ON THE NONPARAMETRIC IDENTIFICATION OF NONLINEAR SIMULTANEOUS EQUATIONS MODELS:

Comment on B. Brown (1983) and Roehrig (1988)

By

C. Lanier Benkard and Steven Berry

October 2004

COWLES FOUNDATION DISCUSSION PAPER NO. 1482

COWLES FOUNDATION FOR RESEARCH IN ECONOMICS

YALE UNIVERSITY

Box 208281

New Haven, Connecticut 06520-8281

http://cowles.econ.yale.edu/

On the Nonparametric Identification of Nonlinear Simultaneous Equations Models: Comment on B. Brown (1983) and Roehrig (1988)*

C. Lanier Benkard
Stanford University and NBER

Steven Berry
Yale University and NBER

September 7, 2004

Abstract

This note revisits the identification theorems of B. 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, contrary to the theorems, the reduced form of a nonseparable simultaneous equations model is not identified even under the assumptions of those papers. 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.

*We have had very helpful conversations with Pat Bayer, Don Brown, Yossi Feinberg, Guido Imbens, Yuliy Sannikov, Andy Skrzypacz, and Chris Timmons. Any remaining errors are our own.

1 Introduction

In this note, we reconsider the nonparametric identification of nonlinear simultaneous equations models, as in B. Brown (1983) and Roehrig (1988). We believe the identification results in these papers to be potentially much more powerful than has been widely recognized to date, applying in some cases to systems whose identification has only been shown very recently, including examples in Chernozhukov and Hansen (2001), Chesher (2003), Imbens and Newey (2002), Matzkin (2003), and others. Despite this, the Brown/Roehrig identification theorems have thus far remained largely as a topic in the literature on mathematical economics, and have been used infrequently and/or awkwardly in other areas.

Our intention in starting this project was to show the strength of these papers' results and to show how they could be applied quite generally in empirical work, extending even to cases where identification is not currently known to hold. However, in revisiting this literature, we have also discovered that a key lemma (B. Brown (1983), pp 180-181, cited by Roehrig (1988), p. 438) in the proof of the primary theorems of both Brown and Roehrig is false. This finding is substantive. An important implication of this lemma, 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 for a counterexample.

The counterexample also contradicts the main theorems in both Brown and Roehrig, and 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. We view these theorems as having important implications for empirical work in economics and we therefore hope that this shortcoming will be rectified soon in future research.

Note that Brown (1983) and Roehrig (1988) are widely cited in the literature on nonparametric identification (some recent examples include Newey and Powell (1999), Angrist, Graddy, and Imbens (2000), Guerre, Perrigne, and Vuong (2000), Athey and Haile (2002), Imbens and Newey (2002), Chesher (2003), Matzkin (2003), and Newey and Powell (2003)). Their identification theorems also play a key role in a few important papers in this literature, including D. Brown and Wegkamp (2002), and D. 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 Brown's lemma and a summary of the flaw in the proof of the lemma, as well as some intuition behind the failure of the lemma. Next we present the counterexample to the lemma on the identification of the model's reduced form. We conclude with a discussion of why we believe that the main identification theorems of Brown and Roehrig may still hold with some modifications to the assumptions, and 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 assumed not to be observed. The endogenous variables, $Y \in \mathbb{R}^G$, are assumed to be observed and are defined below. Realizations of the 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, (Y, X), 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,

$$Y = m(X, Y, U; \theta), \tag{1}$$

where $m(\cdot)$ is a known function and θ is an unknown parameter vector.[1] We emphasize that our results hold in a more general setting in which $m(\cdot)$ is unknown. We denote the true data generating process by the parameter vector θ_0.

[1]Note that since 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 structure is the nonseparable supply and demand model,

$$\begin{aligned} Q &= D(Z, P, \epsilon_D; \theta_D) && (2) \\ P &= S(W, Q, \epsilon_S; \theta_S), && (3) \end{aligned}$$

where in the general notation, $Y = (Q, P)$, $X = (Z, W)$, $U = (\epsilon_D, \epsilon_S)$ and the structural model is $m = (D, S)$.

We assume for purposes of identification that the joint distribution of the endogenous variables and the observed exogenous shifters, $F(x, y)$, is known. We maintain the assumptions that the structural equations are continuously differentiable, and that the data comes from a continuous distribution.[2] These are largely technical assumptions. All that is required is differentiability almost everywhere, and it is our belief that even this could be relaxed.

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 lemma in dispute. The remaining assumptions, which we omit from this note, are rank conditions similar to those used in identifying the structural equations in a linear model.

2.1 The Three Basic Assumptions

Assumption 1 [Reduced Form]: The data is generated by a continuously differentiable reduced-form,

$$Y = f(X, U; \theta_0), \tag{4}$$

where θ_0 is the index of the true structural model.

Brown and Roehrig assume that the structural model gives a unique reduced-form, but their proofs require only that the data is generated by a differentiable function that maps X and U into Y.

Assumption 2 [Solution for U]: There is a unique solution to the structural equations giving the unobservables as a function of the observables. That is, for every θ, there exists a $\rho(Y, X; \theta)$ such that $Y = m(X, Y, \rho(Y, X; \theta); \theta)$. Further, the function $U = \rho(Y, X; \theta)$ is continuously differentiable.

[2]The data is generated by a continuous distribution with positive density everywhere on its support.

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, while the error may be non-separable in the structural equations, there is a transformation of the structural equations that is linear in the errors.

In addition to the above 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 Brown's Lemma

In considering identification of the model, an important insight of Brown (1983) was to focus on the residual function, ρ. For any alternative model under consideration, θ, we can substitute the true reduced form into the residual function to obtain a mapping from (x, u) into u:

$$\tilde{u}(x, u; \theta, \theta_0) = \rho(f(x, u; \theta_0), x; \theta), \tag{5}$$

where $\tilde{u}(x, u; \theta, \theta_0)$ is the error implied by the model evaluated at θ (and where $\tilde{u}(x, u; \theta_0, \theta_0) = u$).[3] Brown (1983), followed by Roehrig (1988), uses this relationship to provide conditions under which imposing independence on the errors in the alternative model, $\tilde{U} \equiv \tilde{u}(X, U; \theta, \theta_0)$, will identify the true model θ_0.

The argument of Brown's lemma is quite simple. Assumptions 1-2 require that the relationship in (5) holds for any candidate model, θ, for all (x, u) pairs in the support of (X, U). The theorem then asks what restrictions, if any, independence places on the candidate model.

Lemma 1 (B. Brown (1983), pp 180-181). *The residuals from the candidate model, $\tilde{U}$, are independent of X if and only if the derivative of the*

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

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

mapping $\tilde{u}(x, u; \theta, \theta_0)$ with respect to x is everywhere zero, that is if:

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

for all (x, u) pairs in the support of (X, U).

The lemma seems very intuitive at first and, as we will see below, if true it would be incredibly powerful. The lemma is quite obviously true in one direction: If the derivative of the residual mapping with respect to x is everywhere zero, then the mapping is not a function of x. Therefore, the errors in the alternative model, $\tilde{U}$, can be written as a function only of the errors from the true model, U. Since U is independent of X, it must be that $\tilde{U}$ is as well.

The problem with the lemma is in the other direction. While the lemma is true if the errors are univariate[4], in multidimensional spaces it is easy to generate mappings with nonzero derivatives with respect to x but that still generate independent errors. We provide a graphical intuition for this below, and a more interesting class of counterexamples in the next section. Here, we first replicate the error in the proof of the lemma.

Proof from Brown, p. 181:
Suppose that, for a particular candidate model, $\bar{\theta}$, the total derivative of the mapping is not zero everywhere. Then there exists a point, (x^0, u^0), at which the derivative of the mapping is not zero. Without loss of generality, suppose

$$\frac{\partial \tilde{u}_i(x^0, u^0; \bar{\theta}, \theta_0)}{\partial x_j} > 0.$$

By continuity of the derivative, $\frac{\partial \tilde{u}_i}{\partial x_j} > 0$ for all $(x, u) \in \mathcal{N}_x \times \mathcal{N}_u$, where $\mathcal{N}_x$ and $\mathcal{N}_u$ represent neighborhoods of x^0 and u^0 respectively. Let $\mathcal{N}_{\tilde{u}}$ denote the image set generated by the mapping $\tilde{u}(x, u; \bar{\theta}, \theta_0)$ for $x \in \mathcal{N}_x$ and $u \in \mathcal{N}_u$ and define

$$g(\tilde{u}) = \begin{cases} \tilde{u}_i(x, u; \bar{\theta}, \theta_0) & \text{for } u \in \mathcal{N}_u, \\ 0 & \text{otherwise.} \end{cases} \tag{6}$$

[4]There are numerous ways of showing this but perhaps the simplest is that of Matzkin (2003).

Then

$$Eg(\tilde{u}(x, U; \bar{\theta}, \theta_0)) = \int_{u \in \mathcal{N}_u} \tilde{u}_i(x, u; \bar{\theta}, \theta_0) f(u) du \tag{7}$$

exists and is finite. Consider a point $x^1 \in \mathcal{N}_x$ such that $x^1 = x^0$ except that $x_{1,j} > x_{0,j}$. Since the derivative of the mapping is positive everywhere on $\mathcal{N}_x \times \mathcal{N}_u$, it must be that $\tilde{u}_i(x^1, u; \bar{\theta}, \theta_0) > \tilde{u}_i(x^0, u; \bar{\theta}, \theta_0)$ for all $u \in \mathcal{N}_u$. Therefore

$$Eg(\tilde{u}(x^1, U; \bar{\theta}, \theta_0)) > Eg(\tilde{u}(x^0, U; \bar{\theta}, \theta_0)). \tag{8}$$

Brown, p. 181, then concludes that this last line implies that $g(\tilde{U})$, and hence $\tilde{U}$, is stochastically dependent upon X. $\square$

The error in the proof is in the definition of the function g in equation (6). The proof uses the fact that g is stochastically dependent on X to infer that $\tilde{U}$ is stochastically dependent on X. This inference would be correct if g were a function only of $\tilde{u}$, as is suggested by the left hand side of the definition in (6). However, the right side of (6) is written in terms of x and u. Unfortunately, this is not just a simple error in notation. In general, the right hand side of the definition can not be written in terms of $\tilde{u}$ alone because it must also depend separately on x.

To see this more clearly, consider the simple univariate example $\tilde{u} = u + x$, with $\mathcal{N}_u = (0, 1)$. There are many combinations of u and x that yield the same value for $\tilde{u}$ but different values for $g(\cdot)$ as defined in (6). For example, $(x, u) = (2, -1)$ gives $\tilde{u} = 1$ and $g = 0$, whereas $(x, u) = (0.5, 0.5)$ gives the same $\tilde{u}$ but now $g = 1$. The function g therefore can not be written as a function of $\tilde{u}$ alone, but also depends separately on x: holding $\tilde{u}$ fixed and varying x gives different values for g. Because g is a function of both $\tilde{u}$ and x, the fact that the expectation of g changes with x, as is shown in (8), does not contradict the independence of $\tilde{U}$ and X.

Note also what happens if we try to fix the definition of g. For example, we could redefine g as a true indicator function of $\tilde{u}$. Holding x fixed at x^0 we can think of the set, $\mathcal{N}_{\tilde{u}}$, of values of $\tilde{u}$ that keeps u in the original neighborhood $\mathcal{N}_u$. Following the proof through, in the analog to (7) one finds that both the integrand and the region of integration change with x, in possibly off-setting ways, and so the inequality in (8) is no longer guaranteed. Furthermore, below we have counterexamples to the lemma, so larger changes to the proof also will not rescue the result.

3.0.1 A Graphical Intuition

There is also a simple graphical intuition as to why the lemma is false. Suppose that the true structural errors are two-dimensional and uniformly distributed on the unit circle, independent of X. Suppose that the $\tilde{u}$ mapping is such that the true errors are simply rotated about the origin by an amount determined by x. This would generate a new set of errors, $\tilde{U}$, that are still uniformly distributed on the unit circle for every outcome $X = x$, and therefore independent of X. However, the mapping has nonzero derivatives with respect to x almost everywhere.

In a single dimension, such rotations are not possible, and it is easy to show that Brown's lemma holds in a single dimension (obtaining a result similar to Matzkin (2003)). 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 lemma fails. We provide a more substantive example in the next section.

4 A Counterexample: Identification of the Reduced Form

The power of Brown's lemma can be seen in the following Corollary:

Corollary 2 (Corollary to Brown's Lemma: Identification of the Reduced Form Derivatives). *At every point (y, x) in the support of (Y, X), any model, θ, that satisfies the basic assumptions and is observationally equivalent*[5] *to the true model, θ_0, must imply the same reduced form derivatives with respect to the exogenous shifters as the true reduced form. That is,*

$$\frac{\partial f(x, \rho(y, x; \theta), \theta)}{\partial x} = \frac{\partial f(x, \rho(y, x; \theta_0), \theta_0)}{\partial x} \tag{9}$$

Proof. Since θ is observationally equivalent to the true model,

$$f(x, \tilde{u}; \theta) = y = f(x, u; \theta_0)$$

[5] Here we are using Roehrig's (p. 435) definition of observational equivalence.

for every point (x, u) in the support of (X, U), where $\tilde{u}$ is given by (5).

Since the reduced form functions, $f(\cdot)$, are differentiable, it must be the case that in the support of (X, U),

$$\frac{df(x, \tilde{u}; \theta)}{dx} = \frac{df(x, u; \theta_0)}{dx}$$

Brown's lemma implies that $\frac{\partial \tilde{u}}{dx} = 0$. Therefore,

$$\frac{\partial f(x, \tilde{u}; \theta)}{\partial x} = \frac{\partial f(x, u; \theta_0)}{\partial x}$$

□

If Brown's lemma held, then we would automatically have the result that the reduced form was identified in any model satisfying the three basic assumptions. The remaining rank conditions assumptions in the main theorems of Brown and Roehrig could then be seen as being analogous to the rank conditions needed to identify the structure in linear models, with the distinction that they need to hold at every point. However, the following counterexample shows that, in fact, the basic assumptions (1-3) are not sufficient for identification of the reduced form.

If the corollary above were true, in order to estimate the reduced form we would only need to find a model that satisfied assumptions 1 and 2 and generated errors that were independent of X. One such model is the following triangular construction,[6]

$$\begin{aligned}
\tilde{U}_1 &= F(Y_1|X) \\
\tilde{U}_2 &= F(Y_2|X, Y_1) \\
\tilde{U}_3 &= F(Y_3|X, Y_1, Y_2) \\
\vdots &= \vdots \\
\tilde{U}_G &= F(Y_G|X, Y_1, ..., Y_{G-1})
\end{aligned} \tag{10}$$

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

[6]This representation assumes that the dimension of U equals the dimension of Y. While this was not assumed explicitly above, it is essentially required by assumption 2.

If Brown's theorem held, then this system should retrieve the true reduced form. However, it is easy to see that the system above could not in general retrieve the true reduced form. Consider the derivative of y_1 with respect to x implied by the triangular system (obtained via the implicit function theorem),

$$D_x y_1 = -[D_x F(y_1|x)]^{-1} \frac{\partial F(y_1|x)}{\partial y_1}$$

This expression does not change with y_2 (nor any of the other y's). What this means is that the error from the second equation (and also from the other equations) does not change the derivative of y_1 with respect to x. However, a simple inspection of the reduced form in equation (4) shows that, in general, the derivative of y_1 with respect to x might change with the error from the second equation. In fact, the restriction that the derivative of y_1 with respect to x be constant would only be true in some very special cases, such as if the reduced form were additively separable in the error term, in which case the derivative of y_1 with respect to x depends only on x, or if the error terms from all equations entered the reduced form as a single index, in which case the derivative of y_1 with respect to x depends only on y_1 and x. Such restrictions may hold for some systems, but do not hold in general for nonseparable systems. Note also that we could have defined the triangular system starting with any equation so there are many potentially different ways to construct the reduced form using this approach.

The triangular construction fails to retrieve the true reduced form for exactly the same reason that Brown's lemma fails above. The errors in the triangular construction, $\tilde{U}$, are independent of the exogenous variables, X, but could have been transformed from the original errors in a way that depends on X. For example, they could have been rotated in a way that depends on X. Such a transformation could change the implicit functional relationship between Y and X and thus would not necessarily recover the true reduced form.

This example shows that the three basic assumptions above are *not* sufficient to identify the reduced form. An important but negative implication of the result is that, in general, without further assumptions, a projection of the endogenous variables on the exogenous variables 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 depending on the values of the error terms in the other structural equations. A projection of y_1 onto x that ignores the values of $y_2, ..., y_G$ (or alternatively, $u_2, ..., u_G$) recovers not the true reduced form, but something akin to the average reduced form derivative weighted over the distribution of the left out variables.

Finally, the main theorems in Brown and Roehrig add rank conditions assumptions that serve to identify the structural model from the reduced form. Since the reduced form always satisfies these rank conditions, the example above is also a counterexample to these theorems.

5 Conclusions/Areas for Future Research

So far we have shown that the 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 in order 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. 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 only utilize the rank conditions assumptions in identifying the structure from the reduced form, but it is possible that similar exclusion restrictions may help to identify the reduced form as well.

One set of 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 (10) and suppose that this system represented the form of the true structural model. Then the first equation can be shown to be identified using single equation methods. Similarly, each successive row can be shown to be identified conditional on all previous rows. Imbens and Newey (2002) consider a two-dimensional system of this kind, and Chesher (2003) considers a similar multi-dimensional 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 Brown's lemma 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.

The identification theorems of Brown and Roehrig 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.

6 References

Angrist, J. D., Graddy, K., 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*, 67:3, 499-527.

Athey, S., and P. A. Haile (2002), "Identification of Standard Auction Models," *Econometrica*, 70:6, 2107-2140.

Brown, B. W. (1983), "The Identification Problem in Systems Nonlinear in the Variables," *Econometrica*, **51:1**, 175-196.

Brown, D. J. and R. Matzkin (1998), "Estimation of Nonparametric Functions in Simultaneous Equations Models, with an Application to Consumer Demand," Cowles Foundation Discussion Paper #1175.

Brown, D. J. and M. H. Wegkamp (2002), "Weighted Mean-Square Minimum Distance from Independence Estimation," *Econometrica*, **70:5**, 2035-2051.

Chernozhukov, V. and C. Hansen (2001), "An IV Model of Quantile Treatment Effects," MIT Working Paper # 02-06.

Chesher, A. (2003), "Identification in Nonseparable Models," *Econometrica,* **71:5**, 1405-1441.

Guerre, E., Perrigne, I., and Q. Vuong (2000), "Optimal Nonparametric Estimation of First-Price Auctions," *Econometrica*, 68:3, 525-574.

Imbens, G. and W. K. Newey (2002), "Identification and Estimation of Triangular Simultaneous Equations Models Without Additivity," Mimeo.

Matzkin, R. (2003), "Nonparametric Estimation of Nonadditive Random Functions," Econometrica, 71:5, 1332-1375.

Newey, W. K., and J. L. Powell (1999) "Nonparametric Estimation of Triangular Simultaneous Equations Models," *Econometrica*, 67:3, 565-603.

Newey, W. K., and J. L. Powell (2003) "Instrumental Variable Estimation of Nonparametric Models," *Econometrica*, 71:5, 1565-1578.

Roehrig, C. S. (1988), "Conditions for Identification in Nonparametric and Parametric Models," *Econometrica*, **56:2**, 433-447.