Econometrica, Vol. 74, No. 5 (September, 2006), 1429–1440

ON THE NONPARAMETRIC IDENTIFICATION OF NONLINEAR SIMULTANEOUS EQUATIONS MODELS: COMMENT ON BROWN (1983) AND ROEHRIG (1988)

BY C. LANIER BENKARD AND STEVEN BERRY[1]

This note revisits the identification theorems of Brown (1983) and Roehrig (1988). We describe an error in the proofs of the main identification theorems in these papers, and provide an important counterexample to the theorems on the identification of the reduced form. Specifically, the reduced form of a nonseparable simultaneous equations model is not identified even under the assumptions of these papers. We provide conditions under which the reduced form is identified and is recoverable using the distribution of the endogenous variables conditional on the exogenous variables. However, these conditions place substantial limitations on the structural model. We conclude the note with a conjecture that it may be possible to use classical exclusion restrictions to recover some of the key implications of the theorems in more general settings.

KEYWORDS: Econometric theory, simultaneity, structural model, reduced form, triangular systems.

1. INTRODUCTION

IN THIS NOTE, we reconsider the nonparametric identification of nonlinear simultaneous equations models as in Brown (1983) and Roehrig (1988). In revisiting this literature we have discovered that a key condition (Brown (1983, pp. 180–181), cited by Roehrig (1988, p. 438)) used in the proofs of the primary theorems of both Brown and Roehrig is false. This finding is substantive. An important implication of this condition, that the model's reduced form is identified under assumptions much weaker than the structural model, can be shown to be false: see Section 4.1 for a counterexample.

The counterexample also contradicts the main theorems in both Brown and Roehrig. We have as yet been unable to correct the theorems ourselves. However, we remain optimistic that some essential features of the theorems may still be true and that the theorems may be able to be corrected with some modifications to the assumptions. Indeed, Matzkin (2005) makes considerable progress in this direction. We view these theorems as having important implications for empirical work in economics and we therefore hope that future research will recover many of their implications.

Note that Brown (1983) and Roehrig (1988) are widely cited in the literature on nonparametric identification (some recent examples include Newey, Powell, and Vella (1999), Angrist, Graddy, and Imbens (2000), Guerre, Perrigne, and Vuong (2000), Athey and Haile (2002), Imbens and Newey (2003), Chesher (2003), Matzkin (2003), and Newey and Powell (2003)). Their identification

[1]We have had helpful conversations with Pat Bayer, Don Brown, Yossi Feinberg, Guido Imbens, Yuliy Sannikov, Andy Skrzypacz, and Chris Timmons. We also thank the co-editor and four anonymous referees for their useful comments. Any remaining errors are our own.

theorems also play a key role in a few important papers in this literature, including Brown and Wegkamp (2002) and Brown and Matzkin (1998).

We begin the note with an outline of the model and a statement of the primary assumptions of Brown (1983). We continue with a statement of the false derivative condition, as well as some intuition behind its failure. Next we outline the role of the derivative condition in identification and present an important counterexample to the identification theorems regarding identification of the model's reduced form. We then outline a substantial set of restrictions under which the reduced form is identified and is recoverable using the distribution of the endogenous variables conditional on the exogenous shifters. We conclude with a discussion of why we believe that the main identification theorems of Brown and Roehrig may still hold more generally with some modifications to the assumptions, and we suggest directions for future research.

2. MODEL AND ASSUMPTIONS

The model is characterized by a set of exogenous variables (X, U) and a set of endogenous variables Y. The random vector $X \in \mathbb{R}^K$ denotes the observed exogenous variables, while $U \in \mathbb{R}^L$ are not observed. We assume that (X, U) is generated according to a continuous distribution function Φ with positive density everywhere on its support. The endogenous variables $Y \in \mathbb{R}^G$ are observed and are subsequently defined. Realizations of the random variables are denoted using lowercase letters, e.g., (x, y, u).

Brown considers a parametric system of structural equations that is nonlinear in the observed variables (X, Y), but that can be written as linear in an unknown parameter vector and an additively separable error term U. Roehrig relaxes Brown's framework to allow for a nonparametric system of structural equations with a nonseparable error. Our results apply equally to both frameworks. Indeed, they apply to any system of structural equations that is sufficiently general as to generate a reduced form that is nonseparable in the errors. Thus, because it makes the exposition cleaner, we will present our results in the context of a nonseparable parametric structural model. In the model, the endogenous variables Y are defined implicitly as the solution to a set of structural equations

$$Y = m(X, Y, U; \theta), \tag{2.1}$$

where $m(\cdot)$ is a known function and θ is an unknown parameter vector.[2] We emphasize that our results hold in a more general setting in which $m(\cdot)$ is unknown (i.e., θ could be infinite dimensional).

[2]Note that because the function $m(\cdot)$ could contain the term Y, this specification is nonrestrictive. We write the model in this form because it matches the form of many commonly used econometric models.

A leading example of such a system is the nonseparable supply and demand model

$$Q = D(Z, P, \varepsilon_D; \theta_D), \tag{2.2}$$

$$P = S(W, Q, \varepsilon_S; \theta_S), \tag{2.3}$$

where, in the general notation, $Y = (Q, P)$, $X = (Z, W)$, $U = (\varepsilon_D, \varepsilon_S)$, $\theta = (\theta_D, \theta_S)$, and the structural model is $m = (D, S)$.

We maintain the assumption that the structural equations are continuously differentiable. We assume for purposes of identification that the joint distribution of the endogenous variables and the observed exogenous shifters, $\Psi(X, Y)$, is known.

Following Roehrig (1988), we will call a structure S a pair (θ, Φ) that together define the data generating process. We denote the true data generating process by the structure $S_0 \equiv (\theta_0, \Phi_0) \in \Omega$, where the set Ω consists of all structures S that have the characteristics known a priori to apply to S_0. These definitions will allow us to define identification formally. Again following Roehrig (1988), we have the following definitions:

DEFINITION 2.1: Let Ψ and Ψ' be the distribution functions for (X, Y) implied by the structures S and S'. Then S and S' are observationally equivalent if $\Psi = \Psi'$.

DEFINITION 2.2: Structure S is identified in Ω if there is no other $S' \in \Omega$ that is observationally equivalent to S.

We are also interested in cases where the structure S is not identified, but some characteristics of the structure are identified. Let $C(S)$ be a set of characteristics of S. For example, $C(S)$ may consist of functionals of the model defined by θ.

DEFINITION 2.3: The characteristics $C(S)$ are identified if all $S' \in \Omega$ that are observationally equivalent to S have characteristics $C(S)$.

2.1. *The Three Basic Assumptions*

We divide Brown and Roehrig's assumptions into two groups. The first three assumptions, which we call the basic assumptions, are those used by Brown to prove the derivative condition in dispute. The remaining assumptions, which we omit from this note, are rank conditions similar to those used to identify the structural equations in a linear model.

ASSUMPTION 1—Reduced Form: *The model generates a unique continuously differentiable reduced form*

$$Y = f(X, U; \theta). \tag{2.4}$$

Assumption 1 implies that the joint distribution of the observed variables, $\Psi(X, Y)$, is uniquely determined by any structure $S = (\theta, \Phi)$. Note that Brown and Roehrig assume that the structural model gives a unique reduced form, but their proofs require only the assumption that the data are generated by a differentiable function that maps X and U into Y. Such a generalization would be important for models with multiple equilibria.[3]

ASSUMPTION 2—Solution for U: *There is a unique solution to the structural equations that gives the unobservables as a function of the observables. That is, given a structure* $S = (\theta, \Phi)$, *there exists a* $\rho(Y, X; \theta)$ *such that* $Y = m(X, Y, \rho(Y, X; \theta); \theta)$. *Furthermore, the function* $\rho(Y, X; \theta)$ *is continuously differentiable.*

Assumption 2 places strong restrictions on the way the unobservables enter the model and on the relationship between the dimension of the error (L) and the dimension of Y (G). In particular, it would typically require that $L \leq G$. Existence of a residual function ρ also implies that, although the error may be nonseparable in the structural equations, there is a transformation of the structural equations that is linear in the errors.

In addition to the foregoing assumptions about the model, a stochastic restriction on the errors is required:

ASSUMPTION 3—Independence: *The observed exogenous shifters X are independent of the unobserved errors U.*

3. THE DERIVATIVE CONDITION

The following derivative condition, which we will show to be false, plays a key role in the proofs of the main theorems in Brown (1983) and Roehrig (1988).

LEMMA 3.1—Derivative Condition (Brown (1983, pp. 180–181)): *Let X and U be independent random vectors and let* $\tilde{U} = T(X, U)$, *where* $T : \mathbb{R}^K \times \mathbb{R}^L \to \mathbb{R}^L$ *is everywhere differentiable. Then $\tilde{U}$ is independent of X if and only if* $\partial T(x, u)/\partial x = 0$ *everywhere.*

The derivative condition seems intuitive at first and, as we will see subsequently, if it is true, it would be powerful. It is obviously true in one direction: if the derivative of T with respect to x is everywhere zero, then the mapping is not a function of x. Therefore, $\tilde{U}$ can be written as a function only of U. Because U is independent of X, it must be that $\tilde{U}$ is as well.

[3] However, if Assumption 1 were relaxed in this way, then each structure may yield more than one possible observed distribution Ψ. Thus, it would also be necessary to change our definition of identification to account for this possibility.

The problem with the derivative condition is in the other direction. Although the condition is true if the errors are univariate (see subsequent text), in multidimensional spaces it is easy to generate mappings T that have nonzero derivatives with respect to x but that still generate independent errors. We now provide a specific example of this with a clear graphical intuition. In the following section, we provide more substantive economic examples in the context of the identification arguments.

3.1. *Counterexample to the Derivative Condition*

There is a simple graphical intuition as to why the derivative condition fails. Suppose that U is two-dimensional and distributed $N(0, I)$, independent of X. Suppose that the T mapping is such that U is simply rotated about the origin by an amount determined by X. This would generate a new set of errors $\tilde{U}$ that are still $N(0, I)$ for every outcome $X = x$ and, therefore, are independent of X. However, the mapping has nonzero derivatives with respect to X almost everywhere.

More precisely, let $\Gamma(x)$ denote an orthonormal matrix that is a smooth function of x over the support of X.[4] Let $\tilde{U} = T(X, U)$, where $T(x, u) = \Gamma(x)'u$. Then $\tilde{U}$ is $N(0, \Gamma(X)'\Gamma(X)) = N(0, I)$, which does not depend on X. However,

$$\frac{\partial T(x, u)}{\partial x} = \frac{\partial \Gamma(x)'}{\partial x} u \neq 0.$$

In a single dimension, such rotations are not possible, and it is easy to show that the derivative condition holds in a single dimension.[5] However, in multiple dimensions, there are many transformations that can fold, reflect, or rotate the errors in such a way as to conserve independence. If the transformation is a function of the exogenous variables, then the derivative condition fails.

4. THE ROLE OF THE DERIVATIVE CONDITION IN IDENTIFICATION

In considering identification, an important insight of Brown (1983) was to focus on the residual function ρ.[6] For any candidate structure $S = (\theta, \Phi) \in \Omega$

[4]One example is

$$\Gamma(x) = \begin{pmatrix} \cos x & \sin x \\ -\sin x & \cos x \end{pmatrix}.$$

[5]Proving the univariate version of the derivative condition is analogous to showing identification of a single equation model that is nonseparable in the errors. In that case, Assumptions 1–3 are equivalent to the assumptions used in Matzkin (2003), which proves the result.

[6]Note that Matzkin (2005) has pursued a similar approach in deriving a corrected set of identification conditions.

that is observationally equivalent to S_0 and satisfies the three basic assumptions, we can substitute the true reduced form into the residual function to obtain a mapping from (x, u) into u,

$$\tilde{u}(x, u; \theta, \theta_0) = \rho(f(x, u; \theta_0), x; \theta), \tag{4.1}$$

where $\tilde{u}(x, u; \theta, \theta_0)$ is the error implied by the model evaluated at θ (and where $\tilde{u}(x, u; \theta_0, \theta_0) = u$).[7] Brown (1983), followed by Roehrig (1988), uses this relationship to provide conditions under which imposing independence on the errors in the candidate structure $\tilde{U} \equiv \tilde{u}(X, U; \theta, \theta_0)$ will identify certain characteristics of the true structure S_0.

According to the derivative condition above, the residuals from the candidate structure $\tilde{U}$ are independent of X if and only if the derivative of the mapping $\tilde{u}(x, u; \theta, \theta_0)$ with respect to x is everywhere zero, that is, if

$$\begin{aligned} &D_x\tilde{u}(x, u; \theta, \theta_0) \\ &\quad = D_y\rho(f(x, u; \theta_0), x; \theta)D_x f(x, u; \theta_0) + D_x\rho(f(x, u; \theta_0), x; \theta) \\ &\quad = 0 \end{aligned}$$

for all (x, u) pairs in the support of (X, U). This set of equations implies a set of restrictions to the candidate structure that help to identify important characteristics of the true structure S_0. For example, an important implication of the derivative condition is that the reduced form is identified, in a sense that we now make precise.

4.1. *Counterexample to Identification of the Reduced Form*

Consider any alternative structure $S = (\theta, \Phi) \in \Omega$ that is observationally equivalent to S_0 and satisfies the three basic assumptions, with reduced form

$$Y = f(X, \tilde{U}; \theta),$$

where $\tilde{U}$ is independent of X. The derivative condition implies that this alternative reduced form must have the same partial derivatives with respect to x at each point (x, y) as the true reduced form. The following corollary makes this statement precise:

[7]For example, in a single equation linear model, $y = x\beta_0 + u$ is the true functional relationship and the errors from an alternative structure β are given by

$$\tilde{u}(x, u; \beta, \beta_0) = y - x\beta = x(\beta_0 - \beta) + u.$$

COROLLARY 4.1: *If the derivative condition were true, then under the three basic assumptions, the characteristics*

$$\left.\frac{\partial f(x,u;\theta)}{\partial x}\right|_{x=x_0,u=\rho(y_0,x_0;\theta)}$$

would be identified for every point (x_0, y_0) *in the support of* (X, Y).

PROOF: Consider any $S = (\theta, \Phi) \in \Omega$ that is observationally equivalent to S_0 and satisfies the three basic assumptions. By the definition of $\tilde{u}$, we have that $f(x, \tilde{u}(x, u; \theta, \theta_0); \theta) = f(x, u; \theta_0)$. By independence of X and $\tilde{U}$, the derivative condition implies that $\partial\tilde{u}(x, u; \theta, \theta_0)/\partial x = 0$ everywhere. Therefore, the chain rule gives

$$\begin{aligned}\frac{\partial f(x,u;\theta_0)}{\partial x}&\\ &= \left.\frac{\partial f(x,u;\theta)}{\partial x}\right|_{u=\tilde{u}(x,u;\theta,\theta_0)} + \left.\frac{\partial f(x,u;\theta)}{\partial u}\right|_{u=\tilde{u}(x,u;\theta,\theta_0)} \frac{\partial \tilde{u}(x,u;\theta,\theta_0)}{\partial x}\\ &= \left.\frac{\partial f(x,u;\theta)}{\partial x}\right|_{u=\tilde{u}(x,u;\theta,\theta_0)}.\end{aligned}$$

Evaluating at $x = x_0$ and $u = \rho(y_0, x_0; \theta_0)$ gives the result, because $\tilde{u}(x_0, \rho(y_0, x_0; \theta_0); \theta, \theta_0) = \rho(y_0, x_0; \theta)$. *Q.E.D.*

The corollary implies that if the derivative condition held, then we would automatically have the result that the reduced form derivatives with respect to x are identified in any model that satisfies the three basic assumptions. However, using the example from Section 3.1, it is easy to see that, in fact, the basic assumptions (Assumption 1–3) are not sufficient for identification of the reduced form derivatives. Suppose that Y and U have dimension 2, and the reduced form is given by

$$Y = \Gamma(X)'U,$$

where, as in Section 3.1, $\Gamma(x)$ is an orthonormal matrix that is a smooth function of x and U is distributed $N(0, I)$. Then the basic assumptions are satisfied for any $\Gamma(x)$, but the reduced form derivative at the point (x_0, y_0) is given by

$$\left.\frac{\partial f(x,u;\theta)}{\partial x}\right|_{x=x_0,u=\rho(y_0,x_0;\theta)} = \frac{\partial \Gamma(x_0)}{\partial x}\Gamma(x_0)'y_0,$$

which clearly depends on the function Γ. For example, if $\Gamma(x) = I$, then the reduced form derivative with respect to x is always zero, but in general $\partial\Gamma(x_0)/\partial x$ would not be zero.

4.2. *Recovering the Reduced Form Using the Triangular Construction*

If the derivative condition and its corollary above were true, to recover the reduced form, we would need only to find a model that was observationally equivalent to S_0, satisfied Assumptions 1 and 2, and generated errors that were independent of X. One such model is the triangular construction. Consider the supply and demand model in (2.2) and (2.3), and let $\Psi(Q, P, Z, W)$ be the (known) joint distribution of the endogenous variables and exogenous shifters. Suppose we let

$$\tilde{U}_1 = \Psi(Q|Z, W), \qquad (4.2)$$
$$\tilde{U}_2 = \Psi(P|Z, W, Q),$$

where the $\tilde{U}$'s are constructed such that they are independent of one another as well as independent of (Z, W), satisfying Assumption 3. The construction itself satisfies Assumption 2, and it is easy to invert the system to retrieve a reduced form that satisfies Assumption 1.

If Brown and Roehrig's theorems were true, then this system should retrieve the true reduced form. However, it is easy to see that this cannot be the case in general. Consider the reduced form derivative of Q with respect to Z implied by the triangular system (obtained via the implicit function theorem)

$$\frac{\partial Q}{\partial Z} = -\frac{1}{\partial \Psi(Q|Z, W)/\partial Q} \frac{\partial \Psi(Q|Z, W)}{\partial Z}.$$

This expression is a function only of (Q, Z, W), but is not a function of P. This implies a restriction to the reduced form that need not hold in general. Specifically, the triangular construction implies that the reduced form takes the form

$$Q = f(Z, W, \tilde{U}_1),$$

whereas in the general model the reduced form takes the form

$$Q = f(Z, W, \varepsilon_D, \varepsilon_S).$$

The distinction here is that, in general, the two structural error terms enter the reduced form nonseparably. Note that this distinction is important because it is easy to come up with examples where this would be the case. The form implied by the triangular system suggests that the two errors enter only as a single index. In the general model, it would be possible to hold (Z, W) fixed and move $(\varepsilon_D, \varepsilon_S)$ together in such a way as to hold Q fixed (but changing P), changing the reduced form derivatives with respect to (Z, W). The triangular system does not allow for this possibility and, therefore, it cannot retrieve the true reduced form in general.

The triangular construction fails to retrieve the true reduced form for exactly the same reason that the derivative condition fails to hold. The errors in the triangular construction $\tilde{U}$ are independent of the exogenous variables (Z, W), but could have been transformed from the original errors in a way that depends on (Z, W). For example, they could have been rotated, as in the example from Section 3.1. Such transformations change the implicit functional relationship between Q and (Z, W), leading to incorrect reduced form derivatives.

An important but negative implication of these results is that, in general, without further assumptions, a projection of the endogenous variables on the exogenous variables (given by the conditional distribution $\Psi(Y|X)$) does not recover the true reduced form derivatives, even under the relatively strict Assumptions 1–3. The intuition for this failure is as follows: in a nonseparable system, for given values of Y_1 and X, there may be different derivatives of Y_1 with respect to X that depend on the values of the error terms in the other structural equations. A projection of Y_1 onto X that ignores the values of $Y_2, \ldots, Y_G$ (or alternatively, $U_2, \ldots, U_G$) recovers not the true reduced form, but something akin to the average reduced form derivative weighted over the distribution of the left out variables. As the example in Section 3.1 shows, this problem can even be so severe that the effect of X on Y disappears completely.

5. USING THE CONDITIONAL DISTRIBUTION TO RECOVER THE TRUE REDUCED FORM DERIVATIVES

We have shown that the triangular system does not necessarily recover the true reduced form of a nonseparable system and that, therefore, the conditional distribution $\Psi(Y|X)$ does not necessarily recover the true reduced form derivatives. Here we present restrictions under which the conditional distribution does recover the true reduced form derivatives.

Suppose that in the true reduced form the errors enter each equation as a single index,

$$
\begin{aligned}
&Y_1 = f_1(X, g_1(U); \theta_0), \\
&Y_2 = f_2(X, g_2(U); \theta_0), \\
&\vdots \\
&Y_G = f_G(X, g_G(U); \theta_0),
\end{aligned}
\tag{5.1}
$$

where each function $g_i : \mathbb{R}^L \to \mathbb{R}$ and where each function f_i is increasing in its last argument. Under these assumptions, the reduced form functions f_i are identified and their derivatives with respect to x can be recovered using the conditional distribution $\Psi(Y|X)$ (the proof is an application of Matzkin (2003)).

The question remains as to what structural models would lead to this single index property. We have found it difficult to characterize the entire class of models that have this property. Furthermore, there are likely to be special cases where the errors enter the reduced form as a single index only for certain parameter values. However, there is a simple class of models that always have this property. Suppose that the structural model takes the form

$$A(\theta)\begin{bmatrix} h_1(Y_1, X; \theta) \\ h_2(Y_2, X; \theta) \\ \vdots \\ h_G(Y_G, X; \theta) \end{bmatrix} - r(U; \theta) = 0, \tag{5.2}$$

where θ is a parameter vector, $\theta = \theta_0$ is the parameter vector that denotes the true data generating process, $A(\theta)$ is a $G \times G$ matrix with $|A(\theta_0)| \neq 0$, $r(U; \theta)$ is a known $G \times 1$ vector function, and the functions h_i are known and strictly monotonic in Y_i at $\theta = \theta_0$. The true reduced form (the data generating process) of this system is

$$Y = h^{-1}\big(X, A(\theta_0)^{-1} r(U; \theta_0); \theta_0\big),$$

which satisfies the assumptions above.

This class of models includes linear systems as a special case. It also includes systems that are linear in logs, some polynomial systems (that maintain the required equation-by-equation monotonicity in Y), as well as many other nonlinear models, and systems that consist of mixtures of all of these types. However, the system represents a substantial restriction to the system considered by Brown (1983) because the system restricts the way that the endogenous variables Y_i can enter across different equations. Each variable Y_i must enter every equation in essentially the same way. For example, in the supply and demand example, if the demand function specifies that the logarithm of Q is a linear function of P, then the foregoing class requires that the supply function has P as a linear function of the logarithm of Q. This is a relatively narrow class of models compared with the general case in (2.1).

In all systems in this class, the conditional distribution recovers the true reduced form derivatives. However, recall that it is not necessary to assume that the structural model has this form to be able to obtain identification of the reduced form. It is only necessary to make the weaker assumption that the reduced form has the single index property.

6. CONCLUSIONS AND AREAS FOR FUTURE RESEARCH

So far we have shown that the Brown/Roehrig identification theorems are incorrect as stated. We have also shown that one consequence is that additional assumptions beyond those listed in Section 2 are required to be able to obtain identification of the reduced form.

However, despite these results, we remain optimistic that some version of the Brown/Roehrig theorems can be established, perhaps under stronger conditions (see also Matzkin (2005)). The spirit of Brown's and Roehrig's rank conditions is that exclusion restrictions can be used to obtain identification of the system. We have been unable to contradict this notion. Brown's and Roehrig's proofs utilize only the rank conditions assumptions to identify the structure from the reduced form, but it is possible that similar exclusion restrictions may help to identify the reduced form as well. In particular, exclusion restrictions provide restrictions *across points*. Brown's and Roehrig's proofs of identification of the reduced form do not rely on any cross-point restrictions, which is the primary reason that the theorems fail.

One set of exclusion restrictions that does identify the system is if the true model is triangular. For the sake of brevity we provide only an outline of the proof. Consider the triangular system shown in (4.2) and suppose that this system represents the true form of the structural model, i.e., suppose that price was excluded from the demand equation in the true model. Then the first equation can be shown to be identified using single equation methods. Similarly, the second equation can be shown to be identified conditional on the first. This logic also easily generalizes to higher-dimensional systems. Note that Imbens and Newey (2003) consider a two-dimensional system of this kind, and Chesher (2003) considers a similar multidimensional triangular system.

The triangular system uses exclusion restrictions only on the endogenous variables. However, it is possible that traditional exclusion restrictions on the exogenous variables or on groups of both endogenous and exogenous variables might also yield identification of the system. It seems likely that such restrictions would rule out the kinds of transformations that cause the derivative condition to fail. However, we have as yet been unable to verify or contradict this conjecture ourselves, and its seems likely that a proof may require more complex arguments than those used in the original papers.

A corrected set of identification theorems similar to those of Brown and Roehrig would provide a simple yet powerful method for proving identification for a large class of structural models. Thus, our hope is that these issues will be sorted out in future research.

Graduate School of Business, Stanford University, 518 Memorial Way, Stanford, CA 94305-5015, U.S.A.; lanierb@stanford.edu

and

Dept. of Economics, Yale University, 37 Hillhouse Ave., New Haven, CT 06520-8264, U.S.A.; steven.berry@yale.edu.

Manuscript received September, 2004; final revision received April, 2006.

REFERENCES

ANGRIST, J. D., K. GRADDY, AND G. W. IMBENS (2000): "'The Interpretation of Instrumental Variables Estimators in Simultaneous Equations Models with an Application to the Demand for Fish," *Review of Economic Studies*, 67, 499–527. [1429]

ATHEY, S., AND P. A. HAILE (2002): "Identification of Standard Auction Models," *Econometrica*, 70, 2107–2140. [1429]

BROWN, B. W. (1983): "'The Identification Problem in Systems Nonlinear in the Variables," *Econometrica*, 51, 175–196. [1429,1430,1432-1434,1438]

BROWN, D. J., AND R. MATZKIN (1998): "Estimation of Nonparametric Functions in Simultaneous Equations Models, with an Application to Consumer Demand," Discussion Paper 1175, Cowles Foundation. [1430]

BROWN, D. J., AND M. H. WEGKAMP (2002): "Weighted Mean-Square Minimum Distance from Independence Estimation," *Econometrica*, 70, 2035–2051. [1430]

CHESHER, A. (2003): "Identification in Nonseparable Models," *Econometrica*, 71, 1405–1441. [1429,1439]

GUERRE, E., I. PERRIGNE, AND Q. VUONG (2000): "Optimal Nonparametric Estimation of First-Price Auctions," *Econometrica*, 68, 525–574. [1429]

IMBENS, G., AND W. K. NEWEY (2003): "Identification and Estimation of Triangular Simultaneous Equations Models Without Additivity," Mimeo, MIT. [1429,1439]

MATZKIN, R. (2003): "Nonparametric Estimation of Nonadditive Random Functions," *Econometrica*, 71, 1332–1375. [1429,1433,1437]

——— (2005): "Identification of Nonparametric Simultaneous Equations," Mimeo, Northwestern University. [1429,1433,1439]

NEWEY, W. K., AND J. L. POWELL (2003): "Instrumental Variable Estimation of Nonparametric Models," *Econometrica*, 71, 1565–1578. [1429]

NEWEY, W. K., J. L. POWELL, AND F. VELLA (1999): "Nonparametric Estimation of Triangular Simultaneous Equations Models," *Econometrica*, 67, 565–603. [1429]

ROEHRIG, C. S. (1988): "Conditions for Identification in Nonparametric and Parametric Models," *Econometrica*, 56, 433–447. [1429,1431,1432,1434]