Econometrica, Vol. 76, No. 5 (September, 2008), 945–978

IDENTIFICATION IN NONPARAMETRIC SIMULTANEOUS EQUATIONS MODELS

BY ROSA L. MATZKIN[1]

This paper provides conditions for identification of functionals in nonparametric simultaneous equations models with nonadditive unobservable random terms. The conditions are derived from a characterization of observational equivalence between models. We show that, in the models considered, observational equivalence can be characterized by a restriction on the rank of a matrix. The use of the new results is exemplified by deriving previously known results about identification in parametric and nonparametric models as well as new results. A stylized method for analyzing identification, which is useful in some situations, is also presented.

KEYWORDS: Nonparametric methods, nonadditive models, nonseparable models, identification, simultaneous equations, endogeneity.

1. INTRODUCTION

THE INTERPLAY BETWEEN ECONOMETRICS AND ECONOMIC THEORY comes to its full force when analyzing the identification of underlying functions and distributions in structural models. Identification in structural models that are linear in variables and parameters and have additive unobservable variables has been studied for a long time. On the other hand, identification in structural models that do not impose parametric assumptions in the functions and distributions in the model, or do not impose additivity in the unobservable variables, has not yet been completely understood. The objective of this paper is to provide insight into these latter cases. Starting from a characterization of observational equivalence, this paper provides new conditions that can be used to determine the identification of the underlying functions and distributions in simultaneous equations models.

The study of identification is a key element in the econometric analysis of many structural models. Such study allows one to determine conditions under which, from the distribution of observable variables, one can recover fea-

[1]The support of NSF through grants SES 0551272, BCS 0433990, and SES 0241858 is gratefully acknowledged. I am very grateful to Whitney Newey for his invaluable input in this paper and to three anonymous referees for their insightful and useful reports. I have benefited from the comments and suggestions of participants at the 2005 Berkeley Conference in Honor of Daniel McFadden, the 2005 BIRS workshop in the series "Mathematical Structures in Economic Theory and Econometrics," the 2006 Winter Meeting of the Econometric Society, seminar participants at Caltech, Columbia University, Harvard/MIT, Northwestern University, NYU, Stanford University, UC Berkeley, UC Irvine, UCLA, UCSD, Universidad Torcuato Di Tella, University College London, University of Chicago, University of Maryland, University of Wisconsin at Madison, USC, and Yale University, research assistants David M. Kang and Yong Hyeon Yang, and from conversations with Richard Blundell, Donald Brown, Steven Berry, Lanier Benkard, Andrew Chesher, Jinyong Hahn, Jerry Hausman, James Heckman, Joel Horowitz, Guido Imbens, Arthur Lewbel, Daniel McFadden, Lars Nesheim, Elie Tamer, and Chris Sims.

 DOI: 10.3982/ECTA5940

tures of the primitive functions and distributions in the model. These features are needed, for example, for the analysis of counterfactuals, where one wants to calculate the outcomes that would result when some of the elements in the model change. The analysis of identification dates back to the works of H. Working (1925), E. J. Working (1927), Tinbergen (1930), Frisch (1934, 1938), Haavelmo (1943, 1944), Hurwicz (1950a), Koopmans and Reiersol (1950), Koopmans, Rubin, and Leipnik (1950), Wald (1950), Fisher (1959, 1961, 1965, 1966), Wegge (1965), Rothenberg (1971), and Bowden (1973). While the importance of identification in models with nonparametric functions and with nonadditive unobservable random terms has been recognized since the early years (see Hurwicz (1950a, 1950b)), most works at the time concentrated on providing conditions for linear models with additive unobservable random terms or for nonlinear parametric models.

More recently, nonparametric models with nonadditive unobservable variables have received increasing attention, with new theoretical developments and application possibilities frequently appearing. This has motivated researchers to revisit older studies armed with new tools. In the context of identification in simultaneous equations models, Benkard and Berry (2006) recently revisited the path-breaking results by Brown (1983), and their extension by Roehrig (1988), on the identification of nonlinear and nonparametric simultaneous equations, and found some arguments to be controversial. A contribution of this paper is to provide a different set of conditions for the identification of such models.

The current literature on identification of nonparametric models with endogenous regressors is very large. Within this literature, Ng and Pinske (1995), 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, 2005) 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 derivatives and average derivatives in triangular systems with nonadditive random terms. Matzkin (2003, 2004) considered estimation under conditional independence, with normalizations and restrictions on nonadditive functions. Vytlacil and Yildiz (2007) studied estimation of average effects in models with weak separability and dummy endogenous variables. For nontriangular systems, Newey and Powell (1989, 2003), Darolles, Florens, and Renault (2003), and Hall and Horowitz (2005) considered estimation using conditional moment conditions between additive unobservables and instruments. Brown and Matzkin (1998), Ai and Chen (2003), Chernozhukov and Hansen (2005), and Chernozhukov, Imbens, and Newey (2007) allowed for the unobservable variables to be nonadditive. The latter two articles exploited an independence assumption between the unobservable variables and an instrument, to study identification. Matzkin (2004) considered identification using instruments and an independence condition across the unobservable variables. Blundell and Powell (2003) analyzed several nonparametric and semiparametric models, and provided many references. Matzkin (2007a) provided a partial survey of recent results on nonparametric identification. A parallel approach has considered partial identification in structural triangular and nontriangular models (see Chesher (2005, 2007)).

The outline of the paper is as follows. In the next section, we describe the model and its main assumptions. In Section 3, we derive several characterizations of observational equivalence. We demonstrate how these characterizations can be used to determine identification in a linear and an additively separable model, in Section 4, and how they can be used to obtain the already known results for single and triangular nonadditive models, in Section 5. A more stylized method for analyzing identification is presented in Section 6. Section 7 presents the main conclusions of the paper.

2. THE MODEL

We consider a system of structural equations, described as

$$U = r(Y, X), \tag{2.1}$$

where $r: R^{G+K} \to R^G$ is an unknown, twice continuously differentiable function, Y is a vector of G observable endogenous variables, X is a vector of K observable exogenous variables, and U is a vector of G unobservable variables, which is assumed to be distributed independently of X. Let f_U denote the density of U, assumed to be continuously differentiable. Our objective is to determine conditions under which the function r and the density f_U are identified within a set of functions and densities to which r and f_U belong. We will assume that the vector X has a continuous, known density f_X that has support R^K. Assuming that f_X is known does not generate a loss of generality, for the purpose of the analysis of identification, because X is observable.

A typical example of a system (2.1) is a demand and supply model,

$$\begin{aligned} Q &= D(P, I, U_1), \\ P &= S(Q, W, U_2), \end{aligned} \tag{2.2}$$

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 $Y = (P, Q)$, $X = (I, W)$, and $U = (U_1, U_2)$, 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:

$$U_1 = r_1(Q, P, I),$$
$$U_2 = r_2(Q, P, W).$$

We will assume that the system of structural equations (2.1) possesses a unique reduced form system

$$Y = h(X, U), \tag{2.3}$$

where $h: R^{K+G} \to R^G$ is twice continuously differentiable. In particular, conditional on X, r is one-to-one in Y. In the supply and demand example, this reduced form system is expressed as

$$\begin{aligned} Q &= h_1(I, W, U_1, U_2), \\ P &= h_2(I, W, U_1, U_2), \end{aligned} \tag{2.4}$$

where the values of Q and P are the unique values satisfying (2.2).

To determine conditions for identification, we start out from a characterization of observational equivalence, within a set of functions and distributions to which r and f_U are, respectively, assumed to belong. We will let Γ denote the set of functions to which r belongs and let Φ denote the set of densities to which f_U belongs.

The functions $\widetilde{r}: R^{G+K} \to R^G$, in the set Γ, satisfy the following properties.

(i) $\widetilde{r}$ is twice continuously differentiable on R^{G+K}.

(ii) For each $x \in R^K$, $\widetilde{r}(\cdot, x): R^G \to R^G$ is one-to-one and onto R^G.

(iii) For each $(y, x) \in R^{G+K}$, the Jacobian determinant $|\partial\widetilde{r}(y, x)/\partial y|$ is strictly positive.

Note that to each such $\widetilde{r}$ there corresponds a function $\widetilde{h}: R^{K+G} \to R^G$, which assigns to each value $(x, u) \in R^{K+G}$ the unique value y satisfying $u = \widetilde{r}(y, x)$. The function $\widetilde{h}$ is twice continuously differentiable on R^{K+G}. For each $x \in R^K$, $\widetilde{h}(x, \cdot): R^G \to R^G$ is one-to-one and onto R^G.

The set Φ will be defined to be the set of densities $f_{\widetilde{U}}: R^G \to R$ such that (i) $f_{\widetilde{U}}$ is continuously differentiable on R^G and (ii) the support of $f_{\widetilde{U}}$ is R^G.

The differentiability of $\widetilde{r}$ and $f_{\widetilde{U}}$ will allow us to express conditions in terms of derivatives. The support conditions on $f_{\widetilde{U}}$, the density of X, and on the density of Y conditional on X will allow us to guarantee that all densities converge to 0 as the value of one of their arguments tends to infinity. The condition on the sign of the Jacobian determinant $|\partial\widetilde{r}(y, x)/\partial y|$ is a normalization.

Given f_X, we can derive, for each $(\widetilde{r}, f_{\widetilde{U}}) \in (\Gamma \times \Phi)$, a unique distribution function, $F_{Y,X}(\cdot; (\widetilde{r}, f_{\widetilde{U}}))$ for the vector of observable variables (Y, X). Under our conditions, if f_X is differentiable, $F_{Y,X}(\cdot; (\widetilde{r}, f_{\widetilde{U}}))$ is characterized by a differentiable density $f_{Y,X}(\cdot; (\widetilde{r}, f_{\widetilde{U}}))$, which is defined, for all y, x, by

$$\begin{aligned} f_{Y,X}(y, x; (\widetilde{r}, f_{\widetilde{U}})) &= f_{Y|X=x}(y; (\widetilde{r}, f_{\widetilde{U}}))f_X(x) \\ &= f_{\widetilde{U}}(\widetilde{r}(y, x))\left|\frac{\partial\widetilde{r}(y, x)}{\partial y}\right|f_X(x). \end{aligned}$$

3. OBSERVATIONAL EQUIVALENCE

Following the standard definition, we will say that two elements of $(\Gamma \times \Phi)$ are observationally equivalent if they generate the same distribution of the observable variables. Formally, this is stated as follows.

DEFINITION 3.1: $(\widetilde{r}, f_{\widetilde{U}}), (\overline{r}, f_{\overline{U}}) \in (\Gamma \times \Phi)$ are *observationally equivalent*, if for all $(y, x) \in R^{G+K}$,

$$F_{Y,X}(y, x; (\widetilde{r}, f_{\widetilde{U}})) = F_{Y,X}(y, x; (\overline{r}, f_{\overline{U}})).$$

The standard, closely related, definition of identification is given by the next statement:

DEFINITION 3.2: $(r, f_U) \in (\Gamma \times \Phi)$ is *identified* in $(\Gamma \times \Phi)$, if for all $(\widetilde{r}, f_{\widetilde{U}}) \in (\Gamma \times \Phi)$ such that $(r, f_U) \neq (\widetilde{r}, f_{\widetilde{U}})$, $(\widetilde{r}, f_{\widetilde{U}})$ is not observationally equivalent to (r, f_U).

More generally, we might be interested in the identification of the value of some functional, $\mu(r, f_U)$. Let $\Omega = \{\mu(\widetilde{r}, f_{\widetilde{U}})|(\widetilde{r}, f_{\widetilde{U}}) \in (\Gamma \times \Phi)\}$ denote the set of all possible values that μ can attain over pairs $(\widetilde{r}, f_{\widetilde{U}})$ in $(\Gamma \times \Phi)$, given f_X. Then we have the following definition.

DEFINITION 3.3: The value $\widetilde{\omega} \in \Omega$ is *observationally equivalent to* $\overline{\omega} \in \Omega$ if there exist $(\widetilde{r}, f_{\widetilde{U}}), (\overline{r}, f_{\overline{U}}) \in (\Gamma \times \Phi)$ such that $\widetilde{\omega} = \mu(\widetilde{r}, f_{\widetilde{U}})$, $\overline{\omega} = \mu(\overline{r}, f_{\overline{U}})$, and $(\widetilde{r}, f_{\widetilde{U}})$ is observationally equivalent to $(\overline{r}, f_{\overline{U}})$.

The value ω of any functional μ at (r, f_U) is identified if all pairs $(\widetilde{r}, f_{\widetilde{U}}) \in (\Gamma \times \Phi)$ that are observationally equivalent to (r, f_U) are assigned, by μ, the same value, ω; that is, $\mu(\widetilde{r}, f_{\widetilde{U}}) = \omega$. The formal statement follows.

DEFINITION 3.4: The value $\omega = \mu(r, f_U) \in \Omega$ is *identified within* Ω, with respect to $(\Gamma \times \Phi)$, if for any $(\widetilde{r}, f_{\widetilde{U}}) \in (\Gamma \times \Phi)$ and $\widetilde{\omega} \in \Omega$ such that $\widetilde{\omega} = \mu(\widetilde{r}, f_{\widetilde{U}}) \neq \mu(r, f_U) = \omega$, $(\widetilde{r}, f_{\widetilde{U}})$ is not observationally equivalent to (r, f_U); that is,

$$F_{Y,X}(\cdot; (\widetilde{r}, f_{\widetilde{U}})) \neq F_{Y,X}(\cdot; (r, f_U)).$$

Since, in our model, the continuous marginal density f_X, whose support is R^K, does not depend on $(\widetilde{r}, f_{\widetilde{U}})$ or $(\overline{r}, f_{\overline{U}})$, we can state that $(\widetilde{r}, f_{\widetilde{U}}), (\overline{r}, f_{\overline{U}}) \in$

$(\Gamma \times \Phi)$ are *observationally equivalent* if for all y, x,

$$f_{\overline{U}}(\overline{r}(y,x))\left|\frac{\partial \overline{r}(y,x)}{\partial y}\right| = f_{\widetilde{U}}(\widetilde{r}(y,x))\left|\frac{\partial \widetilde{r}(y,x)}{\partial y}\right|. \tag{3.1}$$

Note that, under our conditions, if r is identified, so is f_U. This is easy to see, since for any u,

$$f_U(u) = f_{Y|X=x}(h(x,u))\left|\frac{\partial h(x,u)}{\partial u}\right|$$

and if r is identified, so is h.

We will analyze the identification of functionals, μ, of (r, f_U). The approach used will be to first determine conditions for observational equivalence between (r, f_U) and any $(\widetilde{r}, f_{\widetilde{U}}) \in (\Gamma \times \Phi)$, and then verify that for any such $(\widetilde{r}, f_{\widetilde{U}})$ that is observationally equivalent to (r, f_U),

$$\mu(\widetilde{r}, f_{\widetilde{U}}) = \mu(r, f_U).$$

Our starting point is equation (3.1). Given the true (r, f_U) and an alternative function $\widetilde{r} \in \Gamma$, this equation can be used to derive a density, $f_{\widetilde{U}}$, such that $(\widetilde{r}, f_{\widetilde{U}})$ is observationally equivalent to (r, f_U). For this, we study the relationship between the mapping which assigns to any (y, x), the value $\widetilde{u}$ of $\widetilde{U}$, satisfying

$$\widetilde{u} = \widetilde{r}(y,x) \tag{3.2}$$

and the mapping which assigns to that same (y, x), the value u of U, satisfying

$$u = r(y,x). \tag{3.3}$$

Since $\widetilde{r} \in \Gamma$, (3.2) implies that

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

where $\widetilde{h}$ is the reduced form function corresponding to $\widetilde{r}$. Substituting in (3.3), we get that[2]

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

Hence, we can write (3.1) as

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

[2] Brown (1983) and Roehrig (1988) used a mapping like this. They analyzed the restrictions that independence between the observable and unobservable explanatory variables imposes on this mapping, deriving different results than the ones we derive in this paper.

or, after dividing both sides by the first determinant, as

$$f_{\widetilde{U}}(\widetilde{u}) = f_U\big(r(\widetilde{h}(x,\widetilde{u}),x)\big)\left|\frac{\partial r(\widetilde{h}(x,\widetilde{u}),x)}{\partial y}\right|\left|\frac{\partial \widetilde{r}(\widetilde{h}(x,\widetilde{u}),x)}{\partial y}\right|^{-1}. \tag{3.4}$$

Two important implications can be derived from expression (3.4). First, (3.4) implies that $f_{\widetilde{U}}(\widetilde{u})$ is completely determined by $\widetilde{r}$, r, and f_U. That is, once we know $\widetilde{r}$, r, and f_U, we can, when it exists, determine the distribution of $\widetilde{U} = \widetilde{r}(Y,X)$ such that $(\widetilde{r}, f_{\widetilde{U}})$ is observationally equivalent to (r, f_U). Second, since the left-hand side of (3.4) does not depend on x, the right-hand side should not depend on x either. As we next show, the latter is a condition for independence between $\widetilde{U}$ and X.

3.1. *Independence*

Consider deriving, for each x, the conditional density, $f_{\widetilde{U}|X=x}$ of $\widetilde{U} = \widetilde{r}(Y,X)$ given $X = x$. Under our assumptions, this conditional density always exists and belongs to Φ. Since $U = r(\widetilde{h}(X,\widetilde{U}),X)$, and $\widetilde{h}$ and r are one-to-one and onto, conditional on $X = x$, it follows by the standard formula for transformation of variables that

$$f_{\widetilde{U}|X=x}(\widetilde{u}) = f_U\big(r(\widetilde{h}(x,\widetilde{u}),x)\big)\left|\frac{\partial r(\widetilde{h}(x,\widetilde{u}),x)}{\partial \widetilde{u}}\right|. \tag{3.5}$$

Differentiating with respect to $\widetilde{u}$ the identity

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

one gets that

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

Hence, the density of $\widetilde{U}$ conditional on $X = x$ is given by

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

Under our assumptions, the random variable $\widetilde{U}$ is independent of X if and only if for all x,

$$f_{\widetilde{U}|X=x}(\widetilde{u}) = f_{\widetilde{U}}(\widetilde{u}).$$

Note that this is exactly the same condition as in (3.4). Hence, requiring observational equivalence between $(\widetilde{r}, f_{\widetilde{U}})$ and (r, f_U) is equivalent to requiring that $f_{\widetilde{U}|X=x}$ in (3.5) equals, for all x, the marginal density of $\widetilde{U}$. Making use of our support and differentiability assumptions, the condition for independence between $\widetilde{U}$ and X can be expressed as the condition that for all $x, \widetilde{u}$,

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

3.2. *Characterization of Independence*

To obtain a more practical characterization of the independence condition, we proceed to express it in terms of the derivatives of the functions $\widetilde{r}$ and r. Let $\partial \log f_U(r(y,x))/\partial u$ denote the $G \times 1$ gradient of $\log(f_U(u))$ with respect to u, evaluated at $u = r(y,x)$. Since $f_{\widetilde{U}|X=x}(\widetilde{u}) > 0$, the condition that for all $x, \widetilde{u}$, $\partial f_{\widetilde{U}|X=x}(\widetilde{u})/\partial x = 0$ is equivalent to the condition that for all $x, \widetilde{u}$,

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

Since

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

the above is equivalent to the condition that

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

or, substituting $\widetilde{h}(x,\widetilde{u})$ by y, to

$$\left(\frac{\partial \log f_U(r(y,x))}{\partial u}\right)'\left[\frac{\partial r(y,x)}{\partial x}+\frac{\partial r(y,x)}{\partial y}\frac{\partial \widetilde{h}(x,\widetilde{u})}{\partial x}\right] \tag{3.6}$$
$$+\left(\frac{\partial}{\partial x}\left(\log\left|\frac{\partial r(y,x)}{\partial y}\right|\right)\right)'+\left(\frac{\partial}{\partial y}\left(\log\left|\frac{\partial r(y,x)}{\partial y}\right|\right)\right)'\frac{\partial \widetilde{h}(x,\widetilde{u})}{\partial x}$$
$$=\left(\frac{\partial}{\partial x}\log\left|\frac{\partial \widetilde{r}(y,x)}{\partial y}\right|\right)'+\left(\frac{\partial}{\partial y}\left(\log\left|\frac{\partial \widetilde{r}(y,x)}{\partial y}\right|\right)\right)'\frac{\partial \widetilde{h}(x,\widetilde{u})}{\partial x}.$$

This may 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 $\widetilde{h}$, has to equal the proportional change in the value of the determinant determined by $\widetilde{r}$ when X changes and Y responds to that change according to $\widetilde{h}$.

To obtain an equivalent expression for (3.6) in terms of only the structural functions, r and $\widetilde{r}$, and the density f_U, we note that differentiating the identity

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

with respect to x gives

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

Using the relationship, derived above, that

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

and substituting y for $\widetilde{h}(x,\widetilde{u})$ gives an expression for the derivative with respect to x of the reduced form function $\widetilde{h}$, in terms of derivatives with respect to y and x, of the structural function $\widetilde{r}$:

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

where $\widetilde{h}(x,\widetilde{u})$ is the reduced form function of the alternative model evaluated at $\widetilde{u}=\widetilde{r}(y,x)$. Hence, a different way of writing condition (3.6), in terms of the structural functions r and $\widetilde{r}$ of the observable variables, and the density f_U, is

$$\left(\frac{\partial \log f_U(r(y,x))}{\partial u}\right)'\left[\frac{\partial r(y,x)}{\partial x}-\frac{\partial r(y,x)}{\partial y}\left(\frac{\partial \widetilde{r}(y,x)}{\partial y}\right)^{-1}\frac{\partial \widetilde{r}(y,x)}{\partial x}\right] \tag{3.7}$$

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

We can express condition (3.7) in a more succinct way. Define, for any y, x, the $G \times K$ matrix $A(y, x; \partial r, \partial\widetilde{r})$, the $K \times 1$ vector $b(y, x; \partial r, \partial\widetilde{r}, \partial^2 r, \partial^2\widetilde{r})$, and the $G \times 1$ vector $s(y, x; f_U, r)$ by

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

and

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

Condition (3.7) can then be expressed as stating that for all y, x,

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

We index the $G \times K$ matrix $A(y, x)$ by $(\partial r, \partial\widetilde{r})$, the $K \times 1$ vector $b(y, x)$ by $(\partial r, \partial\widetilde{r}, \partial^2 r, \partial^2\widetilde{r})$, and the $G \times 1$ vector $s(y, x)$ 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 order derivatives of the functions r and $\widetilde{r}$, and the value of s depends on the function f_U and the value of the function r. Our arguments above lead to the following result:

THEOREM 3.1: *Suppose that* $(r, f_U) \in (\Gamma \times \Phi)$ *and that* $\widetilde{r} \in \Gamma$. *Define the density of* $\widetilde{U}$ *conditional on* $X = x$ *as in* (3.5). *Then*

$$\frac{\partial f_{\widetilde{U}|X=x}(\widetilde{u})}{\partial x} = 0 \quad \textit{for all } x, \widetilde{u}$$

if and only if for all y, x,

$$s(y, x; f_U, r)' A(y, x; \partial r, \partial\widetilde{r}) = b(y, x; \partial r, \partial\widetilde{r}, \partial^2 r, \partial^2\widetilde{r})'. \tag{3.8}$$

3.3. *Characterization of Observational Equivalence as an Independence Condition*

Making use of the connection between independence and observational equivalence, we can use (3.8) to provide a characterization of observational equivalence. This is established in the next theorem.

THEOREM 3.2: *Suppose that* $(r, f_U) \in \Gamma \times \Phi$ *and* $\widetilde{r} \in \Gamma$. *There exists* $f_{\widetilde{U}} \in \Phi$ *such that* $(\widetilde{r}, f_{\widetilde{U}})$ *is observationally equivalent to* (r, f_U) *if and only if for all* y, x, (3.8) *is satisfied*.

The proof of Theorem 3.2 follows, again, by the previous arguments. Observational equivalence between $(\widetilde{r}, f_{\widetilde{U}})$ and (r, f_U), as in (3.1) and (3.4), implies that $f_{\widetilde{U}|X=x}$ defined by (3.5) satisfies $\partial f_{\widetilde{U}|X=x}(\widetilde{u})/\partial x = 0$ for all $\widetilde{u}, x$. By Theorem 3.1, this implies (3.8). Conversely, given (r, f_U) and $\widetilde{r}$, define $f_{\widetilde{U}|X=x}$ by (3.5). The condition in (3.8) implies, by Theorem 3.1, that $\partial f_{\widetilde{U}|X=x}(\widetilde{u})/\partial x = 0$ for all $\widetilde{u}, x$. Hence, $\widetilde{U}$ is independent of X. This together with (3.4) and (3.5) implies that $(\widetilde{r}, f_{\widetilde{U}})$ and (r, f_U) are observational equivalent.

We next provide some intuition about condition (3.8) by means of a particular example. Note that (3.8) is a set of K restrictions on the density f_U, the function r, and the alternative function $\widetilde{r}$. These restrictions highlight the power of the density f_U to restrict the set of observationally equivalent values of functionals. Suppose, for example, that the model has the form

$$U = m(Y, Z) + BX,$$

where Y is the vector of observable endogenous variables and $(Z, X) \in R^{K_1+K_2}$ is a vector of observable exogenous variables ($K_2 \geq G$; K_1 may be 0), and where B is a $G \times K_2$ matrix of constants. Let an alternative model be

$$\widetilde{U} = \widetilde{m}(Y, Z) + \widetilde{B}X.$$

Consider determining the implications of observational equivalence for the relationship between $\partial m(y, z)/\partial(y, z)$ and $\partial\widetilde{m}(y, z)/\partial(y, z)$ at some specified value (y, z) of (Y, Z). Assume that the range of the function

$\partial \log(f_U(r(y,z,\cdot)))/\partial u : R^{K_2} \to R^G$ contains an open neighborhood. Note that as y, z stay fixed and x varies, the matrix $A(y, z, x; \partial r, \partial \widetilde{r})$ and the vector $b(y, z, x; \partial r, \partial \widetilde{r}, \partial^2 r, \partial^2 \widetilde{r})$ stay constant, since the derivatives of m and $\widetilde{m}$ do not depend on x, and the derivatives B and $\widetilde{B}$ with respect to X are constant. On the other hand, the value of $\partial \log(f_U(r(y,z,\cdot)))/\partial u$ will, by assumption, vary. When multiplied by nonzero elements of $A(y, x; \partial r, \partial \widetilde{r})$, these different values of $\partial \log(f_U(r(y,z,\cdot)))/\partial u$ should cause the equality in (3.8) not to be satisfied. Hence, observational equivalence will force elements of the matrix $A(y, z, x; \partial r, \partial \widetilde{r})$ to be zero. Let a_{ij} denote the element in the ith row and $(K_1 + j)$th column of $A(y, z, x; \partial r, \partial \widetilde{r})$. It is possible to show (see, e.g., Brown (1983), Roehrig (1988), or Matzkin (2005)) that $a_{ij} \neq 0$ if and only if the rank of the matrix

$$\begin{pmatrix} \dfrac{\partial r^i(y,z,x)}{\partial(y,z)} & \dfrac{\partial r^i(y,x)}{\partial x_j} \\ \dfrac{\partial \widetilde{r}(y,z,x)}{\partial(y,z)} & \dfrac{\partial \widetilde{r}(y,x)}{\partial x_j} \end{pmatrix}$$

is $G+1$, where $r = (r^1, \ldots, r^G)$. Hence, observational equivalence together with variation in the value of the vector $\partial \log(f_U(r(y,z,x)))/\partial u$ will imply restrictions on the rank of matrices whose elements are derivatives of $\widetilde{r}$ and of r. The next subsection provides a rank condition on a matrix that depends also on the vector $\partial \log(f_U(r(y,z,x)))/\partial u$ and on $b(y, z, x; \partial r, \partial \widetilde{r}, \partial^2 r, \partial^2 \widetilde{r})$, and from which all particular cases can be derived.

3.4. *Rank Conditions for Observational Equivalence*

The condition for independence between $\widetilde{U}$ and X, or alternatively, the condition for observational equivalence, can be expressed in terms of a condition about the rank of a matrix. To see this, recall the equation determining the distribution of $\widetilde{U}$ conditional on $X = x$. By our assumptions, this distribution always exists. Its density is defined by the condition that for all y, x,

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

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

$$\text{(3.9)} \qquad \left(\frac{\partial \log f_{\widetilde{U}|X=x}(\widetilde{r}(y,x))}{\partial \widetilde{u}} \right)' \frac{\partial \widetilde{r}(y,x)}{\partial y} + \left(\frac{\partial}{\partial y} \log \left| \frac{\partial \widetilde{r}(y,x)}{\partial y} \right| \right)'$$
$$= \left(\frac{\partial \log(f_U(r(y,x)))}{\partial u} \right)' \frac{\partial r(y,x)}{\partial y} + \left(\frac{\partial}{\partial y} \log \left| \frac{\partial r(y,x)}{\partial y} \right| \right)'$$

and

$$\left(\frac{\partial \log f_{\widetilde{U}|X=x}(\widetilde{r}(y,x))}{\partial \widetilde{u}}\right)' \frac{\partial \widetilde{r}(y,x)}{\partial x} + \left(\frac{\partial}{\partial x}\log\left|\frac{\partial \widetilde{r}(y,x)}{\partial y}\right|\right)' + \left(\left.\frac{\partial \log f_{\widetilde{U}|X=x}(t)}{\partial x}\right|_{t=\widetilde{r}(y,x)}\right)' = \left(\frac{\partial \log(f_U(r(y,x)))}{\partial u}\right)' \frac{\partial r(y,x)}{\partial x} + \left(\frac{\partial}{\partial x}\log\left|\frac{\partial r(y,x)}{\partial y}\right|\right)', \tag{3.10}$$

where $\partial \log \widetilde{f}_U(\widetilde{r}(y,x))/\partial \widetilde{u}$ and $\partial \log f_U(r(y,x))/\partial u$ are $G \times 1$ vectors, $\partial \widetilde{r}(y, x)/\partial y$ and $\partial r(y,x)/\partial y$ are $G \times G$ matrices, whose i, jth entries are, respectively, $\partial \widetilde{r}^i(y,x)/\partial y_j$ and $\partial r^i(y,x)/\partial y_j$; $\partial \widetilde{r}(y,x)/\partial x$ and $\partial r(y,x)/\partial x$ are $G \times K$ matrices, whose i, jth entries are, respectively, $\partial \widetilde{r}^i(y,x)/\partial x_j$ and $\partial r^i(y,x)/\partial x_j$; $\partial \log(|\partial \widetilde{r}(y,x)/\partial y|)/\partial y$, $\partial \log(|\partial r(y,x)/\partial y|)/\partial y$ are $G \times 1$ vectors, and $\partial \log(|\partial \widetilde{r}(y,x)/\partial y|)/\partial x$, $\partial \log(|\partial r(y,x)/\partial y|)/\partial x$ are $K \times 1$ vectors, where $\widetilde{r} = (\widetilde{r}^1, \dots, \widetilde{r}^G)$ and $r = (r^1, \dots, r^G)$.

The critical term in these expressions, whose value determines the dependence between $\widetilde{U}$ and X, is $\partial \log f_{\widetilde{U}|X=x}(t)/\partial x$. Given r, f_U, and $\widetilde{r}$, one can view (3.9) and (3.10) as a system of equations with unknown vectors

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

We may ask under what conditions a solution exists and satisfies for all t,

$$\left.\frac{\partial \log f_{\widetilde{U}|X=x}(t)}{\partial x}\right|_{t=\widetilde{r}(y,x)} = 0.$$

The following theorem establishes a rank condition that guarantees this, and hence it provides an alternative characterization of observational equivalence. Let

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

and

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

THEOREM 3.3: *Suppose that* $(r, f_U) \in \Gamma \times \Phi$ *and* $\widetilde{r} \in \Gamma$. *There exists* $f_{\widetilde{U}} \in \Phi$ *such that* $(\widetilde{r}, f_{\widetilde{U}})$ *is observationally equivalent to* (r, f_U) *if and only if for all* y, x, *the rank of the matrix*

$$\left(\begin{array}{cc} \left(\dfrac{\partial \widetilde{r}(y,x)}{\partial y}\right)' & \begin{array}{c} \Delta_y(y,x;\partial r, \partial^2 r, \partial \widetilde{r}, \partial^2 \widetilde{r}) \\ + \left(\dfrac{\partial r(y,x)}{\partial y}\right)' \dfrac{\partial \log(f_U(r(y,x)))}{\partial u} \end{array} \\ \left(\dfrac{\partial \widetilde{r}(y,x)}{\partial x}\right)' & \begin{array}{c} \Delta_x(y,x;\partial r, \partial^2 r, \partial \widetilde{r}, \partial^2 \widetilde{r}) \\ + \left(\dfrac{\partial r(y,x)}{\partial x}\right)' \dfrac{\partial \log(f_U(r(y,x)))}{\partial u} \end{array} \end{array} \right) \tag{3.11}$$

is G.

PROOF: Let

$$\widetilde{r}_y = \frac{\partial \widetilde{r}(y,x)}{\partial y}, \quad r_y = \frac{\partial r(y,x)}{\partial y}, \quad \widetilde{r}_x = \frac{\partial \widetilde{r}(y,x)}{\partial x}, \quad r_x = \frac{\partial r(y,x)}{\partial x},$$

$$\widetilde{s}_u = \frac{\partial \log f_{\widetilde{U}|X=x}(\widetilde{r}(y,x))}{\partial u}, \quad s_u = \frac{\partial \log f_U(r(y,x))}{\partial u},$$

$$\widetilde{s}_x = \left.\frac{\partial \log f_{\widetilde{U}|X=x}(t)}{\partial x}\right|_{t=\widetilde{r}(y,x)},$$

$$\Delta_y = \frac{\partial \log\left(\left|\dfrac{\partial r(y,x)}{\partial y}\right|\right)}{\partial y} - \frac{\partial \log\left(\left|\dfrac{\partial \widetilde{r}(y,x)}{\partial y}\right|\right)}{\partial y}, \quad \text{and}$$

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

Equations (3.9) and (3.10) can be written as

$$\widetilde{s}_u' \widetilde{r}_y = s_u' r_y + \Delta_y',$$
$$\widetilde{s}_u' \widetilde{r}_x + \widetilde{s}_x' = s_u' r_x + \Delta_x'$$

or, after transposing, as

$$\widetilde{r}_y' \widetilde{s}_u = r_y' s_u + \Delta_y, \tag{3.12}$$

$$\widetilde{r}_x' \widetilde{s}_u + \widetilde{s}_x = r_x' s_u + \Delta_x. \tag{3.13}$$

Equation (3.12) states that $r_y' s_u + \Delta_y$ is a linear combination of the columns of $\widetilde{r}_y'$, with the coefficients given by $\widetilde{s}_u$. Since $\widetilde{r}_y$ is invertible, this vector of coefficients is unique. Suppose that $\widetilde{s}_x = 0$. Then equation (3.13) states that $r_x' s_u + \Delta_x$

is a linear combination of the columns of $\widetilde{r}'_x$ and that the vector of coefficients is $\widetilde{s}_u$ also. Consider the $(G+K)\times(G+1)$ matrix

$$\begin{pmatrix} \widetilde{r}'_y & r'_y s_u + \Delta_y \\ \widetilde{r}'_x & r'_x s_u + \Delta_x \end{pmatrix}. \tag{3.14}$$

The rank of this matrix must be at least G, because $\widetilde{r}'_y$ is invertible. When $\widetilde{s}_x = 0$, the last column is a linear combination of the other G columns. Hence, when $\widetilde{s}_x = 0$, the rank of this matrix is G. But observational equivalence implies that $\widetilde{s}_x = 0$. (This can also be seen by using (3.12) to solve for $\widetilde{s}_u$, substituting the result in (3.13), and obtaining then that

$$\widetilde{s}_x = (r'_x - \widetilde{r}'_x(\widetilde{r}'_y)^{-1}r'_y)s_u + \Delta_x - \widetilde{r}'_x(\widetilde{r}'_y)^{-1}\Delta_y,$$

which is exactly the transpose of the expression in (3.7).) Hence, observational equivalence implies that the rank of the matrix in (3.11) and (3.14) is G.

Conversely, suppose that the matrix in (3.11) and (3.14) has rank G for all y, x. Then since $\widetilde{r}_y$ is invertible, it must be that the last column is a linear combination of the first G columns. Let $\lambda \in R^G$ be the vector of coefficients such that

$$\widetilde{r}'_y \lambda = r'_y s_u + \Delta_y. \tag{3.15}$$

Note that λ is unique. Since $\widetilde{s}_u$ satisfies (3.12), it must be that $\lambda = \widetilde{s}_u$, and since the rank of the matrix being G implies that λ satisfies

$$\widetilde{r}'_x \lambda = r'_x s_u + \Delta_x, \tag{3.16}$$

it must be also that

$$\widetilde{r}'_x \widetilde{s}_u = r'_x s_u + \Delta_x. \tag{3.17}$$

This implies that $\widetilde{s}_x = 0$, which, as shown above, is just (3.7). Hence, if the rank of the matrix is G, $(\widetilde{r}, f_{\widetilde{U}})$ is observationally equivalent to (r, f_U). *Q.E.D.*

4. IDENTIFICATION IN LINEAR AND SEPARABLE MODELS

We next provide examples that use the results derived in the previous sections to determine the identification of functionals of (r, f_U). Recall that a functional of (r, f_U) is identified if whenever $(\widetilde{r}, f_{\widetilde{U}})$ is observationally equivalent to (r, f_U), the value of the functional at $(\widetilde{r}, f_{\widetilde{U}})$ equals its value at (r, f_U).

4.1. *A Linear Simultaneous Equations Model*

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

$$r(y, x) = By + Cx \quad \text{and} \quad \widetilde{r}(y, x) = \widetilde{B}y + \widetilde{C}x,$$

where B and $\widetilde{B}$ are $G \times G$ nonsingular matrices, and C and $\widetilde{C}$ are $G \times K$ matrices. Since the functions are linear, for all y, x, $\Delta_y = \Delta_x = 0$. Let $F = [B, C]$ and $\widetilde{F} = [\widetilde{B}, \widetilde{C}]$ denote the matrices of all coefficients. Consider identification of the first row, F_1, of F. Suppose that there exists a value of (y, x) such that the gradient of $\log f_U$ evaluated at $r(y, x)$ is $(s_1(r(y, x), \ldots, s_G(r(y, x)))' = (1, 0, \ldots, 0)'$; that is, $\partial \log f_U(r(y, x))/\partial u_1 \neq 0$ and for $j = 2, \ldots, G$, $\partial \log f_U(r(y, x))/\partial u_j = 0$. Observational equivalence then implies that the matrix

$$(\widetilde{F}' \quad F_1')$$

has rank G. Consider linear restrictions on F_1, denoted by $\phi F_1' = 0$, where ϕ is a constant matrix. The rank condition for identification is then[3]

$$\operatorname{rank}(\phi F') = G - 1.$$

To see this, note that since the rank of $\widetilde{F}'$ is G, $F_1' = \widetilde{F}'c$ for some c. Then, premultiplying by ϕ gives

$$\phi \widetilde{F}' c = \phi F_1' = 0.$$

The rank condition for $\widetilde{F}$ says that $\operatorname{rank}(\phi \widetilde{F}') = G - 1$. Since the first column of $\phi \widetilde{F}'$ is zero, this rank condition implies that the other $G - 1$ columns of $\widetilde{F}$ must be linearly independent, so that all the elements of c other than the first element, c_1, must be zero. Therefore,

$$F_1 = c_1 \widetilde{F}_1.$$

By the usual normalization that one of the elements of F_1 and $\widetilde{F}_1$ is equal to 1, we have $c = 1$. That is, we must have that $F_1 = \widetilde{F}_1$. Hence, if Γ is the set of linear functions whose coefficients are characterized by F, the linear restrictions on the first row, $\phi F_1' = 0$, satisfy $\operatorname{rank}(\phi F') = G - 1$, one coefficient of F_1 is normalized to 1, and for some (y, x), $s_1(r(y, x)) \neq 0$ while for $j = 2, \ldots, G$, $s_j(r(y, x)) = 0$, then F_1 is identified.

4.2. *A Demand and Supply Example*

Consider a demand and supply model specified as

$$u_1 = D(p, q) + m(I),$$
$$u_2 = S(p, q) + v(w)$$

[3] I am grateful to Whitney Newey for detailed comments on this and the example in Section 4.2, which included the following new result with its proof. See Matzkin (2005, 2007a, 2007b, 2008) for other sets of conditions.

and an alternative model specified as

$$\widetilde{u}_1 = \widetilde{D}(p, q) + \widetilde{m}(I),$$
$$\widetilde{u}_2 = \widetilde{S}(p, q) + \widetilde{v}(w),$$

where p and q are, respectively, price and quantity, and I and w are, respectively, income and wages. Suppose that for all I, $m_I(I) = \partial m(I)/\partial I > 0$, and for all w, $v_W(w) = \partial v(w)/\partial w > 0$. Assume that (I, w) is independent of (u_1, u_2) and the supports of $(m(I), v(w))$ and of $(D(p, q), S(p, q))$ are R^2. Further assume that there exists (u_1^0, u_2^0), and that for all u_1^1, u_2^2, there exist u_2^1, u_1^2 such that

$$\frac{\partial f_U(u_1^0, u_2^0)}{\partial u_1} = \frac{\partial f_U(u_1^0, u_2^0)}{\partial u_2} = 0,$$
$$\frac{\partial f_U(u_1^1, u_2^1)}{\partial u_1} \neq 0, \quad \frac{\partial f_U(u_1^1, u_2^1)}{\partial u_2} = 0,$$
$$\frac{\partial f_U(u_1^2, u_2^2)}{\partial u_1} = 0, \quad \frac{\partial f_U(u_1^2, u_2^2)}{\partial u_2} \neq 0.$$

We will show that the derivatives of the demand and supply functions are identified up to scale. That is, for any alternative function $\widetilde{r} = (\widetilde{D} + \widetilde{m}, \widetilde{S} + \widetilde{v})$ for which there exists $f_{\widetilde{U}}$ such that $(\widetilde{r}, f_{\widetilde{U}})$ is observationally equivalent to (r, f_U), there exists $\lambda_1, \lambda_2 \in R$ such that for all p, q, I, w,

$$D_p(p, q) = \lambda_1 \widetilde{D}_p(p, q),$$
$$D_q(p, q) = \lambda_1 \widetilde{D}_q(p, q),$$
$$m_I(I) = \lambda_1 \widetilde{m}_I(I)$$

and

$$S_p(p, q) = \lambda_2 \widetilde{S}_p(p, q),$$
$$S_q(p, q) = \lambda_2 \widetilde{S}_q(p, q),$$
$$v_W(w) = \lambda_2 \widetilde{v}_W(w).$$

Note that, because of the additive separability, Δ_p and Δ_q depend only on (p, q), and $\Delta_I = \Delta_w = 0$. (We suppress arguments for simplicity.) Let p, q be arbitrary. By our assumptions, there exists (I^0, w^0), and for all values I_1, w_2, there exist values I_2 and w_1 such that

$$\frac{\partial f_U(D(p, q) + m(I_0), S(p, q) + v(w_0))}{\partial u_1} = 0, \tag{4.1}$$
$$\frac{\partial f_U(D(p, q) + m(I_0), S(p, q) + v(w_0))}{\partial u_2} = 0,$$

$$\frac{\partial f_U(D(p,q)+m(I_1),S(p,q)+v(w_1))}{\partial u_1}\neq 0,$$
$$\frac{\partial f_U(D(p,q)+m(I_1),S(p,q)+v(w_1))}{\partial u_2}=0,$$
$$\frac{\partial f_U(D(p,q)+m(I_2),S(p,q)+v(w_2))}{\partial u_1}=0,$$
$$\frac{\partial f_U(D(p,q)+m(I_2),S(p,q)+v(w_2))}{\partial u_2}\neq 0.$$

Observational equivalence implies that for any values of (I,w),

$$\operatorname{rank}\begin{pmatrix}\widetilde{D}_p(p,q) & \widetilde{S}_p(p,q) & \Delta_p+s_1D_p(p,q)+s_2S_p(p,q)\\ \widetilde{D}_q(p,q) & \widetilde{S}_q(p,q) & \Delta_q+s_1D_q(p,q)+s_2S_q(p,q)\\ \widetilde{m}_I(I) & 0 & s_1m_I(I)\\ 0 & \widetilde{v}_W(w) & s_2v_W(w)\end{pmatrix}=2,$$

where

$$s_1=\partial\log f_U(D(p,q)+m(I),S(p,q)+v(w))/\partial u_1$$

and

$$s_2=\partial\log f_U(D(p,q)+m(I),S(p,q)+v(w))/\partial u_2.$$

Letting $(I,w)=(I_0,w_0)$, we get that

$$\operatorname{rank}\begin{pmatrix}\widetilde{D}_p(p,q) & \widetilde{S}_p(p,q) & \Delta_p\\ \widetilde{D}_q(p,q) & \widetilde{S}_q(p,q) & \Delta_q\\ \widetilde{m}_I(I_0) & 0 & 0\\ 0 & \widetilde{v}_W(w_0) & 0\end{pmatrix}=2.$$

Since, by assumption, the matrix

$$\begin{pmatrix}\widetilde{D}_p(p,q) & \widetilde{S}_p(p,q)\\ \widetilde{D}_q(p,q) & \widetilde{S}_q(p,q)\end{pmatrix}$$

is invertible, the third column must be a linear combination of the first two. It follows that for some $\lambda_1^0=\lambda_1(p,q,I_0,w_0)$ and $\lambda_2^0=\lambda_2(p,q,I_0,w_0)$,

$$\lambda_1^0\widetilde{m}_I(I_0)=0\quad\text{and}\quad\lambda_2^0\widetilde{v}_W(w_0)=0.$$

Since $\widetilde{m}_I(I_0) \neq 0$ and $\widetilde{v}_W(w_0) \neq 0$, it must be that $\lambda_1^0 = \lambda_2^0 = 0$. Hence, $\Delta_p(p,q) = \Delta_q(p,q) = 0$. Since $\Delta_p(p,q)$ and $\Delta_q(p,q)$ do not depend on I, w, it follows that for all (I, w),

$$\operatorname{rank}\begin{pmatrix} \widetilde{D}_p(p,q) & \widetilde{S}_p(p,q) & s_1 D_p(p,q) + s_2 S_p(p,q) \\ \widetilde{D}_q(p,q) & \widetilde{S}_q(p,q) & s_1\, D_q(p,q) + s_2 S_q(p,q) \\ \widetilde{m}_I(I) & 0 & s_1 m_I(I) \\ 0 & \widetilde{v}_W(w) & s_2 v_W(w) \end{pmatrix} = 2.$$

Letting $(I, w) = (I_1, w_1)$, for arbitrary I_1 and for w_1 as in (4.1), the matrix becomes

$$\begin{pmatrix} \widetilde{D}_p(p,q) & \widetilde{S}_p(p,q) & s_1 D_p(p,q) \\ \widetilde{D}_q(p,q) & \widetilde{S}_q(p,q) & s_1 D_q(p,q) \\ \widetilde{m}_I(I_1) & 0 & s_1 m_I(I_1) \\ 0 & \widetilde{v}_W(w_1) & 0 \end{pmatrix}.$$

Again, linear independence of the first two columns and the matrix having rank 2 implies that the third column is a linear combination of the first two. The zeroes in the fourth row imply that the coefficient of the second column is zero. Hence, for some $\lambda_1^1 = \lambda_1(p, q, I_1, w_1)$,

$$\lambda_1^1(p, q, I_1, w_1)\widetilde{D}_p(p,q) = D_p(p,q),$$
$$\lambda_1^1(p, q, I_1, w_1)\widetilde{D}_q(p,q) = D_q(p,q),$$
$$\lambda_1^1(p, q, I_1, w_1)\widetilde{m}_I(I_1) = m_I(I_1).$$

Since I_1 was arbitrary and $\widetilde{D}$ and D are not functions of I, w, the first two equations imply that λ_1^1 is not a function of I, w. Likewise, reaching these equations by fixing I_1, varying (p, q) arbitrarely, and letting w_1 satisfy (4.1), the third equation implies that λ_1^1 is not a function of p, q. Hence, λ_1^1 is a constant. It follows that the derivatives D_p, D_q, and m_I are identified up to scale.

An analogous argument can be used to show that S_p, S_q, and v_W are also identified up to scale.

5. IDENTIFICATION IN NONPARAMETRIC NONSEPARABLE MODELS

We next apply our results to two standard nonparametric models with nonadditive unobservable random terms. We first consider the single equation model, with $G = 1$, considered in Matzkin (1999, 2003):

$$y = m(x, u).$$

We show below that in this model, with m strictly increasing in u, application of our theorems implies the well known result that for all u,

$$\frac{\partial m(x,u)}{\partial x},$$

the partial derivative of m with respect to x, for any fixed value of x and u, is identified.

In Section 5.2, we consider the triangular model with nonadditive unobservable random terms considered in Chesher (2003) and Imbens and Newey (2003):

$$y_1 = m_1(y_2, u_1),$$
$$y_2 = m_2(x, u_2).$$

Assuming that X is distributed independently of (u_1, u_2), and that m_1 and m_2 are strictly increasing, respectively, in u_1 and u_2, we derive the well-known result that for all u_1, y_2,

$$\frac{\partial m_1(y_2, u_1)}{\partial y_2}$$

is identified.[4]

5.1. *Single Equation Model*

Consider the model

$$y = m(x, u)$$

with $y, u \in R$, u and x independently distributed, $f_U \in \Phi$, and the inverse of m belonging to Γ. Letting r denote the inverse of m with respect to u, we have the model

$$u = r(y, x)$$

with $\partial r(y,x)/\partial y > 0$. Let $\widetilde{r} \in \Gamma$ be an alternative function, so that $\widetilde{u} = \widetilde{r}(y, x)$. The condition for observational equivalence requires that the matrix

$$\begin{bmatrix} \widetilde{r}_y & s r_y + \Delta_y \\ \widetilde{r}_x & s r_x + \Delta_x \end{bmatrix}$$

has rank 1, for all y, x, where $s = \partial \log f_U(r(y,x))/\partial u$. Hence, for all y, x,

$$s r_x \widetilde{r}_y + \Delta_x \widetilde{r}_y = s r_y \widetilde{r}_x + \Delta_y \widetilde{r}_x$$

[4] See Matzkin (2008) for identification of $\partial m_1(y_2, u_1)/\partial y_2$ when $y_2 = m_2(y_1, x, u_2)$.

or

$$s(r_y\widetilde{r}_x - r_x\widetilde{r}_y) = \Delta_x\widetilde{r}_y - \Delta_y\widetilde{r}_x.$$

Note that

$$\frac{\partial}{\partial y}\left(\frac{r_x}{r_y}\right) = \frac{r_{yx}}{r_y} - \frac{r_{yy}}{r_y}\frac{r_x}{r_y}.$$

Hence,

$$\begin{aligned}\Delta_x &= \frac{r_{yx}}{r_y} - \frac{\widetilde{r}_{yx}}{\widetilde{r}_y}\\ &= \frac{r_{yy}}{r_y}\frac{r_x}{r_y} - \frac{\widetilde{r}_{yy}}{\widetilde{r}_y}\frac{\widetilde{r}_x}{\widetilde{r}_y} + \frac{\partial}{\partial y}\left(\frac{r_x}{r_y} - \frac{\widetilde{r}_x}{\widetilde{r}_y}\right)\end{aligned}$$

and, since

$$\Delta_y = \frac{r_{yy}}{r_y} - \frac{\widetilde{r}_{yy}}{\widetilde{r}_y},$$

$$\frac{-\Delta_x\widetilde{r}_y + \Delta_y\widetilde{r}_x}{\widetilde{r}_y} = \frac{r_{yy}}{r_y}\left(\frac{\widetilde{r}_x}{\widetilde{r}_y} - \frac{r_x}{r_y}\right) + \frac{\partial}{\partial y}\left(\frac{\widetilde{r}_x}{\widetilde{r}_y} - \frac{r_x}{r_y}\right).$$

Hence, the rank condition implies that

$$sr_y\left(\frac{\widetilde{r}_x}{\widetilde{r}_y} - \frac{r_x}{r_y}\right) + \frac{r_{yy}}{r_y}\left(\frac{\widetilde{r}_x}{\widetilde{r}_y} - \frac{r_x}{r_y}\right) + \frac{\partial}{\partial y}\left(\frac{\widetilde{r}_x}{\widetilde{r}_y} - \frac{r_x}{r_y}\right) = 0.$$

Writing explicitly the arguments of all functions and multiplying both sides of the equality by $f_U(r(y,x))r_y(y,x)$ gives

$$\frac{\partial}{\partial y}\left[f_U(r(y,x))r_y(y,x)\left(\frac{\widetilde{r}_x(y,x)}{\widetilde{r}_y(y,x)} - \frac{r_x(y,x)}{r_y(y,x)}\right)\right] = 0.$$

Observational equivalence then implies that the function v defined by

$$v(y,x) = \left[f_U(r(y,x))r_y(y,x)\left(\frac{\widetilde{r}_x(y,x)}{\widetilde{r}_y(y,x)} - \frac{r_x(y,x)}{r_y(y,x)}\right)\right]$$

is a constant function of y. Since for any x, the range of $r(\cdot,x)$ is R and $f_U r_y \to 0$ as $|y| \to \infty$, as long as the ratios $\widetilde{r}_x/\widetilde{r}_y$ and r_x/r_y are uniformly bounded, it must be that for any y, x, $v(y,x) = 0$. Since $r_y > 0$, it follows from these conditions that for all y, x at which $f_U(r(y,x)) > 0$,

$$\text{(5.1)} \qquad \frac{\widetilde{r}_x(y,x)}{\widetilde{r}_y(y,x)} = \frac{r_x(y,x)}{r_y(y,x)}.$$

Hence, observational equivalence implies that the ratio of the derivatives of r is identified. Since

$$u = r(m(x,u),x),$$

it follows that

$$\frac{\partial m(x,u)}{\partial x} = -\frac{r_x(y,x)}{r_y(y,x)}$$

is identified.

5.2. *A Triangular Model*

Consider now the model

$$y_1 = m_1(y_2,u_1),$$
$$y_2 = m_2(x,u_2)$$

with $y_1, y_2, u_1, u_2 \in R$, m_1 strictly increasing in u_1, and m_2 strictly increasing in u_2. Assume that x is distributed independently of (u_1,u_2) and that the density of u belongs to Φ. Let r^1 denote the inverse of m_1 with respect to u_1 and let r^2 denote the inverse of m_2 with respect to u_2. Hence

$$u_1 = r^1(y_1,y_2),$$
$$u_2 = r^2(y_2,x).$$

Consider the alternative model

$$\widetilde{u}_1 = \widetilde{r}^1(y_1,y_2),$$
$$\widetilde{u}_2 = \widetilde{r}^2(y_2,x).$$

Assume that $r = (r^1,r^2) \in \Gamma$ and $\widetilde{r} = (\widetilde{r}^1,\widetilde{r}^2) \in \Gamma$.

Observational equivalence implies that the rank of the matrix

$$\begin{bmatrix} \widetilde{r}^1_{y_1} & 0 & s_1 r^1_{y_1} + \Delta_{y_1} \\ \widetilde{r}^1_{y_2} & \widetilde{r}^2_{y_2} & s_1 r^1_{y_2} + s_2 r^2_{y_2} + \Delta_{y_2} \\ 0 & \widetilde{r}^2_x & s_2 r^2_x + \Delta_x \end{bmatrix}$$

is 2, where

$$s_1 = \partial \log f_{U_1,U_2}(r^1(y_1,y_2), r^2(y_2,x))/\partial u_1$$

and

$$s_2 = \partial \log f_{U_1,U_2}(r^1(y_1,y_2), r^2(y_2,x))/\partial u_2.$$

Since the first two columns are linearly independent, the third column must be a linear combination of the first two. Hence, for some λ_1, λ_2,

$$\lambda_1 \widetilde{r}^1_{y_1} = s_1 r^1_{y_1} + \frac{r^1_{y_1 y_1}}{r^1_{y_1}} - \frac{\widetilde{r}^1_{y_1 y_1}}{\widetilde{r}^1_{y_1}},$$

$$\lambda_1 \widetilde{r}^1_{y_2} + \lambda_2 \widetilde{r}^2_{y_2} = s_1\, r^1_{y_2} + s_2 r^2_{y_2} + \frac{r^1_{y_1 y_2}}{r^1_{y_1}} - \frac{\widetilde{r}^1_{y_1 y_2}}{\widetilde{r}^1_{y_1}} + \frac{r^2_{y_2 y_2}}{r^2_{y_2}} - \frac{\widetilde{r}^2_{y_2 y_2}}{\widetilde{r}^2_{y_2}},$$

$$\lambda_2 \widetilde{r}^2_x = s_2 r^2_x + \frac{r^2_{y_2 x}}{r^2_{y_2}} - \frac{\widetilde{r}^2_{y_2 x}}{\widetilde{r}^2_{y_2}}.$$

Solving for λ_1 and λ_2 from the first two equations, and substituting them into the third, one gets, after rearranging terms, the expression

$$\begin{aligned}
& s_1 r^1_{y_1}\left(\frac{r^1_{y_2}}{r^1_{y_1}} - \frac{\widetilde{r}^1_{y_2}}{\widetilde{r}^1_{y_1}}\right) - s_2 r^2_{y_2}\frac{\widetilde{r}^2_{y_2}}{\widetilde{r}^2_x}\left(\frac{r^2_x}{r^2_{y_2}} - \frac{\widetilde{r}^2_x}{\widetilde{r}^2_{y_2}}\right) \\
&\quad - \frac{\widetilde{r}^2_{y_2}}{\widetilde{r}^2_x}\frac{\partial}{\partial y_2}\left(\frac{r^2_x}{r^2_{y_2}} - \frac{\widetilde{r}^2_x}{\widetilde{r}^2_{y_2}}\right) - \frac{\widetilde{r}^2_{y_2}}{\widetilde{r}^2_x}\frac{r^2_{y_2 y_2}}{r^2_{y_2}}\left(\frac{r^2_x}{r^2_{y_2}} - \frac{\widetilde{r}^2_x}{\widetilde{r}^2_{y_2}}\right) \\
&\quad - \frac{\partial}{\partial y_1}\left(\frac{\widetilde{r}^1_{y_2}}{\widetilde{r}^1_{y_1}} - \frac{r^1_{y_2}}{r^1_{y_1}}\right) + \frac{r^1_{y_1 y_1}}{r^1_{y_1}}\left(\frac{r^1_{y_2}}{r^1_{y_1}} - \frac{\widetilde{r}^1_{y_2}}{\widetilde{r}^1_{y_1}}\right) \\
& = 0.
\end{aligned}$$

Multiplying both sides of the equality by $f_U r^1_{y_1} r^2_{y_2}$ gives

$$\begin{aligned}
& r^2_{y_2}\left(\left.\frac{\partial f_U(r^1(y_1,y_2),u_2)}{\partial y_1}\right|_{u_2=r^2(y_2,x)}\right) r^1_{y_1}\left(\frac{r^1_{y_2}}{r^1_{y_1}} - \frac{\widetilde{r}^1_{y_2}}{\widetilde{r}^1_{y_1}}\right) \\
&\quad + r^2_{y_2} f_U\left(\frac{\partial r^1_{y_1}}{\partial y_1}\right)\left(\frac{r^1_{y_2}}{r^1_{y_1}} - \frac{\widetilde{r}^1_{y_2}}{\widetilde{r}^1_{y_1}}\right) + r^2_{y_2} f_U r^1_{y_1}\, \frac{\partial\left(\frac{r^1_{y_2}}{r^1_{y_1}} - \frac{\widetilde{r}^1_{y_2}}{\widetilde{r}^1_{y_1}}\right)}{\partial y_1} \\
&\quad + \frac{\widetilde{r}^2_{y_2}}{\widetilde{r}^2_x} r^1_{y_1}\left(\left.\frac{\partial f_U(u_1, r^2(y_2,x))}{\partial y_2}\right|_{u_1=r^1(y_1,y_2)}\right) r^2_{y_2}\left(\frac{\widetilde{r}^2_x}{\widetilde{r}^2_{y_2}} - \frac{r^2_x}{r^2_{y_2}}\right) \\
&\quad + \frac{\widetilde{r}^2_{y_2}}{\widetilde{r}^2_x} r^1_{y_1} f_U\left(\frac{\partial r^2_{y_2}}{\partial y_2}\right)\left(\frac{\widetilde{r}^2_x}{\widetilde{r}^2_{y_2}} - \frac{r^2_x}{r^2_{y_2}}\right) + \frac{\widetilde{r}^2_{y_2}}{\widetilde{r}^2_x} f_U r^1_{y_1} r^2_{y_2}\frac{\partial\left(\frac{\widetilde{r}^2_x}{\widetilde{r}^2_{y_2}} - \frac{r^2_x}{r^2_{y_2}}\right)}{\partial y_2} \\
& = 0
\end{aligned}$$

or, after gathering terms,

$$\begin{aligned}
&r^2_{y_2}\frac{\partial}{\partial y_1}\left[f_U(r^1(y_1,y_2),u_2)\big|_{u_2=r^2(y_2,x)}r^1_{y_1}\left(\frac{r^1_{y_2}}{r^1_{y_1}}-\frac{\widetilde{r}^1_{y_2}}{\widetilde{r}^1_{y_1}}\right)\right]\\
&\qquad+\frac{\widetilde{r}^2_{y_2}}{\widetilde{r}^2_x}r^1_{y_1}\frac{\partial}{\partial y_2}\left[\left(f_U(u_1,r^2(y_2,x))\big|_{u_1=r^1(y_1,y_2)}\right)r^2_{y_2}\left(\frac{\widetilde{r}^2_x}{\widetilde{r}^2_{y_2}}-\frac{r^2_x}{r^2_{y_2}}\right)\right]\\
&\quad=0.
\end{aligned}$$

Hence, after dividing by $r^2_{y_2}r^1_{y_1}$, it follows that observational equivalence implies that

$$\begin{aligned}
&\frac{1}{r^1_{y_1}}\frac{\partial}{\partial y_1}\left[f_U(r^1(y_1,y_2),u_2)\big|_{u_2=r^2(y_2,x)}r^1_{y_1}\left(\frac{r^1_{y_2}}{r^1_{y_1}}-\frac{\widetilde{r}^1_{y_2}}{\widetilde{r}^1_{y_1}}\right)\right]\\
&\qquad+\frac{\widetilde{r}^2_{y_2}}{\widetilde{r}^2_x}\frac{1}{r^2_{y_2}}\frac{\partial}{\partial y_2}\left[\left(f_U(u_1,r^2(y_2,x))\big|_{u_1=r^1(y_1,y_2)}\right)r^2_{y_2}\left(\frac{\widetilde{r}^2_x}{\widetilde{r}^2_{y_2}}-\frac{r^2_x}{r^2_{y_2}}\right)\right]\\
&\quad=0.
\end{aligned}\tag{5.2}$$

Note that the first term does not depend on x, other than through u_2, and the second term does not depend on y_1, other than through u_1. Since the independence between x and (u_1,u_2) implies independence between x and u_2, the result, (5.1), derived in the single equation model, applied to r^2, can be used to prove that

$$\frac{\widetilde{r}^2_x}{\widetilde{r}^2_{y_2}}=\frac{r^2_x}{r^2_{y_2}}.$$

This means that the ratio of the derivatives of the structural function r^2 can be identified. This also implies that the second term in (5.2) equals zero. So (5.2) becomes

$$\frac{\partial}{\partial y_1}\left[f_U(r^1(y_1,y_2),u_2)\big|_{u_2=r^2(y_2,x)}r^1_{y_1}\left(\frac{\widetilde{r}^1_{y_2}}{\widetilde{r}^1_{y_1}}-\frac{r^1_{y_2}}{r^1_{y_1}}\right)\right]=0.$$

In other words, the function v defined by

$$v(y_1,y_2,u_2)=\left[f_U(r^1(y_1,y_2),u_2)\big|_{u_2=r^2(y_2,x)}r^1_{y_1}\left(\frac{\widetilde{r}^1_{y_2}}{\widetilde{r}^1_{y_1}}-\frac{r^1_{y_2}}{r^1_{y_1}}\right)\right]$$

must be constant in y_1. Since for any y_2 and u_2, $f_U(r^1(y_1,y_2),u_2)r^1_{y_1}\to 0$ as $|y_1|\to\infty$, as long as the ratios of the derivatives of r^1 and $\widetilde{r}^1$ are uniformly bounded, we can conclude that for all (y_1, y_2, u_2), $v(y_1, y_2, u_2) = 0$. Hence, under these conditions, observational equivalence implies that

$$\frac{\widetilde{r}^1_{y_2}}{\widetilde{r}^1_{y_1}} = \frac{r^1_{y_2}}{r^1_{y_1}}.$$

This shows that the ratio of the derivatives of the structural function r^1 can be identified.

Since

$$u_1 = r^1(y_1, y_2) = r^1(m_1(y_2, u_1), y_2)$$

and $r^1_{y_1} \neq 0$, the implicit function theorem implies that

$$\frac{\partial m_1(y_2, u_1)}{\partial y_2} = -\frac{\dfrac{\partial r^1(m_1(y_2, u_1), y_2)}{\partial y_2}}{\dfrac{\partial r^1(m_1(y_2, u_1), y_2)}{\partial y_1}} = -\left.\frac{\dfrac{\partial r^1(y_1, y_2)}{\partial y_2}}{\dfrac{\partial r^1(y_1, y_2)}{\partial y_1}}\right|_{y_1 = m_1(y_2, u_1)}.$$

Hence, the partial derivative of m_1 with respect to y_2 is identified.

6. OBSERVATIONAL EQUIVALENCE OF TRANSFORMATIONS OF STRUCTURAL FUNCTIONS

A stylized way of analyzing observational equivalence can be derived by considering directly the mapping from the vectors of observable and unobservable explanatory variables, X and U, to an alternative vector of unobservable explanatory variables, $\widetilde{U}$, generated by an alternative function, $\widetilde{r}$.[5] To define such a relationship, we note that given the function r and an alternative function $\widetilde{r} \in \Gamma$, we can express $\widetilde{r}$ as a transformation $\widetilde{g}$ of U and X by defining $\widetilde{g}$ for all x, u as

$$\widetilde{g}(u, x) = \widetilde{r}(h(x, u), x), \tag{6.1}$$

where $h(x, u)$ is the reduced form function derived from the structural function r. By our assumptions on r and $\widetilde{r}$, it follows that $\widetilde{g}$ is invertible in u and that

$$\left|\frac{\partial \widetilde{g}(u, x)}{\partial u}\right| = \left|\frac{\partial \widetilde{r}(h(x, u), x)}{\partial y}\right| \left|\frac{\partial h(x, u)}{\partial u}\right| > 0. \tag{6.2}$$

[5] See Brown (1983) for an earlier development of this approach.

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

$$\widetilde{r}(y,x) = \widetilde{g}(r(y,x),x). \tag{6.3}$$

Recall that, for any given x, $(\widetilde{r}, f_{\widetilde{U}|X=x})$ generates the same distribution of Y given $X = x$, as (r, f_U) does, if and only if for all y,

$$f_{\widetilde{U}|X=x}(\widetilde{r}(y,x))\left|\frac{\partial \widetilde{r}(y,x)}{\partial y}\right| = f_U(r(y,x))\left|\frac{\partial r(y,x)}{\partial y}\right|. \tag{6.4}$$

Hence, using (6.1)–(6.3), we can state that the transformation $\widetilde{g}(U,X)$ generates the same distribution of Y given $X = x$ as (r, f_U) generates if and only if for all u,

$$f_{\widetilde{U}|X=x}(\widetilde{g}(u,x))\left|\frac{\partial \widetilde{g}(u,x)}{\partial u}\right| = f_U(u). \tag{6.5}$$

The analogous results to Theorems 3.1 and 3.3 are Theorems 6.1 and 6.2 below.

THEOREM 6.1: *Suppose that $(r, f_U) \in (\Gamma \times \Phi)$ and that $\widetilde{r} \in \Gamma$. Define the transformation $\widetilde{g}$ by (6.1) and let $\widetilde{U} = \widetilde{g}(U,X)$ be such that for all x, $f_{\widetilde{U}|X=x} \in \Phi$. Then $\partial f_{\widetilde{U}|X=x}(\widetilde{u})/\partial x = 0$ for all x and $\widetilde{u}$ if and only if for all u and x,*

$$\left[-\frac{\partial \log(f_U(u))}{\partial u} + \frac{\partial}{\partial u}\log\left(\left|\frac{\partial \widetilde{g}(u,x)}{\partial u}\right|\right)\right]\left[\left(\frac{\partial \widetilde{g}(u,x)}{\partial u}\right)^{-1}\frac{\partial \widetilde{g}(u,x)}{\partial x}\right] = \frac{\partial}{\partial x}\log\left(\left|\frac{\partial \widetilde{g}(u,x)}{\partial u}\right|\right). \tag{6.6}$$

THEOREM 6.2: *Suppose that $(r, f_U) \in (\Gamma \times \Phi)$ and that $\widetilde{r} \in \Gamma$. Define the transformation $\widetilde{g}$ by (6.1) and let $\widetilde{U} = \widetilde{g}(U,X)$. Suppose further that for all x, $f_{\widetilde{U}|X=x} \in \Phi$. Then $(f_{\widetilde{U}}, \widetilde{g}(r(y,x),x))$ is observationally equivalent to (r, f_U) if and only if for all u, x, the rank of the matrix*

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

is G.

The proofs of Theorems 6.1 and 6.2 use arguments similar to the ones used to derive Theorems 3.1 and 3.3, and are given in the Appendix. To provide an example of the usefulness of these results, suppose that $X \in R$ and $G = 2$, and consider evaluating the implications of observational equivalence when the relationship between $\widetilde{r}$ and r is given by

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

where $\widetilde{B}(x)$ is a 2×2 matrix of functions. One such example for $\widetilde{B}(x)$ could be[6]

$$\widetilde{B}(x) = \begin{pmatrix} \cos(x) & \sin(x) \\ -\sin(x) & \cos(x) \end{pmatrix}. \tag{6.7}$$

Application of Theorem 6.2 yields the result that observational equivalence implies that for all x, u, the matrix

$$\begin{pmatrix} \cos(x) & -\sin(x) & \dfrac{\partial \log f_U(u)}{\partial u_1} \\ \sin(x) & \cos(x) & \dfrac{\partial \log f_U(u)}{\partial u_2} \\ -u_1 \sin(x) + u_2 \cos(x) & -u_1 \cos(x) - u_2 \sin(x) & 0 \end{pmatrix}$$

must have rank 2. This holds if and only if for all u_1, u_2,

$$\frac{\partial f_U(u_1, u_2)/\partial u_1}{\partial f_U(u_1, u_2)/\partial u_2} = \frac{u_1}{u_2}.$$

Note that this condition is satisfied by the bivariate independent standard normal density. Hence, if U is distributed $N(0, I)$, $\widetilde{B}(x)$ is as specified above, and

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

it follows by Theorem 6.2 that $(\widetilde{r}, f_{\widetilde{U}})$ is observationally equivalent to (r, f_U).

7. CONCLUSIONS

We have developed several characterizations of observational equivalence for nonparametric simultaneous equations models with nonadditive unobservable variables.

The models that we considered can be described as

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

[6] This is the example in Benkard and Berry (2006, p. 1433, footnote 4).

where $U \in R^G$ is a vector of unobservable exogenous variables, distributed independently of X, $X \in R^K$ is a vector of observable exogenous variables, Y is a vector of observable endogenous variables, and r is a function such that, conditional on X, r is one-to-one.

Our characterizations were developed by considering an alternative function, $\widetilde{r}$, and analyzing the density of $\widetilde{U} = \widetilde{r}(Y, X)$. We asked what restrictions on $\widetilde{r}$, r, and the density, f_U, of U are necessary and sufficient to guarantee that $\widetilde{U}$ is distributed independently of X. We showed that these restrictions characterize observational equivalence and we provided an expression for them in terms of a restriction on the rank of a matrix.

The use of the new results was exemplified by deriving known results about identification in nonadditive single equation models and triangular equations models. An example of a separable demand and supply model provided insight into the power of separability restrictions.

We also developed a simplified approach to characterize observational equivalence, which is useful when the alternative function, $\widetilde{r}$, is defined as a transformation of the function r.

APPENDIX

PROOF OF THEOREM 6.1: Define the function $g : R^{G+K} \to R^G$ by

$$u = g(\widetilde{u}, x),$$

where $\widetilde{u}$ is defined, as in Section 6, by

$$\widetilde{u} = \widetilde{g}(u, x).$$

Then, since

$$u = g(\widetilde{g}(u, x), x),$$

we get, by differentiating this expression with respect to u and with respect to x, that

$$\text{(A.1)} \qquad I = \frac{\partial g(\widetilde{g}(u, x), x)}{\partial \widetilde{u}} \frac{\partial \widetilde{g}(u, x)}{\partial u}$$

and

$$\text{(A.2)} \qquad 0 = \frac{\partial g(\widetilde{u}, x)}{\partial \widetilde{u}} \frac{\partial \widetilde{g}(u, x)}{\partial x} + \frac{\partial g(\widetilde{u}, x)}{\partial x}.$$

To derive (6.6), we rewrite (6.5) as

$$\text{(A.3)} \qquad f_{\widetilde{U}|X=x}(\widetilde{u}) = f_U(g(\widetilde{u}, x)) \left| \frac{\partial g(\widetilde{u}, x)}{\partial \widetilde{u}} \right|.$$

Independence between $\widetilde{U}$ and X is equivalent to requiring that for all $\widetilde{u}$, x,

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

Taking logs on both sides of (A.3) and differentiating the resulting expressions with respect to x, we get

$$\text{(A.4)} \qquad \frac{\partial \log f_{\widetilde{U}|X=x}(\widetilde{u})}{\partial x} = \frac{\partial \log f_U(g(\widetilde{u},x))}{\partial u}\,\frac{\partial g(\widetilde{u},x)}{\partial x} + \frac{\partial \log\left|\dfrac{\partial g(\widetilde{u},x)}{\partial \widetilde{u}}\right|}{\partial x}.$$

We will get expressions for the terms in the right-hand side of (A.4) in terms of $\widetilde{g}$, u, and x. By (A.1) and (A.2),

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

Differentiating with respect to $\widetilde{u}$ the expression

$$\widetilde{u} = \widetilde{g}(g(\widetilde{u},x),x)$$

one gets

$$I = \frac{\partial \widetilde{g}(g(\widetilde{u},x),x)}{\partial u}\frac{\partial g(\widetilde{u},x)}{\partial \widetilde{u}}.$$

Hence,

$$\text{(A.5)} \qquad 1 = \left|\frac{\partial g(\widetilde{g}(u,x),x)}{\partial \widetilde{u}}\right|\left|\frac{\partial \widetilde{g}(u,x)}{\partial u}\right|.$$

Taking logs and differentiating both sides of (A.5) with respect to x and with respect to u we get

$$0 = \frac{\partial \log\left|\dfrac{\partial g(\widetilde{g}(u,x),x)}{\partial \widetilde{u}}\right|}{\partial \widetilde{u}}\,\frac{\partial \widetilde{g}(u,x)}{\partial x} + \frac{\partial \log\left|\dfrac{\partial g(\widetilde{g}(u,x),x)}{\partial \widetilde{u}}\right|}{\partial x} + \frac{\partial \log\left|\dfrac{\partial \widetilde{g}(u,x)}{\partial u}\right|}{\partial x}$$

and

$$0 = \frac{\partial \log \left| \dfrac{\partial g(\widetilde{g}(u,x),x)}{\partial \widetilde{u}} \right|}{\partial \widetilde{u}} \frac{\partial \widetilde{g}(u,x)}{\partial u} + \frac{\partial \log \left| \dfrac{\partial \widetilde{g}(u,x)}{\partial u} \right|}{\partial u}.$$

Hence,

$$\begin{aligned} &\frac{\partial \log \left| \dfrac{\partial g(\widetilde{u},x)}{\partial \widetilde{u}} \right|}{\partial x} \\ &\quad = \frac{\partial \log \left| \dfrac{\partial \widetilde{g}(u,x)}{\partial u} \right|}{\partial u} \left(\frac{\partial \widetilde{g}(u,x)}{\partial u} \right)^{-1} \frac{\partial \widetilde{g}(u,x)}{\partial x} - \frac{\partial \log \left| \dfrac{\partial \widetilde{g}(u,x)}{\partial u} \right|}{\partial x}. \end{aligned}$$

Substituting into (A.4), we get

$$\begin{aligned} &\frac{\partial \log f_{\widetilde{U}|X=x}(\widetilde{u})}{\partial x} \\ &\quad = \frac{\partial \log f_U(g(\widetilde{u},x))}{\partial u} \frac{\partial g(\widetilde{u},x)}{\partial x} + \frac{\partial \log \left| \dfrac{\partial g(\widetilde{u},x)}{\partial \widetilde{u}} \right|}{\partial x} \\ &\quad = -\frac{\partial \log f_U(g(\widetilde{u},x))}{\partial u} \left(\frac{\partial \widetilde{g}(u,x)}{\partial u} \right)^{-1} \left(\frac{\partial \widetilde{g}(u,x)}{\partial x} \right) \\ &\qquad + \frac{\partial \log \left| \dfrac{\partial \widetilde{g}(u,x)}{\partial u} \right|}{\partial u} \left(\frac{\partial \widetilde{g}(u,x)}{\partial u} \right)^{-1} \frac{\partial \widetilde{g}(u,x)}{\partial x} \\ &\qquad - \frac{\partial \log \left| \dfrac{\partial \widetilde{g}(u,x)}{\partial u} \right|}{\partial x}. \end{aligned}$$

Substituting $g(\widetilde{u},x)$ by u, and setting $\partial \log f_{\widetilde{U}|X=x}(\widetilde{u})/\partial x = 0$, we get (6.6). *Q.E.D.*

PROOF OF THEOREM 6.2: Let

$$\widetilde{s}_u = \left(\frac{\partial \log f_{\widetilde{U}|X=x}(\widetilde{u})}{\partial \widetilde{u}} \bigg|_{\widetilde{u}=\widetilde{g}(u,x)} \right), \quad \widetilde{s}_x = \left(\frac{\partial \log f_{\widetilde{U}|X=x}(\widetilde{u})}{\partial x} \bigg|_{\widetilde{u}=\widetilde{g}(u,x)} \right),$$

$$s_u = \frac{\partial \log f_U(u)}{\partial u},$$

$$\widetilde{g}_u = \frac{\partial \widetilde{g}(u,x)}{\partial u}, \quad \widetilde{g}_x = \frac{\partial \widetilde{g}(u,x)}{\partial x},$$

$$\Delta_u = \frac{\partial \log\left(\left|\frac{\partial \widetilde{g}(u,x)}{\partial u}\right|\right)}{\partial u}, \quad \Delta_x = \frac{\partial \log\left(\left|\frac{\partial \widetilde{g}(u,x)}{\partial u}\right|\right)}{\partial x}.$$

Taking logs of both sides of (6.5) and differentiating with respect to u and x, we get

$$\widetilde{s}'_u \widetilde{g}_u + \Delta'_u = s'_u,$$
$$\widetilde{s}'_x + \widetilde{s}'_u \widetilde{g}_x + \Delta'_x = 0$$

or, after transposing,

$$\text{(A.6)} \qquad \widetilde{g}'_u \widetilde{s}_u + \Delta_u = s_u,$$
$$\text{(A.7)} \qquad \widetilde{s}_x + \widetilde{g}'_x \widetilde{s}_u + \Delta_x = 0.$$

Using the first equality to solve for $\widetilde{s}_u$ and substituting the result into the second equality, one gets

$$\widetilde{s}_x = -\widetilde{g}'_x (\widetilde{g}'_u)^{-1} s_u + \widetilde{g}'_x (\widetilde{g}'_u)^{-1} \Delta_u - \Delta_x.$$

By Theorem 6.1, $\widetilde{U}$ is independent of X if and only if $\widetilde{s}_x = 0$. Consider the matrix

$$\text{(A.8)} \qquad \begin{pmatrix} \widetilde{g}'_u & s_u - \Delta_u \\ \widetilde{g}'_x & -\Delta_x \end{pmatrix}.$$

Observational equivalence implies independence between $\widetilde{U}$ and X, and, by Theorem 6.1, that $\widetilde{s}_x = 0$. When $\widetilde{s}_x = 0$, equations (A.6) and (A.7) imply that $s_u - \Delta_u$ is a linear combination of the columns of $\widetilde{g}'_u$, and that $-\Delta_x$ is that same linear combination, but of the columns of $\widetilde{g}'_x$. Since $\widetilde{g}_u$ is invertible, this implies that the rank of the matrix must be G. Hence, observational equivalence implies that the rank of (A.8) is G.

Conversely, suppose that the rank of the matrix in (A.8) is G. It follows by the invertibility of $\widetilde{g}_u$ that there exists a unique $\lambda \in R^G$ such that

$$\widetilde{g}'_u \lambda = s_u - \Delta_u.$$

By (A.6),

$$\widetilde{g}'_u \widetilde{s}_u = s_u - \Delta_u.$$

Hence,

$$\lambda = \widetilde{s}_u.$$

Since the matrix in (A.8) has rank G, it must be that

$$-\Delta_x = \widetilde{g}'_x \lambda.$$

By solving for λ from the first equation, we get that

$$\lambda = (\widetilde{g}'_u)^{-1}(s_u - \Delta_u).$$

Then it follows that

$$\begin{aligned} -\Delta_x &= \widetilde{g}'_x \lambda \\ &= \widetilde{g}'_x (\widetilde{g}'_u)^{-1}(s_u - \Delta_u). \end{aligned}$$

This implies that

$$\widetilde{s}_x = \widetilde{g}'_x (\widetilde{g}'_u)^{-1}(s_u - \Delta_u) + \Delta_x = 0.$$

By Theorem 6.1, it follows that $\widetilde{U}$ is independent of X. By the equivalence between (6.4) and (6.5) and the definition of observational equivalence in (3.1), $(\widetilde{g}(r(y,x),x), f_{\widetilde{U}})$ is observationally equivalent to (r, f_U). *Q.E.D.*

REFERENCES

AI, C., AND X. CHEN (2003): "Efficient Estimation of Models With Conditional Moments Restrictions Containing Unknown Functions," *Econometrica*, 71, 1795–1843. [946]

ALTONJI, J. G., AND H. ICHIMURA (2000): "Estimating Derivatives in Nonseparable Models With Limited Dependent Variables," Mimeo, Northwestern University. [946]

ALTONJI, J. G., AND R. L. MATZKIN (2001): "Panel Data Estimators for Nonseparable Models With Endogenous Regressors," Working Paper T0267, NBER. [946]

——— (2003): "Cross Section and Panel Data Estimators for Nonseparable Models With Endogenous Regressors," Mimeo, Northwestern University. [946]

——— (2005): "Cross Section and Panel Data Estimators for Nonseparable Models With Endogenous Regressors," *Econometrica*, 73, 1053–1102. [946]

BENKARD, C. L., AND S. BERRY (2006): "On the Nonparametric Identification of Nonlinear Simultaneous Equations Models: Comment on B. Brown (1983) and Roehrig (1988)," *Econometrica*, 74, 1429–1440. [946,971]

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*, Vol. II, ed. by M. Dewatripont, L. P. Hansen, and S. J. Turnovsky. Cambridge, U.K.: Cambridge University Press, 312–357. [947]

BOWDEN, R. (1973): "The Theory of Parametric Identification," *Econometrica*, 41, 1069–1074. [946]

BROWN, B. W. (1983): "The Identification Problem in Systems Nonlinear in the Variables," *Econometrica*, 51, 175–196. [946,950,956,969]

BROWN, D. J., AND R. L. MATZKIN (1998): "Estimation of Nonparametric Functions in Simultaneous Equations Models, With and Application to Consumer Demand," CFDP 1175, Cowles Foundation for Research in Economics, Yale University. [946]

CHERNOZHUKOV, V., AND C. HANSEN (2005): "An IV Model of Quantile Treatment Effects," *Econometrica*, 73, 245–261. [946]

CHERNOZHUKOV, V., G. IMBENS, AND W. NEWEY (2007): "Instrumental Variable Estimation of Nonseparable Models," *Journal of Econometrics*, 139, 4–14. [946]

CHESHER, A. (2003): "Identification in Nonseparable Models," *Econometrica*, 71, 1405–1441. [946,964]

——— (2005): "Nonparametric Identification Under Discrete Variation," *Econometrica*, 73, 1525–1550. [947]

——— (2007): "Endogeneity and Discrete Outcomes," Mimeo, CEMMAP. [947]

DARROLLES, S., J. P. FLORENS, AND E. RENAULT (2003): "Nonparametric Instrumental Regression," IDEI Working Paper 228, University of Toulouse I. [946]

FISHER, F. M. (1959): "Generalization of the Rank and Order Conditions for Identifiability," *Econometrica*, 27, 431–447. [946]

——— (1961): "Identifiability Criteria in Nonlinear Systems," *Econometrica*, 29, 574–590. [946]

——— (1965): "Identifiability Criteria in Nonlinear Systems: A Further Note," *Econometrica*, 33, 197–205. [946]

——— (1966): *The Identification Problem in Econometrics*. New York: McGraw–Hill. [946]

FRISCH, R. A. K. (1934): "Statistical Confluence Analysis by Means of Complete Regression Systems," Publication 5, Universitets Okonomiske Institutt, Oslo. [946]

——— (1938): "Statistical versus Theoretical Relations in Economic Macrodynamics," Memorandum prepared for a conference at Cambridge, England, July 18–20, 1938, to discuss drafts of Tinbergen's League of Nations publications. [946]

HAAVELMO, T. M. (1943): "The Statistical Implications of a System of Simultaneous Equations," *Econometrica*, 11, 1. [946]

——— (1944): "The Probability Approach in Econometrics," *Econometrica*, 12, Supplement (July). [946]

HALL, P., AND J. L. HOROWITZ (2005): "Nonparametric Methods for Inference in the Presence of Instrumental Variables," *Annals of Statistics*, 33, 2904–2929. [946]

HURWICZ, L. (1950a): "Generalization of the Concept of Identification," in *Statistical Inference in Dynamic Economic Models*, Cowles Commission Monograph, Vol. 10, ed. by T. C. Koopmans. New York: Wiley, 245–257. [946]

——— (1950b): "Systems With Nonadditive Disturbances," in *Statistical Inference in Dynamic Economic Models*, Cowles Commission Monograph, Vol. 10, ed. by T. C. Koopmans. New York: Wiley, 410–418. [946]

IMBENS, G. W., AND W. K. NEWEY (2003): "Identification and Estimation of Triangular Simultaneous Equations Models Without Additivity," Mimeo, MIT. [946,964]

KOOPMANS, T. C., AND O. REIERSOL (1950): "The Identification of Structural Characteristics," *Annals of Mathematical Statistics*, 21, 165–181. [946]

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, Vol. 10, ed. by T. C. Koopmans. New York: Wiley, 53–237. [946]

MATZKIN, R. L. (1999): "Nonparametric Estimation of Nonadditive Random Functions," Mimeo, Northwestern University. [963]

——— (2003): "Nonparametric Estimation of Nonadditive Random Functions," *Econometrica*, 71, 1339–1375. [946,963]

——— (2004): "Unobservable Instruments," Mimeo, Northwestern University. [946,947]

——— (2005): "Identification in Nonparametric Simultaneous Equations," Mimeo, Northwestern University. [956,960]

——— (2007a): "Nonparametric Identification," in *Handbook of Econometrics*, Vol. 6B, ed. by J. J. Heckman and E. E. Leamer. Amsterdam: North-Holland, 5307–5368. [947,960]

——— (2007b): "Heterogeneous Choice," in *Advances in Economics and Econometrics, Theory and Applications, Ninth World Congress*, Vol. III, ed. by R. Blundell, W. Newey, and T. Persson. Cambridge, U.K.: Cambridge University Press, 75–110. [960]

——— (2008): "Nonparametric Estimation in Simultaneous Equations Models," Mimeo, UCLA. [960,964]

NEWEY, W. K., AND J. L. POWELL (1989): "Instrumental Variables Estimation for Nonparametric Models," Mimeo, Princeton University. [946]

——— (2003): "Instrumental Variables Estimation for Nonparametric Models," *Econometrica*, 71, 1557–1569. [946]

NEWEY, W. K., J. L. POWELL, AND F. VELLA (1999): "Nonparametric Estimation of Triangular Simultaneous Equations Models," *Econometrica*, 67, 565–603. [946]

NG, S., AND J. PINSKE (1995): "Nonparametric Two-Step Estimation of Unknown Regression Functions When the Regressors and the Regression Error Are Not Independent," Mimeo, University of Montreal. [946]

PINSKE, J. (2000): "Nonparametric Two-Step Regression Estimation When Regressors and Errors Are Dependent," *Canadian Journal of Statistics*, 28, 289–300. [946]

ROEHRIG, C. S. (1988): "Conditions for Identification in Nonparametric and Parametric Models," *Econometrica*, 56, 433–447. [946,950,956]

ROTHENBERG, T. J. (1971): "Identification in Parametric Models," *Econometrica*, 39, 577–592. [946]

TINBERGEN, J. (1930): "Bestimmung und Deutung con Angebotskurven: Ein Beispiel," *Zeitschrift fur Nationalokonomie*, 70, 331–342. [946]

WALD, A. (1950): "Note on Identification of Economic Relations," in *Statistical Inference in Dynamic Economic Models*, Cowles Commission Monograph, Vol. 10, ed. by T. C. Koopmans. New York: Wiley, 238–244. [946]

WEGGE, L. L. (1965): "Identifiability Criteria for a System of Equations as a Whole," *The Australian Journal of Statistics*, 7, 67–77. [946]

WORKING, E. J. (1927): "What Do Statistical 'Demand Curves' Show?" *Quarterly Journal of Economics*, 41, 212–235. [946]

WORKING, H. (1925): "The Statistical Determination of Demand Curves," *Quarterly Journal of Economics*, 39, 503–543. [946]

VYTLACIL, E., AND N. YILDIZ (2007): "Dummy Endogenous Variables in Weakly Separable Models," *Econometrica*, 75, 757–779. [946]

Dept. of Economics, 8283 Bunche Hall, University of California, Los Angeles, Los Angeles, CA 90095, U.S.A.; matzkin@econ.ucla.edu.

Manuscript received July, 2005; final revision received March, 2008.