Conditions for Identification in Nonparametric Systems of Equations

Rosa L. Matzkin
Department of Economics
Northwestern University
March 2005
Preliminary Version

Abstract

This paper considers identification in systems of nonparametric structural equations. It provides two sets of conditions characterizing observational equivalence. These conditions can be used to determine sets of parametric, semiparametric, or nonparametric functions in which the structural function is identified. We provide several such examples, as well as conditions that are easier to verify.

1 Introduction

Many economic models, specially those concerning optimization and equilibrium conditions, involve several equations. In such situations, it is often the case that some of the observable explanatory variables are not distributed independently of some of the unobservable variables in the system. Identification becomes then of concern. A large literature exists, which provides methods to analyze the identification of such systems. In the past, such analyses specified a parametric structure for the equations in the system. More recently, there has been a lot of interest in analyzing identification and estimation in nonparametric systems of equations. Newey, Powell and Vella (1999), Ng and Pinske (1995), and Pinske (2000) considered nonparametric triangular systems with additive unobservable random terms; Chesher (2003) and Imbens and Newey (2003) considered triangular systems with nonadditive random terms. Newey and Powell (1989, 2003), Darolles, Florens, and Renault (2003), Ai and Chen (2003), and Hall and

Horowitz (2003) considered estimation using a mean independence assumption between an additive unobservable and an instrument. Chernozhukov and Hansen (2005) used a rank invariance assumption on a nonadditive random term, conditional on an instrument and another variable. Altonji and Matzkin (2001) considered two methods to identify models with nonadditive unobservables and endogenous regressors. Their first method estimated an average derivative using conditional independence, which was justified using an exchangeability condition. Their second method used a local exchangeability condition to identify and estimate a nonparametric function. Roehrig (1988) considered identification in general nonparametric systems. D. Brown and Matzkin (1998) considered estimation of nonparametric systems, such as those considered by Roehrig (1988), using a "minimum distance from independence" criterion.

The nonparametric identification conditions developed by Roehrig (1988) were extensions of B. Brown's (1983) seminal work, on the identification of systems of equations nonlinear in variables. B. Brown's conditions characterized systems of equations that were observationally equivalent to each other, and showed that these were extensions of the familiar rank conditions used in the identification of systems linear in variables. Roehrig (1988) extension of B. Brown's ideas to nonparametric systems developed similar rank conditions for the more general models. Recently, Benkard and Berry (2004) discovered a major flaw in the proof of a Lemma that B. Brown (1983) used to derive his result. Moreover, Benkard and Berry argued that B. Brown and Roehrig's conditions for observational equivalence implied the impossible result that the derivatives of the reduced form equations were identified.

In this paper, we provide two conditions. Each of these conditions is necessary and sufficient for the observational equivalence of two nonparametric systems of equations. The first condition is given in terms of the true function, an alternative function, and the true density of the unobservable random terms. This contrast the conditions derived by B. Brown and Roehrig, which did not involve the density of the unobservable random terms. Our new conditions can be used to restrict the set of functions to which the true function may belong. As long as any two functions in such set do not satisfy the conditions for observationally equivalence, the true function will be identified within that set. Consistent estimation of the true function can then proceed as in, for example, the method described in D. Brown and Matzkin (1998). We present several examples of such sets of functions. We also describe situations where simpler sets of conditions can be used to determine identification. In some of these situations, these are the rank conditions developed by B. Brown (1983) and Roehrig (1988), together with some additional condition on the density of the unobservable variables.

Our second condition is given in terms of the density of the unobservable random terms and a transformation of the true function. This second condition characterizes transformations of the true function that are observationally equivalent to the true function. When one is interested in studying the type of deviations from the true function that cannot be identified from the observable variables, the second condition that we derive might be more useful than the first condition.

The outline of the paper is as follows. In the next section, we present the model and the first characterization of observational equivalence. In this same section, we present several examples where the characterization is used to determine sets of functions where the true function is identified. In section 3, we present the second characterization.. Section 5 concludes and discusses extensions.

2 Observational equivalence of two structural functions

We consider a system of equations, described as

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

where $r : R^{G+K} \to R^G$ is an unknown, twice 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. 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 $h : R^{K+G} \to R^G$, so that

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

We will assume that for all $y, x,\quad |\partial r\,(y,x)\,/\partial y| > 0$. Let f_U denote the density of U, assumed to be differentiable 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

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

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

$$(2.4) \qquad \widetilde{U} = \widetilde{r}\,(Y, X)$$

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

$$(2.5) \qquad Y = \widetilde{h}\left(X, \widetilde{U}\right)$$

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

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

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

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

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

We can now ask what are the restrictions on $\widetilde{r}$ guaranteeing that $\widetilde{U}$ is independent of X, where (2.7) is satisfied for all y, x. Under our assumptions, this is equivalent to asking the restrictions on $\widetilde{r}$ guaranteeing that for all $\widetilde{u}$ and all x

$$(2.8) \qquad \frac{\partial f_{\widetilde{U}|X=x}\,(\widetilde{u})}{\partial x} = 0$$

when (2.7) is satisfied.

The following theorem, which is proved in the Appendix, provides such restrictions:

Theorem 1: *Under the assumptions made above, $\widetilde{U}$ is distributed independently of X iff for all y, x*

$$(T1.1) \quad \frac{\partial \log\left(f_U\left(r\left(y,x\right)\right)\right)}{\partial u}\left[\frac{\partial r\left(y,x\right)}{\partial x}-\frac{\partial r\left(y,x\right)}{\partial y}\left(\frac{\partial \widetilde{r}\left(y,x\right)}{\partial y}\right)^{-1}\frac{\partial \widetilde{r}\left(y,x\right)}{\partial x}\right]$$

$$+\left[\left(\frac{\partial}{\partial x}\log\left(\left|\frac{\partial r(y,x)}{\partial y}\right|\right)-\frac{\partial}{\partial x}\log\left(\left|\frac{\partial \widetilde{r}(y,x)}{\partial y}\right|\right)\right)\right]$$

$$-\left[\left(\frac{\partial}{\partial y}\log\left(\left|\frac{\partial r(y,x)}{\partial y}\right|\right)-\frac{\partial}{\partial y}\log\left(\left|\frac{\partial \widetilde{r}(y,x)}{\partial y}\right|\right)\right)\left(\frac{\partial \widetilde{r}\left(y,x\right)}{\partial y}\right)^{-1}\frac{\partial \widetilde{r}\left(y,x\right)}{\partial x}\right]$$

$$= \quad 0$$

Note that the conditions in $(T1.1)$ are expressed in terms of the true function, r, the alternative function, $\widetilde{r}$, and the density f_U. In our model, the conclusion of B. Brown's Lemma implied that $\widetilde{U}$ is independent of X iff for all (y,x),

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

The arguments in Benkard and Berry (2004) demonstrated that this is not a necessary condition for $\widetilde{U}$ to be independent of X. The condition is sufficient, as they noted, because it implies that $\widetilde{U}$ is a function of only U and not X, and U is independent of X. Theorem 1 provides further evidence for Benkard and Berry's arguments. It is easy to see from the statement of the theorem that it can be possible that for some (y,x)

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

and still have that the random variable defined by $\widetilde{U}=\widetilde{r}\left(Y,X\right)$ be distributed independently of X. This could happen because the change in the value of one coordinate of this vector is undone by the effect in the change of another coordinate, or, this effect may be also undone by the change in the value of the determinants.

Theorem 1 can be used to determine sets of functions $\widetilde{r}$ in which r is identified. If for any $\widetilde{r} \neq r$ in that set one can find (y, x) at which the equality in the statement of the theorem does not hold, then r is identified in that set. In such a case, an estimation method based on a "minimum distance from independence" as, for example, in D. Brown and Matzkin (1998), can be used to estimate r consistently. To determine such sets, one will need to consider different sets of restrictions on the functions and the densities. We provide several such examples below.

Since verifying whether $(T1.1)$ either holds or does not hold, might look as a cumbersome task, we next discuss the result of Theorem 1 further and provide some lemmas that can be used to determine non-observational equivalence. For any y, x, define the $G \times K$ matrix $A(y, x; \partial r, \partial \widetilde{r})$, the $1 \times K$ vector $b(y, x; \partial r, \partial \widetilde{r}, \partial^2 r, \partial^2 \widetilde{r})$ and the $1 \times G$ vector $\gamma\left(y, x; f_U, r\right)$ by

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

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

and

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

We index $A(y,x)$ by $(\partial r, \partial \widetilde{r})$, $b(y,x)$ by $(\partial r, \partial \widetilde{r}, \partial^2 r, \partial^2 \widetilde{r})$, 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 $\widetilde{r}$, the value of b depends on the first and second derivatives of the functions r and $\widetilde{r}$, and the value of γ depends of the function f_U and the value of the function r. Theorem 1 states that $\widetilde{r}$ is observationally equivalent to r iff for all y, x

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

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 non-observational equivalence, it suffices to show that for some j and some y, x

$$\gamma(y,x;f_U,r)\;\; A_j(y,x;\partial r,\partial \widetilde{r}) \neq b_j(y,x;\partial r,\partial \widetilde{r},\partial^2 r,\partial^2 \widetilde{r}) \tag{2.9}$$

Consider the following condition:

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

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

or (1.ii)

$$\begin{aligned}
A_j(y,x;\partial r,\partial \widetilde{r}) &= A_j(\widetilde{y},\widetilde{x};\partial r,\partial \widetilde{r}) \neq 0,\\
b_j(y,x;\partial r,\partial \widetilde{r},\partial^2 r,\partial^2 \widetilde{r}) &= b_j(\widetilde{y},\widetilde{x};\partial r,\partial \widetilde{r},\partial^2 r,\partial^2 \widetilde{r})\\
\gamma_g(y,x;f_U,r) &< \gamma_g(\widetilde{y},\widetilde{x};f_U,r) \quad \textit{for all } g \textit{ such that } \; a_{gj}(y,x;\partial r,\partial \widetilde{r}) \geq 0,\\
\gamma_g(y,x;f_U,r) &> \gamma_g(\widetilde{y},\widetilde{x};f_U,r) \quad \textit{for all } g \textit{ such that } \; a_{gj}(y,x;\partial r,\partial \widetilde{r}) < 0,
\end{aligned}$$

This condition on two points, (y, x) and $(\widetilde{y}, \widetilde{x})$, requires that the $j-th$ column of A and the $j-th$ coordinate of b attain the same values at (y, x) and $(\widetilde{y}, \widetilde{x})$, 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 $\widetilde{r}$, 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 1 is satisfied, $\widetilde{r}$ is not observationally equivalent to r. Since, if $(1.i)$ is satisfied, then $(T1.1)$ is not satisfied at (y, x). If, on the other hand, $(1.i)$ is not satisfied, so that

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

then, $(1.ii)$ implies that

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

Hence, $(T1.1)$ is not satisfied at $(\widetilde{y}, \widetilde{x})$.

A condition that is stronger than Condition 1, but which may be easier to verify is the following:

Condition 2: *For some* (y, x), *and some* j $(j = 1, ..., K)$, *there exist a set* $\Theta \times \Xi$ *such that*

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

$$
\begin{aligned}
A_j(y,x;\partial r,\partial \widetilde{r}) &= A_j(\widetilde{y},\widetilde{x};\partial r,\partial \widetilde{r}) \qquad \& \\
b_j(y,x;\partial r,\partial \widetilde{r},\partial^2 r,\partial^2 \widetilde{r}) &= b_j(\widetilde{y},\widetilde{x};\partial r,\partial \widetilde{r},\partial^2 r,\partial^2 \widetilde{r})
\end{aligned}
$$

and

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

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

For requirement $(2.i)$ 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. For requirement $(2.ii)$ 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 $(\widetilde{y},\widetilde{x})$ satisfying Condition 1. When either Condition 1 or Condition 2 is satisfied, $\widetilde{r}$ is not observationally equivalent to r. The next lemma considers the situation where for some (y,x), some i $(i \in \{1,...,G\})$ and some j $(j = 1,...,K)$, $a_{ij}(y,x) \neq 0$.

Lemma 1: *Suppose that for some* (y,x), *some* i $(i \in \{1,...,G\})$ *and some* j $(j = 1,...,K)$, $a_{ij}(y,x) \neq 0$. *If such* (y,x) *and* j *satisfy either Condition 1, for some* $(\widetilde{y},\widetilde{x})$, *or Condition 2 for some* $\Theta \times \Xi$, *then,* $\widetilde{r}$ *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 3: *There exist* (y,x), $(\widetilde{y},\widetilde{x})$ *and* j *such that either (i)*

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

or (ii)

$$\begin{aligned} A_j(y,x;\partial r,\partial\widetilde{r}) &= A_j(\widetilde{y},\widetilde{x};\partial r,\partial\widetilde{r}), \\ \gamma\left(y,x;f_U,r\right) &= \gamma\left(\widetilde{y},\widetilde{x};f_U,r\right), \text{ and} \\ b_j(y,x;\partial r,\partial\widetilde{r},\partial^2 r,\partial^2\widetilde{r}) &\neq b_j(\widetilde{y},\widetilde{x};\partial r,\partial\widetilde{r},\partial^2 r,\partial^2\widetilde{r}) \end{aligned}$$

Clearly, if Condition 3 is satisfied, then (2.9) is satisfied, for either (y,x) or for $(\widetilde{y},\widetilde{x})$, and hence, by Theorem 1, r and $\widetilde{r}$ are not observationally equivalent. Since b_j depends on the second order derivatives of r and $\widetilde{r}$, while A_j and γ depend on the derivatives of r and $\widetilde{r}$, the density f_U, and the magnitude of r, one may consider points (y,x) and $(\widetilde{y},\widetilde{x})$ at which only the curvatures of r and $\widetilde{r}$ are different, while the derivatives r and $\widetilde{r}$ 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 3 will be satisfied.

The condition that $a_{ij}(y,x)\neq 0$, which was used in Lemma 1, can be expressed in terms of a condition on the rank of a matrix, following the analysis in B. Brown and Roehrig. The difference between the matrices used by them and ours is that in ours, the matrices are formed by one of the functions of the true model and all the functions of the alternative model. In theirs, all the functions of the true model and only one function of the alternative model is used. Of course, in most cases, the true and alternative models are treated symmetrically, making this distinction unimportant. But, in some cases, one may want to impose some particular restrictions on the true model (such as, strict monotonicity) and consider the identification of it within a set of functions that satisfy a weaker set of restrictions (such as, weak monotonicity).

Lemma 2: The condition $a_{ij}(y,x) \neq 0$ is equivalent to the condition that the rank of the matrix

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

is $G+1$

Lemmas 1-2 can be used to determine whether any particular set of restrictions on the functions $\widetilde{r}, r$ and on the density f_U imply that r and f_U are identified. We provide examples of their use.

Example 1: Suppose that r and $\widetilde{r}$ are specified to be linear:

$$U = CY + BX \qquad \text{and} \qquad \widetilde{U} = \widetilde{C}Y + \widetilde{B}X$$

Suppose that the range of the function $\partial \log(f_U(r(\cdot,\cdot))/\partial u : R^{G+K} \to R^G$ contains an open neighborhood. For each i, Let c_i and b_i denote the $i-th$ row of C and B, respectively. Then, $\widetilde{U}$ is not independent of X if for some i, the rank of the matrix

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

is $G+1$.

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

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

Moreover,

$$A(y,x) = \left[B - C \ \widetilde{C}^{-1} \ \widetilde{B} \right].$$

Hence, $b(y,x)$ and $A(y,x)$ are constant over (y,x). This, together with the assumption on the density imply that Condition 2 is satisfied. The result then follows by Lemmas 1 and 2.

Example 2: Suppose that r and $\widetilde{r}$ are specified as:

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

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 $\partial \log(f_U(r(y,z,\cdot))/\partial u : R^K \to R^G$ 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, $\widetilde{U}$ is not independent of X if for some i and y, z, the rank of the matrix

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

is $G+1$.

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

$$A(y,z,x) = \left[\begin{pmatrix} \frac{\partial m(y,z)}{\partial z} & B \end{pmatrix} - \frac{\partial m\,(y,z)}{\partial y} \left(\frac{\partial \widetilde{m}\,(y,z)}{\partial y} \right)^{-1} \begin{pmatrix} \frac{\partial \widetilde{m}(y,z)}{\partial z} & \widetilde{B} \end{pmatrix} \right]$$

and

$$b(y,z,x) = -\left[\left(\frac{\partial}{\partial z} \log \left| \frac{\partial m(y,z)}{\partial y} \right| - \frac{\partial}{\partial z} \log \left(\left| \frac{\partial \widetilde{m}(y,z)}{\partial y} \right| \right) \right) \right]$$

$$+ \left[\left(\frac{\partial}{\partial y} \log \left| \frac{\partial m(y,z)}{\partial y} \right| - \frac{\partial}{\partial y} \log \left(\left| \frac{\partial \widetilde{m}(y,z)}{\partial y} \right| \right) \right) \left(\frac{\partial \widetilde{m}\,(y,z)}{\partial y} \right)^{-1} \begin{pmatrix} \frac{\partial \widetilde{m}(y,z)}{\partial z} & \widetilde{B} \end{pmatrix} \right]$$

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 $\left\{ (y^*,z^*,x) |\ x \in R^K \right\}$. This and the assumption on the density imply that Condition 2 is satisfied. Hence, the result follows by Lemmas 1 and 2.

Example 3: Suppose that r and $\widetilde{r}$ are homogenous of degree one functions. Let

$$U = r(Y,X) \qquad \text{and} \qquad \widetilde{U} = \widetilde{r}(Y,X)$$

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

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

is $G+1$, and for some λ

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

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

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

To see this, note that the rank condition implies, by Lemma 2, that

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

If

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

then, by Theorem 1, $\widetilde{r}$ is not observationally equivalent to r. Otherwise, consider the set $\Theta \times \Xi = \{(\widetilde{y},\widetilde{x})|$ for some $\lambda > 0$, $(\widetilde{y},\widetilde{x}) = \lambda\left(y,x\right)\}$. Since the functions are homogenous of degree one, for any such $(\widetilde{y},\widetilde{x})$, we have that

$$r(\widetilde{y},\widetilde{x}) = \lambda r\left(y,x\right), \quad \frac{\partial r(\widetilde{y},\widetilde{x})}{\partial y} = \frac{\partial r\left(y,x\right)}{\partial y}, \quad \frac{\partial r(\widetilde{y},\widetilde{x})}{\partial x} = \frac{\partial r\left(y,x\right)}{\partial x},$$

$$\frac{\partial^2 r(\widetilde{y},\widetilde{x})}{\partial y^2} = \left(\frac{1}{\lambda}\right)\frac{\partial^2 r\left(y,x\right)}{\partial y^2}, \; \textit{and} \quad \frac{\partial^2 r(\widetilde{y},\widetilde{x})}{\partial y \partial x} = \left(\frac{1}{\lambda}\right)\frac{\partial^2 r\left(y,x\right)}{\partial y \partial x}.$$

Hence,

$$\begin{aligned} A_j(\widetilde{y}, \widetilde{x}; \partial r, \partial \widetilde{r}) &= A_j(y, x; \partial r, \partial \widetilde{r}) \qquad \& \\ b_j(\widetilde{y}, \widetilde{x}; \partial r, \partial \widetilde{r}, \partial^2 r, \partial^2 \widetilde{r}) &= \left(\frac{1}{\lambda}\right) b_j(y, x; \partial r, \partial \widetilde{r}, \partial^2 r, \partial^2 \widetilde{r}) \end{aligned}$$

It follows that for some $(\widetilde{y}, \widetilde{x}) \in \Theta \times \Xi$ to violate (2.9) it suffices to show that

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

when

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

Note that

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

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

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

Hence, by the arguments in Lemma 1, (2.9) will be violated at some $(\widetilde{y}, \widetilde{x})$ if for some g

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

and for some λ

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

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

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

3 Observational equivalence of transformations of structural functions

In some cases, we may be interested in considering conditions under which particular variations around the true function will not be identified. For such an analysis, it might be easier to represent an alternative function as a particular transformation of the true function. We next develop such an analysis and obtain a characterization of transformations that are observationally equivalent to the true function.

Let $\widetilde{r}$ be given. We will make the same assumptions on r and $\widetilde{r}$ as in the previous section. Define the function $\widetilde{g}(u,x)$ by

$$\widetilde{g}(u,x)=\widetilde{r}(h(x,u),x) \tag{3.1}$$

By our assumptions, it follows that

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

The representation of $\widetilde{r}$ in terms of the new function $\widetilde{g}$ implies that for all y,x

$$\widetilde{r}(y,x)=\widetilde{g}(r(y,x),x) \tag{3.3}$$

The next theorem, which is proved in the Appendix, establishes necessary and sufficient conditions on the function $\widetilde{g}$ and on f_U, for $\widetilde{r}$ to be observationally equivalent to r.

Theorem 2: *Under the same assumptions as in Theorem 1,* $\widetilde{U}=\widetilde{g}(r(Y,X),X)$ *is distributed independently of* X *iff for all* u,x

$$(T2.1)\qquad \frac{\partial\log(f_U(u))}{\partial u}\left[\left(\frac{\partial\widetilde{g}(u,x)}{\partial u}\right)^{-1}\frac{\partial\widetilde{g}(r(y,x),x)}{\partial x}\right]$$

$$=\frac{\partial}{\partial x}\log\left(\left|\frac{\partial\widetilde{g}(u,x)}{\partial u}\right|\right)-\frac{\partial}{\partial u}\log\left(\left|\frac{\partial\widetilde{g}(u,x)}{\partial u}\right|\right)\left[\left(\frac{\partial\widetilde{g}(r(y,x),x)}{\partial u}\right)^{-1}\frac{\partial\widetilde{g}(r(y,x),x)}{\partial x}\right]$$

Note that condition $(T2.1)$ is equivalent to

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

$$= \frac{\partial}{\partial x}\left(\left|\frac{\partial \widetilde{g}\,(u,x)}{\partial u}\right|\right)$$

Note also that if $\partial\widetilde{g}(u,x)/\partial x = 0$, then the right and left hand side of the equality equal 0. Hence, any function that is not observationally equivalent to r will have to satisfy the condition that

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

Example 4: Consider the model

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

Suppose that we wanted to study under what conditions on a function $v : R^K \to R^G$ and on f_U, it is the case that r is observationally equivalent to

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

Define

$$\widetilde{g}\,(U,X) = U + v(X)$$

Then, for all u, x

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

Condition $(T2.1)$ is then satisfied if for all u, x

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

or, equivalently, if

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

So, for example, if f_U is constant, then the function v is not identified.

To determine general conditions under which $(T2.1)$ is not satisfied, one can follow a procedure similar to that used to determine when condition $(T1.1)$ is not satisfied. Specifically, for any u, x, define the $G \times K$ matrix $D(u, x; \partial\widetilde{g})$, the $1 \times K$ vector $d(u, x; \partial\widetilde{g}, \partial^2\widetilde{g})$ and the $1 \times G$ vector $\delta\left(u; f_U\right)$ by

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

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

and

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

Condition $(T2.1)$ can then be stated as

$$\delta\left(u; f_U\right) D(u, x; \partial\widetilde{g}) = d(u, x; \partial\widetilde{g}, \partial^2\widetilde{g})$$

When for some u, x this condition is satisfied with equality, one may consider finding a different $\widetilde{u}, \widetilde{x}$ that attains the same value of two of the elements but different value of the third one, so that the equality is violated.

4 Conclusions and extensions

We have provided conditions to characterize observational equivalence in systems of structural equations. We have used the results to develop conditions under which the rank conditions of B. Brown (1983) and Roehrig (1988) can be used to determine identification of parametric, semiparametric, and nonparametric systems of equations.

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 $\widetilde{u}$ and x

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

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

$$f_{U|X=x}\left(\widetilde{u}\right) = f_{U|X=x'}\left(\widetilde{u}\right)$$

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, under some additional functional restrictions, one can extend the results to the case where the the model is

$$U = r\left(Y^*, X\right)$$

and the observable variables are (Y, X) where Y is a transformation of Y^*. Fourth, 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).

5 Appendix

Proof of Theorem 1: By condition (2.7) it follows that for all y, x

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

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

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

Let y be the unique value for which

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

Then,

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

Substituting in (A.1), we get

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

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

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

Differentiating with respect to x the relationship

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

we get

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

Hence

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

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

$$\left(\begin{array}{ccccc}\frac{\partial r_g(y,x)}{\partial y_1\partial x_j} & \frac{\partial r_g(y,x)}{\partial y_2\partial x_j} & . & . & . & \frac{\partial r_g(y,x)}{\partial y_G\partial x_j}\end{array}\right)$$

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

$$\left(\begin{array}{ccccc}\frac{\partial r_g(y,x)}{\partial y_1\partial y_i} & \frac{\partial r_g(y,x)}{\partial y_2\partial y_i} & . & . & . & \frac{\partial r_g(y,x)}{\partial y_G\partial y_i}\end{array}\right)$$

It can be shown that:

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

and similarly

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

where, for each j

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

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

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

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

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

where

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

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

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

Substituting $y = \widetilde{h}\left(x,\widetilde{u}\right)$, we obtain the desired result.

Proof of Theorem 2: Since $\widetilde{r}(y,x) = \widetilde{g}\left(r\left(y,x\right),x\right)$,

$$
\frac{\partial \widetilde{r}\left(y,x\right)}{\partial y} = \frac{\partial \widetilde{g}\left(r\left(y,x\right),x\right)}{\partial u} \; \frac{\partial r\left(y,x\right)}{\partial y},
$$

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

$$
\frac{\partial \widetilde{r}\left(y,x\right)}{\partial x} = \frac{\partial \widetilde{g}\left(r\left(y,x\right),x\right)}{\partial u} \; \frac{\partial r\left(y,x\right)}{\partial x} + \frac{\partial \widetilde{g}\left(r\left(y,x\right),x\right)}{\partial x}
$$

Hence

$$\frac{\partial r(y,x)}{\partial x} - \frac{\partial r(y,x)}{\partial y}\left(\frac{\partial \widetilde{r}(y,x)}{\partial y}\right)^{-1}\frac{\partial \widetilde{r}(y,x)}{\partial x}$$

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

$$-\frac{\partial r(y,x)}{\partial y}\left(\frac{\partial r(y,x)}{\partial y}\right)^{-1}\left(\frac{\partial \widetilde{g}(r(y,x),x)}{\partial u}\right)^{-1}\left[\frac{\partial \widetilde{g}(r(y,x),x)}{\partial u}\ \frac{\partial r(y,x)}{\partial x} + \frac{\partial \widetilde{g}(r(y,x),x)}{\partial x}\right]$$

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

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

$$\left[\left(\frac{\partial}{\partial x}\log\left(\left|\frac{\partial r(y,x)}{\partial y}\right|\right)-\frac{\partial}{\partial x}\log\left(\left|\frac{\partial \widetilde{r}(y,x)}{\partial y}\right|\right)\right)\right]$$

$$= \frac{\frac{\partial}{\partial x}\left|\frac{\partial r(y,x)}{\partial y}\right|}{\left|\frac{\partial r(y,x)}{\partial y}\right|}-\frac{\frac{\partial}{\partial x}\left|\frac{\partial \widetilde{r}(y,x)}{\partial y}\right|}{\left|\frac{\partial \widetilde{r}(y,x)}{\partial y}\right|}$$

$$= \frac{\frac{\partial}{\partial x}\left|\frac{\partial r(y,x)}{\partial y}\right|}{\left|\frac{\partial r(y,x)}{\partial y}\right|}-\frac{\frac{d}{dx}\left(\left|\frac{\partial \widetilde{g}(r(y,x),x)}{\partial u}\right|\left|\frac{\partial r(y,x)}{\partial y}\right|\right)}{\left|\frac{\partial \widetilde{g}(r(y,x),x)}{\partial u}\right|\left|\frac{\partial r(y,x)}{\partial y}\right|}$$

$$= \frac{\frac{\partial}{\partial x}\left|\frac{\partial r(y,x)}{\partial y}\right|}{\left|\frac{\partial r(y,x)}{\partial y}\right|}-\frac{\frac{d}{dx}\left(\left|\frac{\partial \widetilde{g}(r(y,x),x)}{\partial u}\right|\right)}{\left|\frac{\partial \widetilde{g}(r(y,x),x)}{\partial u}\right|}-\frac{\frac{d}{dx}\left|\frac{\partial r(y,x)}{\partial y}\right|}{\left|\frac{\partial r(y,x)}{\partial y}\right|}$$

$$= -\frac{\frac{d}{dx}\left(\left|\frac{\partial \widetilde{g}(r(y,x),x)}{\partial u}\right|\right)}{\left|\frac{\partial \widetilde{g}(r(y,x),x)}{\partial u}\right|}$$

$$= -\frac{\frac{\partial}{\partial u}\left(\left|\frac{\partial \widetilde{g}(r(y,x),x)}{\partial u}\right|\right)\frac{\partial r(y,x)}{\partial x}}{\left|\frac{\partial \widetilde{g}(r(y,x),x)}{\partial u}\right|}-\frac{\frac{\partial}{\partial x}\left(\left|\frac{\partial \widetilde{g}(r(y,x),x)}{\partial u}\right|\right)}{\left|\frac{\partial \widetilde{g}(r(y,x),x)}{\partial u}\right|}$$

$$\left[\left(\frac{\partial}{\partial y}\log\left(\left|\frac{\partial r(y,x)}{\partial y}\right|\right)-\frac{\partial}{\partial y}\log\left(\left|\frac{\partial\widetilde{r}(y,x)}{\partial y}\right|\right)\right)\right]$$

$$= \frac{\frac{\partial}{\partial y}\left|\frac{\partial r(y,x)}{\partial y}\right|}{\left|\frac{\partial r(y,x)}{\partial y}\right|}-\frac{\frac{\partial}{\partial y}\left|\frac{\partial\widetilde{r}(y,x)}{\partial y}\right|}{\left|\frac{\partial\widetilde{r}(y,x)}{\partial y}\right|}$$

$$= \frac{\frac{\partial}{\partial y}\left|\frac{\partial r(y,x)}{\partial y}\right|}{\left|\frac{\partial r(y,x)}{\partial y}\right|}-\frac{\frac{d}{dy}\left(\left|\frac{\partial\widetilde{g}(r(y,x),x)}{\partial u}\right|\left|\frac{\partial r(y,x)}{\partial y}\right|\right)}{\left|\frac{\partial\widetilde{g}(r(y,x),x)}{\partial u}\right|\left|\frac{\partial r(y,x)}{\partial y}\right|}$$

$$= \frac{\frac{\partial}{\partial y}\left|\frac{\partial r(y,x)}{\partial y}\right|}{\left|\frac{\partial r(y,x)}{\partial y}\right|}-\frac{\frac{d}{dy}\left(\left|\frac{\partial\widetilde{g}(r(y,x),x)}{\partial u}\right|\right)}{\left|\frac{\partial\widetilde{g}(r(y,x),x)}{\partial u}\right|}-\frac{\frac{\partial}{\partial y}\left|\frac{\partial r(y,x)}{\partial y}\right|}{\left|\frac{\partial r(y,x)}{\partial y}\right|}$$

$$= -\frac{\frac{\partial}{du}\left(\left|\frac{\partial\widetilde{g}(r(y,x),x)}{\partial u}\right|\right)\ \frac{\partial r(y,x)}{\partial y}}{\left|\frac{\partial\widetilde{g}(r(y,x),x)}{\partial u}\right|}$$

and

$$\left(\frac{\partial\widetilde{r}\,(y,x)}{\partial y}\right)^{-1}\frac{\partial\widetilde{r}\,(y,x)}{\partial x}$$

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

$$= \left(\frac{\partial r\,(y,x)}{\partial y}\right)^{-1}\frac{\partial r\,(y,x)}{\partial x}+\left(\frac{\partial r\,(y,x)}{\partial y}\right)^{-1}\left(\frac{\partial\widetilde{g}\,(r\,(y,x)\,,x)}{\partial u}\right)^{-1}\frac{\partial\widetilde{g}\,(r\,(y,x)\,,x)}{\partial x}$$

Substituting into condition (T1.1) of Theorem 1, we get that $\widetilde{U}$ is observationally equivalent to U iff

$$
\frac{\partial \log\left(f_U\left(r\left(y,x\right)\right)\right)}{\partial u}\left[\left(\frac{\partial \widetilde{g}\left(r\left(y,x\right),x\right)}{\partial u}\right)^{-1}\frac{\partial \widetilde{g}\left(r\left(y,x\right),x\right)}{\partial x}\right]
$$

$$
+\left[-\frac{\frac{\partial}{\partial u}\left(\left|\frac{\partial \widetilde{g}(r(y,x),x)}{\partial u}\right|\right)\ \frac{\partial r(y,x)}{\partial x}}{\left|\frac{\partial \widetilde{g}(r(y,x),x)}{\partial u}\right|}-\frac{\frac{\partial}{\partial x}\left(\left|\frac{\partial \widetilde{g}(r(y,x),x)}{\partial u}\right|\right)}{\left|\frac{\partial \widetilde{g}(r(y,x),x)}{\partial u}\right|}\right]
$$

$$
+\left[\frac{\frac{\partial}{du}\left(\left|\frac{\partial \widetilde{g}(r(y,x),x)}{\partial u}\right|\right)\ \frac{\partial r(y,x)}{\partial y}}{\left|\frac{\partial \widetilde{g}(r(y,x),x)}{\partial u}\right|}\left(\frac{\partial r\left(y,x\right)}{\partial y}\right)^{-1}\frac{\partial r\left(y,x\right)}{\partial x}\right]
$$

$$
+\left[\frac{\frac{\partial}{du}\left(\left|\frac{\partial \widetilde{g}(r(y,x),x)}{\partial u}\right|\right)\ \frac{\partial r(y,x)}{\partial y}}{\left|\frac{\partial \widetilde{g}(r(y,x),x)}{\partial u}\right|}\left[\left(\frac{\partial r\left(y,x\right)}{\partial y}\right)^{-1}\left(\frac{\partial \widetilde{g}\left(r\left(y,x\right),x\right)}{\partial u}\right)^{-1}\frac{\partial \widetilde{g}\left(r\left(y,x\right),x\right)}{\partial x}\right]\right]
$$

$$
= \quad 0
$$

or, equivalent, iff

$$
\frac{\partial \log\left(f_U\left(r\left(y,x\right)\right)\right)}{\partial u}\left[\left(\frac{\partial \widetilde{g}\left(r\left(y,x\right),x\right)}{\partial u}\right)^{-1}\frac{\partial \widetilde{g}\left(r\left(y,x\right),x\right)}{\partial x}\right]
$$

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

$$
+\frac{\frac{\partial}{\partial u}\left(\left|\frac{\partial \widetilde{g}(r(y,x),x)}{\partial u}\right|\right)}{\left|\frac{\partial \widetilde{g}(r(y,x),x)}{\partial u}\right|}\left[\frac{\partial r\left(y,x\right)}{\partial x}+\left(\frac{\partial \widetilde{g}\left(r\left(y,x\right),x\right)}{\partial u}\right)^{-1}\frac{\partial \widetilde{g}\left(r\left(y,x\right),x\right)}{\partial x}\right]
$$

$$
= \quad 0
$$

Hence, condition (T1.1) is satisfied iff

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

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

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

Proof of Lemma 1: If (y,x) , j, and some $(\widetilde{y},\widetilde{x})$ 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 $(\widetilde{y},\widetilde{x}) \in \Theta \times \Xi$, $A_j(y,x) = A_j(\widetilde{y},\widetilde{x})$ and $b_j(y,x) = b_j(\widetilde{y},\widetilde{x})$, and for $t^* = \partial \log(f_U(r(y,x))/\partial u$ and some $\delta > 0$, $N(t^*;\delta)$ is included in $\left\{w \in R^G | \text{ for some } (\widetilde{y},\widetilde{x}) \in \Theta \times \Xi,\ \partial \log(f_U(r(\widetilde{y},\widetilde{x}))/\partial u = w\right\}$. If

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

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

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

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

Proof of Lemma 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

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

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

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

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

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

By our uniqueness assumptions,

$$(5.1) \quad u = p\left(x, \widetilde{u}\right) = r\left(\widetilde{h}\left(x, \widetilde{u}\right), x\right)$$

By the Implicit Function Theorem and Cramer's Rule,

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

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

$$|M_{ij}| = \left|\begin{array}{cc} \frac{\partial r_i(y,x)}{\partial y} & \frac{\partial r_i(y,x)}{\partial x_j} \\ \frac{\partial \widetilde{r}(y,x)}{\partial y} & \frac{\partial \widetilde{r}(y,x)}{\partial x_j} \end{array}\right|$$

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

$$\left|\begin{array}{cc} \frac{\partial r_i(y,x)}{\partial y} & \frac{\partial r_i(y,x)}{\partial x_j} \\ \frac{\partial \widetilde{r}(y,x)}{\partial y} & \frac{\partial \widetilde{r}(y,x)}{\partial x_j} \end{array}\right|$$

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

6 References

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

ALTONJI, J.G. AND R.L. MATZKIN (2001): “Panel Data Estimators for Nonseparable Models with Endogenous Regressors,” mimeo, Northwestern University, forthcoming in *Econometrica*.

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.

BROWN B.W. (1983) "The Identification Problem in Systems Nonlinear in the Variables," *Econometrica*, 51:1, 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.

CHERNOZHUKOV, V. AND C. HANSEN (2005): "An IV Model of Quantile Treatment Effects," *Econometrica*.

CHESHER, A. (2003): "Identification in Nonseparable Models," *Econometrica*, Vol. 71, No. 5.

DAROLLES, S., J.P. FLORENS, and E. RENAULT (2002), "Nonparametric Instrumental Regression," mimeo, IDEI, Toulouse.

HALL, P. and J.L. HOROWITZ (2003), "Nonparametric Methods for Inference in the Presence of Instrumental Variables," mimeo, Northwestern University.

IMBENS, G.W. AND W.K. NEWEY (2001) "Identification and Estimation of Triangular Simultaneous Equations Models Without Additivity," mimeo, UCLA.

MATZKIN, R.L. (2003) "Nonparametric Estimation of Nonadditive Random Functions," Econometrica, 71, 1339-1375.

MATZKIN, R.L. (2004) "Unobservable Instruments," mimeo, Northwestern University.

NEWEY, W.K. and J.L. POWELL (1989): "Instrumental Variables Estimation for Nonparametric Models," mimeo, Princeton University.

NEWEY, W.K. and J.L. POWELL (2003): "Instrumental Variables Estimation for Nonparametric Models," *Econometrica*, 71, 1557-1569..

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

NG, S. and J. PINKSE (1995): "Nonparametric Two Step Estimation of Unknown Regression Functions when the Regressors and the Regressor Error are not Independent," mimeo, University of Montreal.

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.