JOURNAL OF APPLIED ECONOMETRICS, VOL. 9, 255–268 (1994)

ESTIMATION OF ECONOMETRIC MODELS OF SOME DISCRETE GAMES

PETER KOOREMAN
Wageningen University, Department of Household and Consumer Studies, POB 8060, 6700 DA Wageningen, The Netherlands

SUMMARY

This paper discusses, estimates and compares some microeconometric models for simultaneous discrete endogenous variables. The models are based on the assumption that observed endogenous variables represent the outcome of a static discrete game. I discuss models based on non-cooperative equilibrium concepts (Nash, Stackelberg), as well as models which presume Pareto optimality of observed outcomes. The models are estimated using data on the joint labor force participation decisions of husbands and wives in a sample of Dutch households.

1. INTRODUCTION

During the last two decades models for qualitative and limited dependent variables have evolved from subjects on the frontier of econometrics to well-established and widely used research tools. This is witnessed not only by the existence of several reviews on the subject but also by the availability of numerous computer packages for estimation.

In addition to univariate and multivariate models, simultaneous equations models for discrete and limited dependent variables have been proposed. An example is the simultaneous probit model:

$$\begin{cases} y_1^* = x'\beta_1 + \gamma_1 y_2 + \varepsilon_1 & \text{(1a)} \\ y_2^* = x'\beta_2 + \gamma_2 y_1 + \varepsilon_2 & \text{(1b)} \\ y_i = 1 \text{ if } y_i^* > 0, \quad i = 1,2 & \text{(1c)} \\ \quad 0 \text{ otherwise} & \end{cases}$$

Models like (1) were introduced as being the ostensibly natural adaptations of the classical linear simultaneous equations model to discrete endogenous variables. However, a well-known difficulty with model (1) and similar models is that they require some coherency restriction on parameters in order to be statistically meaningful. The root of the coherency problem is that the relationship between $(\varepsilon_1, \varepsilon_2)$ and (y_1, y_2) defined by model (1) is not one to one. For model (1) the coherency condition is $\gamma_1\gamma_2 = 0$, which essentially eliminates simultaneity from the model (see e.g. Heckman, 1978).

In applications where the discrete endogenous variables represent actions of two agents, y_i^* can often be interpreted as the difference between the utility player i attaches to $y_i = 1$ and the utility attached to $y_i = 0$, given y_j, i.e. given the action of player j ($i \neq j$). Therefore if the

CCC 0883–7252/94/030255–14

Received April 1992
Revised January 1994

utility player i derives from each of the four possible combinations of actions is denoted by $U^i(y_1, y_2)$, we have $y_1^* = U^1(1, y_2) - U^1(0, y_2)$ and $y_2^* = U^2(y_1, 1) - U^2(y_1, 0)$.

Bjorn and Vuong (1984, 1985) noted that given this utility interpretation of the latent variables, one can use the Nash or Stackelberg equilibria of a game between the two agents to relate the latent variables to the observed discrete variables, rather than using (1c). Although their papers present an interesting generalization of traditional discrete choice models, the equilibrium concepts they adopted suffer from the fact that they may yield allocations which are not Pareto optimal. Especially in applications to household behaviour, it seems more appropriate to employ an equilibrium concept which yields Pareto optimal outcomes only. Recently, Bresnahan and Reiss (1991) have discussed identification and estimation issues of a much wider range of models of discrete games, including sequential-moves games, co-operative games and games with mixed strategies.

The present paper actually estimates some models of discrete games employing data on household labour force participation. Both co-operative and nonco-operative equilibrium concepts are used. The paper is organized as follows. Section 2 presents the econometric models, based on four different equilibrium concepts. Section 3 discusses and compares the estimation results. Section 4 concludes.

2. ECONOMETRIC MODELS OF SOME DISCRETE GAMES

Consider two players who each can take one of two actions. The action player i ($i = m, f$) takes is denoted by a dummy variable y_i. The utility player i derives from each of the four possible combinations of actions is denoted by U^i (y_m, y_f). (In the sequel a combination of actions will be called an *allocation*.) The preferences of the ith player in this model are completely characterized by the ranking of his/her four utility levels. For each player 4! different rankings are possible. Therefore for the two players there are $(4!)^2 = 576$ possible combinations of utility rankings.

For empirical implementation, we parameterize the utility levels as follows:

$$\begin{array}{ll} U^m(1,1) = x'\beta_1^m + \alpha_1^m + \varepsilon_1^m & U^f(1,1) = x'\beta_1^f + \alpha_1^f + \varepsilon_1^f \\ U^m(1,0) = x'\beta_1^m + \varepsilon_1^m & U^f(1,0) = x'\beta_0^f + \alpha_0^f + \varepsilon_0^f \\ U^m(0,1) = x'\beta_0^m + \alpha_0^m + \varepsilon_0^m & U^f(0,1) = x'\beta_1^f + \varepsilon_1^f \\ U^m(0,0) = x'\beta_0^m + \varepsilon_0^m & U^f(0,0) = x'\beta_0^f + \varepsilon_0^f \end{array} \tag{2}$$

Specification (2) follows McFadden's random utility hypothesis. It decomposes $U^i(y_1, y_2)$ into a deterministic component which depends upon a vector x of observed exogenous variables, and a random component ε which follows some probability distribution; the α's and β's are fixed parameters (see e.g. McFadden, 1981). Specification (2) assumes that the change in utility of player i caused by a change of action of player j does not depend on x; for example $U^m(1,1) - U^m(1,0) = \alpha_1^m$ and $U^f(1,0) - U^f(0,0) = \alpha_0^f$. The latter assumption reduces the number of possible utility rankings per player from 24 to 6. For example, if α_1^m is positive and α_0^m is negative, then utility rankings with $U^m(1,1) - U^m(1,0) < 0$ and $U^m(0,1) - U^m(0,0) > 0$ cannot occur. The specification of constant terms in model (2) is similar to that in the 'simultaneous-move' games as proposed by Bresnahan and Reiss (1991).

In the sequel, the following additional notation will prove to be useful:

$$\beta^i = \beta_1^i - \beta_0^i; \qquad \alpha^i = \alpha_1^i - \alpha_0^i; \qquad \varepsilon^i = \varepsilon_1^i - \varepsilon_0^i; \qquad \text{for } i = m, f$$

Given the preferences of both players, we can define an equilibrium concept. The aim of this is to attach an allocation to each combination of utility rankings. We will consider the Nash equilibrium, the Stackelberg equilibrium, equilibria based on Pareto optimality only, and three mixtures of Pareto optimal equilibria and Nash equilibria. The first two of these have been proposed earlier in Bjorn and Vuong (1984) and Bjorn and Vuong (1985), respectively.

Model 1. Nash

Each player is assumed to maximize his or her utility function, given the action of the other player. Both players then adjust their actions until their decisions are mutually consistent. More formally, allocation (k, l) is a Nash equilibrium if

$$U^m(k, l) > U^m(1-k, l) \qquad \text{and} \qquad U^f(k, l) > U^f(k, 1-l) \qquad k, l = 0,1$$

Therefore, the Nash equilibrium (NE) is determined by the signs of the following utility differences (*reaction functions*): for player m:

$$U^m(1,1) - U^m(0,1) = x'\beta^m + \alpha^m\varepsilon^m \qquad \text{and} \qquad U^m(1,0) - U^m(0,0) = x'\beta^m + \varepsilon^m \qquad (3a)$$

and for player f:

$$U^f(1,1) - U^f(1,0) = x'\beta^f + \alpha^f + \varepsilon^f \qquad \text{and} \qquad U^f(0,1) - U^f(0,0) = x'\beta^f + \varepsilon^f \qquad (3b)$$

Table AI in the Appendix gives the Nash equilibria corresponding to each of the sixteen possible combinations of signs of reaction functions. In some cases there are two Nash equilibria, whereas in others there does not exist a Nash equilibrium. In case of multiple equilibria we assume, following Bjorn and Vuong (1984), that players choose one of the equilibria at random, such that each equilibrium is chosen with equal probabilities. If there is no Nash equilibrium players are assumed to choose one of the four allocations with equal probabilities. As has been noted by Bresnahan and Reiss (1991), there are several other ways of responding to the nonuniqueness problem. One possibility is to treat some combination of outcomes as one event. However, if the total number of outcomes is already small such an approach is not useful. Another possibility is to restrict the support of the error terms. Basically, this is equivalent to assigning in an *ad hoc* way smaller or even zero probabilities to some outcomes and larger probabilities to others. Given that the models offer no further basis for distinguishing between multiple equilibria, we feel that assuming random choice with equal probabilities is a natural way to proceed. Nevertheless, this assumption is *ad hoc* as well. Additional data, for example direct information on each player's preferences, is needed to investigate the validity of this assumption. Finally, it should be kept in mind that in all cases there is always a positive probability of the existence of pure strategies due to the parametric specification in model (2).

From Table AI, the likelihood contributions can be derived straightforwardly (see Bjorn and Vuong, 1984). As will be clear from Table AI, only α^m and α^f are identified, not α_1^m, α_0^m, α_1^f and α_0^f separately; the same holds true for the β's.

It is well known that the Nash equilibrium is generally not Pareto optimal. For example, consider the case where we have the reaction functions $U^m(1,1) - U^m(0,1) > 0$ and $U^m(1,0) - U^m(0,0) > 0$ for player m and the reaction functions $U^f(1,1) - U^f(1,0) < 0$ and $U^f(0,1) - U^f(0,0) > 0$ for player f. Then the NE is (1,0), but is it perfectly possible that (0,1) is Pareto more efficient than (1,0). To determine whether this is the case, we would need to know the signs of $U^m(1,0) - U^m(0,1) = x'\beta^m - \alpha_0^m + \varepsilon^m$ and $U^f(1,0) - U^f(0,1) = -x'^f\beta + \alpha_0^f - \varepsilon^f$. However, as noted earlier, α_0^m and α_0^f are not identified in the Nash model.

Model 2. Stackelberg

In a Stackelberg game the role of the players is asymmetric. One of the players (the leader), is assumed to maximize his or her utility anticipating the reaction of the other player (the follower). Formally, allocation (k, l) is a Stackelberg equilibrium (SE) with player m being the leader and player f being the follower if

$$\begin{cases} U^f(k,l) > U^f(k,1-l) \\ U^f(1-k,l) > U^f(1-k,1-l) \end{cases} \quad \text{and} \quad U^m(k,l) > U^m(1-k,l)$$

or

$$\begin{cases} U^f(k,l) > U^f(k,1-l) \\ U^f(1-k,l) < U^f(1-k,1-l) \end{cases} \quad \text{and} \quad U^m(k,l) > U^m(1-k,1-l)$$

Table AII in the Appendix gives the Stackelberg equilibrium for all possible configurations. First note that, as opposed to the NE, the Stackelberg equilibrium (SE) is always defined uniquely. In the Stackelberg model β^m, α_1^m and α_0^m are identified for the leader m, whereas for the follower f only β^f and α^f are identified. Like a NE, a SE need not be Pareto optimal. For example, consider the first case in Table AII, i.e. $U^f(1,1) > U^f(1,0)$, $U^f(0,1) > U^f(0,0)$ and $U^m(1,1) > U^m(0,1)$. Here the SE (1,1), but (0,0) might be Pareto more efficient than (1,1). The likelihood function is derived along the same lines as in case of the NE (see Bjorn and Vuong, 1985).

Model 3. Pareto Optimality Only

Allocation (k, l) is Pareto optimal if

$$\begin{cases} [U^m(k,l) > U^m(k,1-l) & \text{or } U^f(k,l) > U^f(k,1-l)] \\ & \text{and} \\ [U^m(k,l) > U^m(1-k,l) & \text{or } U^f(k,l) > U^f(1-k,l)] \\ & \text{and} \\ [U^m(k,l) > U^m(1-k,1-l) & \text{or } U^f(k,l) > U^f(1-k,1-l)] \end{cases}$$

For the case $\alpha_1^m > 0$, $\alpha_0^m > 0$, $\alpha_1^f > 0$ and $\alpha_0^f > 0$, Table AIII in the Appendix gives the Pareto optimal allocations for each of the 36 possible combinations of utility rankings. In many cases several allocations are Pareto optimal. As before, we assume that the players then choose one of the Pareto optimal allocations at random, such that each Pareto optimal allocation is chosen with equal probabilities.

From Table AIII, the likelihood contributions for the Pareto optimality model can be derived straightforwardly. For example, if allocation (0,0) is observed, the likelihood contribution becomes:

$$
\begin{aligned}
\Pr(0,0) = {} & \tfrac{1}{3}\Pr[U_{10}^m < U_{11}^m < U_{00}^m < U_{01}^m \quad \text{and} \quad U_{01}^f < U_{00}^f < U_{10}^f < U_{11}^f] \\
& + \tfrac{1}{4}\Pr[U_{10}^m < U_{11}^m < U_{00}^m < U_{01}^m \quad \text{and} \quad U_{01}^f < U_{00}^f < U_{11}^f < U_{10}^f] \\
& + \tfrac{1}{3}\Pr[U_{10}^m < U_{11}^m < U_{00}^m < U_{01}^m \quad \text{and} \quad U_{01}^f < U_{11}^f < U_{00}^f < U_{10}^f] \\
& + \tfrac{1}{4}\Pr[U_{10}^m < U_{00}^m < U_{11}^m < U_{01}^m \quad \text{and} \quad U_{01}^f < U_{11}^f < U_{00}^f < U_{10}^f] \\
& + \tfrac{1}{3}\Pr[U_{10}^m < U_{00}^m < U_{01}^m < U_{11}^m \quad \text{and} \quad U_{01}^f < U_{11}^f < U_{00}^f < U_{10}^f] \\
= {} & \tfrac{1}{3}\Pr[\varepsilon^m < -x'\beta^m - \alpha_1^m \quad \text{and} \quad -x'\beta^f - \alpha_1^f + \alpha_0^f < \varepsilon^f < -x'\beta^f] \\
& + \tfrac{1}{4}\Pr[\varepsilon^m < -x'\beta^m - \alpha_1^m \quad \text{and} \quad -x'\beta^f - \alpha_1^f < \varepsilon^f < -x'\beta^f - \min(0, \alpha_1^f - \alpha_0^f)] \\
& + \tfrac{1}{3}\Pr[\varepsilon^m < -x'\beta^m - \alpha_1^m \quad \text{and} \quad \varepsilon^f < -x'\beta^f - \alpha_1^f] \\
& + \tfrac{1}{4}\Pr[-x'\beta^m - \alpha_1^m < \varepsilon^m < -x'\beta^m - \min(0, \alpha_1^m - \alpha_0^m) \quad \text{and} \quad \varepsilon^f < -x'\beta^f - \alpha_1^f] \\
& + \tfrac{1}{3}\Pr[-x'\beta^m - \alpha_1^m + \alpha_0^m < \varepsilon^m < -x'\beta^m \quad \text{and} \quad \varepsilon^f < -x'\beta^f - \alpha_1^f].
\end{aligned} \tag{4}
$$

The second equality is based on specification (2). Note that some of the probability terms in model (4) may be zero, depending on the signs of $\alpha_1^m - \alpha_0^m$ and $\alpha_1^f - \alpha_0^f$.

Model 4. Mixed Pareto Optimality/Nash

From the point of view of predictability of a model, a large number of cases with multiple solutions as in model 3 is undesirable. It seems reasonable to choose an equilibrium concept that minimizes the number of cases where multiple solutions arise, and when they arise, that produces as few solutions as possible. The following model uses the Nash principle to reduce the number of multiple Pareto optimal solutions. Three cases can be distinguished:

- *Case 1* (One Nash equilibrium). If the game has exactly one NE and if this NE is Pareto optimal, we assume this to be the outcome of the game (see Example 1).

$$
\begin{gathered}
U^m(0,0) < U^m(1,0) < U^m(1,1) < U^m(0,1) \\
U^f(0,0) < U^f(0,1) < U^f(1,0) < U^f(1,1)
\end{gathered}
$$

Example 1. The unique Nash equilibrium (0,1) is Pareto optimal

If the unique NE is not Pareto optimal, there exists exactly one allocation at which both players are better off as compared to the Nash equilibrium.[1] The players are then assumed to choose this Pareto efficient allocation (see Example 2).

$$
\begin{gathered}
U^m(1,0) < U^m(0,0) < U^m(1,1) < U^m(0,1) \\
U^f(0,1) < U^f(0,0) < U^f(1,1) < U^f(1,0)
\end{gathered}
$$

Example 2. Allocation (1,1) is Pareto more efficient than the NE (0,0)

- *Case 2* (Two Nash equilibria). If the game has two Nash equilibria at least one of these will

[1] Let (k, l) be a Nash equilibrium (NE) which is not Pareto optimal $(k, l = 0,1)$. Because it is a NE we have $U^m(k,l) > U^m(1-k,l)$ and $U^f(k,l) > U^f(k,1-l)$. Therefore the allocation at which both players are better off than at (k,l) must be $(1-k, 1-l)$.

be Pareto optimal.[2] If only one NE is Pareto optimal, we assume this to be the outcome of the game (Example 3).

$$U^m(1,0) < U^m(0,0) < U^m(0,1) < U^m(1,1)$$
$$U^f(0,1) < U^f(0,0) < U^f(1,0) < U^f(1,1)$$

Example 3. Two Nash equilibria ((1,1) and (0,0)); only (1,1) is Pareto optimal

If both NE are Pareto optimal, the players are again assumed to choose one of these with equal probabilities (Example 4).

$$U^m(0,0) < U^m(1,0) < U^m(1,1) < U^m(0,1)$$
$$U^f(0,0) < U^f(0,1) < U^f(1,1) < U^f(1,0)$$

Example 4. Two Nash equilibria ((1,0) and (0,1)), both Pareto optimal

- *Case 3* (No Nash equilibrium). If the game does not have a NE, there may be two, three or four Pareto optimal allocations (Examples 5, 6, and 7, respectively). In such a case the players are assumed to choose one of the Pareto optimal allocations with equal probabilities.

$$U^m(0,0) < U^m(1,0) < U^m(1,1) < U^m(0,1)$$
$$U^f(0,1) < U^f(0,0) < U^f(1,0) < U^f(1,1)$$

Example 5. No Nash equilibrium; two Pareto optimal allocations ((1,1) and (0,1))

$$U^m(0,0) < U^m(1,1) < U^m(1,0) < U^m(0,1)$$
$$U^f(0,1) < U^f(0,0) < U^f(1,0) < U^f(1,1)$$

Example 6. No Nash equilibrium; three Pareto optimal allocations ((1,1), (0,1) and (1,0))

$$U^m(0,0) < U^m(1,1) < U^m(0,1) < U^m(1,0)$$
$$U^f(1,0) < U^f(0,1) < U^f(1,1) < U^f(0,0)$$

Example 7. No Nash equilibrium; all allocations Pareto optimal

---

[2] Suppose (k,l) is one of the two NE $(k,l=0,1)$. Then we have

$$U^m(k,l) > U^m(1-k,l) \tag{F1}$$

and

$$U^f(k,l) > U^f(k,1-l) \tag{F2}$$

so that the other NE must be $(1-k,1-l)$. This implies

$$U^m(1-k,1-l) > U^m(k,1-l) \tag{F3}$$

and

$$U^f(1-k,1-l) > U^f(1-k,l) \tag{F4}$$

Allocations $(1-k,l)$ and $(k,1-l)$ cannot be Pareto more efficient that (k,l) in view of equations (F1) and (F2), whereas $(1-k,l)$ and $(k,1-l)$ cannot be Pareto more efficient than $(1-k,1-l)$ in view of equations (F3) and (F4). If

$$U^i(k,l) > U^i(1-k,1-l) \qquad i=m,f \tag{F5}$$

then only (k,l) is Pareto optimal ($(1-k,1-l)$ if the inequality is reversed). If

$$U^i(k,l) > U^i(1-k,1-l) \qquad \text{and} \qquad U^j(k,l) < U^j(1-k,1-l) \qquad i \neq j$$

then both (k,l) and $(1-k,1-l)$ are Pareto optimal.

Table I. Identifiability of α's

Nash	$(\alpha_1^m - \alpha_0^m)$, $(\alpha_1^f - \alpha_0^f)$
Stackelberg, male leader	α_1^m, α_0^m, $(\alpha_1^f - \alpha_0^f)$
Stackelberg, female leader	$(\alpha_1^m - \alpha_0^m)$, α_1^f, α_0^f
Pareto optimality only	α_1^m, α_0^m, α_1^f, α_0^f
Mixed Pareto optimality/Nash	α_1^m, α_0^m, α_1^f, α_0^f

For the case $\alpha_1^m > 0$, $\alpha_0^m > 0$, $\alpha_1^f > 0$ and $\alpha_0^f > 0$, Table AV in the Appendix gives the outcomes of the game for each of the 36 possible combinations of utility rankings. (Note that Examples 6 and 7 do not arise when all α's are positive.) In order to facilitate the derivation of Table AV, table AIV gives the corresponding Nash equilibria.

Obviously, there are several other possibilities to define mixed models. One example is to start from Pareto optimality and invoke the Nash property in case of multiple Pareto optimal allocations (see Table AVI). This yields a model which is identical to model 4, except for the case with multiple Pareto optimal allocations and a single Nash equilibrium which is not Pareto optimal (Example 2). In that case the Nash equilibrium cannot help to choose among the Pareto optimal allocations, so that one would make the equal probability assumption. We prefer the approach of model 4 since it yields fewer cases with multiple solutions. A second example is to consider outcomes that are both Pareto optimal and Nash equilibria with an equal probability treatment for the cases where there are many or no Pareto Nash allocations. However, this would also yield more cases with multiple solutions than model 4 (see Table AVII).

The identifiability of the α's is summarized in Table I. In all models, only β^m and β^f can be identified, not β_1^m, β_0^m, β_1^f, and β_0^f separately.

3. AN APPLICATION TO HOUSEHOLD LABOUR FORCE PARTICIPATION

In this section the models described in Section 2 are estimated using data on the labour force participation of males and females in Dutch households. The data stem from a labour mobility survey in the Netherlands, conducted in 1985. The sample contains 849 households.

Since in the Netherlands (and in our sample) the proportion of nonparticipating males is small, it seems more interesting to model the choice between working full-time and working part-time for males rather than their choice between working and not working at all. In our empirical analysis we therefore define the dependent variables y_m and y_f as

$$\begin{aligned} y_m &= 1 \text{ if the male works at least 38 hours per week} \\ &= 0 \text{ if the male works less than 38 hours per week} \end{aligned}$$

$$\begin{aligned} y_f &= 1 \text{ if the female works a positive number of hours per week} \\ &= 0 \text{ if the female does not work} \end{aligned}$$

The 38 hours cut-off point is motivated by the fact that in the Netherlands a full-time job usually stands for a working week of 38 hours. In the total sample of 849 households allocation (1,l) is observed in 249 cases, (1,0) in 377 cases, (0,1) in 82 cases, and (0,0) in 141 cases. Table II lists the variables that have been used in the empirical analysis and gives some sample statistics. The education index ranges from 1 to 5, 1 representing the lowest and 5 the highest

Table II. Sample statistics

Variable	Description	Mean	St. dev.	Min	Max
*K*6	1 if the household contains at least one child younger than 6, 0 otherwise	0.26	0.44	0	1
FS:	# of persons in the household	3.47	1.24	2	10
WAGEM:	Wage rate of the male partner per week (in Dfl.)	15.61	5.72	5.14	59.5
AGEM;	Age of the male partner	39.70	10.01	20	63
EDUCM:	Education index of the male partner	2.78	1.08	1	5
WAGEF:	Wage rate of the female partner per week (in Dfl.)	11.94	3.09	5.19	39.38
AGEF:	Age of the female partner	37.17	9.90	18	61
EDUCF:	Education index of the female partner	2.35	2.01	1	5
NLINC:	Non-labour income of the household	64.58	138.5	0	1054.1

level of education. For nonparticipating persons the (potential) wage rate has been predicted on the basis of selectivity bias corrected wages equations.

To specify the set of explanatory variables, preliminary estimates were obtained using a bivariate probit model. In the male participation probit equation household composition variables (*K*6 and *FS*) were found to be insignificant. Nonlabour income was insignificant in both the male and female labour force participation equation, a result that appears to be common for Dutch data sets. In both equations nonlinear wages effects were found for the own wage but not for the spouse's wage. On the basis of these results, we specified the set of explanatory variables in the game-theoretic models as shown in Table III.

The models have been estimated by maximum likelihood assuming that ε^m and ε^f follow a bivariate normal distribution with zero means, unit variances and correlation ρ. Since in the Nash model and in the preliminary probit results the estimate for the correlation coefficient ρ was not significantly different from zero, we estimated the other models with $\rho = 0$.

Although the model based on Pareto optimality only is theoretically identified, its estimation suffered from a lack of convergence of the likelihood maximization algorithm. The explanation is that there is a large number of multiple equilibria in this case (cf. Table AIII), in which case one of the equilibria is chosen at random. Clearly, this reduces the role of explanatory variables.

A dash in Table III indicates that the parameter is not identified in that particular model (see Section 2, Table I).

As can be seen from equation (2) and Tables AI and AV, the game-theoretic models collapse to the bivariate probit model if $\alpha^m = 0$ and $\alpha^f = 0$. Using likelihood ratio tests we find that the bivariate probit model is rejected against the Mixed model and both Stackelberg models, but not against the Nash model. The estimates for β^m and β^f do not differ much across columns. Note that while the female wage rate does not have a significant influence on male preferences, the male wage rate has a significant negative effect on female preferences for work. The own wage effects are nonlinear for both male and female partners. For most of the wives the own wage effect is positive. For a majority of the husbands it is negative, indicating that they are on the backward-sloping part of their labour supply curve. A part of the nonlinearity of the wage effects, however, may be the result of neglected endogeneity of the wage rates due to the tax system. The estimates of α^m and α^f show some variation across columns. Note that (significant) estimates of α^m are negative whereas the estimates of α^f are positive. This implies

Table III. Estimation results (t-values in parentheses)

	Simultaneous probit	Nash	Stackelberg male leader	Stackelberg female leader	Mixed (Pareto optimality imposed)
β^m					
Constant	−0·347	−0·413	−0·096	−0·256	−0·436
	(−0·4)	(−0·7)	(−0·1)	(−0·3)	(−0·5)
WAGEM	−0·087	−0·086	−0·093	−0·086	−0·084
	(−2·6)	(−2·6)	(−2·7)	(−2·7)	(−2·6)
WAGEM^2	0·0016	0·0016	0·0016	0·0016	0·0016
	(2·2)	(2·2)	(2·1)	(2·2)	(2·1)
WAGEF	−0·0003	−0·0008	−0·0011	−0·0006	−0·0010
	(−0·0)	(−0·0)	(−0·1)	(−0·0)	(−0·1)
AGEM	0·103	0·105	0·110	0·103	0·106
	(2·5)	(3·2)	(2·5)	(2·6)	(2·6)
AGEM^2	−0·0013	−0·0013	−0·0014	−0·0013	−0·0013
	(−2·6)	(−3·4)	(−2·7)	(−2·8)	(−2·8)
β^f					
Constant	10·0	11·0	11·0	11·5	11·3
	(5·4)	(5·7)	(6·7)	(6·4)	(6·7)
K6	−0·426	−0·427	−0·465	−0·559	−0·454
	(−5·0)	(−6·1)	(−5·2)	(−5·7)	(−5·0)
FS	−0·477	−0·478	−0·492	−0·555	−0·488
	(−7·1)	(−7·4)	(−7·6)	(−7·5)	(−7·4)
WAGEF	−2·21	−2·22	−2·20	−2·46	−2·25
	(−6·6)	(−7·2)	(−7·5)	(−7·4)	(−7·6)
WAGEF^2	0·093	0·093	0·092	0·102	0·094
	(6·6)	(7·2)	(7·6)	(7·5)	(7·7)
WAGEM	−0·024	−0·025	−0·026	−0·030	−0·027
	(−1·8)	(−1·9)	(−2·3)	(−2·2)	(−2·3)
AGEF	0·224	0·225	0·238	0·289	0·235
	(3·3)	(3·6)	(4·3)	(4·5)	(4·2)
AGEF^2	−0·0032	−0·0033	−0·0035	−0·0042	−0·0034
	(−3·6)	(−4·0)	(−4·9)	(−5·1)	(−4·8)
α_1^m	—	—	−2·84	—	0·0
			(−1·1)		(fixed)
α_0^m	—	—	−2·63	—	−0·024
			(−1·0)		(−0·1)
α_1^f	—	—	—	−3·69	0·117
				(−1·9)	(0·0)
α_0^f	—	—	—	−4·18	−0·097
				(−2·2)	(0·0)
α^m	—	0·034	−0·209[a]	−0·111	—
		(0·2)	(−1·4)	(−2·4)	
α^f	0·309	0·234	0·205	0·492[b]	—
	(0·3)	(0·3)	(2·4)	(2·4)	
ρ	−0·172	−0·145	0·0	0·0	0·0
	(−0·3)	(−0·3)	(fixed)	(fixed)	(fixed)
loglikelihood	−851·07	−851·06	−844·31	−840·52	−847·58
fit[c]	0·620	0·620	0·618	0·620	0·621

[a] Implied by the estimates for α_1^m and α_0^m.
[b] Implied by the estimates for α_1^f and α_0^f.
[c] Proportion of observations with a correct joint prediction for y_m and y_f.

that, *ceteris paribus*, female participation makes the husband less inclined to work full-time, whereas (full-time) male participation makes females more inclined to participate as well.

In the absence of direct information on individual preferences, the differences between the various models are basically differences in the functional form. A way to compare the empirical performance in this case is to look at goodness-of-fit. Comparing the loglikelihood values, i.e. comparing the goodness-of-fit in terms of the Kullback–Leibler information criterion, we find that the Mixed model performs better than the Nash model, but not as good as both Stackelberg models. The highest loglikelihood is attained for the Stackelberg female leader model. If we use the proportion of correct joint predictions for y_m and y_f as a measure of goodness-of-fit, the Mixed model performs slightly better. The general picture, however, is that the differences between the various models are small. An explanation is that the cross-wage effects pick up most of the interdependence between male and female labour force participation. Apparently, once (cross-) wage effects are controlled for, little room is left for direct structural influence as represented by α^f and α^m.

4. CONCLUSIONS

In this paper we have estimated and compared various econometric models for simultaneous discrete endogenous variables. The models are based on the assumption that observed endogenous variables represent the outcome of a static discrete game. Estimation by means of maximum likelihood we found to be feasible and produced plausible results, except for the model based on Pareto optimality only. Apparently, this concept imposes insufficient structure on the model to be able to obtain sensible results.

The models that have been estimated are still relatively simple. For example, each assumes that the decision process within all households can be described by a single equilibrium concept. It seems, however, that identification and estimation of more subtle models require additional data, particularly on the individual preferences of the players.

With the insights of game theory becoming central in many areas of economic theory, it is now a challenge to exploit these insights in empirical research. This paper has attempted to do so in one particular field of applied econometrics: household labour force participation.

APPENDIX

Table AI. Nash model

	$\begin{cases} U^m(1,1) - U^m(0,1) > 0 \\ U^m(1,0) - U^m(0,0) > 0 \end{cases}$ $\Updownarrow$ $\varepsilon^m > -x'\beta^m - \min(0,\alpha^m)$	$\begin{cases} U^m(1,1) - U^m(0,1) > 0 \\ U^m(1,0) - U^m(0,0) < 0 \end{cases}$ $\Updownarrow$ $-x'\beta^m - \alpha^m < \varepsilon^m < -x'\beta^m$	$\begin{cases} U^m(1,1) - U^m(0,1) < 0 \\ U^m(1,0) - U^m(0,0) > 0 \end{cases}$ $\Updownarrow$ $-x'\beta^m < -x'\beta^m - \alpha^m$	$\begin{cases} U^m(1,1) - U^m(0,1) < 0 \\ U^m(1,0) - U^m(0,0) < 0 \end{cases}$ $\Updownarrow$ $\varepsilon^m < -x'\beta^m - \max(0,\alpha^m)$
$\begin{cases} U^f(1,1) - U^f(1,0) > 0 \\ U^f(0,1) - U^f(0,0) > 0 \end{cases}$ $\Updownarrow$ $\varepsilon^f > -x'\beta^f - \min(0,\alpha^f)$	(1,1)	(1,1)	(0,1)	(0,1)
$\begin{cases} U^f(1,1) - U^f(1,0) > 0 \\ U^f(0,1) - U^f(0,0) < 0 \end{cases}$ $\Updownarrow$ $-x'\beta^f - \alpha^f < \varepsilon^f < -x'\beta^f$	(1,1)	(1,1) or (0,0)	No NE	(0,0)
$\begin{cases} U^f(1,1) - U^f(1,0) < 0 \\ U^f(0,1) - U^f(0,0) > 0 \end{cases}$ $\Updownarrow$ $-x'\beta^f < \varepsilon^f < -x'\beta^f - \alpha^f$	(1,0)	No NE	(1,0) or (0,1)	(0,1)
$\begin{cases} U^f(1,1) - U^f(1,0) < 0 \\ U^f(0,1) - U^f(0,0) < 0 \end{cases}$ $\Updownarrow$ $\varepsilon^f < -x'\beta^f - \max(0,\alpha^f)$	(1,0)	(0,0)	(1,0)	(0,0)

Table AII. Stackelberg model, male leader, female follower

$\left\{\begin{array}{l} U^f(1,1) - U^f(1,0) > 0 \\ U^f(0,1) - U^f(0,0) > 0 \end{array}\right.$ $\Updownarrow$ $\varepsilon^f > -x'\beta^f - \min(0, \alpha^f)$	$\left\{\begin{array}{l} U^m(1,1) - U^m(0,1) > 0 \Leftrightarrow \varepsilon^m > -x'\beta^m - \alpha^m \\ (1,1) \text{ is SE} \end{array}\right.$ $\left\{\begin{array}{l} U^m(1,1) - U^m(0,1) < 0 \Leftrightarrow \varepsilon^m < -x'\beta^m - \alpha^m \\ (0,1) \text{ is SE} \end{array}\right.$
$\left\{\begin{array}{l} U^f(1,1) - U^f(1,0) > 0 \\ U^f(0,1) - U^f(0,0) < 0 \end{array}\right.$ $\Updownarrow$ $-x'\beta^f - \alpha^f < \varepsilon^f < -x'\beta^f$	$\left\{\begin{array}{l} U^m(1,1) - U^m(0,0) > 0 \Leftrightarrow \varepsilon^m > -x'\beta^m - \alpha_1^m \\ (1,1) \text{ is SE} \end{array}\right.$ $\left\{\begin{array}{l} U^m(1,1) - U^m(0,0) < 0 \Leftrightarrow \varepsilon^m < -x'\beta^m - \alpha_1^m \\ (0,0) \text{ is SE} \end{array}\right.$
$\left\{\begin{array}{l} U^f(1,1) - U^f(1,0) < 0 \\ U^f(0,1) - U^f(0,0) > 0 \end{array}\right.$ $\Updownarrow$ $-x'\beta^f < \varepsilon^f < -x'\beta^f - \alpha^f$	$\left\{\begin{array}{l} U^m(1,0) - U^m(0,1) > 0 \Leftrightarrow \varepsilon^m > -x'\beta^m + \alpha_0^m \\ (1,1) \text{ is SE} \end{array}\right.$ $\left\{\begin{array}{l} U^m(1,0) - U^m(0,1) < 0 \Leftrightarrow \varepsilon^m < -x'\beta^m + \alpha_0^m \\ (0,1) \text{ is SE} \end{array}\right.$
$\left\{\begin{array}{l} U^f(1,1) - U^f(1,0) < 0 \\ U^f(0,1) - U^f(0,0) < 0 \end{array}\right.$ $\Updownarrow$ $\varepsilon^f < -x'\beta^f - \max(0, \alpha^f)$	$\left\{\begin{array}{l} U^m(1,0) - U^m(0,0) > 0 \Leftrightarrow \varepsilon^m > -x'\beta^m \\ (1,0) \text{ is SE} \end{array}\right.$ $\left\{\begin{array}{l} U^m(1,0) - U^m(0,0) < 0 \Leftrightarrow \varepsilon^m < -x'\beta^m \\ (0,0) \text{ is SE} \end{array}\right.$

Table AIII. Pareto optimality ($\alpha_1^m > 0$, $\alpha_0^m > 0$, $\alpha_0^f > 0$ and $\alpha_1^f > 0$)

	$U_{00}^f < U_{10}^f < U_{01}^f < U_{11}^f$	$U_{00}^f < U_{01}^f < U_{10}^f < U_{11}^f$	$U_{01}^f < U_{00}^f < U_{10}^f < U_{11}^f$	$U_{00}^f < U_{01}^f < U_{11}^f < U_{10}^f$	$U_{01}^f < U_{00}^f < U_{11}^f < U_{10}^f$	$U_{01}^f < U_{11}^f < U_{00}^f < U_{10}^f$
$U_{00}^m < U_{01}^m < U_{10}^m < U_{11}^m$	(1,1)	(1,1)	(1,1)	(1,1)(1,0)	(1,1)(1,0)	(1,1)(1,0)
$U_{00}^m < U_{10}^m < U_{01}^m < U_{11}^m$	(1,1)	(1,1)	(1,1)	(1,1)(1,0)	(1,1)(1,0)	(1,1)(1,0)
$U_{10}^m < U_{00}^m < U_{01}^m < U_{11}^m$	(1,1)	(1,1)	(1,1)	(1,1)(1,0)	(1,1)(1,0)	(1,1)(1,0) (0,0)
$U_{00}^m < U_{10}^m < U_{11}^m < U_{01}^m$	(1,1)(0,1)	(1,1)(0,1)	(1,1)(0,1)	(1,1)(1,0) (0,1)	(1,1)(1,0) (0,1)	(1,1)(1,0) (0,1)
$U_{10}^m < U_{00}^m < U_{11}^m < U_{01}^m$	(1,1)(0,1)	(1,1)(0,1)	(1,1)(0,1)	(1,1)(1,0) (0,1)	(1,1)(1,0) (0,1)	(1,1)(1,0) (0,1)(0,0)
$U_{10}^m < U_{11}^m < U_{00}^m < U_{01}^m$	(1,1)(0,1)	(1,1)(0,1)	(1,1)(0,1) (0,0)	(1,1)(1,0) (0,1)	(1,1)(1,0) (0,1)(0,0)	(1,0)(0,1) (0,0)

Table AIV. Nash model ($\alpha_1^m > 0$, $\alpha_0^m > 0$, $\alpha_1^f > 0$ and $\alpha_0^f > 0$)

	$U_{00}^f < U_{10}^f < U_{01}^f < U_{11}^f$	$U_{00}^f < U_{01}^f < U_{10}^f < U_{11}^f$	$U_{01}^f < U_{00}^f < U_{10}^f < U_{11}^f$	$U_{00}^f < U_{01}^f < U_{11}^f < U_{10}^f$	$U_{01}^f < U_{00}^f < U_{11}^f < U_{10}^f$	$U_{01}^f < U_{11}^f < U_{00}^f < U_{10}^f$
$U_{00}^m < U_{01}^m < U_{10}^m < U_{11}^m$	(1,1)	(1,1)	(1,1)	(1,0)	(1,0)	(1,0)
$U_{00}^m < U_{10}^m < U_{01}^m < U_{11}^m$	(1,1)	(1,1)	(1,1)	(1,0)	(1,0)	(1,0)
$U_{10}^m < U_{00}^m < U_{01}^m < U_{11}^m$	(1,1)	(1,1)	(1,1)(0,0)	—	(0,0)	(0,0)
$U_{00}^m < U_{10}^m < U_{11}^m < U_{01}^m$	(0,1)	(0,1)	—	(1,0)(0,1)	(1,0)	(1,0)
$U_{10}^m < U_{00}^m < U_{11}^m < U_{01}^m$	(0,1)	(0,1)	(0,0)	(0,1)	(0,0)	(0,0)
$U_{10}^m < U_{11}^m < U_{00}^m < U_{01}^m$	(0,1)	(0,1)	(0,0)	(0,1)	(0,0)	(0,0)

Table AV. Mixed Pareto optimality/Nash ($\alpha_1^m > 0$, $\alpha_0^m > 0$, $\alpha_0^f > 0$ and $\alpha_0^f > 0$)

	$U_{00}^f < U_{10}^f < U_{01}^f < U_{11}^f$	$U_{00}^f < U_{01}^f < U_{10}^f < U_{11}^f$	$U_{01}^f < U_{00}^f < U_{10}^f < U_{11}^f$	$U_{00}^f < U_{01}^f < U_{11}^f < U_{10}^f$	$U_{01}^f < U_{00}^f < U_{11}^f < U_{10}^f$	$U_{01}^f < U_{11}^f < U_{00}^f < U_{10}^f$
$U_{00}^m < U_{01}^m < U_{10}^m < U_{11}^m$	(1,1)	(1,1)	(1,1)	(1,0)	(1,0)	(1,0)
$U_{00}^m < U_{10}^m < U_{01}^m < U_{11}^m$	(1,1)	(1,1)	(1,1)	(1,0)	(1,0)	(1,0)
$U_{10}^m < U_{00}^m < U_{01}^m < U_{11}^m$	(1,1)	(1,1)	(1,1)	(1,1)(1,0)	(1,1)	(0,0)
$U_{00}^m < U_{10}^m < U_{11}^m < U_{01}^m$	(0,1)	(0,1)	(1,1)(0,1)	(1,0)(0,1)	(1,0)	(1,0)
$U_{10}^m < U_{00}^m < U_{11}^m < U_{01}^m$	(0,1)	(0,1)	(1,1)	(0,1)	(1,1)	(0,0)
$U_{10}^m < U_{11}^m < U_{00}^m < U_{01}^m$	(0,1)	(0,1)	(0,0)	(0,1)	(0,0)	(0,0)

Table AVI. Definition of mixed model; alternative 1 ($\alpha_1^m > 0$, $\alpha_0^m > 0$, $\alpha_0^f > 0$ and $\alpha_0^f > 0$)

	$U_{00}^f < U_{10}^f < U_{01}^f < U_{11}^f$	$U_{00}^f < U_{01}^f < U_{10}^f < U_{11}^f$	$U_{01}^f < U_{00}^f < U_{10}^f < U_{11}^f$	$U_{00}^f < U_{01}^f < U_{11}^f < U_{10}^f$	$U_{01}^f < U_{00}^f < U_{11}^f < U_{10}^f$	$U_{01}^f < U_{11}^f < U_{00}^f < U_{10}^f$
$U_{00}^m < U_{01}^m < U_{10}^m < U_{11}^m$	(1,1)	(1,1)	(1,1)	(1,0)	(1,0)	(1,0)
$U_{00}^m < U_{10}^m < U_{01}^m < U_{11}^m$	(1,1)	(1,1)	(1,1)	(1,0)	(1,0)	(1,0)
$U_{10}^m < U_{00}^m < U_{01}^m < U_{11}^m$	(1,1)	(1,1)	(1,1)	(1,1)(1,0)	(1,1)(1,0)	(0,0)
$U_{00}^m < U_{10}^m < U_{11}^m < U_{01}^m$	(0,1)	(0,1)	(1,1)(0,1)	(1,0)(0,1)	(1,0)	(1,0)
$U_{10}^m < U_{00}^m < U_{11}^m < U_{01}^m$	(0,1)	(0,1)	(1,1)(0,1)	(0,1)	(1,1)(1,0) (0,1)	(0,0)
$U_{10}^m < U_{11}^m < U_{00}^m < U_{01}^m$	(0,1)	(0,1)	(0,0)	(0,1)	(0,0)	(0,0)

Table AVII. Definition of mixed model; alternative 2 ($\alpha_1^m > 0$, $\alpha_0^m > 0$, $\alpha_0^f > 0$ and $\alpha_0^f > 0$)

	$U_{00}^f < U_{10}^f < U_{01}^f < U_{11}^f$	$U_{00}^f < U_{01}^f < U_{10}^f < U_{11}^f$	$U_{01}^f < U_{00}^f < U_{10}^f < U_{11}^f$	$U_{00}^f < U_{01}^f < U_{11}^f < U_{10}^f$	$U_{01}^f < U_{00}^f < U_{11}^f < U_{10}^f$	$U_{01}^f < U_{11}^f < U_{00}^f < U_{10}^f$
$U_{00}^m < U_{01}^m < U_{10}^m < U_{11}^m$	(1,1)	(1,1)	(1,1)	(1,0)	(1,0)	(1,0)
$U_{00}^m < U_{10}^m < U_{01}^m < U_{11}^m$	(1,1)	(1,1)	(1,1)	(1,0)	(1,0)	(1,0)
$U_{10}^m < U_{00}^m < U_{01}^m < U_{11}^m$	(1,1)	(1,1)	(1,1)	(1,1)(1,0) (0,1)(0,0)	(1,1)(1,0) (0,1)(0,0)	(0,0)
$U_{00}^m < U_{10}^m < U_{11}^m < U_{01}^m$	(0,1)	(0,1)	(1,1)(1,0) (0,1)(0,0)	(1,0)(0,1)	(1,0)	(1,0)
$U_{10}^m < U_{00}^m < U_{11}^m < U_{01}^m$	(0,1)	(0,1)	(1,1)(1,0) (0,1)(0,0)	(0,1)	(1,1)(1,0) (0,1)(0,0)	(0,0)
$U_{10}^m < U_{11}^m < U_{00}^m < U_{01}^m$	(0,1)	(0,1)	(0,0)	(0,1)	(0,0)	(0,0)

ACKNOWLEDGEMENTS

The helpful comments of anonymous referees are gratefully acknowledged. Financial support has been provided by the Netherlands Organization for Scientific Research (NWO) and the Royal Netherlands Academy of Arts and Sciences (KNAW).

REFERENCES

Bjorn, P., and Q. Vuong (1984), 'Simultaneous models for dummy endogenous variables: a game theoretic formulation with an application to household labor force participation', Working paper, California Institute of Technology.

Bjorn P., and Q. Vuong (1985), 'Econometric modeling of a Stackelberg game with an application to household labor force participation', Working paper, California Institute of Technology.

Bresnahan, T. F., and P. C. Reiss (1991), 'Empirical models of discrete games', *Journal of Econometrics*, **48**, 57–81.

Heckman, J. J. (1978), 'Dummy endogenous variables in a simultaneous equation system, *Econometrica*, **46**, 931–60.

McFadden, D. (1981), 'Econometric models of probabilistic choice', in C. Manski and D. McFadden (eds), *Structural Analysis of Discrete Data: With Econometric Applications*, MIT Press, Cambridge, MA.

References

Dummy Endogenous Variables in a Simultaneous Equation System
James J. Heckman
Econometrica, Vol. 46, No. 4. (Jul., 1978), pp. 931-959.
Stable URL:
http://links.jstor.org/sici?sici=0012-9682%28197807%2946%3A4%3C931%3ADEVIAS%3E2.0.CO%3B2-F