Empirical Models of Discrete Strategic Choices

By PETER C. REISS *

Economists have developed a variety of statistical models to analyze qualitative economic decisions. Most of these models describe the preferences and choices of individual agents, such as consumers deciding whether or not to purchase a car (e.g., Daniel McFadden, 1984). This article discusses *multiple-agent* qualitative-response models. Multiple-agent qualitative-response models describe the preferences and choices of interacting agents. These models provide useful descriptions of bargaining behavior, contract data, and the strategic behavior of firms and consumers. The first section illustrates several existing models and draws on work by Paul Bjorn and Quang Vuong (1984), Timothy Bresnahan and Reiss (1987, 1991), and Steven Berry (1992).

I. Structural Models of Discrete Choices

I will begin with an illustrative estimation problem. Suppose that one wishes to use data on homogeneous firms' entry decisions to estimate their fixed costs of production. Suppose also that one observes a large number of distinct regional markets and that there are available data describing market demand and firm input costs. If the regional markets are competitive, one can model firms' entry decisions using a conventional discrete-choice model. In this model, variations in the size of market demand and the number of incumbent firms identify the magnitude of firms' fixed costs relative to demand. If the firms are entering oligopolistic regional markets, one must use a different approach. In concentrated markets, firms' expectations about their competitors' behavior also affect entry decisions. Competitors' decisions in turn depend on other competitors' entry decisions. To know whether high fixed costs cause market concentration or whether competitor behavior causes concentration, one requires statistical models of interdependent decisions.

The theory of games provides a natural structure for modeling interdependence in consumers' or firms' qualitative decisions. Bjorn and Vuong (1984), Bresnahan and Reiss (1990, 1991), and Berry (1992) demonstrate how one can develop statistical qualitative-response models from game-theoretic models. Their approaches begin by relating discrete data on agents' decisions to game-theoretic models of agents' actions, information, payoffs, and strategies. To this description, the modeler adds an equilibrium solution concept that identifies agents' most preferred strategies. This equilibrium solution concept replaces the axiom of revealed preference used in single-agent models. Once the theoretical model is complete, the modeler postulates a stochastic specification for agents' payoffs. This stochastic specification, together with the game's solution concept, permits the modeler to calculate the probabilities of the game's outcomes and thereby to construct a likelihood function for the observed data.

To see how this approach works in practice, suppose that one would like to model how fixed costs affect the entry decisions of two potential entrants. Let $a_i = 1$ represent the event that firm i enters the market, and let $a_i = 0$ represent the event that the firm does not enter. The following normal form summarizes the firms' actions and payoffs:

Firm 1's Payoffs

	$a_2 = 0$	$a_2 = 1$
$a_1 = 0$	Π^1_{00}	Π^1_{01}
$a_1 = 1$	Π^1_{10}	Π^1_{11}

Firm 2's Payoffs

	$a_2 = 0$	$a_2 = 1$
$a_1 = 0$	Π^2_{00}	Π^2_{01}
$a_1 = 1$	Π^2_{10}	Π^2_{11}

* Graduate School of Business, Stanford University, Stanford, CA 94305-5015, and NBER.

In these payoff matrices, $\Pi^i_{a_1,a_2}$ represents the profit firm i earns if firm 1 chooses action a_1 and firm 2 chooses action a_2.

If one assumes that the observed data are from a one-shot, simultaneous-move entry game, and that one can restrict attention to pure-strategy Nash equilibria, agent's equilibrium strategies (denoted by asterisks) can be represented as a system of threshold conditions:

$$(1) \qquad a_1^* = \begin{cases} 1 & \text{if } \pi^{1*} \geq 0 \\ 0 & \text{if } \pi^{1*} < 0 \end{cases}$$

$$a_2^* = \begin{cases} 1 & \text{if } \pi^{2*} \geq 0 \\ 0 & \text{if } \pi^{2*} < 0. \end{cases}$$

In these inequalities, the π^i represent the added profits firm i obtains from entry; that is, $\pi^{1*} = (1 - a_2)(\Pi^1_{10} - \Pi^1_{00}) + a_2(\Pi^1_{11} - \Pi^1_{01})$ and $\pi^{2*} = (1 - a_1)(\Pi^2_{01} - \Pi^2_{00}) + a_1(\Pi^2_{11} - \Pi^2_{10})$. Because firms only incur fixed costs if they enter, these profit differences have fixed costs as separate terms. The main difference between the threshold conditions in (1) and the threshold conditions of single-agent models is the dependence of the choice indexes, the π^{i*}, on other agents' choices. This dependence complicates the computation of choice probabilities.

To compute the probability that neither, either, or both firms will enter a market, one must first specify the joint distribution of firms' incremental profits. A useful reference model is one in which the firms' profit differences are linear functions of a set of observable exogenous variables, X_i, estimable profit parameters, θ, and unobserved profits, ε^i. This linear structure, when combined with (1), yields the dummy endogenous variable system

$$\pi^1 = X_1\theta_1 + a_2X_1\theta_2 + \eta_1$$

$$\pi^2 = X_2\theta_3 + a_1X_2\theta_4 + \eta_2$$

which is similar to systems proposed by James Heckman (1978). Bresnahan and Reiss (1991) discuss the derivation of outcome probabilities for these systems and conditions for the probability model to be well defined.

A key advantage to deriving statistical choice models from game-theoretic models is that one can study how changes in the game and player behavior affect the probabilities of observed outcomes. To illustrate, suppose that it had been assumed in the previous model that firms moved sequentially, with firm 1 moving first. This change alters the threshold conditions describing the game's outcome. Firm 2's strategies given firm 1's initial decision become

$$a_2^{e*} = \begin{cases} 1 & \text{if } a_1 = 1 \text{ and } \pi^{2e} \geq 0 \\ 0 & \text{if } a_1 = 1 \text{ and } \pi^{2e} < 0 \end{cases}$$

$$a_2^{x*} = \begin{cases} 1 & \text{if } a_1 = 0 \text{ and } \pi^{2x} \geq 0 \\ 0 & \text{if } a_1 = 0 \text{ and } \pi^{2x} < 0 \end{cases}$$

where $\pi^{2e} = \Pi^2_{11} - \Pi^2_{10}$ is firm 2's profit from entering a firm-1 monopoly market and $\pi^{2x} = \Pi^2_{01} - \Pi^2_{00}$ is firm 2's incremental monopoly profit. Firm 1's best strategies are

$$a_1^* = \begin{cases} 1 & \text{if } \pi^{1*} \geq 0 \\ 0 & \text{if } \pi^{1*} < 0 \end{cases}$$

where $\pi^1 = a^{2e}\Pi^1_{11} + (1 - a^{2e})\Pi^1_{10} - a^{2x}\Pi^1_{01} - (1 - a^{2x})\Pi^1_{00}$. One can most easily compare the outcome probabilities for this system to those of the simultaneous-move game if the no-entry profits are set to zero (i.e., $\Pi^1_{00} = \Pi^2_{00} = \Pi^1_{01} = \Pi^2_{10} = 0$) and one assumes that monopoly profits are always greater than duopoly profits (i.e., $\Pi^1_{10} > \Pi^1_{11}$ and $\Pi^2_{01} > \Pi^2_{11}$). These assumptions imply

$$P(\text{no entrants}) = P(\Pi^1_{10} < 0, \Pi^2_{01} < 0)$$

$$P(\text{duopoly}) = P(\Pi^1_{11} \geq 0, \Pi^2_{11} \geq 0)$$

$$P(\text{firm-1 monopoly}) = P(\Pi^1_{10} \geq 0, \Pi^2_{11} < 0)$$

$$P(\text{firm-2 monopoly}) = P(\Pi^1_{11} < 0, \Pi^2_{11} \geq 0) + P(\Pi^1_{10} < 0, \Pi^2_{11} < 0 \leq \Pi^2_{01}).$$

The probability statements for markets with no entrants and duopolies are the same in both the simultaneous-move and sequential-move models. The probability statements for monopolies differ, however, and therein lies the predictive difference between the simultaneous-move and sequential-move models. In each case, fixed costs appear as a separate term in the probability statements, and thus, in principle, they can be estimated.

II. Applications and Practical Issues

The statistical choice models described in the previous section generalize to other choice problems, including ones with different decision criteria, more players, and more complicated strategy spaces. Despite their conceptual advantages, at least two practical problems have limited the use of these models. One reason is familiar to empirical researchers: it is difficult to compile extensive data on the same decision problem. A second reason is the complexity of multiple-agent decision models. This section briefly discusses potential applications of multiple-agent models and how studies have dealt with these problems.

Experimental economists have circumvented the difficulty of collecting appropriate data in the natural economy by conducting controlled laboratory experiments. These experiments have the dual advantages that they hold constant the choice problem faced by players and confounding factors that appear in field data. Despite these advantages, experimentalists usually use nonparametric tests to evaluate categorical choice data. The models described here permit structural tests. Additionally, these models would allow researchers to explain why similar subjects, playing the same game with the same "known" payoffs, may make different decisions.

There also are many areas of nonexperimental economics where one can assemble relatively large data sets that describe maximizing firm or individual behavior. Auctions are an excellent example. Bargaining, contracting, and regulatory applications also offer interesting research opportunities. For instance, the bargaining models discussed by John Kennan and Robert Wilson (1993) provide frameworks for analyzing strike, labor-negotiation, sharecropping, and arbitration data. Public goods and externality problems also are sources for potential applications.

Some of the most natural applications of these models arise in industrial organization. Most industrial-organization studies draw their data from geographically distinct markets. Bresnahan and Reiss (1987, 1990), Reiss and Pablo Spiller (1989), and Berry (1992), among others have estimated discrete game-theoretic entry and exit models. More-recent studies have estimated technological adoption and investment models. One limitation present in many of these studies is they impose strong homogeneity restrictions on firms' observed and unobserved profits. In part, this is because of the computational burden of estimating higher-dimensional multinomial probability models.

III. Computational Issues and Evidence

Multiple-agent discrete-choice models pose great computational challenges. Many maximum-likelihood models, for example, require the evaluation of high-dimensional multinomial probability integrals. The regions of integration describing the choices of agents also are nonrectangular. These computational difficulties have limited the use of maximum-likelihood methods to applications where there are few agents and choices. Berry (1992) proposed replacing the maximum-likelihood estimator in larger problems with a simulated method-of-moments estimator. In his application, he uses moment conditions for the number of firms in the market and the probabilities that individual firms will enter based on their profitability.

Newly developed simulation estimators and advances in computing power are making simulated maximum-likelihood estimators easier to compute. To date, however, there is little evidence on how simulated maximum-likelihood estimators perform on discrete strategic-choice models. Here I provide evidence on a simulated maximum-likelihood estimator for a sequential-move model. This simulated maximum-likelihood estimator employs a probability simulator proposed by Axel Börsch-Supan and Vassilis Hajivassiliou (1993). Table 1 summarizes an analysis of 500 independent data sets and estimations. Each data set consists of 200 independent replications of the two-firm sequential-move entry game described in

TABLE 1—MEDIAN ESTIMATES OF THE PARAMETERS OF A TWO-FIRM SEQUENTIAL-MOVE ENTRY GAME

True parameter value	Median ML estimate (i)	Median SIM estimate (ii)	Median ML − SIM difference (iii)
−0.80	−0.82	−0.84	0.02
	(0.33)	(0.34)	(0.05)
−1.00	−1.05	−1.10	0.04
	(0.45)	(0.42)	(0.09)
1.35	1.37	1.36	0.00
	(0.31)	(0.31)	(0.02)
1.45	1.47	1.47	0.00
	(0.30)	(0.31)	(0.02)
$\gamma = 1.00$	1.01	1.02	−0.01
	(0.18)	(0.17)	(0.02)
−0.90	−0.91	−0.87	−0.04
	(0.38)	(0.39)	(0.10)
$\rho = 0.50$	0.48	0.44	0.06
	(0.41)	(0.39)	(0.13)
Log likelihood	146.54	146.61	0.13
	(13.82)	(13.55)	(0.66)
Elapsed estimation time (sec)	28	72	−43
	(3)	(62)	(62)

Note: The estimates in the table are calculated from 500 replications of 200 paired entry decisions. Numbers in parentheses are interquartile ranges of the estimates. ML = maximum likelihood; SIM = Börsch-Supan and Hajivassiliou simulator.

Section I. These results are part of a larger study (Reiss, 1996). While an analysis of a two-player game does not directly address higher-dimensional computational issues, it does permit comparison of conventional estimation methods and newer simulation methods.

The results summarized in Tables 1 and 2 assume that firm profits have the form

$$\Pi^1_{10} = -0.8 + 1.35x_1 + \gamma x_{12} + \varepsilon_1$$

$$\Pi^1_{11} = \Pi^1_{10} - 0.9$$

$$\Pi^2_{01} = -1.0 + 1.45x_1 + \gamma x_{22} + \varepsilon_2$$

$$\Pi^2_{11} = \Pi^2_{01} - 0.9$$

where x_1 is a covariate that varies across markets, the x_{12} and x_{22} are covariates that vary across firms and markets, γ is a parameter that differs between Tables 1 and 2, and ε_1 and ε_2 are unobserved profits. The unobserved profits are assumed to have a bivariate normal distribution with unit variances and correlation parameter $\rho = 0.5$.

The columns of Table 1 summarize the performance of the two estimators. Column (i) reports the medians of the 500 maximum-likelihood (probit) estimates.[1] Column (ii) reports the medians of the 500 simulated maximum-likelihood estimates. The Borsch-Supan and Hajivassiliou estimator has the advantage that it easily generalizes to higher-dimensional problems. To facilitate comparisons, column (iii) reports the median difference between the maximum-likelihood estimate and the simulation estimate. The numbers in parentheses below each median are the interquartile ranges of the estimates.

The results are remarkable in several respects. The median estimates for both methods are close to the true parameter values. Both algorithms also provide reasonably accurate estimates of the intercept and slope terms. The estimates of the correlation coefficient are less precise but appear to be reasonable. The maximum-likelihood algorithm takes less time to compute on average, but not significantly less time. Trials of other parameter values generally produced similar results. However, several factors do affect the accuracy of the simulation estimator. One is the number of replications of the individual probabilities. These simulations use 30 replications; typically, at least ten are required to provide results comparable to the maximum-likelihood estimates. The number of observations in the four outcome cells also affects the accuracy of the simulation estimator. The parameter values used in these simulations ensure that there are enough observations in each outcome cell. (There are on average 14 markets with no firms, 48 firm-1 monopolies, 43 firm-2 monopolies, and 95 duopolies.) Neither algorithm performed well when less than 5 percent of the

[1] A Sun Sparc-10 Workstation performed the calculations. Numerical Analysis Group computer routines generated the random numbers and optimized the objective function. The estimation routines presume that firm 1 moves first.

TABLE 2—MEDIAN ESTIMATES OF THE PARAMETERS OF A TWO-FIRM SEQUENTIAL-MOVE ENTRY GAME

True parameter value	Median ML estimate (i)	Median SIM estimate (ii)	Median ML − SIM difference (iii)
−0.80	−0.86	−0.82	−0.00
	(0.44)	(0.45)	(0.06)
−1.00	−1.01	−1.15	0.02
	(0.66)	(0.58)	(0.32)
1.35	1.31	1.24	0.01
	(0.44)	(0.40)	(0.21)
1.45	1.43	1.36	0.00
	(0.39)	(0.35)	(0.18)
$\gamma = 0.00$	0.00	−0.00	0.00
	(0.09)	(0.09)	(0.01)
−0.90	−0.77	−0.49	−0.02
	(1.45)	(1.04)	(0.67)
$\rho = 0.50$	0.36	0.11	0.03
	(1.28)	(0.85)	(0.59)
Log likelihood	222.61	222.43	0.01
	(11.18)	(11.63)	(0.57)
Elapsed estimation time (sec)	27	82	−56
	(5)	(66)	(65)

Note: The estimates in the table are calculated from 500 replications of 200 paired entry decisions. Numbers in parentheses are interquartile ranges of the estimates. ML = maximum likelihood; SIM = Börsch-Supan and Hajivassiliou simulator.

sample fell in any one outcome cell. For example, the simulations in Table 1 produced a positive relationship between the sampling bias for the estimated γ and the number of markets with no entrants.

The results in Table 2 draw attention to another situation in which both algorithms fail. This is the case in which there are no firm-specific covariates (here when $\gamma = 0$). Michael Keane (1992) notes that, although the multinomial probit model is identified in this case, the likelihood function is insensitive to changes in the correlation parameter. Although my particular model is different from Keane's multinomial probit model, the results in Table 2 support his findings. The precision of the estimates deteriorates, especially those for the correlation coefficient and the duopoly profit parameter. Additionally, the maximum-likelihood estimator converges to the boundary of the parameter space (usually $\delta = 0$ or $\rho = 1$) in 273 of the estimations, while the simulation estimator does so in 154 cases. These results suggest that researchers should exercise caution when interpreting the parameters of specifications that do not have (meaningful) player or choice-specific covariates.

IV. Conclusion

Economists have recently devoted considerable attention to studying how agents make strategic decisions. Generalizations of conventional discrete-choice models now afford an opportunity to test different models of strategic choices. Progress in this area hinges both on the development of interesting data and on computational advances. The preliminary computational evidence summarized here suggests that simulated maximum-likelihood estimators may do well in estimating high-dimensional choice models.

While multiple-agent discrete-choice models have advantages, these advantages do not come without cost. The most obvious cost is that most real-world problems are complex, and any statistical model will at best be an approximation. With game-theoretic models, empirical researchers face substantial trade-offs between adhering to the theory and developing an estimable model. With additional structure there also are the costs of misspecification. For example, it is unclear how the incorrect choice of a solution concept will affect the resulting estimates. These issues deserve further study.

REFERENCES

Berry, Steven T. "Estimation of a Model of Entry in the U.S. Airline Industry." *Econometrica*, July 1992, *60*(4), pp. 889–917.

Bjorn, Paul A. and Vuong, Quang. "Simultaneous Equations Models for Dummy Endogenous Variables: A Game Theoretic Formulation with an Application to Labor Force Participation." Social Science working paper, California Institute of Technology, 1984.

Börsch-Supan, Axel and Hajivassiliou, Vassilis A. "Smooth Unbiased Multivariate Probability Simulators for Maximum Likelihood Estimation of Limited Dependent Variable

Models." *Journal of Econometrics*, August 1993, *58*(3), pp. 347–68.

Bresnahan, Timothy F. and Reiss, Peter C. "Do Entry Conditions Vary Across Markets?" *Brookings Papers on Economic Activity*, 1987, (3), pp. 833–81.

______. "Entry in Monopoly Markets." *Review of Economic Studies*, October 1990, *57*(4), pp. 531–53.

______. "Empirical Models of Discrete Games." *Journal of Econometrics*, April/May 1991, *48*(1/2), pp. 57–81.

Heckman, James. "Dummy Endogenous Variable Models in a Simultaneous Equation System." *Econometrica*, July 1978, *46*(4), pp. 931–60.

Keane, Michael P. "A Note on the Identification of the Multinomial Probit." *Journal of Business and Economics Statistics*, April 1992, *10*(2), pp. 193–200.

Kennan, John and Wilson, Robert. "Bargaining with Private Information." *Journal of Economic Literature*, March 1993, *31*(1), pp. 45–104.

McFadden, Daniel. "Econometric Analysis of Qualitative Response Models," in Zvi Griliches and Michael Intrilligator, eds., *Handbook of econometrics*. Amsterdam: North-Holland, 1984, pp. 1395–1457.

Reiss, Peter C. "A Multinomial Model of an Entry-Game." Working paper, Stanford Graduate School of Business, 1996.

Reiss, Peter C. and Spiller, Pablo T. "Competition and Entry in Small Airline Markets." *Journal of Law and Economics*, October 1989, *32*(2), pp. S179–S202.