bar{u}} \frac{\partial \tilde{r}(y, x)}{\partial x} + \frac{\partial}{\partial x} \log \left| \frac{\partial \tilde{r}(y, x)}{\partial y} \right| \\ & = \frac{\partial \log (f_U(r(y, x)))}{\partial u} \frac{\partial r(y, x)}{\partial x} + \frac{\partial}{\partial x} \log \left| \frac{\partial r(y, x)}{\partial y} \right| \end{aligned}$$

Differentiating the log of (2.7) with respect to x gives,

$$\begin{aligned} (A.7) \quad & \frac{\partial \log f_{\tilde{U}|X=x}(\tilde{r}(y, x))}{\partial \bar{u}} \frac{\partial \tilde{r}(y, x)}{\partial x} + \frac{\partial}{\partial x} \log \left| \frac{\partial \tilde{r}(y, x)}{\partial y} \right| \\ & + \frac{\partial \log f_{\tilde{U}|X=x}(t)}{\partial x} \Big|_{t=\tilde{r}(y, x)} \\ & = \frac{\partial \log (f_U(r(y, x)))}{\partial u} \frac{\partial r(y, x)}{\partial x} + \frac{\partial}{\partial x} \log \left| \frac{\partial r(y, x)}{\partial y} \right| \end{aligned}$$

>From (A.6) and (A.7) it follows that

$$\frac{\partial \log f_{\tilde{U}|X=x}(t)}{\partial x} \Big|_{t=\tilde{r}(y, x)} = 0,$$

which implies that \tilde{U} is independent of X . Hence, if (T2.1), there exists \tilde{U} independent of X satisfying (2.7).

Conversely, suppose that \tilde{r} is observationally equivalent to r . Let $f_{\tilde{U}}$ denote a density such that \tilde{U} is independent of X and $(\tilde{r}, f_{\tilde{U}})$ generate the same distribution of Y given X as (r, f_U) generate. Then, for all y, x

$$(A.9) \quad f_{\tilde{U}}(\tilde{r}(y, x)) \left| \frac{\partial \tilde{r}(y, x)}{\partial y} \right| = f_U(r(y, x)) \left| \frac{\partial r(y, x)}{\partial y} \right|$$

Taking logs and differentiating this expression first with respect to y and then with respect to x , one gets that $f_{\tilde{U}}(\tilde{r}(y, x))$ satisfies

$$(A.10) \quad \frac{\partial \log f_{\tilde{U}}(\tilde{r}(y, x))}{\partial \tilde{u}} \frac{\partial \tilde{r}(y, x)}{\partial x} + \frac{\partial}{\partial x} \log \left| \frac{\partial \tilde{r}(y, x)}{\partial y} \right| \\ = \frac{\partial \log (f_U(r(y, x)))}{\partial u} \frac{\partial r(y, x)}{\partial x} + \frac{\partial}{\partial x} \log \left| \frac{\partial r(y, x)}{\partial y} \right|$$

and

$$(A.11) \quad \frac{\partial \log f_{\tilde{U}}(\tilde{r}(y, x))}{\partial \tilde{u}} \frac{\partial \tilde{r}(y, x)}{\partial x} + \frac{\partial}{\partial x} \log \left| \frac{\partial \tilde{r}(y, x)}{\partial y} \right| \\ = \frac{\partial \log (f_U(r(y, x)))}{\partial u} \frac{\partial r(y, x)}{\partial x} + \frac{\partial}{\partial x} \log \left| \frac{\partial r(y, x)}{\partial y} \right|$$

Solving for $\partial \log f_{\tilde{U}}(\tilde{r}(y, x)) / \partial \tilde{u}$ in (A.10) and substituting the result in (A.11), (T2.1) follows.

Proof of Theorem 2.3: Define, as in Section 2.1,

$$\tilde{\gamma}_{\tilde{u}}(\tilde{r}(y, x), x) = \frac{\partial \log f_{\tilde{U}|X=x}(\tilde{r}(y, x))}{\partial \tilde{u}},$$

$$\tilde{\gamma}_x(\tilde{r}(y, x), x) = \frac{\partial \log f_{\tilde{U}|X=x}(t)}{\partial x} \Big|_{t=\tilde{r}(y, x)}$$

$$\gamma(r(y, x)) = \frac{\partial \log f_U(r(y, x))}{\partial u},$$

$$\Delta_y(y, x; \partial r, \partial^2 r, \partial \tilde{r}, \partial^2 \tilde{r}) = \frac{\partial}{\partial y} \log \left| \frac{\partial r(y, x)}{\partial y} \right| - \frac{\partial}{\partial y} \log \left| \frac{\partial \tilde{r}(y, x)}{\partial y} \right|$$

$$\Delta_x(y, x; \partial r, \partial^2 r, \partial \tilde{r}, \partial^2 \tilde{r}) = \frac{\partial}{\partial x} \log \left| \frac{\partial r(y, x)}{\partial y} \right| - \frac{\partial}{\partial x} \log \left| \frac{\partial \tilde{r}(y, x)}{\partial y} \right|$$

Then, (2.10) and (2.11) can be written

$$\begin{pmatrix} \left(\frac{\partial \tilde{r}(y,x)}{\partial y}\right)' & 0 \\ \left(\frac{\partial \tilde{r}(y,x)}{\partial x}\right)' & I \end{pmatrix} \begin{pmatrix} \tilde{\gamma}_{\tilde{u}}(\tilde{r}(y,x), x) \\ \tilde{\gamma}_x(\tilde{r}(y,x), x) \end{pmatrix} = \begin{pmatrix} \Delta_y(y, x; \partial r, \partial^2 r, \partial \tilde{r}, \partial^2 \tilde{r}) - \frac{\partial \log(f_U(r(y,x)))}{\partial u} \frac{\partial r(y,x)}{\partial y} \\ \Delta_x(y, x; \partial r, \partial^2 r, \partial \tilde{r}, \partial^2 \tilde{r}) - \frac{\partial \log(f_U(r(y,x)))}{\partial u} \frac{\partial r(y,x)}{\partial x} \end{pmatrix}$$

Since $(\partial \tilde{r}(y, x) / \partial y)$ is invertible, there exists a unique solution $(\tilde{\gamma}_{\tilde{u}}(\tilde{r}(y, x), x), \tilde{\gamma}_x(\tilde{r}(y, x), x))$ to this system of equations. By Cramer's rule, the value of the j -th coordinate, $\tilde{\gamma}_{x_j}(\tilde{r}(y, x), x)$, of $\tilde{\gamma}_x(\tilde{r}(y, x), x)$, for any j , can be calculated by

$$\tilde{\gamma}_{x_j}(\tilde{r}(y, x), x) = \left| \begin{array}{c|c} \left(\frac{\partial \tilde{r}(y,x)}{\partial y}\right)' & 0_j \\ \left(\frac{\partial \tilde{r}(y,x)}{\partial x}\right)' & I_j \end{array} \right| \left[\left| \begin{array}{c|c} \left(\frac{\partial \tilde{r}(y,x)}{\partial y}\right)' & 0 \\ \left(\frac{\partial \tilde{r}(y,x)}{\partial x}\right)' & I \end{array} \right| \right]^{-1}$$

where 0_j is the zero matrix with the j -th column substituted with

$$\Delta_y(y, x; \partial r, \partial^2 r, \partial \tilde{r}, \partial^2 \tilde{r}) - (\partial \log(f_U(r(y, x))) / \partial u) (\partial r(y, x) / \partial y),$$

and where I_j is the identity matrix with its j -th column substituted with

$$\Delta_x(y, x; \partial r, \partial^2 r, \partial \tilde{r}, \partial^2 \tilde{r}) - (\partial \log(f_U(r(y, x))) / \partial u) (\partial r(y, x) / \partial x).$$

Since the term in the denominator is different from 0, it follows that $\tilde{\gamma}_{x_j}(\tilde{r}(y, x), x) = 0$ if and only if the rank of the matrix

$$\begin{pmatrix} \left(\frac{\partial \tilde{r}(y,x)}{\partial y}\right)' & \Delta_y(y, x; \partial r, \partial^2 r, \partial \tilde{r}, \partial^2 \tilde{r}) - \frac{\partial \log(f_U(r(y,x)))}{\partial u} \frac{\partial r(y,x)}{\partial y} \\ \left(\frac{\partial \tilde{r}(y,x)}{\partial x_j}\right)' & \Delta_{x_j}(y, x; \partial r, \partial^2 r, \partial \tilde{r}, \partial^2 \tilde{r}) - \frac{\partial \log(f_U(r(y,x)))}{\partial u} \frac{\partial r(y,x)}{\partial x_j} \end{pmatrix}$$

is G . Since this must hold for all j , we get that $\tilde{\gamma}_x(\tilde{r}(y, x), x) = 0$ iff the rank of the matrix

$$\begin{pmatrix} \left(\frac{\partial \tilde{r}(y,x)}{\partial y}\right)' & \Delta_y(y, x; \partial r, \partial^2 r, \partial \tilde{r}, \partial^2 \tilde{r}) - \frac{\partial \log(f_U(r(y,x)))}{\partial u} \frac{\partial r(y,x)}{\partial y} \\ \left(\frac{\partial \tilde{r}(y,x)}{\partial x}\right)' & \Delta_x(y, x; \partial r, \partial^2 r, \partial \tilde{r}, \partial^2 \tilde{r}) - \frac{\partial \log(f_U(r(y,x)))}{\partial u} \frac{\partial r(y,x)}{\partial x} \end{pmatrix}$$

is G .

Conversely, suppose that the rank is G . Then, this and the assumption that $(\partial\tilde{r}(y, x)/\partial y)$ is invertible imply that there exists a solution

$$(\partial \log f_{\tilde{U}}(\tilde{r}(y, x)) / \partial \tilde{u}, \partial \log f_{\tilde{U}}(\tilde{r}(y, x)) / \partial x)$$

satisfying

$$\begin{aligned} & \frac{\partial \log f_{\tilde{U}}(\tilde{r}(y, x))}{\partial \tilde{u}} \frac{\partial \tilde{r}(y, x)}{\partial y} + \frac{\partial}{\partial y} \log \left| \frac{\partial \tilde{r}(y, x)}{\partial y} \right| \\ = & \frac{\partial \log (f_U(r(y, x)))}{\partial u} \frac{\partial r(y, x)}{\partial y} + \frac{\partial}{\partial y} \log \left| \frac{\partial r(y, x)}{\partial y} \right| \end{aligned}$$

and

$$\begin{aligned} & \frac{\partial \log f_{\tilde{U}}(\tilde{r}(y, x))}{\partial \tilde{u}} \frac{\partial \tilde{r}(y, x)}{\partial x} + \frac{\partial}{\partial x} \log \left| \frac{\partial \tilde{r}(y, x)}{\partial y} \right| \\ = & \frac{\partial \log (f_U(r(y, x)))}{\partial u} \frac{\partial r(y, x)}{\partial x} + \frac{\partial}{\partial x} \log \left| \frac{\partial r(y, x)}{\partial y} \right| \end{aligned}$$

Using the first equation to solve for $\partial \log f_{\tilde{U}}(\tilde{r}(y, x)) / \partial \tilde{u}$ and substituting this into the second, one gets (T2.1). Hence, by Theorem 2.2, \tilde{r} is observationally equivalent to r .

Proof of Theorem 3.1: Since $\tilde{r}(y, x) = g(r(y, x), x)$,

$$\frac{\partial \tilde{r}(y, x)}{\partial y} = \frac{\partial g(r(y, x), x)}{\partial u} \frac{\partial r(y, x)}{\partial y},$$

$$\left| \frac{\partial \tilde{r}(y, x)}{\partial y} \right| = \left| \frac{\partial g(r(y, x), x)}{\partial u} \right| \left| \frac{\partial r(y, x)}{\partial y} \right|, \text{ and}$$

$$\frac{\partial \tilde{r}(y, x)}{\partial x} = \frac{\partial g(r(y, x), x)}{\partial u} \frac{\partial r(y, x)}{\partial x} + \frac{\partial g(r(y, x), x)}{\partial x}$$

Hence

$$\begin{aligned}
& \frac{\partial r(y, x)}{\partial x} - \frac{\partial r(y, x)}{\partial y} \left(\frac{\partial \tilde{r}(y, x)}{\partial y} \right)^{-1} \frac{\partial \tilde{r}(y, x)}{\partial x} \\
&= \frac{\partial r(y, x)}{\partial x} \\
& \quad - \frac{\partial r(y, x)}{\partial y} \left(\frac{\partial r(y, x)}{\partial y} \right)^{-1} \left(\frac{\partial g(r(y, x), x)}{\partial u} \right)^{-1} \left[\frac{\partial g(r(y, x), x)}{\partial u} \frac{\partial r(y, x)}{\partial x} + \frac{\partial g(r(y, x), x)}{\partial x} \right] \\
&= \frac{\partial r(y, x)}{\partial x} - \frac{\partial r(y, x)}{\partial x} - \left(\frac{\partial g(r(y, x), x)}{\partial u} \right)^{-1} \frac{\partial g(r(y, x), x)}{\partial x} \\
&= \left(\frac{\partial g(r(y, x), x)}{\partial u} \right)^{-1} \frac{\partial g(r(y, x), x)}{\partial x}
\end{aligned}$$

$$\begin{aligned}
& \left[\left(\frac{\partial}{\partial x} \log \left(\left| \frac{\partial r(y, x)}{\partial y} \right| \right) - \frac{\partial}{\partial x} \log \left(\left| \frac{\partial \tilde{r}(y, x)}{\partial y} \right| \right) \right) \right] \\
&= \frac{\frac{\partial}{\partial x} \left| \frac{\partial r(y, x)}{\partial y} \right|}{\left| \frac{\partial r(y, x)}{\partial y} \right|} - \frac{\frac{\partial}{\partial x} \left| \frac{\partial \tilde{r}(y, x)}{\partial y} \right|}{\left| \frac{\partial \tilde{r}(y, x)}{\partial y} \right|} \\
&= \frac{\frac{\partial}{\partial x} \left| \frac{\partial r(y, x)}{\partial y} \right|}{\left| \frac{\partial r(y, x)}{\partial y} \right|} - \frac{\frac{d}{dx} \left(\left| \frac{\partial g(r(y, x), x)}{\partial u} \right| \left| \frac{\partial r(y, x)}{\partial y} \right| \right)}{\left| \frac{\partial g(r(y, x), x)}{\partial u} \right| \left| \frac{\partial r(y, x)}{\partial y} \right|} \\
&= \frac{\frac{\partial}{\partial x} \left| \frac{\partial r(y, x)}{\partial y} \right|}{\left| \frac{\partial r(y, x)}{\partial y} \right|} - \frac{\frac{d}{dx} \left(\left| \frac{\partial g(r(y, x), x)}{\partial u} \right| \right)}{\left| \frac{\partial g(r(y, x), x)}{\partial u} \right|} - \frac{\frac{d}{dx} \left| \frac{\partial r(y, x)}{\partial y} \right|}{\left| \frac{\partial r(y, x)}{\partial y} \right|} \\
&= - \frac{\frac{d}{dx} \left(\left| \frac{\partial g(r(y, x), x)}{\partial u} \right| \right)}{\left| \frac{\partial g(r(y, x), x)}{\partial u} \right|} \\
&= - \frac{\frac{\partial}{\partial u} \left(\left| \frac{\partial g(r(y, x), x)}{\partial u} \right| \right) \frac{\partial r(y, x)}{\partial x}}{\left| \frac{\partial g(r(y, x), x)}{\partial u} \right|} - \frac{\frac{\partial}{\partial x} \left(\left| \frac{\partial g(r(y, x), x)}{\partial u} \right| \right)}{\left| \frac{\partial g(r(y, x), x)}{\partial u} \right|}
\end{aligned}$$

$$\begin{aligned}
& \left[\left(\frac{\partial}{\partial y} \log \left(\left| \frac{\partial r(y, x)}{\partial y} \right| \right) - \frac{\partial}{\partial y} \log \left(\left| \frac{\partial \tilde{r}(y, x)}{\partial y} \right| \right) \right) \right] \\
&= \frac{\frac{\partial}{\partial y} \left| \frac{\partial r(y, x)}{\partial y} \right|}{\left| \frac{\partial r(y, x)}{\partial y} \right|} - \frac{\frac{\partial}{\partial y} \left| \frac{\partial \tilde{r}(y, x)}{\partial y} \right|}{\left| \frac{\partial \tilde{r}(y, x)}{\partial y} \right|} \\
&= \frac{\frac{\partial}{\partial y} \left| \frac{\partial r(y, x)}{\partial y} \right|}{\left| \frac{\partial r(y, x)}{\partial y} \right|} - \frac{\frac{d}{dy} \left(\left| \frac{\partial g(r(y, x), x)}{\partial u} \right| \left| \frac{\partial r(y, x)}{\partial y} \right| \right)}{\left| \frac{\partial g(r(y, x), x)}{\partial u} \right| \left| \frac{\partial r(y, x)}{\partial y} \right|} \\
&= \frac{\frac{\partial}{\partial y} \left| \frac{\partial r(y, x)}{\partial y} \right|}{\left| \frac{\partial r(y, x)}{\partial y} \right|} - \frac{\frac{d}{dy} \left(\left| \frac{\partial g(r(y, x), x)}{\partial u} \right| \right)}{\left| \frac{\partial g(r(y, x), x)}{\partial u} \right|} - \frac{\frac{\partial}{\partial y} \left| \frac{\partial r(y, x)}{\partial y} \right|}{\left| \frac{\partial r(y, x)}{\partial y} \right|} \\
&= -\frac{\frac{\partial}{\partial u} \left(\left| \frac{\partial \tilde{g}(r(y, x), x)}{\partial u} \right| \right) \frac{\partial r(y, x)}{\partial y}}{\left| \frac{\partial \tilde{g}(r(y, x), x)}{\partial u} \right|}
\end{aligned}$$

and

$$\begin{aligned}
& \left(\frac{\partial \tilde{r}(y, x)}{\partial y} \right)^{-1} \frac{\partial \tilde{r}(y, x)}{\partial x} \\
&= \left(\frac{\partial r(y, x)}{\partial y} \right)^{-1} \left(\frac{\partial g(r(y, x), x)}{\partial u} \right)^{-1} \left[\frac{\partial g(r(y, x), x)}{\partial u} \frac{\partial r(y, x)}{\partial x} + \frac{\partial g(r(y, x), x)}{\partial x} \right] \\
&= \left(\frac{\partial r(y, x)}{\partial y} \right)^{-1} \frac{\partial r(y, x)}{\partial x} + \left(\frac{\partial r(y, x)}{\partial y} \right)^{-1} \left(\frac{\partial g(r(y, x), x)}{\partial u} \right)^{-1} \frac{\partial g(r(y, x), x)}{\partial x}
\end{aligned}$$

Substituting into condition (T2.1), we get that \tilde{U} is independent of X iff

$$\begin{aligned}
& \frac{\partial \log(f_U(r(y,x)))}{\partial u} \left[\left(\frac{\partial g(r(y,x),x)}{\partial u} \right)^{-1} \frac{\partial g(r(y,x),x)}{\partial x} \right] \\
& + \left[-\frac{\frac{\partial}{\partial u} \left(\left| \frac{\partial g(r(y,x),x)}{\partial u} \right| \right)}{\left| \frac{\partial g(r(y,x),x)}{\partial u} \right|} \frac{\partial r(y,x)}{\partial x} - \frac{\frac{\partial}{\partial x} \left(\left| \frac{\partial g(r(y,x),x)}{\partial u} \right| \right)}{\left| \frac{\partial g(r(y,x),x)}{\partial u} \right|} \right] \\
& + \left[\frac{\frac{\partial}{\partial u} \left(\left| \frac{\partial g(r(y,x),x)}{\partial u} \right| \right)}{\left| \frac{\partial g(r(y,x),x)}{\partial u} \right|} \frac{\partial r(y,x)}{\partial y} \left(\frac{\partial r(y,x)}{\partial y} \right)^{-1} \frac{\partial r(y,x)}{\partial x} \right] \\
& + \left[\frac{\frac{\partial}{\partial u} \left(\left| \frac{\partial g(r(y,x),x)}{\partial u} \right| \right)}{\left| \frac{\partial g(r(y,x),x)}{\partial u} \right|} \frac{\partial r(y,x)}{\partial y} \left[\left(\frac{\partial r(y,x)}{\partial y} \right)^{-1} \left(\frac{\partial g(r(y,x),x)}{\partial u} \right)^{-1} \frac{\partial g(r(y,x),x)}{\partial x} \right] \right] \\
& = 0
\end{aligned}$$

or, equivalently, iff

$$\begin{aligned}
& -\frac{\partial \log(f_U(r(y,x)))}{\partial u} \left[\left(\frac{\partial g(r(y,x),x)}{\partial u} \right)^{-1} \frac{\partial g(r(y,x),x)}{\partial x} \right] \\
& -\frac{\frac{\partial}{\partial u} \left(\left| \frac{\partial g(r(y,x),x)}{\partial u} \right| \right)}{\left| \frac{\partial g(r(y,x),x)}{\partial u} \right|} \frac{\partial r(y,x)}{\partial x} - \frac{\frac{\partial}{\partial x} \left(\left| \frac{\partial g(r(y,x),x)}{\partial u} \right| \right)}{\left| \frac{\partial g(r(y,x),x)}{\partial u} \right|} \\
& + \frac{\frac{\partial}{\partial u} \left(\left| \frac{\partial g(r(y,x),x)}{\partial u} \right| \right)}{\left| \frac{\partial g(r(y,x),x)}{\partial u} \right|} \left[\frac{\partial r(y,x)}{\partial x} + \left(\frac{\partial g(r(y,x),x)}{\partial u} \right)^{-1} \frac{\partial g(r(y,x),x)}{\partial x} \right] \\
& = 0
\end{aligned}$$

Hence, condition (T2.1) is satisfied iff

$$\begin{aligned}
& - \frac{\partial \log (f_U (r (y, x)))}{\partial u} \left[\left(\frac{\partial g (r (y, x), x)}{\partial u} \right)^{-1} \frac{\partial g (r (y, x), x)}{\partial x} \right] \\
& - \frac{\frac{\partial}{\partial x} \left(\left| \frac{\partial g (r (y, x), x)}{\partial u} \right| \right)}{\left| \frac{\partial g (r (y, x), x)}{\partial u} \right|} \\
& + \frac{\frac{\partial}{\partial u} \left(\left| \frac{\partial g (r (y, x), x)}{\partial u} \right| \right)}{\left| \frac{\partial g (r (y, x), x)}{\partial u} \right|} \left[\left(\frac{\partial g (r (y, x), x)}{\partial u} \right)^{-1} \frac{\partial g (r (y, x), x)}{\partial x} \right] \\
& = 0
\end{aligned}$$

Denoting $r (y, x)$ by u , the condition can be expressed as

$$\begin{aligned}
& \frac{\partial \log (f_U (u))}{\partial u} \left[\left(\frac{\partial g (u, x)}{\partial u} \right)^{-1} \frac{\partial g (u, x)}{\partial x} \right] \\
& = \frac{\partial}{\partial x} \log \left(\left| \frac{\partial g (u, x)}{\partial u} \right| \right) - \frac{\partial}{\partial u} \log \left(\left| \frac{\partial g (u, x)}{\partial u} \right| \right) \left[\left(\frac{\partial g (u, x)}{\partial u} \right)^{-1} \frac{\partial g (u, x)}{\partial x} \right]
\end{aligned}$$

Proof of Theorem 3.2: Given f_U and g , equations (3.6) and (3.7) represent a system of equations in $\partial \log f_{\tilde{U}|X=x} (g(u, x)) / \partial \tilde{u}$ and $\partial \log f_{\tilde{U}|X=x} (\tilde{u}) / \partial x|_{\tilde{u}=g(u, x)}$, which can be written in matrix form as

$$\begin{bmatrix} \left(\frac{\partial g(u, x)}{\partial u} \right)' & 0 \\ \left(\frac{\partial g(u, x)}{\partial x} \right)' & I_K \end{bmatrix} \begin{bmatrix} \tilde{\lambda}_{\tilde{u}} (g(u, x), x) \\ \tilde{\lambda}_x (g(u, x), x) \end{bmatrix} = \begin{bmatrix} \frac{\partial \log f_U (u)}{\partial u} - \frac{\partial \log \left| \frac{\partial g(u, x)}{\partial u} \right|}{\partial u} \\ - \frac{\partial \log \left| \frac{\partial g(u, x)}{\partial u} \right|}{\partial x} \end{bmatrix}$$

where the ij -th element in the $G \times G$ matrix $(\partial g(u, x) / \partial u)'$ is $\partial g_j(u, x) / \partial u_i$, the ij -th element in the $K \times G$ matrix $(\partial g(u, x) / \partial x)'$ is $\partial g_j(u, x) / \partial x_i$,

$$\tilde{\lambda}_{\tilde{u}}(g(u, x), x) = \left(\frac{\partial \log f_{\tilde{U}|X=x}(g(u, x))}{\partial \tilde{u}} \right)'$$

and

$$\tilde{\lambda}_x(g(u, x), x) = \left(\frac{\partial \log f_{\tilde{U}|X=x}(\tilde{u})}{\partial x} \Big|_{\tilde{u}=g(u, x)} \right)'$$

Since $(\partial g(u, x)/\partial u)$ is invertible, there exists a unique solution. By Cramer's rule, for each j ($j = 1, \dots, K$), we have that

$$(3.8) \quad \tilde{\lambda}_{x_j}(g(u, x), x) = \frac{\begin{vmatrix} \left(\frac{\partial g(u, x)}{\partial u}\right)' & 0^{(j)} \\ \left(\frac{\partial g(u, x)}{\partial x}\right)' & I^{(j)} \end{vmatrix}}{\begin{vmatrix} \left(\frac{\partial g(u, x)}{\partial u}\right)' & 0 \\ \left(\frac{\partial g(u, x)}{\partial x}\right)' & I_K \end{vmatrix}}$$

where $\tilde{\lambda}_{x_j}(g(u, x), x)$ denotes the j -th coordinate of $\tilde{\lambda}_x(g(u, x), x)$, $0^{(j)}$ is the zero matrix with the j -th column replaced with

$$\frac{\partial \log f_U(u)}{\partial u} - \frac{\partial \log \left| \frac{\partial g(u, x)}{\partial u} \right|}{\partial u}$$

and $I^{(j)}$ is the unit matrix with the j -th column replaced by

$$-\frac{\partial \log \left| \frac{\partial g(u, x)}{\partial u} \right|}{\partial x}$$

Since the determinant in the denominator is nonzero, $\tilde{\lambda}_{x_j}(g(u, x), x) = 0$ if and only if the determinant of

$$\begin{pmatrix} \left(\frac{\partial g(u, x)}{\partial u}\right)' & \frac{\partial \log f_U(u)}{\partial u} - \frac{\partial \log \left| \frac{\partial g(u, x)}{\partial u} \right|}{\partial u} \\ \left(\frac{\partial g(u, x)}{\partial x_j}\right)' & -\frac{\partial \log \left| \frac{\partial g(u, x)}{\partial u} \right|}{\partial x_j} \end{pmatrix}$$

is nonzero. This is equivalent to requiring that the rank of the matrix be G . $\tilde{\lambda}_x(g(u, x), x) = 0$ if and only if this holds for all j . Hence, $\tilde{\lambda}_x(g(u, x), x) = 0$ if and only if for all u, x , the rank of the matrix

$$\begin{pmatrix} \left(\frac{\partial g(u,x)}{\partial u} \right)' & \frac{\partial \log f_U(u)}{\partial u} - \frac{\partial \log \left| \frac{\partial g(u,x)}{\partial u} \right|}{\partial u} \\ \left(\frac{\partial g(u,x)}{\partial x} \right)' & - \frac{\partial \log \left| \frac{\partial g(u,x)}{\partial u} \right|}{\partial x} \end{pmatrix}$$

is G .

Proof of Lemma 4.1: If (y, x) , j , and some (\tilde{y}, \tilde{x}) satisfy Condition 1. Then, that (T1.1) is not satisfied has already been shown. Suppose that (y, x) and some $\Theta \times \Xi$ satisfy Condition 2. Then, for all $(\tilde{y}, \tilde{x}) \in \Theta \times \Xi$, $A_j(y, x) = A_j(\tilde{y}, \tilde{x})$ and $b_j(y, x) = b_j(\tilde{y}, \tilde{x})$, and for $t^* = \partial \log(f_U(r(y, x)))/\partial u$ and some $\delta > 0$, $N(t^*; \delta)$ is included in $\{w \in R^G \mid \text{for some } (\tilde{y}, \tilde{x}) \in \Theta \times \Xi, \partial \log(f_U(r(\tilde{y}, \tilde{x}))/\partial u) = w\}$. If

$$\gamma(y, x) A_j(y, x) \neq b_j(y, x)$$

the result holds with $(\tilde{y}, \tilde{x}) = (y, x)$. Suppose that

$$\gamma(y, x) A_j(y, x) = b_j(y, x).$$

Let $\eta = \delta/(\sqrt{2} G)$. Define $\tilde{\gamma} = (\tilde{\gamma}_1, \dots, \tilde{\gamma}_G)$ by $\tilde{\gamma}_g = \gamma_g(y, x) - \eta$ if $a_{gj}(y, x) < 0$, and $\tilde{\gamma}_g = \gamma_g(y, x) + \eta$ if $a_{gj}(y, x) \geq 0$. Then, since $a_{ij}(y, x) \neq 0$, $\sum_{g=1}^G \tilde{\gamma}_g a_{gj}(y, x) > \sum_{g=1}^G \gamma_g(y, x) a_{gj}(y, x) = b_j(y, x)$. Since $\sum_{g=1}^G (\tilde{\gamma}_g - \gamma_g(y, x))^2 = G \eta^2 = \delta^2/2$, we have that $\tilde{\gamma} \in N(\gamma_g(y, x); \delta)$. Hence, by the definition of $\gamma_g(y, x)$ and Condition 1, it follows that there exist (\tilde{y}, \tilde{x}) such that $A_j(y, x) = A_j(\tilde{y}, \tilde{x})$, $b_j(y, x) = b_j(\tilde{y}, \tilde{x})$, and $\partial \log(f_U(r(\tilde{y}, \tilde{x}))/\partial u) = \tilde{\gamma}$. Since $\sum_{g=1}^G \tilde{\gamma}_g a_{gj}(y, x) > \sum_{g=1}^G \gamma_g(y, x) a_{gj}(y, x) = b_j(y, x)$, it follows that $\sum_{g=1}^G \gamma_g(\tilde{y}, \tilde{x}) a_{gj}(\tilde{y}, \tilde{x}) = \gamma(\tilde{y}, \tilde{x}) A_j(\tilde{y}, \tilde{x}) > b_j(\tilde{y}, \tilde{x})$. Hence, equation (1) of the Theorem is not satisfied with equality at (\tilde{y}, \tilde{x}) .

Proof of Lemma 4.2: (The argument is identical to that in Roehrig (1988). The only difference is that the implicit functions studied are different.) Define the function m by

$$m(u, \tilde{u}, y, x) = \begin{bmatrix} u - r(y, x) \\ \tilde{u} - \tilde{r}(y, x) \end{bmatrix}$$

and such that for any y, x $m(u, \tilde{u}, y, x) = 0$. Since

$$\left| \frac{\partial m(u, \tilde{u}, y, x)}{\partial (u, y)} \right| \neq 0$$

it follows by the Implicit Function Theorem that there exist functions $u = p(x, \tilde{u})$ and $y = \tilde{h}(x, \tilde{u})$ such that

$$m(u, \tilde{u}, y, x) = \begin{bmatrix} p(x, \tilde{u}) - r(\tilde{h}(x, \tilde{u}), x) \\ \tilde{u} - \tilde{r}(\tilde{h}(x, \tilde{u}), x) \end{bmatrix} = 0$$

By our uniqueness assumptions,

$$(5.1) \quad u = p(x, \tilde{u}) = r(\tilde{h}(x, \tilde{u}), x)$$

By the Implicit Function Theorem and Cramer's Rule,

$$\frac{d(p_i(x, \tilde{u}))}{dx_j} = \frac{|M_{ij}|}{\left| \frac{\partial m(u, \tilde{u}, y, x)}{\partial (u, y)} \right|}$$

where M_i is the matrix $\frac{\partial m(u, \tilde{u}, y, x)}{\partial (u, y)}$ with the i -th column replaced by the vector whose first G coordinates are $(\partial r(y, x)/\partial x_j)'$ and $(\partial \tilde{r}(y, x)/\partial x_j)'$. It is easy to see that

$$|M_{ij}| = \begin{vmatrix} \frac{\partial r_i(y, x)}{\partial y} & \frac{\partial r_i(y, x)}{\partial x_j} \\ \frac{\partial \tilde{r}(y, x)}{\partial y} & \frac{\partial \tilde{r}(y, x)}{\partial x_j} \end{vmatrix}$$

Since $\left| \frac{\partial m(u, \tilde{u}, y, x)}{\partial (u, y)} \right| \neq 0$, $\frac{d(p_i(x, \tilde{u}))}{dx_j} \neq 0$ if and only iff

$$\begin{vmatrix} \frac{\partial r_i(y, x)}{\partial y} & \frac{\partial r_i(y, x)}{\partial x_j} \\ \frac{\partial \tilde{r}(y, x)}{\partial y} & \frac{\partial \tilde{r}(y, x)}{\partial x_j} \end{vmatrix}$$

It is easy to verify that $a_{ij}(y, x) = \frac{d(p_i(x, \tilde{u}))}{dx_j}$. Hence, the result follows.

References

AI, C. and X. CHEN (2003), "Efficient Estimation of Models with Conditional Moments Restrictions Containing Unknown Functions," *Econometrica*, 71, 1795-1843.

ALTONJI, J.G. AND H. ICHIMURA (2000): "Estimating Derivatives in Nonseparable Models with Limited Dependent Variables," mimeo.

ALTONJI, J.G. AND R.L. MATZKIN (2001): "Panel Data Estimators for Nonseparable Models with Endogenous Regressors," mimeo, Northwestern University.

ALTONJI, J.G. AND R.L. MATZKIN (2003): "Cross Section and Panel Data Estimators for Nonseparable Models with Endogenous Regressors," *Econometrica*, 73, 3, 1053-1102.

BLUNDELL, R. and J. L. POWELL (2003) "Endogeneity in Nonparametric and Semiparametric Regression Models," in *Advances in Economics and Econometrics, Theory and Applications, Eighth World Congress*, Volume II, edited by M. Dewatripont, L.P. Hansen, and S.J. Turnovsky, Cambridge University Press, Cambridge, U.K.

BENKARD, C.L. and S. BERRY (2004) "On the Nonparametric Identification of Nonlinear Simultaneous Equations Models: Comment on B. Brown (1983) and Roehrig (1988), Cowles Foundation Discussion Paper #1482.

BOWDEN, R. (1973) "The Theory of Parametric Identification," *Econometrica*, 41, 1069-1074.

BRIESCH, R., P. CHINTAGUNTA, and R. MATZKIN (1997) "Nonparametric Discrete Choice Models with Unobserved Heterogeneity," mimeo, Northwestern University.

- BROWN B.W. (1983) "The Identification Problem in Systems Nonlinear in the Variables," *Econometrica*, 51, 175-196.
- BROWN, D.J. and R.L. MATZKIN (1998) "Estimation of Nonparametric Functions in Simultaneous Equations Models, with an Application to Consumer Demand," Cowles Foundation Discussion Paper #1175.
- BROWN, D.J. AND M. WEGKAMP (2002): "Weighted Minimum Mean-Square Distance from Independence Estimation," *Econometrica*, 70 (5), 2035-2051.
- CHERNOZHUKOV, V. AND C. HANSEN (2005): "An IV Model of Quantile Treatment Effects," *Econometrica*, 73, 245-261.
- CHERNOZHUKOV, V., G. IMBENS, AND W. NEWEY (2004): "Instrumental Variable Identification and Estimation of Nonseparable Models via Quatile Conditions," mimeo, Department of Economics, M.I.T.
- CHESHER, A. (2003): "Identification in Nonseparable Models," *Econometrica*, Vol. 71, No. 5.
- CHESHER, A. (2004): "Identification in Additive Error Models with Discrete Endogenous Variables," CEMMAP working paper CWP11/04.
- CHESHER, A. (2005): "Identification with Excess Heterogeneity" mimeo, CEMMAP.
- DAROLLES, S., J.P. FLORENS, and E. RENAULT (2002), "Nonparametric Instrumental Regression," mimeo, IDEI, Toulouse.
- DAS, M. (2004) "Instrumental Variables Estimation of Nonparametric Models with Discrete Endogenous Variables," forthcoming in *Journal of Econometrics*.
- FISHER, F.M. (1961) "Identifiability Criteria in Nonlinear Systems," *Econometrica*, 29, 574-590.
- FISHER, F.M. (1965) "Identifiability Criteria in Nonlinear Systems: A Further Note," *Econometrica*, 33, 197-205.
- FISHER, F.M. (1966) *The Identification Problem in Econometrics*. New York: McGraw-Hill.
- HALL, P. and J.L. HOROWITZ (2003), "Nonparametric Methods for Inference in the Presence of Instrumental Variables," mimeo, Northwestern University.
- HAUSMAN, J.A. (1983) "Specification and Estimation of Simultaneous Equation Models," in *Handbook of Econometrics*, Vol. I, edited by Z. Griliches and M.D. Intriligator.
- HSIAO, C. (1983) "Identification," in *Handbook of Econometrics*, Vol. I, edited by Z. Griliches and M.D. Intriligator.
- IMBENS, G.W. AND W.K. NEWEY (2003) "Identification and Estimation of Triangular Simultaneous Equations Models Without Additivity," mimeo, UCLA.

- KOOPMANS, T.C., A. RUBIN, and R.B. LEIPNIK (1950) "Measuring the Equation System of Dynamic Economics," in *Statistical Inference in Dynamic Equilibrium Models*, Cowles Commission Monograph 10, New York: John Wiley.
- LEWBEL, A. (2000), "Semiparametric Qualitative Response Model Estimation with Unknown Heteroskedasticity and Instrumental Variables," *Journal of Econometrics*, 97, 145-177.
- MANSKI, C.F. (1983) "Closest Empirical Distribution Estimation," *Econometrica*, 51(2), 305-320.
- MATZKIN, R.L. (1992) "Nonparametric and Distribution-Free Estimation of the Threshold Crossing and Binary Choice Models," *Econometrica*, 60, 239-270.
- MATZKIN, R.L. (1994) "Restrictions of Economic Theory in Nonparametric Methods," in *Handbook of Econometrics*, Vol. IV, edited by R.F. Engel and D.L. McFadden.
- MATZKIN, R.L. (1999) "Nonparametric Estimation of Nonadditive Random Functions," mimeo, Northwestern University.
- MATZKIN, R.L. (2003) "Nonparametric Estimation of Nonadditive Random Functions," *Econometrica*, 71, 1339-1375.
- MATZKIN, R.L. (2004) "Unobservable Instruments," mimeo, Northwestern University.
- MATZKIN, R.L. (2005) "Estimation of Nonparametric Simultaneous Equations," in preparation, Northwestern University.
- NEWKEY, W.K. and J.L. POWELL (1989): "Instrumental Variables Estimation for Nonparametric Models," mimeo, Princeton University.
- NEWKEY, W.K. and J.L. POWELL (2003): "Instrumental Variables Estimation for Nonparametric Models," *Econometrica*, 71, 1557-1569..
- NEWKEY, 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 Estimation when Regressors and Errors are Dependent," *Canadian Journal of Statistics*, 28-2, 289-300.
- ROEHRIG, C.S. (1988) "Conditions for Identification in Nonparametric and Parametric Models," *Econometrica*, 56:2, 433-447.
- ROTHENBERG, T.J. (1971) "Identification in Parametric Models," *Econometrica*, 39, 577-592.
- VYTLACIL, E. and N. YILDIZ (2004) "Dummy Endogenous Variables in Weakly Separable Models," mimeo, Department of Economics, Stanford University.