COHERENCY AND COMPLETENESS OF STRUCTURAL MODELS CONTAINING A DUMMY ENDOGENOUS VARIABLE

Arthur Lewbel   Boston College

original Sept. 1997, revised Feb. 2006

Abstract

Let y be a vector endogenous variables and let w be a vector of covariates, parameters and errors or unobservables that together are assumed to determine y. A structural model $y = H(y, w)$ is complete and coherent if it has a well defined reduced form, meaning that for any value of w there exists a unique value for y. Coherence and completeness simplifies identification, and is required for many estimators and many model applications. Incoherency or incompleteness can arise in models with multiple decision makers such as games, or when the decision making of individuals is either incorrectly or incompletely specified. This paper provides necessary and sufficient conditions for the coherence and completeness of simultaneous equation systems where one equation is a binomial response. Examples are dummy endogenous regressor models, regime switching regressions, treatment response models, sample selection models, endogenous choice systems, and determining if a pair of binary choices are substitutes or complements.

Keywords: Binary choice, Binomial Response, Simultaneous equations, Games, Multiple equilibria, Endogeneity, Coherent models, Incompleteness, Identification, Latent Variable models, Substitutes and Complements.

The author wishes to thank Jim Heckman, Chuck Manski, Elie Tamer, Andrew Chesher, Richard Blundell, participants of the Festshrift for Daniel McFadden, and anonymous referees for many helpful comments. Any errors are my own.

Arthur Lewbel, Department of Economics, Boston College, 140 Commonwealth Ave., Chestnut Hill, MA, 02467, USA. (781)-862-3678, lewbel@bc.edu.

1 Introduction

Let y be a vector of endogenous variables, and let $w \in \Omega$ be a vector of observables and unobservables that determine y. Here w could contain unknown parameters, exogenous observed covariates and error terms. Consider a proposed structural model of the form $y = H(y, w)$. Gourieroux, Laffont, and Monfort (1980) define the model to be *coherent* if for each $w \in \Omega$ there exists a unique value for y, which we may denote by the reduced form equation $y = G(w)$, such that $G(w) = H[G(w), w]$. Heckman (1978) refers to this condition as the *principal assumption* and as *conditions for existence of the model*. Other authors that consider coherency of various model specifications include Blundell and Smith (1994) and Dagenais (1997). More recently, Tamer (2003) uses the term *coherency* to only refer to existence of a y that solves $y = H(y, w)$, and calling the model *complete* if the solution is unique. I will adopt this newer terminology. Coherency and completeness of a structural model together imply existence and uniqueness of a reduced form.

Incoherent or incomplete models with dummy endogenous regressors arise in some simultaneous games, e.g., the industry entry game of Bresnahan and Reiss (1991) yields a system of two binary choice equations, each of which depends on the outcome of the other. Tamer (2003) observes that incoherency corresponds to case where the game has no Nash equilibrium, and incompleteness to the case of multiple equilibria. Aradillas-Lopez (2005) removes the incompleteness in these games by showing that a unique Nash equilibrium exists when the player's information sets are incomplete.

Incoherence can be interpreted as a form of model misspecification, since it implies that for some feasible values of w there does not exist a corresponding value of y, whereas in reality some value of y would be observed. We may think of incompleteness as a model that is not fully specified, since for some feasible values of w, the model does not deliver a corresponding unique value for y. Parameters of incoherent or incomplete models can sometimes be point identified and estimated (Tamer 2003 provides examples), however, such models cannot be used to make predictions about y over the space of all values of w, and they cannot be used with any parameter identification scheme or estimator that depends on the existence of a well defined reduced form. Incompleteness may often give rise to models with parameters that are set rather than point identified. See, e.g., Manski and Tamer (2002). Nevertheless, coherency and completeness are certainly desirable and commonly assumed properties of econometric models. If nothing else, it is important to know when a model may or may not be coherent or complete for estimation and interpretation of the model.

Incompleteness or incoherency can arise in models with multiple decision makers, with one example being strategically interacting players. Models of a single optimizing agent will typically be coherent though sometimes incomplete (if, e.g., the same utility or profit level can be attained in more than one way), though more general incoherency or incompleteness can arise in such models when the decision making process is either incorrectly or incompletely specified, or is not characterized by optimizing behavior. Ad hoc equilibrium selection mechanisms or rules for tie breaking in optimization models can be interpreted as techniques for resolving this type of incompleteness. These are relatively harmless when incompleteness, such as ties, can only occur with probability zero. This paper will be concerned with more fundamental incompleteness or incoherence, where no solutions or multiple solutions exist on a positive measure subset of relevant variables' supports.

Let $y = (y_1, y_2)$, where y_1 is a dummy endogenous variable. This paper provides necessary and sufficient conditions for coherence and completeness of

$$
\begin{aligned}
(1) \qquad y_1 &= H_1(y_1, y_2, w) \\
(2) \qquad y_2 &= H_2(y_1, y_2, w)
\end{aligned}
$$

for arbitrary functions H_1 and H_2, where H_1 can only equal zero or one. Structural models of this type are very common in econometrics. Examples include discrete endogenous regressor models, regime shift models, treatment response models, sample selection models, joint continuous-discrete demand models, and simultaneous choice models (in which both y_1 and y_2 are discrete). Heckman (1978), Blundell and Smith (1994), and Dagenais (1997) each provide conditions required for coherence and completeness of different special cases of this class of models, while Gourieroux, Laffont, and Monfort (1980) analyze coherency and completeness for closely related piecewise linear models.

The system of equations (1) and (2) is defined to be triangular, or recursive, if either H_1 does not depend on y_2 or H_2 does not depend on y_1. See, e.g., Maddala and Lee (1976). Triangular systems are generally coherent and complete if the individual equations are separately coherent and complete.

To illustrate the completeness and coherency problems in simultaneous systems, consider the simple model

$$
\begin{aligned}
y_1 &= I(y_2 + e_1 \geq 0) \\
y_2 &= \alpha y_1 + e_2
\end{aligned}
$$

where $w = (\alpha, e_1, e_2)$ and I is the indicator function that equals one if its argument is true and zero otherwise. These equations could be the reaction functions of two players in some game, where player one makes a discrete choice y_1 (such as whether to enter a market), and player two makes some continuous decision y_2 (such as the quantity to produce of a good). Then $y_1 = I(\alpha y_1 + e_1 + e_2 \geq 0)$ so $y_1 = 0$, $y_2 = e_2$ if $0 = I(e_1 + e_2 \geq 0)$, that is, $e_1 + e_2 < 0$, and $y_1 = 1$, $y_2 = a + e_2$ if $1 = I(\alpha + e_1 + e_2 \geq 0)$, meaning $\alpha + e_1 + e_2 \geq 0$. Therefore, the model implies both $y_1 = 0$ and $y_1 = 1$, and so is incomplete, if $-a \leq e_1 + e_2 < 0$. Neither $y_1 = 0$ nor $y_1 = 1$ satisfy the model if $0 \leq e_1 + e_2 < -a$, so in that region the model is incoherent. This model is both coherent and complete only if $a = 0$ or if $e_1 + e_2$ is constrained to not lie between zero and $-a$.

The next section provides general characterizations of conditions for completeness and coherence. This is then followed by examples including endogenous selection and treatment models, dummy endogenous regressor models, and regime switching models. Next, simultaneous systems of binary choice equations are considered in depth. For these models completeness and coherency are obtained both by applying the general characterization theorems and by replacing the simultaneous system with a model of optimizing behavior, and examples of likelihood functions for such models are provided. Application of these models for determining if a pair of binary choices are substitutes or complements is described.

2 Necessary and Sufficient Conditions for Coherence and Completeness

Theorem 1: Assume $y_1 \in \{0, 1\}$, $y_2 \in \Psi$ and $w \in \Omega$ for some support sets Ψ and Ω. The system of equations (1) and (2) is coherent and complete if and only if there exists a function $g : \{0, 1\} \times \Omega \to \Psi$ such that, for all $w \in \Omega$, the following equations hold

$$H_1[1, g(1, w), w] = H_1[0, g(0, w), w] \tag{3}$$
$$y_2 = g(y_1, w) \tag{4}$$

Theorem 1 demonstrates the severity of the completeness and coherency conditions with a dummy endogenous variable. Equation (3) shows that, after substituting out y_2 into the equation for y_1, the right side of the resulting expression must be independent of y_1.

As an example, consider the general selection model in which y_1 indexes whether y_2 is observed.

Corollary 1: The general endogenous selection model

$$\begin{aligned} y_1 &= R(y_2, w) \\ y_2 &= r(w) y_1 \end{aligned}$$

is coherent and complete if and only if R is independent of y_2.

Corollary 1 illustrates the strength of Theorem 1, by showing that no complete, coherent selection model can be endogenous, where endogeneity is defined as having the selection criterion y_1 depend on the observed outcome y_2. Note, however, that completeness is possible using some other notion of endogeneity, such as having y_1 depend on the latent outcome $r(w)$.

Next consider a typical binary choice specification for y_1, with a latent additive error. Replace equation (1) with

$$(5) \qquad y_1 = I[h(y_1, y_2, w) + e_1 \geq 0]$$

for some function h, where $e_1 \in w$. If equation (4) holds, define functions s_0 and s_1 by

$$\begin{aligned} s_0(w) &= h[0, g(0, w), w] \\ s_1(w) &= h[1, g(1, w), w]. \end{aligned}$$

Theorem 2. The system of equations (5) and (2) is coherent and complete if and only if there exists a function g such that equation (4) holds and, for every $w \in \mathbf{\Omega}$, either $s_0(w) = s_1(w)$, or e_1 does not lie in the interval bounded by $-s_0(w)$ and $-s_1(w)$. The equality $s_0(w) = s_1(w)$ holds if and only if there exists a function f such that.

$$(6) \qquad y_1 = I[f[y_2 + [g(0, w) - g(1, w)]y_1, w] + e_1 \geq 0]$$

Alternatively, $s_0(w) = s_1(w)$ holds if and only if there exists a function ϕ and a dummy function d that only takes on the values zero and one, such that

$$\begin{aligned} (7) \qquad y_1 &= I[\phi[(1 - d(w))y_2, w] + e_1 \geq 0] \\ (8) \qquad y_2 &= g[d(w)y_1, w] \end{aligned}$$

Theorem 2 shows there are only two methods to obtain coherence and completeness in the presence of a binary choice equation. One method is to restrict the support of the errors to rule out regions of incoherency or incompleteness. Dagenais (1997) is a special case of this method. It is, however, difficult to motivate such data dependent restrictions on the values that the error can take on.

The only other way to obtain coherence is to restrict attention to the class of models that can be represented by equations (4) and (6), or equivalently by (7) and (8). The usefulness of one of these representations over the other will depend on context, e.g., taking f to be linear in (6) yields a different coherent system than taking ϕ to be linear in (7). The next section provides examples.

Theorem 2 shows that, unless one peculiarly restricts the support e_1, by equations (7) and (8) completeness requires a model that is equivalent to a triangular system, though the direction of dependence (whether y_1 depends on y_2 or y_2 depends on y_1), which is indexed by the binary dummy variable $d(w)$, may vary across observations. It is perhaps surprising that the simple introduction of d to generalize triangular systems has not been proposed before. In contrast, the representation given by equations (4) and (6) can be interpreted as the nonlinear generalization of Blundell and Smith (1994).

Equations (7) and (8) readily extend to provide coherent, complete specifications for endogenous y having any support. The system

$$\begin{aligned} y_1 &= g_1[(1-d(w))y_2, w] \\ y_2 &= g_2[d(w)y_1, w] \end{aligned}$$

is coherent for any functions g_1 and g_2, since it has the well defined reduced form

$$\begin{aligned} y_1 &= g_1[(1-d(w))g_2(0, w), w] \\ y_2 &= g_2[d(w)g_1(0, w), w]. \end{aligned}$$

3 Examples

Let $d = d(w)$ be a dummy variable that only takes the values zero and one. Here d can either be observed or it can be a known function of errors, covariates, and parameters. Let x be a vector of regressors, which can include both d and a constant term. For integers j let each e_j be an unobserved error which may have conditional support equal to the real line, let each β_j be a parameter vector and let each α_j and γ_j be scalar parameters.

3.1 Nonparametric Dummy Endogenous Regressor and Treatment Models

Consider a model of y_2 where y_1 is a dummy endogenous regressor. For some functions G_1 and G_2 let

$$\begin{aligned} y_1 &= G_1(y_2, x, e_1) \\ y_2 &= G_2(y_1, x) + e_2 \end{aligned}$$

where e_1 is independent of x and $E(e_2|x) = 0$. Das (2004) proposes a nonparametric estimator for the function G_2, leaving G_1 unspecified. The function G_2 can be interpreted as the conditional average outcome of an endogenous treatment y_1.

Theorem 1 shows that coherency and completeness of this model requires that $G_1[G_2(y_1, x) + e_2, x, e_1]$ be independent of y_1. Analogous to equation (6), a complete, coherent alternative model is

$$\begin{aligned} y_1 &= G_1[y_2 + [G_2(0, x) - G_2(1, x)]y_1, x, e_1] \\ y_2 &= G_2(y_1, x) + e_2 \end{aligned}$$

which would permit application of the Das estimator to G_2.

Another complete coherent alternative is

$$\begin{aligned} y_1 &= G_1[(1-d)y_2, x, e_1) \\ y_2 &= G_2(dy_1, x) + e_2, \end{aligned}$$

which for $d = 1$ equals a standard treatment effects specification.

3.2 Linear Dummy Endogenous Regressor Models

Consider the linear dummy endogenous regressor system

$$\begin{aligned} y_1 &= I[x'\beta_1 + y_2\alpha_1 + e_1 \geq 0] \\ y_2 &= x'\beta_2 + y_1\alpha_2 + e_2 \end{aligned}$$

Without restricting the support of the errors, Heckman (1978) found that coherency and completeness of this model requires either $\alpha_1 = 0$ or $\alpha_2 = 0$, which are triangular systems. A recent semiparametric estimator for this model is Klein and Vella (2001). Blundell and Smith (1994) proposed the generalization

$$\begin{aligned} y_1 &= I[x'\beta_1 + y_2\alpha_1 + y_1\gamma_1 + e_1 \geq 0] \\ y_2 &= x'\beta_2 + y_1\alpha_2 + e_2 \end{aligned}$$

which they found to be complete and coherent if $\gamma_1 = -\alpha_1\alpha_2$. This model equals the special case of equations (4) and (6) in which the functions f and g are linear.

A new complete coherent system may be obtained by taking the functions ϕ and g in equations (7) and (8) to be linear. This yields the model

$$\begin{aligned} y_1 &= I[x'\beta_1 + (1-d)y_2\alpha_1 + e_1 \geq 0] \\ y_2 &= x'\beta_2 + dy_1\alpha_2 + e_2 \end{aligned}$$

which is complete and coherent without restriction on the coefficients. We could also add $y_2\alpha_3 + dy_1\gamma_1$ to the latent variable determining y_1 and maintain completeness by imposing $\gamma_1 = -\alpha_2\alpha_3$.

3.3 Endogenous Regime Switching Models

The linear regime switching regression specification

$$\begin{aligned} y_1 &= I[x'\beta_1 + y_2\alpha_1 + e_1 \geq 0] \\ y_2 &= x'\beta_2 + e_2 + (x'\beta_3 + e_3)y_1 \end{aligned}$$

will not be coherent except under severe restrictions such as $\alpha_1 = 0$. Paralleling the previous section, Theorem 2 suggests two alternatives. One is

$$\begin{aligned} y_1 &= I[(x'\beta_1 + y_2\alpha_1 + y_1x'\beta_4 + e_4 \geq 0] \\ y_2 &= x'\beta_2 + e_2 + (x'\beta_3 + e_3)y_1 \end{aligned}$$

which is coherent and complete if $\beta_4 = -\alpha_1\beta_3$ and $e_4 = -\alpha_1y_1e_3 + e_1$ for some e_1. Another is

$$\begin{aligned} y_1 &= I[x'\beta_1 + (1-d)y_2\alpha_1 + e_1 \geq 0] \\ y_2 &= x'\beta_2 + e_2 + (x'\beta_3 + e_3)dy_1 \end{aligned}$$

(where d is again a binary dummy variable), which is coherent and complete without restrictions on the coefficients.

4 Simultaneous Systems of Binary Choices

Consider the simultaneous system of binomial responses

$$\begin{aligned} y_1 &= I[h_1(y_1, y_2, w) + e_1 \geq 0] \\ y_2 &= I[h_2(y_1, y_2, w) + e_2 \geq 0] \end{aligned}$$

A practical application of models like this is to determine whether interrelated binary choices are substitutes or complements, e.g., finding out if selecting $y_1 = 1$ increases or decreases the probability of choosing $y_2 = 1$. Dagenais (1997) obtains coherence and completeness in this model by imposing linearity on h_1 and h_2 and restricting the support of (e_1, e_2) to rule out regions of values that result in either no solutions or multiple solutions for y_1 and y_2.

Based on Theorem 2, a coherent and complete simultaneous system of binary choices that does not restrict the error supports is

$$\begin{aligned} y_1 &= I[f_1[y_2 - r(w)y_1, w] + e_1 \geq 0] \\ y_2 &= I[f_2(y_1, w) + e_2 \geq 0] \end{aligned}$$

for arbitrary choices of the functions f_1 and f_2, where the function r is defined by

$$r(w) = I[f_2(1, w) + e_2 \geq 0] - I[f_2(0, w) + e_2 \geq 0].$$

Alternatively, equations (7) and (8) in Theorem 2 suggest a simpler, more symmetric system

$$\begin{aligned} y_1 &= I[\phi_1[(1-d)y_2, w] + e_1 \geq 0] \\ y_2 &= I[\phi_2[dy_1, w] + e_2 \geq 0] \end{aligned}$$

which will be coherent and complete for any choice of the functions ϕ_1, ϕ_2, and binary dummy d. In particular, a nearly linear complete system of binary choice equations is

$$\begin{aligned} &(9) \qquad y_1 = I[x'\beta_1 + (1-d)y_2\alpha_1 + e_1 \geq 0] \\ &(10) \qquad y_2 = I[x'\beta_2 + dy_1\alpha_2 + e_2 \geq 0] \end{aligned}$$

which can be readily estimated with, e.g., jointly normal errors. Here d may be included in the list of regressors x. An example of d is to let $d = 1$ for individuals that make decision y_1 first, otherwise let $d = 0$.

Given a completely specified model, estimation can proceed using maximum likelihood by parameterizing the error distributions and evaluating the probability or density of each value the endogenous variables can take on. For example, in the model of equations (9) and (10), the probability that $y_1 = 1$ and $y_2 = 1$ is the probability that $x'\beta_1 + (1-d)\alpha_1 + e_1 \geq 0$ and $x'\beta_2 + d\alpha_2 + e_2 \geq 0$, while the probability that $y_1 = 0$ and $y_2 = 1$ is the probability that $x'\beta_1 + (1-d)\alpha_1 + e_1 < 0$ and $x'\beta_2 + e_2 \geq 0$ (note the absence of $d\alpha_2$ in this last equation, because $y_1 = 0$). Let $f(e_1, e_2, \lambda)$ denote the joint probability density functions of e_1 and e_2, parameterized by a vector λ, assumed independent of x. Then, conditioning on x, the probability of choosing $y_1 = 1$ and $y_2 = 1$ is

$$P_{11}(\theta \mid x) = \int_{-x'\beta_2 - d\alpha_2}^{\infty} \left(\int_{-x'\beta_1-(1-d)\alpha_1}^{\infty} f(e_1, e_2, \lambda) de_1 \right) de_2$$

where θ denotes the set of parameters $\beta_1, \beta_2, \alpha_1, \alpha_2, \lambda$. Similarly, the probabilities of choosing other values for y, denoted P_{y_1,y_2}, are

$$P_{01}(\theta \mid x) = \int_{-x'\beta_2}^{\infty} \left(\int_{-\infty}^{-x'\beta_1-(1-d)\alpha_1} f(e_1, e_2, \lambda) de_1 \right) de_2$$

$$P_{10}(\theta \mid x) = \int_{-\infty}^{-x'\beta_2 - d\alpha_2} \left(\int_{-x'\beta_1}^{\infty} f(e_1, e_2, \lambda) de_1 \right) de_2$$

$$P_{00}(\theta \mid x) = \int_{-\infty}^{-x'\beta_2} \left(\int_{-\infty}^{-x'\beta_1} f(e_1, e_2, \lambda) de_1 \right) de_2$$

If e_1 and e_2 are independent standard normals, then these expression simplify to

$$\begin{aligned}
P_{11}(\theta \mid x) &= \left[1 - \Phi\left(-x'\beta_1 - (1-d)\alpha_1\right)\right]\left[1 - \Phi\left(-x'\beta_2 - d\alpha_2\right)\right] \\
P_{01}(\theta \mid x) &= \Phi\left(-x'\beta_1 - (1-d)\alpha_1\right)\left[1 - \Phi\left(-x'\beta_2\right)\right] \\
P_{10}(\theta \mid x) &= \left[1 - \Phi\left(-x'\beta_1\right)\right]\Phi\left(-x'\beta_2 - d\alpha_2\right) \\
P_{00}(\theta \mid x) &= \Phi\left(-x'\beta_2\right)\Phi\left(-x'\beta_1\right)
\end{aligned}$$

Using either expression for the P functions, assuming n draws with independent errors, the log likelihood function for this model is then

$$\sum_{i=1}^{n} y_{1i} y_{2i} \ln P_{11}(\theta \mid x_i) + (1 - y_{1i}) y_{2i} \ln P_{01}(\theta \mid x_i)$$

$$+y_{1i}(1-y_{2i})\ln P_{10}(\theta \mid x_i)+(1-y_{1i})(1-y_{2i})\ln P_{00}(\theta \mid x_i)$$

which may be maximized with respect to θ to yield estimates of $\beta_1, \beta_2, a_1, a_2$.

If the temporal order of the decisions is not observed, one might let $d = I(|x'\beta_1 + e_1| > |x'\beta_2 + e_2|)$, so the choice that an individual feels most strongly about (as evidenced by the magnitude of the latent variable) is the decision that is made first. In this case estimation might be facilitated by using simulated moments as in McFadden (1989), though this would also introduce the difficulty of a nondifferentiable objective function.

4.1 Binary Choice Systems Based on Maximizing Behavior

Many of the example models provided so far are somewhat ad hoc, that is, they apply Theorems 1 or 2 to obtain coherency and completeness, but no underlying economic argument is provided to motivate the resulting models. One set of economically rationalizable models arising from Theorem 2 are those that use the dummy regressor d, which can be motivated as a model of sequential decision making, with d being the indicator of which decision an individual makes first, or which player in a sequential game moves first. This generalizes the usual triangular systems that are known to be coherent and complete, by permitting the direction of triangularity to vary across individuals.

In some applications, incompleteness may be eliminated by more fully modeling the behavior of agents, e.g., incompleteness resulting from games having multiple equilibria may be resolved by modeling how agents choose amongst the equilibria.

Consider an individual that makes two simultaneous binary decisions. The naive model in which each decision depends upon the other as a regressor,

$$\begin{aligned} &(11) \qquad y_1 = I[x'\beta_1 + y_2\alpha_1 + e_1 \geq 0] \\ &(12) \qquad y_2 = I[x'\beta_2 + y_1\alpha_2 + e_2 \geq 0] \end{aligned}$$

is incoherent and incomplete without some restrictions on the parameters, the supports of the errors, or both. One way to resolve both incoherency and incompleteness is by sequential decision making, using equations (9) and (10).

Another possibility is to avoid incoherence by restricting the parameters and augmenting the model to resolve incompleteness. For example, suppose the individual acts as if this were a game, choosing a Nash equilibrium that allows for randomness (essentially, flipping weighted coins to construct mixed strategies). If $\alpha_1\alpha_2 \leq 0$ then a unique mixed strategy equilbrium exists resulting in a coherent and complete model. In this case the coin tosses and Nash behavior complete an otherwise incomplete model. See Bresnahan and Reiss (1991), Tamer (2003), and Aradillas-Lopez (2005) for a more detailed analysis of treating (11) and (12) as a two person game.

Yet another way to avoid incoherence and incompleteness is to consider a random utility model, as in McFadden (1973). Equation (11) arises from assuming that the difference in utility U_1 between choosing $y_1 = 1$ vs $y_1 = 0$ is $x'\beta_1 + y_2\alpha_1 + e_1$, and so the individual chooses y_1 to maximize U_1. Similarly, the individual chooses y_2 using equation (12) to maximize the utility U_2 associated with that decision. To eliminate incoherency and incompleteness, we may assume the consumer chooses both y_1 and y_2 to maximize an overall utility function $U(U_1, U_2)$ that depends upon the utilities associated with each choice. Equivalently, if y_1 and y_2 are the actions of two players in a game, this would correspond to removing incoherency and incompleteness by collusion.

Generally, maximizing $U(U_1, U_2)$ results in the structure of a multinomial choice problem, maximizing the overall utility associated with each of the four values that $y = (y_1, y_2)$ can take on. However, specific forms for $U(U_1, U_2)$ will give rise to restricted versions of this model. Suppose

$$U(U_1, U_2) = U_1 + U_2 \tag{13}$$

where

$$\begin{aligned} U_1 &= (x'\beta_1 + y_2\alpha_1 + e_1)y_1 && (14)\\ U_2 &= (x'\beta_2 + y_1\alpha_2 + e_2)y_2 && (15) \end{aligned}$$

For example, if the individual is a firm, $x'\beta_1 + y_2\alpha_1 + e_1$ could be the difference in profit resulting from choosing $y_1 = 1$ versus $y_1 = 0$ holding y_2 fixed, and similarly $x'\beta_2 + y_1\alpha_2 + e_2$ could be the difference in profit resulting from choosing $y_2 = 1$ versus $y_2 = 0$ holding y_1 fixed. Then profit is maximized by choosing the value of $y = (y_1, y_2)$ that maximizes $U_1 + U_2$ with U_1 and U_2 given by equations (14) and (15).

This model has the feature that, conditioning on y_2, the utility maximizing choice for y_1 is given by equation (11), and that conditioning on y_1, the utility maximizing choice for y_2 is given by equation (12). This model is therefore consistent with the logic that gives rise to ordinary binary choice models such as probit or logit for each of the choices considered separately, while avoiding the incoherency or incompleteness of equations (11) and (12) as a system. The potential incoherency or incompleteness is eliminated by simultaneously considering the utilities of both choices. Letting $a = a_1 + a_2$, and letting $V(y)$ denote the utility associated with choice y, the result is

$$\begin{aligned}
V(0,0) &= 0 \\
V(1,0) &= x'\beta_1 + e_1 \\
V(0,1) &= x'\beta_2 + e_2 \\
V(1,1) &= x'(\beta_1 + \beta_2) + a + e_1 + e_2
\end{aligned}$$

where one chooses whichever value of y yields the maximum of these four values of V. This is a special case of ordinary multinomial choice where the utility associated with the last value y is $x'(\beta_1+\beta_2)+a+e_1+e_2$, instead of $x'\beta_3+e_3$.

The above model is coherent and complete (or more formally only has a harmless, probability zero chance of incompleteness) as long as e_1 and e_2 are continuously distributed, which ensures that ties in utility, which make the choice of y indeterminate, happen with probability zero.

Let $f(e_1, e_2, \lambda)$ denote the joint probability density functions of e_1 and e_2, parameterized by a vector λ, assumed independent of x. For example, if e_1 and e_2 are independent normals, then $f(e_1, e_2, \lambda) = \phi(e_1/\sigma_1)\phi(e_2/\sigma_2)/(\sigma_1\sigma_2)$ where $\lambda = (\sigma_1, \sigma_2)$ and ϕ is the standard normal probability density function. Conditioning on x, the probability of choosing $y = (1,1)$ is the probability that $V(1,1)$ is larger than the other values of V, which is

$$P_{11}(\theta \mid x) = \int_{-x'\beta_2-a}^{\infty} \left(\int_{\max[-x'\beta_1-a,-x'(\beta_1+\beta_2)-a-e_2]}^{\infty} f(e_1,e_2)de_1 \right) de_2$$

where θ denotes the set of parameters $\beta_1, \beta_2, a, \lambda$. Similarly, the probabilities of choosing other values for y, denoted P_{y_1,y_2}, are

$$P_{01}(\theta \mid x) = \int_{-\infty}^{-x'\beta_2} \left(\int_{-\infty}^{\min(-x'\beta_1-a,x'(\beta_2-\beta_1)+e_2} f(e_1,e_2)de_1 \right) de_2$$

$$P_{10}(\theta \mid x) = \int_{-\infty}^{-x'\beta_2-a} \left(\int_{\max[-x'\beta_1,x'(\beta_2-\beta_1)+e_2]}^{\infty} f(e_1,e_2)de_1 \right) de_2$$

$$P_{00}(\theta \mid x) = \int_{-\infty}^{-x'\beta_2} \left(\int_{-\infty}^{\min(-x'\beta_1,-x'(\beta_1+\beta_2)-a-e_2} f(e_1,e_2)de_1 \right) de_2$$

Assuming n draws with independent errors, the log likelihood function for this model is then, as before,

$$\sum_{i=1}^{n} y_{1i} y_{2i} \ln P_{11}(\theta \mid x_i) + (1 - y_{1i}) y_{2i} \ln P_{01}(\theta \mid x_i)$$

$$+ y_{1i}(1 - y_{2i}) \ln P_{10}(\theta \mid x_i) + (1 - y_{1i})(1 - y_{2i}) \ln P_{00}(\theta \mid x_i)$$

4.2 Are Binary Choices Substitutes or Complements?

Does engaging in one risky behavior like speeding or smoking make one more or less likely to engage other risky behaviors like not wearing seat belts or gambling? Does adopting a poison pill make firms more or less likely to adopt other antitakeover measures? Let y_1 and y_2 denote two binary choices, such as the decision to smoke and the decision to drink, respectively. If we wanted to know whether drinking makes one more or less likely to smoke, then a standard model is equation (11), where $\alpha_1 > 0$ means that drinking increases the probability of smoking, making it a complement, otherwise it is a substitute. Similarly, the sign of α_2 in equation (12) would show whether smoking increases or decreases the probability of drinking, but both equations together can be incoherent.

One complete, coherent solution is to estimate the system of equations (9) and (10). In this model the signs of α_1 and α_2 still indicate whether each choice is a substitute or a complement for the other. They can have opposite signs, e.g., if $\alpha_1 > 0$ and $\alpha_2 < 0$, then individuals that decide y_1 first, or more generally have $d = 1$, view the choices as substitutes, while for individuals that have $d = 0$, the choices are complements.

Another complete, coherent solution, one that does not require ordering the decisions, is equations (13), (14) and (15) described in the previous section. In that model a_1 and a_2 are not separately identified, since the chosen outcome only depends on their sum a. If a_1 and a_2 are known to have the same sign, then identification of a tells whether the choices are substitutes or complements. Even if a_1 and a_2 have opposite signs, it is still reasonable to say that the choices are complements if the sum a is positive and substitutes if a is negative, because having $a > 0$ in this model increases the utility of (and hence the probability of choosing) $y_1 = y_2 = 1$, relative to other choices

5 Conclusions

Necessary and sufficient conditions for coherency and completeness of simultaneous systems containing a binary choice equation were provided. One interpretation of the results is that coherency and completeness usually requires the model to be triangular or recursive, similar to Heckman's (1978) linear model result, except that nonlinearity permits the direction of causality to vary across observations. Alternatively, coherency and completeness can be obtained by nesting the behavioral models that generate each equation separately into a single larger behavioral model that determines both.

References

[1] ARADILLAS-LOPEZ, A. (2005), "Semiparametric Estimation of a Simultaneous Game with Incomplete Information," unpublished Manuscript.

[2] BLUNDELL, R. AND R. J. SMITH (1994), ”Coherency and Estimation in Simultaneous Models with Censored or Qualitative Dependent Variables,” Journal of Econometrics, 64 355-373.

[3] BRESNAHAN, T. F. AND P. C. REISS (1991), "Empirical Models of Discrete Games." Journal of Econometrics, 48, 57-81

[4] DAGENAIS, M. (1997), ”A Simultaneous Probit Model,” Unpublished Manuscript.

[5] DAS, M., (2004), "Instrumental Variables Estimators of Nonparametric Models With Discrete Endogenous Regressors.," Journal of Econometrics, 124, 335-361.

[6] GOURIEROUX, C., J. J. LAFFONT, AND A. MONFORT (1980), ”Coherency Conditions in Simultaneous Linear Equations Models with Endogenous Switching Regimes,” Econometrica, 48, 675-695.

[7] HECKMAN, J. J. (1978), ”Dummy Endogenous Variables in Simultaneous Equation Systems,” Econometrica, 46 931-960.

[8] KLEIN, R. AND F. VELLA (2001), ”The Semiparametric Binary Treatment Model Under Heteroscedasticity,” unpublished manuscript.

[9] MADDALA, G. S., AND L.-F. LEE (1976), ”Recursive Models With Qualitative Endogenous Variables,” Annals of Economic and Social Measurement, 5, 525-545.

[10] MCFADDEN, D. L., (1973), "Conditional Logit Analysis of Qualitative Choice Behavior," in Frontiers in Econometrics, edited by P. Zarembka, New York: Academic Press

[11] MCFADDEN, D. L. (1989), ”A Method of Simulated Moments for Estimation of Discrete Response Models Without Numerical Integration,” Econometrica, 57, 995-1026

[12] MANSKI, C.F. AND E. TAMER (2002), ”Inference on Regressions With Interval Data on a Regressor or Outcome,” Econometrica, 70, 519-546.

[13] TAMER, E. (2003), "Incomplete Simultaneous Discrete Response Model with Multiple Equilibria," Review of Economics Studies, 70, 147-165.

6 Appendix

Proof of Theorem 1:. Assume first that the system is coherent and complete. Suppose for a given w that $y_1 = 1$. Then for that w the equation $y_2 = H_2(1, y_2, w)$ must be complete, which requires the existence of a uniquely valued function g_{21} such that $g_{21}(w) = H_2[1, g_{21}(w), w]$. Similarly, if for the given w we have $y_1 = 0$ then there exists g_{20} such that $g_{20}(w) = H_2[0, g_{20}(w), w]$. We may then define the function g in equation (4) by $g(y_1, w) = g_{20}(w)(1-y_1)+g_{21}(w)y_1$. Substituting (4) into (1) gives $y_1 = H_1[y_1, g(y_1, w), w]$. Let $I_1(w) = H_1[1, g(1, w), w]$ and $I_0(w) = H_1[0, g(0, w), w]$. If $I_1(w) = 1$ and $I_0(w) = 0$ then equation (1) is satisfied for both $y_1 = 1$ and $y_1 = 0$, which contradicts completeness. If $I_1(w) = 0$ and $I_0(w) = 1$ then equation (1) is not satisfied by either $y_1 = 1$ or $y_1 = 0$, which also contradicts coherence. We therefore require $I_1(w) = I_0(w)$, which is equation (3). It has now been shown that completeness and coherency implies (3) and (4). To show the converse, we have that given (3) and (4) the reduced form model is given by $y_1 = I_1(w) = I_0(w)$ and $y_2 = g[I_1(w), w]$. It can then be verified by the definitions of I_1, I_0, and g that this reduced form defines unique values for y_1 and y_2 that satisfy equations (1) and (2).

Proof of Corollary 1:. Applying equation (3) in Theorem 1 shows that completeness requires $R[r(w), w] = R(0, w)$, and hence that $R(y_2, w) = R(0, w)$

for every value y_2 may take on.

Proof of Theorem 2:. Applying Theorem 1 implies completeness of (2) and (5) if and only if (4) holds and $I[s_0(w) + e_1 \geq 0] = I[s_1(w) + e_1 \geq 0]$. If $s_0(w) \neq s_1(w)$ then this completeness requirement will be violated if and only if e_1 equals the negative of any value between $s_0(w)$ and $s_1(w)$.

Now consider $s_0(w) = s_1(w)$. Define the functions r and $\widetilde{f}$ by $r(w) = g(1, w) - g(0, w)$ and $\widetilde{f}[\psi_1, \psi_2 - r(w)\psi_1, w] = h(\psi_1, \psi_2, w)$. Given equations (3) and (4), we have that $s_0(w) = s_1(w)$ implies $\widetilde{f}[0, g(0, w), w] = \widetilde{f}[1, g(0, w), w]$, so we may define the function f by $f[\psi_2 - r(w)\psi_1, w] = \widetilde{f}[0, \psi_2 - r(w)\psi_1, w]$. This shows that completeness implies existence of a function f satisfying equation (6). To show the converse, observe that for any functions f and g, the system (4) and (6) has the reduced form $y_1 = I[f[g(0, w), w] + e_1 \geq 0]$ and $y_2 = g[I[f[g(0, w), w] + e_1 \geq 0], w]$.

Next, starting from equations (4) and (6) let $d(w) = I[r(w) \neq 0]$, and $\phi[\psi, w] = f[\psi + d(w)g(0, w), w]$. With these definitions, the equivalence of (4) with (8) follows from both being equivalent to $y_2 = g(0, w) + r(w)y_1$. For the equivalence of (6) with (7) observe that $y_2 - r(w)y_1 = [1 - d(w)]y_2 + d(w)[y_2 - r(w)y_1] = [1 - d(w)]y_2 + d(w)g(0, w)$. Substituting this expression for $y_2 - r(w)y_1$ into equation (6) gives equation (7). It has now been shown that given completeness, and hence an f and g, we may construct corresponding functions ϕ and d. To show the converse, observe that the system (7) and (8) has the reduced form $y_1 = I[\phi[(1 - d(w))g(0, w), w] + e_1 \geq 0]$ and $y_2 = g[d(w)I[\phi[(1 - d(w))g(0, w), w] + e_1 \geq 0], w]$.