Independence, Identification and Specification in Separable Econometric Models

Donald J. Brown and Marten H. Wegkamp

June 2003

*This is a revision of Cowles Foundation Discussion Paper No. 1395, "Tests of Independence in Separable Econometric Models," January 2003.

Introduction

Manski, C.F. (1983), "Closest Empirical Distribution Estimation," *Econometrica*, 51, 305–320.

Brown, B.W. (1983), "The Identification Problem in Systems Nonlinear in the Variables," *Econometrica*, 51, 175–196.

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

Brown, D.J. and R. Matzkin (1998), "Estimation of Nonparametric Functions in Simultaneous Equations Models, with an Application to Consumer Demand," Working Paper, Yale University.

Brown, D.J. and M. Wegkamp (2002), "Weighted Minimum Mean-Square Distance from Independence Estimation," *Econometrica*, 70, 2035–2051.

Brown, D.J. and M. Wegkamp (2003), "Tests of Independence in Separable Econometric Models," CFDP No. 1395.

Identification Test

Existence Test $\equiv$ Independence Test (Brown–Wegkamp)

Uiqueness Test $\equiv$ Uniqueness Theorem (Brown–Roehrig)

Specification Test $\equiv$ Existence Test + Uniqueness Test

Our specification test is analogous to the test for overidentifying restrictions in GMM.

Uniqueness Test

Theorem (Brown–Roehrig). If

(a) (X, Y, U) a triple of random vectors in $\mathbb{R}^{k_1} \times \mathbb{R}^{k_2} \times \mathbb{R}^{k_2}$

(b) $U = \rho(X, Y, \theta)$ has a unique global inverse $Y(X, U, \theta)$, i.e., $U \equiv \rho(X, Y(X, U, \theta), \theta)$ for all $\theta \in \Theta \subseteq \mathbb{R}^p$

(c) $(x, y) \to \rho(x, y, \theta)$ is a smooth map for each $\theta \in \Theta$

(d) $\tilde{U} = \rho(X, Y(X, U), \tilde{\theta})$

(e) U is independent of X

(f) The DGP is $\overline{\mathbb{P}}_X \times \overline{\mathbb{P}}_U \cdot (X, Y(X, U, \bar{\theta}))^{-1}$. Then $\tilde{U}$ is independent of X iff $\frac{\partial \rho}{\partial x}(X, Y(X, U, \bar{\theta}), \tilde{\theta}) \equiv 0$.

Brown–Roehrig Theorem (B.W. Brown's Proof). Suppose $(X, U, \tilde{U})$ is a triple of random vectors in $\mathbb{R}^{K_1} \times \mathbb{R}^{K_2} \times \mathbb{R}^{K_2}$, where $\tilde{U} = h(U, X)$ and h is C^1, i.e., $\partial h_j / \partial x_i|_{(u,x)}$ is continuous for all (u, x) and for all i, j. If the map $(u, x) \to (h(u, x), x) \equiv (\tilde{u}, x)$ has a global inverse $(\tilde{u}, x) \to (\ell(\tilde{u}, x), x) = (u, x)$ for some C^1 map $\ell : \mathbb{R}^{K_2} \times \mathbb{R}^{K_1} \to \mathbb{R}^{K_2}$. Call this property $(*)$.

Lemma. *Under Assumptions* (b) *and* (c) *the map* $h(u, x) \equiv (\rho(x, y(x, u, \bar{\theta}), \tilde{\theta}), x)$ *has property* $(*)$ $\forall(\bar{\theta}, \tilde{\theta}) \in \Theta \times \Theta$. *Moreover,* $\ell(\tilde{u}, x) = \rho(x, y(x, \tilde{u}, \tilde{\theta}), \bar{\theta})$.

Proof. Suppose $\exists(u_1, x_1)$ and (u_2, x_2) s.t. $h(u_1, x_1) = h(u_2, x_2)$. Hence $x_1 = x_2 \equiv x$ and $\rho(x_1, y(x_1, u_1, \bar{\theta}), \tilde{\theta}) = \rho(x_2, y(x_2, u_2, \bar{\theta}), \tilde{\theta}) \equiv \tilde{u} \Rightarrow \rho(x, y(x, u_1, \bar{\theta}), \tilde{\theta}) = \rho(x, y(x, u_2, \bar{\theta}), \tilde{\theta}) = \tilde{u}$. Therefore, $y(x, \tilde{u}, \tilde{\theta}) = y(x, u_1, \bar{\theta})$ and $y(x, \tilde{u}, \tilde{\theta}) = y(x, u_2, \bar{\theta})$. But $y(x, u_1, \bar{\theta}) = y(x, u_2, \bar{\theta})$ $\Rightarrow u_1 = u_2$. Hence the map $(u, x) \to (\rho(x, y(x, u, \bar{\theta}), \tilde{\theta}), x)$ $\equiv (\tilde{u}, x)$ is one-to-one $\forall(\bar{\theta}, \tilde{\theta}) \in \Theta \times \Theta$.

Now consider the map $\ell(\tilde{u}, x) \equiv \rho(x, y(x, \tilde{u}, \tilde{\theta}), \bar{\theta})$. If $\tilde{u} = \rho(x, y(x, u, \bar{\theta}), \tilde{\theta})$ then $y(x, \tilde{u}, \tilde{\theta}) = y(x, u, \bar{\theta})$ and $\rho(x, y(x, \tilde{u}, \tilde{\theta}), \bar{\theta}) = \rho(x, y(x, u, \bar{\theta}), \bar{\theta})$. But $\rho(x, y(x, u, \bar{\theta}), \bar{\theta}) = u$. Hence $\ell(\tilde{u}, x) = u$, i.e. $(\tilde{u}, x) \to (\ell(\tilde{u}, x), x)$ is the inverse of $(u, x) \to (h(u, x), x) \equiv (\tilde{u}, x)$.

Property $(*)$ $\Rightarrow$ Jacobian of the map $h(u, x)$, i.e., $\partial \tilde{u}/\partial u \equiv \partial h/\partial u|_{(u,x)}$ is nonsingular $\forall (u, x)$. Let $\tilde{h}(\tilde{u}, x) \equiv h{\cdot}(\ell(\tilde{u}, x), x)$, then $(h(u, x), x) \equiv (\tilde{u}, x) = (\tilde{h}(\tilde{u}, x), x)$. If $\partial h/\partial x \not\equiv 0$, then $\exists i, j$ s.t. $\partial \tilde{h}_j/\partial x_i|_{(\tilde{\bar{u}}, \bar{x})} > 0$ for some $(\bar{u}, \bar{x})$, where $\tilde{\bar{u}} = h(\bar{u}, \bar{x})$. By continuity of $\partial \tilde{h}_j/\partial x_i$, $\exists$ open balls $\mathbb{B}_\varepsilon(\bar{x})$ and $\mathbb{B}_\delta(\tilde{\bar{u}})$ s.t. $\forall (\hat{u}, \hat{x}) \in \mathbb{B}_\delta(\tilde{\bar{u}}) \times \mathbb{B}_\varepsilon(\bar{x})$: $\partial \tilde{h}_j/\partial x_i|_{(\tilde{\bar{u}}, \bar{x})} > 0$. Let $\bar{\bar{x}} = \bar{x} + \alpha e_i$, where $\alpha > 0$ and $\bar{\bar{x}} \in \mathbb{B}_\varepsilon(\bar{x})$. Then by the Mean Value Theorem: If $\tilde{u} \in \mathbb{B}_\delta(\tilde{\bar{u}})$: $\tilde{h}_j(\tilde{u}, \bar{\bar{x}}) - \tilde{h}_j(\tilde{u}, \bar{x}) = \partial \tilde{h}_j/\partial x_i|_{(\tilde{u}, \bar{x} + \zeta e_i)} \left\| \bar{\bar{x}} - \bar{x} \right\| > 0$, where $0 \leq \zeta \leq \alpha$, and $\left\| \bar{\bar{x}} - \bar{x} \right\| = \alpha$.

Hence $\tilde{h}_j(\tilde{u}, \bar{\bar{x}}) > \tilde{h}_j(\tilde{u}, \bar{x})\ \forall \tilde{u} \in \mathbb{B}_\delta(\bar{\tilde{u}})$. Therefore, $\int_{B_\delta(\bar{\tilde{u}})} \tilde{h}_j(\tilde{u}, \bar{\bar{x}}) d\tilde{u} > \int_{B_\delta(\bar{\tilde{u}})} \tilde{h}_j(\tilde{u}, \bar{x}) d\tilde{u}$. If $\tilde{h}(\tilde{U}, X)$ is independent of X, then any function of $h(\tilde{U}, X)$ is independent of X. In particular, the random variable

$$\gamma(\tilde{U}, X) = \begin{cases} \tilde{h}_j(\tilde{U}, X), & \text{if } (\tilde{u}, x) \in \mathbb{B}_\delta(\bar{\tilde{u}}) \times \mathbb{B}_\varepsilon(\bar{x}) \\ 0, & \text{otherwise} \end{cases}$$

must then be independent of X. Consequently the conditional distribution of γ, conditioning on $X = \hat{x}$, must be the same for all $\hat{x} \in \mathbb{B}_\varepsilon(\bar{x})$. Of course, then the means of these distributions must be the same for all $\hat{x} \in \mathbb{B}_\varepsilon(\bar{x})$. But we have shown above that this is not the case, i.e., $E_{\tilde{u}}[\gamma(\tilde{U}, X)|X = \bar{\bar{x}}] > E_{\tilde{u}}[\gamma(\tilde{U}, X)|X = \bar{x}]$. Hence $\tilde{U}$ is dependent on X.

Brown–Roehrig Theorem: An Example. We consider a consumer with a random demand function $Y(P, I, U, \theta_0)$ derived from maximizing a random utility function $V(Y, U, \theta_0)$ subject to her budget constraint $P \cdot Y = I$. First, the consumer draws U from a fixed and known distribution. Then nature draws $X = (P, I)$, from a fixed but unknown distribution. The main model assumption is that U and X are stochastically independent. The consumer solves the following optimization problem:

$$\max V(y, u, \theta_0) \text{ over } y \text{ such that } p \cdot y = i$$

The econometrician knows $V(y, u, \theta)$ and Θ, the set of all possible values for the parameter θ, but does not know θ_0, the true value of θ. Nor does the econometrician observe U or know the distribution of U. The econometrician does observe $X = (P, I)$. The econometrician's problem is to estimate θ_0 and the distribution of U from a sequence of observations $Z_i = (X_i, Y_i)$ for $i = 1, 2, ..., n$. The structural equations for this model are simply the first-order conditions of the consumer's optimization problem, after she has observed u and $x = (p, i)$.

These conditions define an implicit nonlinear simultaneous equations model of the form $U = \rho(X, Y, \theta)$, where the reduced form function is the consumer's random demand function $Y(P, I, U, \theta_0)$ for the specification of $V(Y, U, \theta)$ proposed by Brown and Matzkin (1998), i.e., $V(Y, U, \theta) = U(Y, \theta) + U \cdot Y$. They assume that for all $\theta \in \Theta$, $W(Y, \theta)$ is a smooth monotone strictly concave utility function on the positive orthant of $\mathbb{R}^k$, i.e., $DW(Y, \theta) \in \mathbb{R}^k_{++}$ and $D^2W(Y, \theta)$ is negative definite for all θ. Moreover, they assume $U \in \mathbb{R}^k_+$.

Our example is suggested by their model, where we consider

$$
\begin{aligned}
& V(Y, U, \theta) \\
= \; & \ln Y_0 + \theta_1 \ln Y_1 + \theta_2 \ln Y_2 + U_1 Y_1 + U_2 Y_2,
\end{aligned}
$$

for $\theta_1, \theta_2 \in (0,1)$ and Y_0 is the numeraire good. Then the first-order conditions for this optimization problem can be written as $U = \rho(X, Y, \theta)$, where $X = (P_1, P_2, I)$, $Y = (Y_0, Y_1, Y_2)$ and $\theta = (\theta_1, \theta_2)$

(i) $U_1 = P_1(I - P_1Y_1 - P_2Y_2)^{-1} - \theta_1 Y_1^{-1}$

(ii) $U_2 = P_2(I - P_1Y_1 - P_2Y_2)^{-1} - \theta_2 Y_2^{-1}$

Equations (i) and (ii) can (in principle) be solved uniquely for the random demand functions $Y_1(X, U, \theta)$ and $Y_2(X, U, \theta)$. This verifies that there exists a unique reduced form $Y = \gamma(X, U, \theta)$ such that $U = \rho(X, \gamma(X, U, \theta), \theta)$.

The implicit function theorem implies that $\partial y(x,u,\theta)/\partial x$ can be computed from the structural equations $u = \rho(x,y,\theta)$ if $\partial\rho/\partial y$ has full rank, where

$$\left(\frac{\partial\rho}{\partial y}\right) = \begin{bmatrix} \frac{\partial\rho_1}{\partial y_1} & \frac{\partial\rho_1}{\partial y_2} \\ \frac{\partial\rho_2}{\partial y_1} & \frac{\partial\rho_2}{\partial y_2} \end{bmatrix} \text{ and } \left(\frac{\partial y}{\partial x}\right) = -\left[\frac{\partial\rho}{\partial y}\right]^{-1}\left[\frac{\partial\rho}{\partial x}\right]$$

We show that $\partial\rho/\partial x$ has rank 2 and is independent of θ and u and that $\partial\rho/\partial y$ has full rank and is independent of u. Therefore

$$\frac{\partial y(x,u,\theta_0)}{\partial x} = \frac{\partial y(x,\tilde{u},\tilde{\theta})}{\partial x} \text{ a.e. if and only if}$$

$$\frac{\partial\rho(x,y,\theta_0)}{\partial y} = \frac{\partial\rho(x,y,\tilde{\theta})}{\partial y} \text{ a.e.}$$

But

(iii) $\frac{\partial\rho(x,y,\theta_0)}{\partial y} = \frac{\partial\rho(x,y,\tilde{\theta})}{\partial y}$ a.e. if and only if $\theta = \theta_0$.

Hence all the conditions of the Brown–Roehrig Theorem are satisfied.