Journal of Econometrics 48 (1991) 57–81. North-Holland

Empirical models of discrete games

Timothy F. Bresnahan and Peter C. Reiss*
Stanford University, Stanford, CA 94305, USA

Received May 1986, final version received January 1990

This paper develops econometric models for discrete games. Specifically, we model the payoffs of games where a researcher observes qualitative or censored information about agents' decisions and payoffs. These models extend single-person qualitative choice models introduced by McFadden (1974) and others to multiple-person choice problems. The equations describing players' equilibrium strategies depend on the game's structure and the equilibrium solution concept. We show that one can describe the equilibria of a simultaneous-move Nash game with a linear system of dummy endogenous variables. We also show that sequential-move and cooperative models have different, but related, econometric structures. A series of applied examples address identification and estimation issues. These examples include models of market entry, technology adoption, tax auditing, and cooperative family labor supply.

1. Introduction

The theory of games provides a general analytical framework for modeling interrelated economic decisions. Applied researchers have used this theory to test several different models of strategic behavior. Industrial organization economists, for instance, have modeled the extent of oligopolistic competition using game-theoretic models of price and quantity competition. [See Bresnahan (1989).] To date, almost all applied game-theoretic models presume that economic agents choose continuous-valued strategies. Often, however, agents make discrete decisions or one only observes qualitative information about players' actions. These games pose new empirical issues, as one must now draw inferences about players' incentives from qualitative data.

This paper develops econometric models of games where players have discrete strategies. Specifically, we consider the following estimation prob-

*We received generous financial support from Marvin Bower, Fletcher Jones, and NBER–Olin Fellowships. We also received helpful comments from Paul Milgrom, Steven Klepper, the Editorial Board, and two anonymous referees.

lem: How can one draw inferences about agents' unobserved payoffs from qualitative data describing their actions? Our answer to this question follows the approach of *single-person* discrete choice models [e.g., McFadden (1974, 1982) and Hausman and Wise (1978)]. In single-person choice models, one estimates parameters of agents' preferences using threshold models of consumer behavior. In *multiple-person* choice models, one estimates parameters of agents' payoffs using threshold models of games.

In contrast to single-person choice models, game-theoretic models can have many different forms. In this paper, we illustrate our general approach to modeling players' decisions using one-shot perfect-information games. We focus on these simple games because they raise most of the modeling issues present in more complicated games. We begin by showing how multiple-person choice models differ from single-person choice models. In multiple-person choice models, the presence of simultaneity among players' choices complicates the econometric model of players' strategies. In some cases, this simultaneity makes it impossible to define probability statements for players' decisions. These problems can occur *even when the underlying game has well-defined equilibria*. We illustrate these problems and our resolution of them using several common two-by-two games. We show that one can represent the equilibria of a simultaneous-move Nash game with a linear dummy endogenous variable system. Thus, simultaneous-move games provide an economic rationale for dummy endogenous variable models. We also describe the structure of sequential-move and cooperative games. These games do not have simple dummy endogenous variable representations, but do have structures related to the simultaneous-move Nash model. We compare the econometric structures of these different games using a series of economic games. These examples include models of market entry, technology adoption, tax auditing, and family labor supply. Our examples also address identification and estimation issues that arise when players use mixed strategies. The conclusion discusses practical issues associated with the application of these models.

2. Stochastic specifications of games

Following McFadden's (1974) analysis of individual discrete choice models, we would like to use the theory of games to derive probability statements for players' choices (i.e., equilibrium strategies). To understand how our multiple-person choice models differ from single-person choice models, consider how an applied economist might model the entry decisions of potential producers. Suppose the economist observes a large sample of separate markets, each with a different set of potential entrants. Assume that these firms can either 'Enter' or 'Not enter' a market.[1] Single-person models of entry would estimate the parameters of firms' profit functions from threshold conditions that presume firm i enters when

$$\Pi^i = \bar{\Pi}^i(X_i, \theta) - \varepsilon^i \geq 0. \tag{1}$$

Here, ε^i represents profits the econometrician does not observe and $\bar{\Pi}^i(X_i, \theta)$ represents the conditional mean of firm i's profits. In this model, the conditional mean of firm i's profits depends on an exogenous vector of observed covariates, X_i, and a set of unknown parameters, θ. The econometrician includes in X any observed variables that explain differences in firms' profits within and across markets. The error term, ε, summarizes any remaining differences in firms' fixed costs or variable profits. [See Bresnahan and Reiss (1990).]

Multiple-person choice models differ from single-person models because they include the actions of other players in (1). That is, they recognize that entrants' profits depend on competitors' entry decisions. The presence of other players' strategies suggests that one should estimate (1) as part of a system of threshold equations. The structure of this system depends on the economic structure of the entry game and the assumptions one makes about the distribution of firms' unobserved profits. In the games we study, we follow the single-person discrete choice models discussed above by assuming that each player receives an additive payoff of the form:

$$\begin{aligned}\Pi^i(a_1, \ldots, a_N) &= \bar{\Pi}^i(a_1, \ldots, a_N, X_i, \theta) - \varepsilon^i(a_1, \ldots, a_N) \\ &= \bar{\Pi}^i_a(X_i, \theta) - \varepsilon^i_a.\end{aligned} \tag{2}$$

Here, a equals an N-dimensional vector of integers describing all players' actions. We summarize the payoffs of the game by the vector $\Pi = \bar{\Pi}(X, \theta) - \varepsilon$. Although our Π notation suggests that the players receive monetary payoffs, the Π's also can represent utility functions. In single-person discrete choice models, ε represents unobserved heterogeneity in individual tastes or characteristics. Commonly, the econometrician assumes that these unobserved tastes have a normal or logistic distribution in the population of individuals. We interpret the vector ε in a similar way. Players of the game observe ε and have complete information about other players' actions. The econometrician, however, does not observe players' payoffs and therefore treats them as random variables.

[1]These firms also may choose continuous-valued strategies, such as how much to produce. Here we assume that the econometrician only observes whether the firm produced.

Although the games we model assume players have complete information, one also can use our framework to model games of incomplete information. Incomplete-information games have more complicated error structures however. In the simplest incomplete-information games, players have rational expectations and common knowledge about the uncertainty each player faces. These games fit most easily into our complete-information framework. Indeed, one can often model these games using the profit functions in (2), provided one interprets Π^i as player i's expected payoff given i's knowledge. The error terms, the ε, then represent that part of players' expected profits one does not observe. Games with private information pose much more complicated estimation issues. For example, in games where only player i observes ε^i [e.g., Harsanyi (1973)], one must model differences in player information sets and expectations. Similar issues arise in games with mixed-strategy equilibria where players condition their strategies on private signals [e.g., Milgrom and Weber (1986)]. We do not consider these games here.

In addition to payoff functions, a game contains an equilibrium solution concept. Game theorists use the solution concept to derive players' optimal decision rules from players' payoffs. Here, we derive our econometric models in much the same way. The equilibrium solution concept allows us to infer players' decision rules from the observed and unobserved variables. These rules have the general form

$$a_i^* = a_i^*\left(\overline{\Pi}, \varepsilon, a_{\{i\}}\right), \tag{3}$$

where $a_{\{i\}}$ contains all strategies other than player i's. To illustrate how one computes a_i^*, consider the widely used Nash solution concept for simultaneous-move games. The Nash equilibrium concept imposes a series of threshold conditions on players' payoffs. Formally, the strategy a_i^* is a Nash best response to $a_{\{i\}}^*$ if

$$\Pi^i(a_1^*, \ldots, a_i^*, \ldots, a_N^*) \geq \Pi^i(a_1^*, \ldots, a_i, \ldots, a_N^*), \tag{4}$$

for all feasible a_i. These equilibrium conditions have the same revealed preference structure as those in single-person discrete choice models, except that they depend on others' choices. To determine exactly which actions form an equilibrium, one solves these relations to obtain a series of reduced-form equations. These reduced-form equations, which we denote by $a_i^{**} = a_i^{**}(X, \theta, \varepsilon)$, relate the observed and unobserved payoff components to players' actions. In solving these systems, theorists sometimes impose equality or inequality restrictions on players' payoffs. These restrictions insure that equilibrium strategies exist and are perhaps unique. In an entry game, for example, economists often assume that firms' profits never increase when competing firms enter. This economic assumption insures the existence of pure strategies and rules out some nonunique equilibria. (We illustrate this point in section 3.1.)

Generally, games do not always have unique pure-strategy equilibria. We denote the vector of all possible equilibrium strategies by the vector a^{**}. When the game has a unique pure-strategy equilibrium for admissible values of Π, a^{**} is a single-valued function of X, θ, and ε. When a pure-strategy equilibrium exists but is not unique, a^{**} is a correspondence of the form $\{a, a', a'', \dots\} \in a^{**}(\bar{\Pi}(X,\theta),\varepsilon)$. When there are no pure-strategy equilibria, a^{**} is empty. Finally, when the game has mixed-strategy equilibria, we express the outcomes of the game in terms of probability distributions over the equilibrium strategies.[2]

Games that only have unique pure-strategy equilibria are the easiest to model econometrically. In these games, one uses the stochastic structure of the model to define outcome probabilities, much as one would in single-person choice models. For example, let $I(a,X,\theta,\varepsilon)$ equal 1 when $a^{**}(\bar{\Pi}(X,\theta),\varepsilon)=a$, and 0 otherwise. The probability of a given outcome a' equals

$$\Pr(a';X,\theta) = \int_{R_\varepsilon} I(a',X,\theta,\varepsilon)\,\mathrm{d}F(\varepsilon), \tag{5}$$

where R_ε is the relevant range of the ε's and $\mathrm{d}F(\varepsilon)$ is the multivariate probability density function of the unobserved payoffs.

Games with unique mixed-strategy equilibria pose somewhat more complicated analytical issues, but raise few additional conceptual problems. We discuss the treatment of mixed strategies below in our examples. Much more difficult issues arise when the structural equations $a_i^*(\bar{\Pi},\varepsilon,a_{\{i\}})$ have more than one solution for an admissible value of X, θ, and ε. In these games, a^{**} has multiple values. When a game has several outcomes for a single value of ε, the probability that any one a occurs in equilibrium is not well-defined. This is not a trivial theoretical concern.

2.1. Pitfalls in stochastic models of games

When a game has multiple equilibria, there is no longer a unique relation between players' observed strategies and those predicted by the theory. Although it might seem as though this would rarely happen, it always happens in simultaneous-move Nash models when the errors have sufficiently wide supports. In particular, it occurs when one assumes that the errors in (2) have normal or logistic distributions.

[2]Bjorn and Vuong (1985) take a somewhat different approach to games by defining distributions over reaction functions. It is also possible to model other types of strategies, such as the correlated strategies discussed by Aumann (1974).

To see how arbitrary assumptions about the unobserved heterogeneities in players' payoffs can create modeling problems, suppose that the errors in (2) have the following conditional density:

$$
\begin{aligned}
&f\left(\varepsilon^{i}(a) \mid \varepsilon^{(i)}(a), X, \theta\right)>0, \\
&f\left(\varepsilon^{i}\left(a_{i}^{\prime}, a^{(i)}\right) \mid \varepsilon^{i}\left(a_{i}, a^{(i)}\right), X, \theta\right)>0,
\end{aligned} \tag{6}
$$

for all a, $a_i' \neq a_i$, and $\varepsilon^i(a) \in (-\infty, \infty)$. In this equation, we use the notation $\varepsilon^{(i)}(a)$ to represent a vector of errors that does not include $\varepsilon^i(a)$. The first line of (6) maintains that there is some independence of the errors across players, while the second maintains there is some independence across outcomes of the game. Such a specification holds, for example, if we assume that for each outcome of the game players' profits have independent normal distributions. We define any two outcomes of the game, a and a', as *adjacent* if no more than one player's action differs between a and a'. Similarly, two outcomes are *nonadjacent* when $a^i \neq a'^i$ and $a^j \neq a'^j$ for players i and j.

The following proposition states that nonunique equilibria always occur in simultaneous-move games with nonadjacent outcomes and errors generated from (6):

Proposition 1. Let a and a' be nonadjacent outcomes of a simultaneous-move game. The assumptions in (2) and (6) imply that

$$
P_{\text{both}} \equiv \Pr\left(\{a, a'\} \in a^{**}\left(\bar{\Pi}(X, \theta), \varepsilon\right)\right) > 0.
$$

See the appendix for a proof.

Proposition 1 establishes that 'probabilistic nonuniqueness' occurs in multiple-person choice models whose errors have infinite support. This nonuniqueness occurs for any value of $\bar{\Pi}$. It can also occur when the errors do not have arbitrarily wide supports. Intuitively, the proposition states that if one does not restrict the errors on players' payoffs, then the errors can take on values that would preclude the existence of equilibrium strategies. Thus, the econometrician's model of payoffs can violate economic restrictions that guarantee unique strategies in the underlying game.

When the econometric model of the game permits nonunique outcomes, one cannot define probability statements for the events a and a'. While one can define the probability of the events 'only a is an equilibrium', P_a, and, 'only a' is an equilibrium', $P_{a'}$, there is a third event: 'either a or a' is an equilibrium'.[3] Thus, without additional assumptions, one cannot construct

[3]The region that has two pure-strategy equilibria also contains a mixed-strategy equilibrium. One could model the use of mixed strategies, but these mixed strategies complicate the nonuniqueness problem.

probability statements for the events 'a is the observed equilibrium' or 'a' is the observed equilibrium'.

One could respond to this nonuniqueness problem by changing the model in an *ad hoc* way. For example, one could arbitrarily assume that players randomly choose among nonunique equilibria. Suppose, for instance, that a and a' are both equilibria. The econometrician could assume that a occurs with probability $\lambda_{aa'}$ and a' with probability $(1-\lambda_{aa'})$. While this procedure solves the nonuniqueness problem, it adds many nuisance parameters to the model. In a 2×3 game, for example, it would add 18 separate λ parameters. The introduction of these nuisance parameters often complicates estimation and makes it difficult to interpret players' behavior.

Instead of introducing nuisance parameters, one also could resolve nonuniqueness problems by treating the events a and a' as one event. This approach changes the model of players' strategies to one that predicts only the outcome $a \cup a'$. The aggregation of nonunique equilibrium outcomes clearly involves some loss of information and thereby restricts the class of games that can be studied. To see this, suppose that we divide all simultaneous-move games into two categories: 2×2 games, in which each of two players chooses one of two actions, and games with two or more players, each with two or more strategies. Without further restrictions on the game, the observational equivalence of nonunique equilibria and aggregation implies:

Proposition 2. In games other than 2×2 games, all outcomes are observationally equivalent under the assumptions in (6). Further, in 2×2 games there are two observational equivalence classes: $(0,0),(1,1)$ *and* $(0,1),(1,0)$.

See the appendix for a proof.

Thus, in the absence of further restrictions, nearly all outcomes are observationally equivalent. One can also show:

Proposition 3. The probability that there is no pure-strategy equilibrium under the assumptions in (2) and (6) is positive.

To summarize these three propositions, one cannot construct econometric models of simultaneous-move games when players' payoffs have the unbounded supports in (6). While other equilibrium concepts are less troubling in this regard, the presence of this problem in simultaneous-move models means that we cannot simply assume that the errors in (2) have unrestricted normal or logistic distributions.

Amemiya (1974) and Heckman (1978) have also noted that similar nonexistence and nonuniqueness problems occur in dummy endogenous dependent variable models. These papers suggest that one can resolve these problems by making the model recursive. Below we show that this approach often rules out interesting interactions among players. This leads us to modify the error structure of these models so that we can preclude nonunique and nonexistent equilibria.

3. Models of simple games

This section illustrates the specification and identification issues raised above using 2×2 games. The first example discusses nonexistence and nonuniqueness problems in simultaneous-move Nash games. It also shows that by making a game recursive, one makes the game trivial. This leads us to consider different restrictions on (2). Later subsections explore the econometric structure of different games forms and solution concepts.

3.1. Simultaneous-move games

This subsection constructs a latent variable model of a 2×2 simultaneous-move game. In this game, the two players choose one of two strategies, $a^i = \{0, 1\}$. The players have complete information about each other. Any pair of players only plays the game once.

We assume the payoffs to each player are:

Player 1's payoffs

	$a_2 = 0$	$a_2 = 1$
$a_1 = 0$	Π^1_{00}	Π^1_{01}
$a_1 = 1$	$\Pi^1_{00} + \Delta^1_0$	$\Pi^1_{01} + \Delta^1_0 + \Delta^1_1$

Player 2's payoffs

	$a_2 = 0$	$a_2 = 1$
$a_1 = 0$	Π^2_{00}	$\Pi^2_{00} + \Delta^2_0$
$a_1 = 1$	Π^2_{10}	$\Pi^2_{10} + \Delta^2_0 + \Delta^2_1$

To fix ideas, suppose this game models the decisions of two potential market entrants. Each potential entrant's profits depends on action of their competitor. The subscripts on the firms' profit functions, the Π, denote each firm's action. To interpret the structure of the payoff matrices, let $a_i = 0$ indicate firm i stays out and $a_i = 1$ denote the event that firm i enters. The Δ's have natural economic interpretations. The quantity Δ^i_0 equals the incremental profit firm i obtains from being the only firm to enter. The term Δ^i_1 is the fall in firm i's profits when firm j enters i's monopoly market. Economic theory suggests that entry by a competitor always lowers profits, i.e., $\Delta^i_1 \leq 0$.[4]

[4]Economic theory also suggests that $\Pi^1_{00} = \Pi^1_{01}$ and $\Pi^2_{00} = \Pi^2_{10}$. These restrictions do not, however, affect the players' decisions.

The Nash equilibrium conditions for this game form a two-equation system. From (4) one can derive

$$
\begin{aligned}
a^1 = 0 &\Leftrightarrow \Delta_0^1 + a^2\Delta_1^1 \leq 0, \\
a^2 = 0 &\Leftrightarrow \Delta_0^2 + a^1\Delta_1^2 \leq 0.
\end{aligned}
\tag{7}
$$

The econometrician adds to these equations a stochastic specification for firms' unobserved profits. One could, for example, treat the Δ_1^i as constants and the Δ_0^i as random variables that vary across firms and markets (i.e., games). If we assume that the incremental profit firm i receives as a monopolist in any given game is a linear function of observables and unobservables, i.e., $\Delta_0^i = X\beta_0^i - \varepsilon^i$, then the structural equations determining the Nash equilibria of this 2×2 game have the form

$$y_1^* = X_1\beta_0^1 + a^2\Delta_1^1 - \varepsilon^1,$$

$$y_2^* = X_2\beta_0^2 + a^1\Delta_1^2 - \varepsilon^2,$$

with

$$a^i = \begin{cases} 0 & \text{if} \quad y_i^* < 0 \\ 1 & \text{if} \quad y_i^* \geq 0 \end{cases} \quad \text{for} \quad i = 1,2. \tag{8}$$

These two equations form a discrete endogenous variable system for the unobserved y_i^*. This system is a version of the systems considered by Heckman (1978). Thus, the theory of simultaneous-move games provides one possible economic justification for linear dummy endogenous variable models. Notice that the linearity of the Δ's in observable and unobservable variables is a maintained assumption. One also could create more general models by making the Δ_1^i functions of both observables and unobservables.

Amemiya (1974), Heckman (1978), Maddala and Lee (1976), Schmidt (1981), and others have shown that if the errors in (8) have unbounded support, then the reduced form of (8) is not well-defined. More generally, Propositions 1 through 3 suggest that more complicated dummy endogenous variable models also have this problem. A necessary and sufficient condition for model (8) to have a well-defined reduced form for the outcomes: no entrants, monopoly for firm 1, monopoly for firm 2, and duopoly, is that the system (8) be recursive, i.e., $\Delta_1^1 \times \Delta_1^2 = 0$.[5] It might seem natural to impose

[5]See, for example, Amemiya (1974) and Heckman (1978). To see that recursivity is necessary, consider realizations of ε^i such that Δ_0^i is negative and $\Delta_0^i + \Delta_1^i$ is positive. Such events occur with positive probability if, for example, the range of the ε's is the real line. Under these conditions, both ($a^1 = 0$, $a^2 = 0$) and ($a^1 = 1$, $a^2 = 1$) solve (7). In other words, the mapping from exogenous variables and errors to endogenous variables is not single-valued. Setting one of these Δ_1^i to zero prevents this problem.

this restriction here by assuming that

$$y_1^* = X\beta_0^1 + a_2\Delta_1^1 - \varepsilon_1,$$

$$y_2^* = X\beta_0^2 - \varepsilon_2,$$

with

$$a_i = \begin{cases} 0 & \text{if} \quad y_i^* < 0 \\ 1 & \text{if} \quad y_i^* \geq 0 \end{cases} \quad \text{for} \quad i = 1,2.$$

This convenient econometric assumption produces a very unattractive model of entry. It presumes firm 2's profits do not depend on the presence of firm 1 in the market. Thus, recursive models make very special assumptions about players' interactions.

The previous section also proposed resolving nonexistence and nonuniqueness problems by insuring that the stochastic assumptions of the econometric model match the assumptions of the economic game. In entry games, the Δ_i^1 are always less than or equal to zero. By imposing this assumption, we obtain an econometric model with the following five equilibrium outcomes: the four unique solutions (0,0), (0,1), (1,0), and (1,1) and the nonunique outcome (0,1) or (1,0). Notice that *both* the underlying theoretical entry game and the econometric model have the same nonunique equilibrium solution: (0,1) *or* (1,0). By assuming that Δ_1^i is negative, however, we remove two types of problems. First, this assumption insures that a pure-strategy equilibrium exists. Second, it precludes the nonunique equilibrium (0,0) or (1,1) from occurring.

Figs. 1, 2, and 3 illustrate these points graphically. Each figure assumes without loss of generality that $X\beta_0^i = 0$, each ε^i has infinite support, and the Δ_1^i are constants. Fig. 1 represents an entry game where entry by a competitor is costly (i.e., $\Delta_1^i < 0$). Fig. 2 assumes the opposite, namely that entry by another firm is beneficial (i.e., $\Delta_1^i > 0$). Fig. 3 represents an intermediate case where entry by firm 1 is costly for firm 2 (i.e., $\Delta_1^2 < 0$), but entry by firm 2 helps firm 1 (i.e., $\Delta_1^1 > 0$). The assumptions underlying figs. 2 and 3 make little or no economic sense in a strategic entry game, and they are precisely the cases ruled out by imposing the constraint $\Delta_1^i \leq 0$. The center rectangle in each figure is of special note. In fig. 1, either firm could maintain a monopoly in this region. Thus, this region contains nonunique monopoly outcomes. The probability integrals in (5) count the probability mass in this region twice: once when firm 1 could have a monopoly and once when firm 2 could have a monopoly. This double counting makes the outcome probabilities sum to more than one. In fig. 2, a similar nonuniqueness problem occurs in the center rectangle. Here, because entry by one firm encourages entry by the

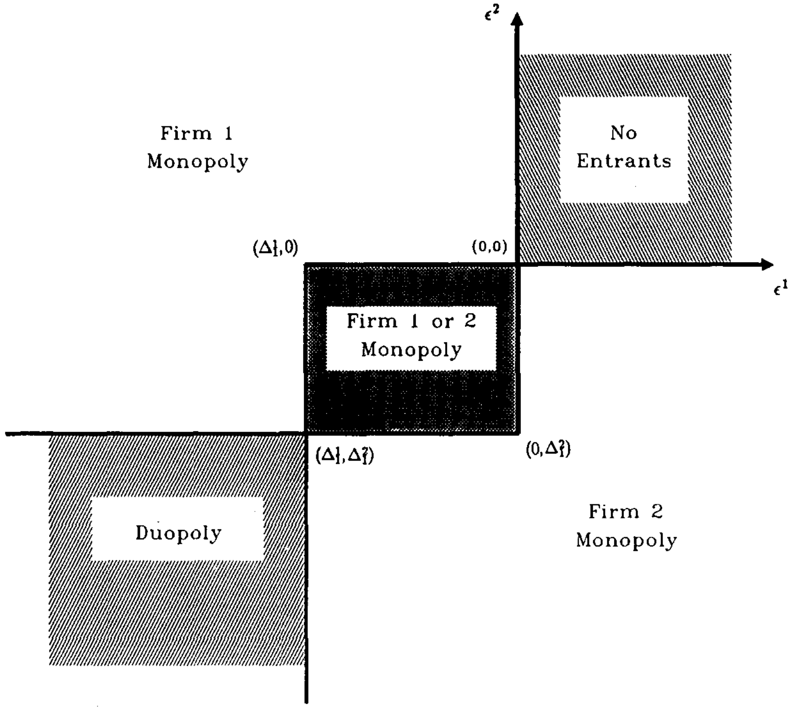

Fig. 1. The simultaneous-move entry game with costly entry (i.e., $\Delta_1^1 < 0$ and $\Delta_1^2 < 0$).

other firm, the center region supports either no firms producing or both firms producing as equilibria. In fig. 3, there are no pure-strategy Nash equilibria in the center rectangle. Because (5) does not include this center region in the outcome probabilities, the outcome probabilities will sum to less than one.[6]

To summarize, figs. 1 through 3 show why restrictions on the errors are useful. By assuming that $\Delta_1^i < 0$ we insure the existence of pure-strategies and preclude one set of nonunique strategies. One region of nonunique outcomes remains. This region with nonunique outcomes, however, can occur in the underlying theoretical game. We now consider whether this residual nonuniqueness prevents the identification of the model's parameters.

3.2. A specification strategy for the entry game

Proposition 2 suggests why the econometrician may want to treat nonunique outcomes as observationally equivalent. If we aggregate the monopoly out-

[6]Regions in which pure-strategy equilibria do not exist often support mixed-strategy equilibria. One can extend the model to include mixed strategies. See section 3.5 below.

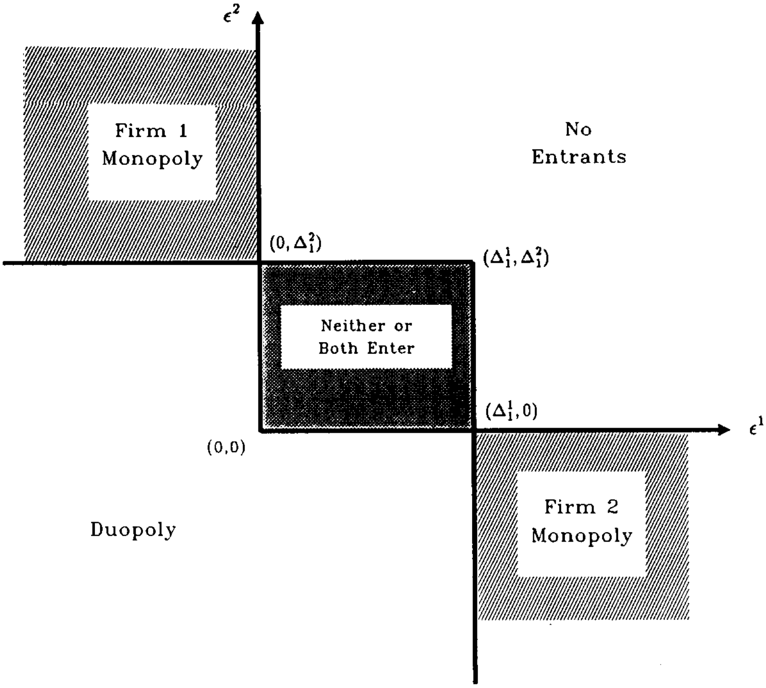

Fig. 2. The simultaneous-move entry game with beneficial entry (i.e., $\Delta_1^1 > 0$ and $\Delta_1^2 > 0$).

comes in the entry model, we transform the model from one that explains each entrant's strategy to one that predicts the number of entrants: $N = 0$, 1, or 2. The likelihood function for this model contains the following probability statements:

$$\Pr(a_1 = 0,\ a_2 = 0) = \Pr(\text{no entrants}) = \Pr\left(X\beta_0^1 < \varepsilon^1,\ X\beta_0^2 < \varepsilon^2\right),$$

$$\Pr(a_1 = 1,\ a_2 = 1) = \Pr(\text{duopoly})$$

$$= \Pr\left(X\beta_0^1 + \Delta_1^1 > \varepsilon^1,\ X\beta_0^2 + \Delta_1^2 > \varepsilon^2\right),$$

$$\Pr(\text{monopoly}) = 1 - \Pr(\text{no entrants}) - \Pr(\text{duopoly}). \tag{9}$$

The joint distribution of the Δ_1^i determines the specific functional form for these probability statements. To estimate the parameters of this model using maximum-likelihood methods, one must select a joint distribution for the Δ_1^i that respects the economic constraints on players' payoffs and permits

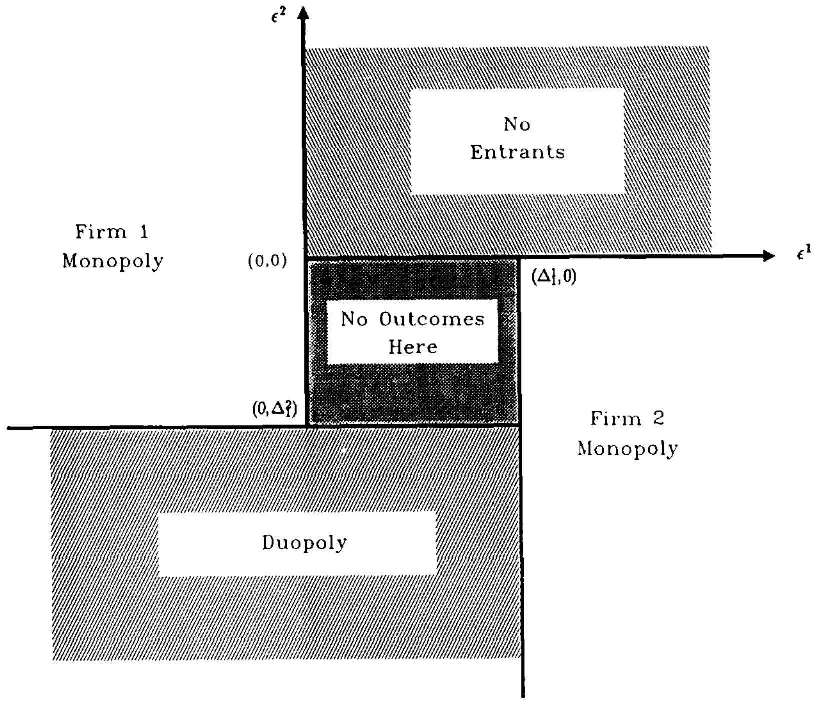

Fig. 3. The simultaneous-move entry game with asymmetric entry costs (i.e., $\Delta_1^1 > 0$ and $\Delta_1^2 < 0$).

identification of key parameters. An expedient restriction for the entry model is to assume that the Δ_1^i are negative constants. More general specifications also might allow for unobserved heterogeneities among players. For example, we could set

$$\Delta_1^i = g(Z\gamma^i) + \eta^i,$$

where $g(\cdot)$ is a function that is everywhere negative and η^i is a random variable with an upper bound of zero. Bresnahan and Reiss (1990) estimate an entry model with this structure.

The practice of constraining payoff distributions may sometimes complicate both estimation and inference procedures in these models. By aggregating outcomes, one may also introduce an identification problem. In our entry model where we combine the monopoly outcomes, for example, we must typically impose restrictions on Δ_0^i and Δ_1^i to identify the model. In general, identification must proceed on a case-by-case basis. The following subsection illustrates this point using an alternative equilibrium solution concept.

3.3. Sequential-move games

In many games, players move sequentially rather than simultaneously. In these sequential-move games the possibility of preemption arises. Preemption occurs when a prior mover takes an action that forecloses options to later movers. In an entry game, for example, an early entrant could forestall later entrants. Similarly, in a technology adoption game, early adopters may preempt later adopters. [See, for example, Gilbert and Newbery (1982).] In this subsection, we model preemption in a sequential technology adoption game. In this game, each player can either choose to 'adopt' or 'not adopt' a new technology. The first firm to adopt is the 'innovator'. The second firm adopting is an 'imitator'. Preemption occurs when the first mover blocks the second firm from adopting.

Suppose that firm 2 moves first. Solving for firm 1's optimal strategy conditional on firm 2's possible moves, we obtain the same structural equation for firm 1 as we did for the simultaneous-move game,

$$a^1 = 0 \Leftrightarrow \Delta_0^1 + a^2\Delta_1^1 \leq 0. \tag{10}$$

Here, Δ_0^i is the value of being the only firm to innovate and Δ_1^1 is the effect on firm 1's profit when firm 2 also innovates. If $\Delta_1^1 > 0$ (< 0), then innovation by the leader, firm 2, raises (lowers) firm 1's adoption profits. Eq. (10) determines a single value of a^1 for any distribution of the Δ's and any value of a^2. In a subgame-perfect equilibrium, firm 2's structural equation follows from its best response to eq. (10). That is, firm 2 maximizes $\Pi^2(a^1,(a^2,\Delta_0^1,\Delta_1^1),a^2)$. The solution to firm 2's problem is $a^2 = 1$ when

Region	Conditions	
(1)	$\Delta_0^2 > \Delta_2^2,\quad \Delta_0^1 + \Delta_1^1 < 0,\quad \Delta_0^1 \geq 0,$ firm 1 is preempted;	
(2)	$\Delta_0^2 > 0,\quad \Delta_0^1 + \Delta_1^1 < 0,\quad \Delta_0^1 < 0,$ firm 1 never adopts;	(11)
(3)	$\Delta_2^2 + \Delta_0^2 + \Delta_1^2 > 0,\quad \Delta_0^1 + \Delta_1^1 \geq 0,\quad \Delta_0^1 < 0,$ firm 1 imitates;	
(4)	$\Delta_0^2 + \Delta_1^2 > 0,\quad \Delta_0^1 + \Delta_1^1 \geq 0,\quad \Delta_0^1 \geq 0,$ firm 1 always adopts;	

where $\Delta_2^2 = \Pi_{01}^2 - \Pi_{00}^2$. In region (1), both firms can individually innovate, but by moving first firm 2 establishes itself as the innovator. In the second region, firm 1 never finds adoption profitable, no matter what firm 2 does. Area (3) is one where firm 1 can adopt only after firm 2 innovates. The last region is one where firm 1 always adopts, no matter what firm 2 does.

The conditions defining these regions do not have a convenient dummy endogenous variable representation. However, if the Δ's have continuous distributions, then relation (11) almost surely gives a unique strategy for player 2. Similarly, relation (10) almost surely generates a unique value of a^1 conditional on a^2. Thus, the sequential-move game has a unique equilibrium even when the errors have infinite support. From the conditions in (11), one can compute the probability of the event $a^2 = 1$. From (10), one can compute conditional probability statements for the event $a^1 = 1$. One can similarly determine the probabilities of the remaining strategies. From these probability statements, one can construct and maximize a likelihood function for firms' technology choices. One also can estimate the parameters of firms' profits using two-stage estimation methods. Two-stage methods provide more flexible models of the unobserved errors by exploiting the sequential structure of decisions. In the first stage, one would estimate relation (11). In the second stage, one would estimate relation (10). While two-stage methods may sometimes be less complicated than maximum-likelihood methods, two-stage methods have their own complications. For instance, one cannot easily

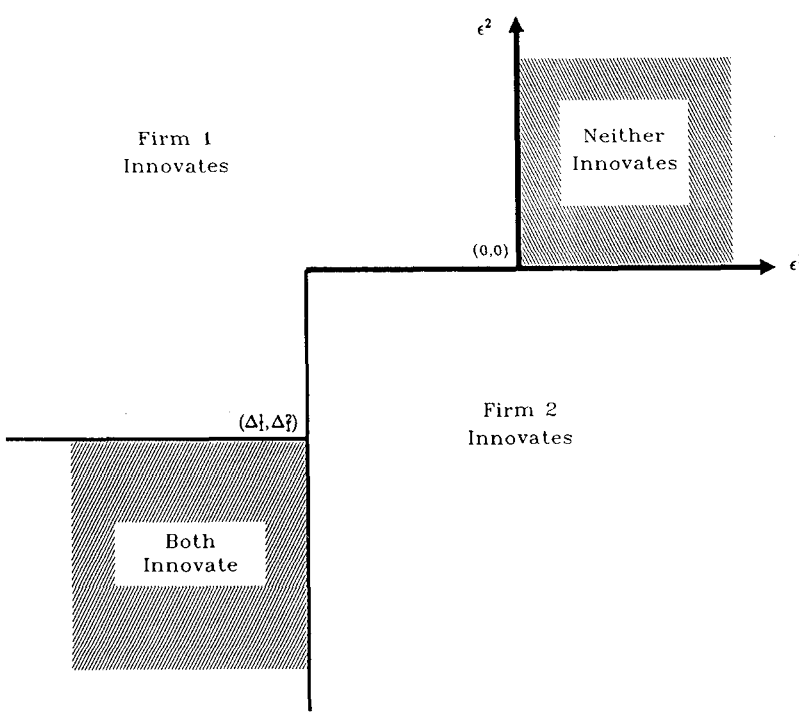

Fig. 4. The sequential-move entry game.

impose cross-equation constraints on the threshold conditions using two-stage methods.

Finally, depending on the functional forms of the Δ's, not all the parameters in this model may be identified. Identification again must be determined on a case-by-case basis. Economic theory may occasionally help with identification. In a sequential-move entry game, for instance, when a firm does not enter, its profits do not depend on the other firm's action. Fig. 4 shows how this restriction partitions the associated payoff space. (This figure uses the same assumptions as fig. 1.) The center region of this figure corresponds to the profit outcomes where firm 2 preempts firm 1 from entering. Comparing fig. 4 to fig. 1, we see that the ability of firm 2 to move first assigns all nonunique monopoly markets in the simultaneous-move game to firm 2.

The uniqueness of players' strategies in sequential-move games seems to favor the use of sequential-move models over simultaneous-move models. Notice, however, that sequential-move games impose their own special requirements. If the econometrician does not know the order in which the players moved, the sequential-move econometric model is indistinguishable from the simultaneous-move model. To see this, recall that the center region in fig. 1 contains the nonunique equilibria of the simultaneous-move game. In the sequential-move entry game, the first mover claims these monopoly markets. If we do not know which firm moved first, we cannot assign these markets to either firm. In this case, the sequential-move game has the same indeterminant outcomes as the simultaneous-move game.

3.4. *Cooperative games*

In many multi-person decision problems, players cooperate. This subsection develops the implications of a simple cooperative game where a husband and wife jointly make their labor force participation decisions. Unlike many previous econometric models of labor force participation, this model examines how the discrete participation decision of each spouse affects the participation decision of the other spouse.[7]

Suppose that we observe the labor force participation decisions of many two-person households. Assume that in addition we also observe characteristics of the household members, such as members' ages, education levels, incomes, and so on. Denote the utility functions of the husband and the wife by U^H and U^W, respectively, and let $a^i = 0$ when member i does not work and $a^i = 1$ when i works. Although we do not observe members' hours of work, we know each member's utility depends on the hours worked by both

[7]Early examples of discrete family labor supply models include Kniessner (1976) and Wales and Woodland (1976), among others. Later work by Heckman and MaCurdy (1980) and Ransom (1987) recognize the endogeneity of household decisions, but do not explicitly model household decision-making.

members, i.e., $U^W = U^W(L^H, L^W, X)$ and $U^H = U^H(L^W, L^H, X)$, where L^H and L^W equal the hours worked by the husband and wife and X represents exogenous characteristics of the household. By maximizing the sum of these two utility functions subject to the household budget constraint, we obtain two reduced-form labor supply equations as a function of X. These cooperative labor supply schedules have four pieces or segments, depending on which of the following constraints bind: $\{L^H > 0,\ L^W > 0\}$, $\{L^H = 0,\ L^W > 0\}$, $\{L^H > 0,\ L^W = 0\}$, or $\{L^H = 0,\ L^W = 0\}$. In other words, each member's labor supply schedule varies, depending upon whether both work, only the husband works, only the wife works, or neither works. Substituting these reduced-form labor supply schedules into the spouses' utility functions, we obtain the indirect utility functions

$$U^H = \bar{U}_0^H + a^H \Delta_1^H + a^W \Delta_2^H + a^H a^W \Delta_3^H, \tag{12}$$

$$U^W = \bar{U}_0^W + a^W \Delta_1^W + a^H \Delta_2^W + a^H a^W \Delta_3^W. \tag{13}$$

By construction, $\bar{U}_0^i$ equals each member's baseline utility should neither work. The variable Δ_1^i equals the change in i's utility when only i works. Consumer demand theory suggests that Δ_1^i is positive if wage income more than compensates for foregone leisure. The variable Δ_2^i equals the incremental utility (disutility) of having a spouse work. This change in utility does not depend on whether the individual works. Finally, the variable Δ_3^i represents the individual utility (disutility) from having both work. The signs and magnitudes of the Δ_1^i, Δ_2^i, and Δ_3^i determine the household preferences for work. Both the $\bar{U}_0^i$ and the Δ's depend on the exogenous variables in X.

To see how cooperative behavior affects the estimation of the indirect utility functions U^H and U^W, consider the simple case of a balkanized household. In a balkanized household, each member's indirect utility function depends on whether they alone work. Household members maximize household utility by working whenever $\Delta_1^i > 0$. Fig. 5 depicts this solution under the assumption that $\bar{U}_0^i = 0$. Following the specifications we adopted in our earlier models, we assume that the Δ_1^i contain unobserved differences in tastes. If we assume that these differences have a logistic or normal distribution, then we obtain a conventional probability model for each of the four household labor supply outcomes depicted in fig. 5.

The balkanized household model resembles previous labor force participation models because individual utilities do not depend on their spouse's labor supply decision. One obvious extension of the previous model that would allow for some interdependence is

$$U^H = \bar{U}_0^H + a^H \Delta_1^H + a^W \Delta_2^H,$$

$$U^W = \bar{U}_0^W + a^W \Delta_1^W + a^H \Delta_2^W.$$

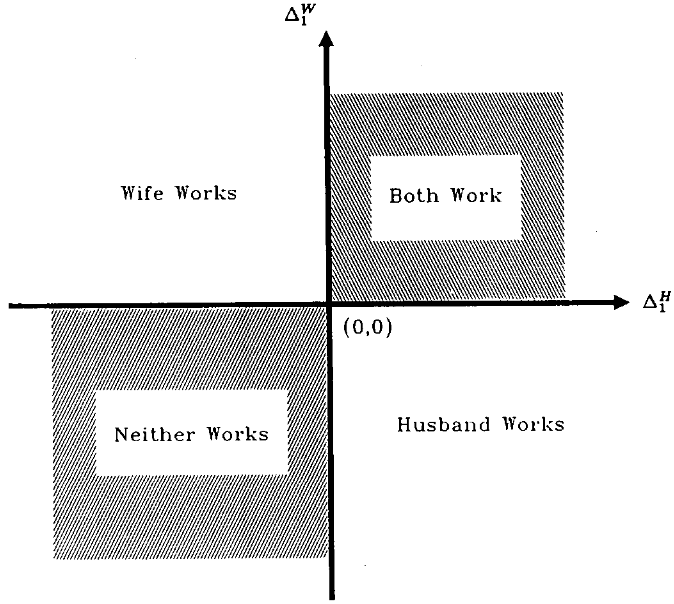

Fig. 5. The cooperative labor supply model with no interpersonal effects.

The cooperative solution for these utilities depends on both Δ_1^i and Δ_2^i. Notice, however, that the probability statements for the discrete labor supply decisions have the same form as those in the previous model. (To see this, define two new random variables $\Delta_1 = \Delta_1^H + \Delta_2^W$ and $\Delta_2 = \Delta_2^H + \Delta_1^W$.) Thus, in the absence of identifying restrictions on individual utility functions, we cannot distinguish between the disutility of having a spouse work and the own utility of working.

To obtain interesting externality effects in labor supply decisions, we must allow the indirect utility function to have interaction effects. Suppose without loss of generality that $\Delta_2^i = 0$,

$$U^H = \bar{U}_0^H + a^H \Delta_1^H + a^H a^W \Delta_3^H,$$

$$U^W = \bar{U}_0^W + a^W \Delta_1^W + a^W a^H \Delta_3^W.$$

In this model, own and spouse labor supply decisions interact, and not necessarily in the same way for each member. Figs. 6 and 7 illustrate the consequences of these interaction effects. Fig. 6 graphs the equilibrium configurations using the assumption that $\bar{U}_0^i = 0$. This figure assumes that Δ_3^i is fixed and uses the notation $\Delta_3 = \Delta_3^H + \Delta_3^W$. Fig. 6 also assumes that $\Delta_3 > 0$,

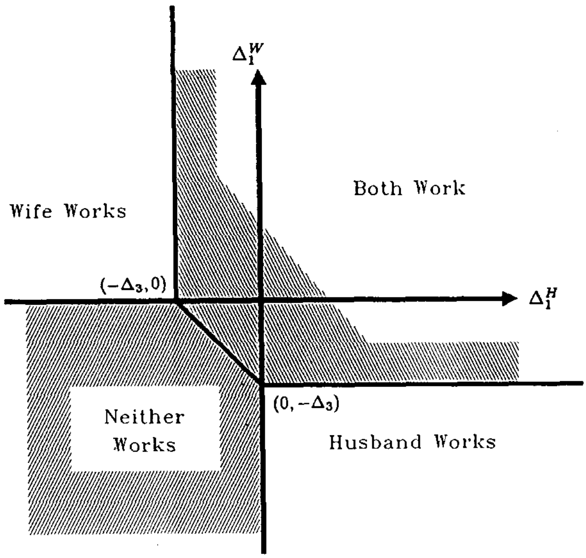

Fig. 6. The cooperative labor supply model with $\Delta_3 > 0$.

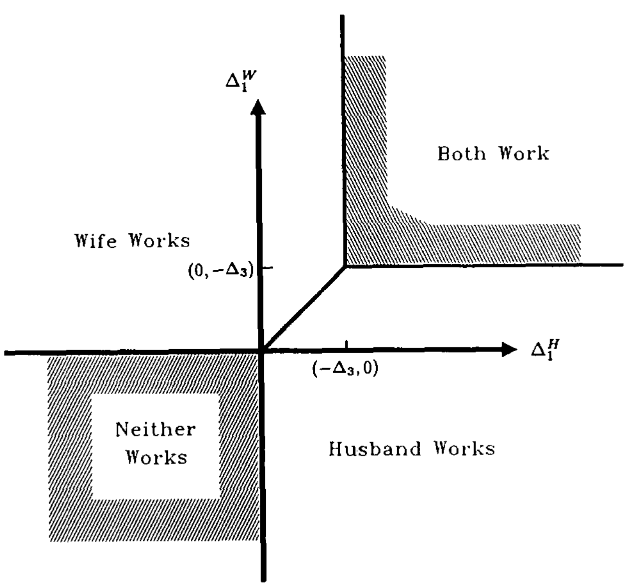

Fig. 7. The cooperative labor supply model with $\Delta_3 < 0$.

while fig. 7 assumes the opposite, namely $\Delta_3 < 0$. In other words, figs. 6 and 7 differ according to the degree of complementarity between joint home and work production. In each figure, whether only one member works depends on the other member's relative disutility of being the only one to work. While in principle one might like to identify the individual Δ_3^i terms, one usually cannot unless one has more information about each spouse's preferences.

One can easily extend the household utility models considered here to more general models of household behavior. For example, one might include the decisions of other family members and also model household members' hours of work. These models present estimation problems similar to those in the individual continuous/discrete choice models discussed by Hanemann (1984) and Dubin and McFadden (1984). Reiss and Spiller (1989) discuss some of these problems in an oligopoly model where firms make discrete entry and continuous price decisions.

3.5. *Games with mixed strategies*

So far we have considered only the pure-strategy equilibria of discrete games. We now consider a simple example where the underlying game has a mixed-strategy equilibrium. Games with mixed-strategy equilibria pose difficult inference problems because the researcher does not know whether an action reflects a randomized or a pure-strategy equilibrium outcome. In this example, we relate the randomization probabilities and the threshold conditions to a common set of variables that identify the pure-strategy equilibrium.

Consider a 2×2 game played between a tax auditor and a taxpayer.[8] The auditor decides whether to audit the taxpayer for cheating. The taxpayer decides whether to cheat. Assume the payoffs to each are

	Auditor's payoffs		Taxpayer's payoffs	
	No cheating	Cheating	No cheating	Cheating
No auditing	0	0	0	S
Auditing	$-C$	$P+S-C$	0	$-P-S$

where C equals the auditor's cost of auditing, P represents the taxpayer's penalty if caught cheating, and S equals the taxpayer's potential tax savings if he or she cheats.

It is easy to see that the only pure-strategy equilibrium of this game is 'no auditing' and 'cheating'. Further, this pure-strategy equilibrium exists only when $C \geq P+S$. When $C < P+S$ (i.e., it is profitable to audit), only a mixed-strategy equilibrium exists. The mixed-strategy equilibrium is: audit

[8] Others have studied the auditing inference problem from slightly different perspectives. See Dubin and Wilde (1986), Shibano (1986), and Alexander and Feinstein (1986).

with probability p and cheat with probability q, where

$$p = \frac{S}{P+2S} \quad \text{and} \quad q = \frac{C}{P+S}.$$

If we do not know the costs of auditing and cheating, then we must also estimate p and q. To see how one might do this, consider the event 'do not audit' and 'cheat'. This event occurs with probability one when $C \geq P + S$ and with probability $q(1-p) = C/(P+2S)$ when $C < P + S$. Because we do not observe the ratio $C/(P+2S)$, we must define the event probabilities in terms of their expected probability. Let B represent the event $P + S > C$, and let $\mathrm{E}_B[\cdot]$ represent the econometrician's expectation operator conditional on the event $P + S > C$. The probability of observing the outcome 'do not audit' and 'cheat' equals

$$\begin{aligned}\Pr(\text{cheating, no auditing}) = {} & \mathrm{E}_B[q(1-p)]\Pr(P+S>C) \\ & + \Pr(P+S<C). \qquad (14)\end{aligned}$$

The probabilities of observing each of the other three outcomes equal

$$\begin{aligned}&\Pr(\text{no cheating, no auditing}) = \mathrm{E}_B[(1-q)(1-p)]\Pr(P+S>C), \\ &\Pr(\text{no cheating, auditing}) = \mathrm{E}_B[(1-q)p]\Pr(P+S>C), \qquad (15) \\ &\Pr(\text{cheating, auditing}) = \mathrm{E}_B[qp]\Pr(P+S>C).\end{aligned}$$

The complexity of the likelihood function for this model depends on the distribution of the unobserved costs and penalties. As in our previous models, the researcher must exercise care when selecting these distributions. If one assumes the payoffs are normally distributed, then C, P, and S can be negative, in which case the expectations for the randomization probabilities do not exist. Just as there are assumptions that complicate estimation of the model, there are also assumptions that can simplify estimation. For instance, if one knows or assumes that the auditor penalizes cheaters in proportion to the amount they underreport (i.e., $P = \alpha S$), then the event probabilities simplify to $p = 1/(2+\alpha)$ and $q = C/S(1+\alpha)$. If one also assumes C and S are proportional to independent chi-squared random variables, then q is proportional to an F random variable. This latter assumption makes it easy to compute the expectations and probabilities in (14) and (15).

Although it appears that the likelihood function for this problem is no more complicated than previous likelihood functions, it may be difficult to estimate this type of model. In most applications, the investigator will require a large amount of cross-section or panel data to estimate the randomization probabilities with any reasonable degree of precision. Large amounts of data, however, also increase the computational burdens associated with estimating the expectation parameters and the payoff functions. By collecting more data the investigator may also introduce other problems. For example, by repeatedly sampling the same auditor one introduces correlation among the observed outcomes. One also may raise new modeling issues, since the same auditor may use strategies designed to affect more than one taxpayer. Such a possibility requires restructuring the empirical model to recognize both the common auditor problem and the sequencing of audits.

4. Conclusion

This paper developed econometric models of finite outcome games from assumptions about the distribution of players' payoffs. These models extended conventional single-person discrete choice models to situations involving interrelated choice. When agents make decisions noncooperatively and simultaneously, unrestricted stochastic specifications of players' payoffs produce ill-defined probability models. In response, we developed several alternative models of choice. Our simultaneous-move models had linear dummy endogenous variable representations. We also showed that sequential-move and cooperative games had related, but more complicated structures.

In concluding, we would like to comment on the relevance of our methods for practical applied work. We have written (2) and (4) as though all players have complete information. We also assumed that players only played these games once. In many practical applications, agents have imperfect information and interact repeatedly. One can, however, extend our approach to consider these applications, although the econometric models become more complicated. An important class of games that our basic framework does not address are games in which players' actions affect each others' information sets. Such situations arise in signalling and reputation games. [See, for example, Roberts (1987) and Tirole (1988).] In these games, the econometrician must model both players' priors about their competitors and their true types. Just as the players in these games often have trouble drawing inferences about their opponents, so too will the econometrician.

We regard the insights of game theory as central to many applied problems in economics, especially those involving with agency, moral hazard, and self-selection problems. To apply the models we have developed here, one will need detailed data on a series of independent, yet related games. One also will have to match the structure of the theoretical game to the hypothetical distributions of the observed and unobserved variables. These tasks pose challenges, but they are not insurmountable challenges. Elsewhere [Bresnahan and Reiss (1990)], we have used cross-section data on retail markets to test structural models of entry. We also believe that these models will prove useful in the study of such strategic practices as bargaining, contracting, and auditing.

Appendix

Proof of Proposition 1. Let $A(a)$ denote the set of outcomes that are adjacent to a. Let a' denote an outcome that is not adjacent to a. It is sufficient to prove that the event 'both a and a' are Nash equilibria' occurs with positive probability. That is,

$$\Pi^i(a) \geq \Pi^i(\alpha),$$

$$\Pi^i(a') \geq \Pi^i(\alpha)$$

occurs with positive probability for all $\alpha \in \{A(a) \cup A(a')\}$. From (2), this event occurs when

$$\bar{\Pi}^i(a, X, \theta) - \bar{\Pi}^i(\alpha, X, \theta) \geq \varepsilon^i(a) - \varepsilon^i(\alpha),$$

$$\bar{\Pi}^i(a', X, \theta) - \bar{\Pi}^i(\alpha, X, \theta) \geq \varepsilon^i(a') - \varepsilon^i(\alpha)$$

for each α. The conditional probability assumptions (6) imply that this event has positive probability.

Proof of Proposition 2. In a finite outcome game with the error structure (6), any outcome can be an equilibrium with positive probability. Without loss of generality, suppose that there are two observationally distinct outcome sets, A and B. We know from Proposition 1 that all the elements of A are adjacent to all the elements of B. That is, if

$$a_{ij} = \{a_1, \ldots, a_i, a_j, \ldots, a_N\} \in A,$$

then if $j' \neq j$ and $i' \neq i$, $a_{i'j'} \in A$.

Suppose first that the game allows either i or j to have more than two strategies. Assume without loss of generality that this is player i. Then there exists some $a_{i'k} \in A$, where $k \neq j$ or j'. However, $a_{i'k}$ is nonadjacent to $a_{ij'} \in B$. By Proposition 1, A and B therefore cannot be observationally distinct.

We now consider the multiple-player case where each player has only two strategies. Without loss of generality, suppose that there are only three players. Each outcome of this game is nonadjacent to three other outcomes. Because this game has at most eight possible outcomes, there are at most two observationally distinct sets of outcomes. If the two sets are observationally distinct, then all cells in each set are mutually nonadjacent. Consider then the outcomes $(0,0,1)$, $(0,1,0)$, and $(1,0,1)$. Although the first two outcomes are nonadjacent, it is not true that both outcomes are simultaneously nonadjacent to the third. Thus, all outcomes are observationally equivalent.

The second part of the proposition is immediate given the above proofs. The two observationally distinct outcomes are $A=\{(0,0),(1,1)\}$ and $B=\{(1,0),(0,1)\}$.

Proof of Proposition 3. Immediate from the proof of Proposition 1.

References

Alexander, C. and J.S. Feinstein, 1986, A microeconomic analysis of income tax evasion (MIT, Department of Economics, Cambridge, MA).

Amemiya, T., 1974, Multivariate regression and simultaneous equation models when the dependent variables are truncated normal, Econometrica 42, 999–1012.

Aumann, R.J., 1974, Subjectivity and correlation in randomized strategies, Journal of Mathematical Economics 1, 67–96.

Bjorn, P.A. and Q.H. Vuong, 1985, Simultaneous equations models for dummy endogenous variables: A game theoretic formulation with an application to labor force participation, Working paper no. 537 (California Institute of Technology, Pasadena, CA).

Bresnahan, T.F., 1989, Empirical studies of industries with market power, in: R. Schmalensee and R. Willig, eds., Handbook of industrial organization (North-Holland, Amsterdam).

Bresnahan, T.F. and P.C. Reiss, 1990, Entry in monopoly markets, Review of Economic Studies, forthcoming.

Dubin, J.A. and D. McFadden, 1984, An econometric analysis of residential electric appliance holdings and consumption, Econometrica 52, 345–362.

Dubin, J.A. and L.L. Wilde, 1986, An empirical analysis of federal income tax auditing and compliance, Working paper no. 615 (California Institute of Technology, Pasadena, CA).

Gilbert, R.J. and D.M. Newbery, 1982, Preemptive patenting and the persistence of monopoly, American Economic Review 72, 514–526.

Hanemann, W.M., 1984, Discrete/continuous models of consumer demand, Econometrica 52, 541–561.

Harsanyi, John, 1973, Games with randomly distributed payoffs: A new rationale for mixed strategy equilibrium points, International Journal of Game Theory 2, 1–23.

Hausman, J.A. and D.A. Wise, 1978, A conditional probit model for qualitative choice: Discrete decisions recognizing interdependence and heterogeneous preferences, Econometrica 46, 403–426.

Heckman, J.J., 1978, Dummy endogenous variables in a simultaneous equation system, Econometrica 46, 931–959.

Heckman, J.J. and T.E. MaCurdy, 1980, A life cycle model of female labor supply, Review of Economic Studies 47, 47–74.

Kniessner, T., 1976, An indirect test of complementarity in a family labor supply model, Econometrica 44, 651–659.

McFadden, D., 1974, Conditional logit analysis of qualitative choice behavior, in: P. Zarembka ed., Frontiers of econometrics (Academic Press, New York, NY) 105–142.

McFadden, D., 1982, Qualitative response models, in: Werner Hildenbrand, ed., Advances in econometrics (Cambridge University Press, Cambridge) 1–37.

Maddala, G.S. and L.F. Lee, 1976, Recursive models with qualitative endogenous variables, Annals of Economic and Social Measurement 5, 525–545.

Ransom, M., 1987, The labor supply of married men: A switching regressions model, Journal of Labor Economics 5, 63–75.

Milgrom, P. and R.J. Weber, 1986, Distributional strategies for games with incomplete information, Mathematics of Operations Research 10, 619–632.

Reiss, P.C. and P.T. Spiller, 1989, Competition and entry in small airline markets, Journal of Law and Economics 32, S179–S202.

Roberts, D.J., 1987, Battles for market share: Incomplete information, aggressive strategic pricing and competitive dynamics, in: T. Bewley, ed., Advances in economic theory (Cambridge University Press, Cambridge) 157–195.

Schmidt, P., 1981, Constraints on the parameters in simultaneous Tobit and probit models, in: C. Manski and D. McFadden, eds., Structural analysis of discrete data with econometric applications (MIT Press, Cambridge, MA) 422–434.

Shibano, T., 1986, Inference in verification games, Working paper (Stanford University, Graduate School of Business, Stanford, CA).

Tirole, J., 1988, The theory of industrial organization (MIT Press, Cambridge, MA).

Wales, T. and A. Woodland, 1976, Estimation of household utility functions and the labor supply response, International Economic Review 17, 397–410.