ECONOMETRIC MODELING OF A STACKELBERG GAME WITH AN APPLICATION TO LABOR FORCE PARTICIPATION

Paul A. Bjorn and Quang H. Vuong
California Institute of Technology
ECONOMETRIC MODELING OF A STACKELBERG GAME WITH AN APPLICATION TO LABOR FORCE PARTICIPATION

Paul A. Bjorn and Quang H. Vuong
California Institute of Technology

ABSTRACT

Following Bjorn and Vuong (1984), a model for dummy endogenous variables is derived from a game theoretic framework where the equilibrium concept used is that of Stackelberg. A distinctive feature of our model is that it contains as a special case the usual recursive model for discrete endogenous variables (see e.g., Maddala and Lee (1976)). A structural interpretation of this latter model can then be given in terms of a Stackelberg game in which the leader is indifferent to the follower’s action. Finally, the model is applied to a study of husband/wife labor force participation.
1. INTRODUCTION

Over the last few decades, economists have become increasingly interested in the modeling of choice over a finite number of alternatives (see, e.g., Manski and McFadden (1981), Maddala (1983)). Although the first models were essentially single equation in nature, the literature on discrete variable models has rapidly evolved into simultaneous modeling (see e.g., Amemiya (1978) and Heckman (1978)).

In an earlier paper (Bjorn and Vuong (1984)), we proposed an alternative simultaneous model for discrete endogenous variables. A distinctive feature of our model is that no logical consistency constraints on the parameters need to be imposed. In addition, our simultaneous model was derived from optimizing behavior as an outcome of a game between two players. The equilibrium concept used was that of Nash.

Following this game theoretic formulation, we shall still assume that each player maximizes his own utility. The model proposed in the present paper is, however, different from our earlier simultaneous model since the equilibrium concept used here will be that of Stackelberg. Though it may appear that the model is recursive, it will be seen that the model in fact generalizes recursive models for discrete endogenous variables that have been considered up to now in the literature (see, e.g., Maddala and Lee (1976)). As before, our model becomes stochastic by adopting the random utility framework introduced by McFadden (1974, 1981).

As an empirical application, we shall study the joint decision of a husband and wife whether or not to participate in the labor force. The statistical model is derived by assuming the husband is the Stackelberg leader and his wife the follower. That is, we assume the husband knows what action his wife will take conditional upon his action and he thus optimizes accordingly.

The paper is organized as follows. In Section 2, we derive the statistical model where the outcomes are generated as Stackelberg equilibria of a game played between two players. Section 3 compares the usual formulation of the problem in terms of recursive models with our alternative formulation. In particular, it is shown that the usual recursive model is nested in our more general model. In Section 4, we discuss identification and estimation of our model. In Section 5, the empirical example of husband/wife labor force participation is presented. Section 6 concludes the paper. Proofs of all propositions are presented in the Appendix and the construction of the data can be found in Bjorn and Vuong (1984).
2. THE MODEL

For ease of exposition, assume that the husband is the Stackelberg leader and the wife is the follower. Let \( U_h(i,j) \) be the payoff to the husband when he takes action \( i \) and his wife takes action \( j \), \( i,j \in \{0,1\} \). Analogously, let \( U_w(j,i) \) be the payoff to the wife. Then we have the extensive form:

\[
\begin{align*}
\text{Husband} & : & \text{Wife} \\
1 & : & 1,1 \\
0 & : & 0,0 \\
1 & : & 1,0 \\
0 & : & 0,1
\end{align*}
\]

The husband, in making his decision whether to take action 1 or 0 must take the wife's payoffs into account. That is, the husband must take action 1 such that when the wife takes action \( j \), conditional on \( i \), \( U_h(i,j) \) gives the husband the greatest possible payoff. There are four possible cases, \( W_1, W_2, W_3, \) and \( W_4 \) for the husband to consider before taking his action \( i \):

\[
\begin{align*}
W_1 & : U_w(1,0) \geq U_w(0,0) \neq U_w(1,1) \neq U_w(0,1) \\
W_2 & : U_w(1,0) \leq U_w(0,0) \neq U_w(1,1) \neq U_w(0,1) \\
W_3 & : U_w(1,0) \leq U_w(0,0) \neq U_w(1,1) \leq U_w(0,1) \\
W_4 & : U_w(1,0) \leq U_w(0,0) \neq U_w(1,1) < U_w(0,1)
\end{align*}
\]

The four cases \( W_1, W_2, W_3, \) and \( W_4 \) are the wife's reaction functions as given in Figure 2. For example, reaction function \( W_1 \) says that whether the husband chooses action 1 or 0, the wife always chooses action 1. Conditional on the reaction function chosen by the wife, the husband then takes that action which maximizes his payoff. For example, if the wife follows reaction function \( W_1 \), the husband will choose action 1 when \( U_h(1,1) \geq U_h(0,1) \), while choosing action 0 when the inequality is reversed. Thus, each reaction function \( W_i \) for the wife calls for a payoff comparison \( C_i \) for the husband. Therefore we define:
Let $\bar{C}_1$ indicate the negation of $C_1$.

Now that the reaction functions for the wife $W_1$ and the payoff comparisons for the husband $C_1$ have been defined, we can readily find the Stackelberg outcomes of this game, as indicated in Table 1. Note that for each outcome, the first number in each ordered pair refers to the husband while the second number refers to the wife.

Table 1: Stackelberg Equilibria

| $W_1$ & $C_1$ | (1,1) |
| $W_1$ & $C_2$ | (0,1) |
| $W_2$ & $C_1$ | (1,1) |
| $W_2$ & $C_2$ | (0,0) |

To introduce a stochastic structure, we shall follow McFadden (1974, 1981). The utilities $\bar{U}_h(1,1)$ and $\bar{U}_w(1,1)$ are then treated as random, and decomposed into deterministic components and random components. Further, we shall allow for the possibility that the utility the husband receives depends on the wife's decision whether or not to work. We make a similar allowance for the wife. Then formally we have the following set of four equations:

\[
\begin{align*}
U_h(1,Y_w) &= U^1_h + a^1_h Y_w + \eta_h \\
U_h(0,Y_w) &= U^0_h + a^0_h Y_w + \eta_h \\
U_w(1,Y_h) &= U^1_w + a^1_w Y_h + \eta_w \\
U_w(0,Y_h) &= U^0_w + a^0_w Y_h + \eta_w
\end{align*}
\]

where

\[
Y_h = \begin{cases} 
1 & \text{if the husband works} \\
0 & \text{otherwise}
\end{cases}
\]

\[
Y_w = \begin{cases} 
1 & \text{if the wife works} \\
0 & \text{otherwise}
\end{cases}
\]

To illustrate, the utility that the husband receives from working when his wife also works ($Y_w = 1$) is given by

\[
U_h(1,Y_w) = U^1_h + a^1_h Y_w + \eta_h
\]

As can be seen from the wife's reaction functions $W_i$ and the husband's utility (payoff) comparisons $C_i$, $i = 1, 2, 3, 4$, only differences in utilities are relevant in the husband's and wife's respective decisions whether or not to work. As a result, we define $\epsilon_h = \eta_h - \eta_h$ and $\epsilon_w = \eta_w - \eta_w$. It is assumed thereafter that the pair $(\epsilon_h, \epsilon_w)$ is normally distributed with zero means, unit variances and correlation $\rho$.

The distribution of the random components $(\epsilon_h, \epsilon_w)$ then induces a probabilistic structure on the observed decisions $(Y_h, Y_w)$. Indeed, each reaction function $W_i$ for the wife will occur if some conditions on the random component $\epsilon_w$ are satisfied. For instance, reaction function $W_1$ arises if and only if $U^1_w - U^0_w + \epsilon_w \geq 0$ and $U^1_w - U^0_w + a^1_w - a^0_w - \epsilon_w \geq 0$. Once a reaction function for the wife is
determined, a utility comparison for the husband is also determined; that is, if the wife's reaction function is given by \( W_1 \), the husband makes utility comparison \( C_1, i \in \{1, \ldots, 4\} \). As with the wife, each utility comparison \( C_i \) will occur if a certain condition on the random component \( \varepsilon_h \) is satisfied. For instance \( C_1 \) holds if and only if 
\[
U_{h}^1 - U_{h}^0 + \alpha_{h}^1 - \alpha_{h}^0 + \varepsilon_h \geq 0.
\]
As shown in the Appendix, the conditions that must be satisfied by \( \varepsilon_w \) and \( \varepsilon_h \) are given by the following two tables.

### Table 2: Conditions for Wife's Reaction Functions

| \( W_1 \) | \( \varepsilon_w \) & (\( U_{w}^1 - U_{w}^0 \)) & min(\( 0, a_{w}^1 - a_{w}^0 \)) |
| --- | --- | --- | --- |
| \( W_2 \) | \( \varepsilon_w \) & (\( U_{w}^1 - U_{w}^0 \)) & max(\( 0, a_{w}^1 - a_{w}^0 \)) |
| \( W_3 \) | \( \varepsilon_w \) & (\( U_{w}^1 - U_{w}^0 \)) & (\( U_{w}^1 - U_{w}^0 \)) if \( a_{w}^1 - a_{w}^0 \leq 0 \) |
| \( W_4 \) | \( \varepsilon_w \) & (\( U_{w}^1 - U_{w}^0 \)) & (\( U_{w}^1 - U_{w}^0 \)) if \( a_{w}^1 - a_{w}^0 \leq 0 \); otherwise cannot occur |

### Table 3: Conditions for Husband's Utility Comparisons

| \( C_1 \) | \( \varepsilon_h \) & (\( U_{h}^1 - U_{h}^0 \)) & (\( U_{h}^1 - U_{h}^0 \)) |
| --- | --- | --- | --- |
| \( C_2 \) | \( \varepsilon_h \) & (\( U_{h}^1 - U_{h}^0 \)) & (\( U_{h}^1 - U_{h}^0 \)) |
| \( C_3 \) | \( \varepsilon_h \) & (\( U_{h}^1 - U_{h}^0 \)) & (\( U_{h}^1 - U_{h}^0 \)) |
| \( C_4 \) | \( \varepsilon_h \) & (\( U_{h}^1 - U_{h}^0 \)) & (\( U_{h}^1 - U_{h}^0 \)) |

Now that randomness has been introduced into the model, we can derive the joint probabilities on the part of both the husband and wife whether or not to work. Let \( \Pr(i,j) \) be the probability that the random variables \( Y_h \) and \( Y_w \) take on the values \( i \) and \( j \). To see \( i,j \in \{0,1\} \).

From Table 1, we have

\[
\Pr(0,0) = \Pr(W_2 \& C_2) + \Pr(W_4 \& C_4) \\
\Pr(1,0) = \Pr(W_2 \& C_2) + \Pr(W_4 \& C_4) \\
\Pr(0,1) = \Pr(W_1 \& C_1) + \Pr(W_4 \& C_4) \\
\Pr(1,1) = \Pr(W_1 \& C_1) + \Pr(W_2 \& C_2)
\]

Using Tables 2 and 3 and Equations (5)-(8) we can derive the probabilities in terms of the unknown parameters. Let \( F(a,b,p) \) be the c.d.f. evaluated at \((a,b)\) of a bivariate normal distribution with zero means, unit variances, and correlation \( p \). Moreover, let \( I(a,b,c,d,p) \) be the integral corresponding to a bivariate density over the range \( a \leq \varepsilon_h \leq c, b \leq \varepsilon_w \leq d \). As can be seen from Table 2, the probabilities \( \Pr(i,j) \) will depend on the sign of \( \Delta_{w} = (a_{w}^1 - a_{w}^0) \). We then have:

**PROPOSITION 1:**

\[
\Pr(0,0) = F(-A_{h},-A_{w},-p) - I_{+}^{B} \\
= F(-A_{h},-A_{w},-p) \\
\Pr(1,0) = F(A_{h},-A_{w} - \Delta_{w},-p) \\
= F(A_{h},-A_{w} - \Delta_{w},-p) + I_{-}^{B} \\
\Pr(0,1) = F(-A_{h} - \Delta_{w} - a_{h}^0 + a_{w}^1, A_{w},-p) \\
= F(-A_{h} - \Delta_{w} - a_{h}^0 + a_{w}^1, A_{w},-p) + I_{-}^{A} \\
\Pr(1,1) = F(-A_{h} - a_{h}^0 + a_{w}^1, A_{h},-p)
\]

if \( \Delta_{w} \geq 0 \)

otherwise
\[
\begin{align*}
\Pr(1, 1) &= F(\Delta U_h + a_h^1 - a_h^0, \Delta U_w + \Delta a_w, \rho) - I_+^A \text{ if } \Delta a_w \geq 0 \\
&= F(\Delta U_h + a_h^1 - a_h^0, \Delta U_w + \Delta a_w, \rho) \text{ otherwise }
\end{align*}
\]

where

\[
\begin{align*}
I^A &= I(-\Delta U_h - a_h^1, -\Delta U_w - \Delta a_h^1 - a_h^0, -\Delta U_w - \Delta a_w, \rho) \\
I^B &= I(-\Delta U_h - \Delta a_h^1 - a_h^0, -\Delta U_w - \Delta a_w, -\Delta U_h - a_h^1, -\Delta U_w - \Delta a_w, \rho) \\
I^A_+ &= I(-\Delta U_h - a_h^1, -\Delta U_w - \Delta a_h^1 - a_h^0, -\Delta U_w - \Delta a_w, \rho) \\
I^B_+ &= I(-\Delta U_h - \Delta a_h^1 - a_h^0, -\Delta U_w - \Delta a_w, -\Delta U_h - a_h^1, -\Delta U_w - \Delta a_w, \rho) \\
\Delta a_h &= u_h^1 - u_h^0 \text{ and } \Delta a_w = u_w^1 - u_w^0.
\end{align*}
\]

It is of interest to know the direction of change in the probabilities that the husband \(\Pr(1, \cdot)\) and the wife \(\Pr(\cdot, 1)\) will work as the parameters vary. We then have:

**PROPOSITION 2:**

1. An increase in \(a_h^1\) or \(\Delta U_h\) always increases the probability that the husband will work, \(\Pr(1, \cdot)\);
2. An increase in \(a_h^0\) always decreases the probability that the husband will work;
3. An increase in \(\Delta a_w\) or \(\Delta U_w\) always increases the probability that the wife will work, \(\Pr(\cdot, 1)\).

As expected, an increase in \(\Delta U_h\) increases the probability that the husband will work, whether or not the wife chooses to work; a similar remark holds for an increase in \(\Delta U_w\). Also, as can be seen from equation (1), an increase in \(a_h^1\) increases the probability that the husband will work when he knows his wife wishes to work, while having no effect on his propensity to work when he knows his wife chooses not to work. From equation (2), it is clear that an increase in \(a_h^0\) increases the husband’s utility of not working. Finally using equations (3) and (4), it is seen that an increase in \(\Delta U_w\) increases the wife’s utility of joining the labor market. (Included with the proof of Proposition 2 in the Appendix is a table indicating the direction on change in the probabilities \(\Pr(i, j)\) as all parameters are allowed to vary).

3. A COMPARISON OF MODELS

Now that we have developed a model in which the outcomes of the sequential decision-making problem are generated as Stackelberg equilibria of a game between two players, we are in a position to compare it to the usual recursive probability model for dichotomous variables (see e.g., Maddala and Lee (1976)). According to the usual formulation, a recursive equation system is described in terms of latent continuous variables, where the observed dichotomous variables are generated using a dichotomization. In our case, the corresponding recursive probability model is

\[
\begin{align*}
Y_w^* &= \Delta_w + \beta_w Y_h + \varepsilon_w \\
Y_h^* &= \Delta_h + \varepsilon_h
\end{align*}
\]

for some \(\Delta_h\) and \(\Delta_w\), and

\[
\begin{align*}
Y_h &= \begin{cases} 
1 & \text{if } Y_h^* > 0, \\
0 & \text{otherwise}
\end{cases} \\
Y_w &= \begin{cases} 
1 & \text{if } Y_w^* > 0, \\
0 & \text{otherwise}
\end{cases}
\end{align*}
\]
The purpose of this section is to show that this recursive probability model is nested in our model of Section 2.

Suppose that
\[ a_h^1 = a_h^0 = 0; \tag{16} \]
then from equations (1) and (2) defining the husband's utilities, we have:
\[
\begin{align*}
U_h(1,Y_w) &= U_h^1 + \eta_h^1 \\
U_h(0,Y_w) &= U_h^0 + \eta_h^0
\end{align*}
\]

Thus the restrictions (16) can be interpreted as imposing that the utilities derived by the husband from working or not working do not depend on the wife's decision whether or not to work.

But now note that if the restrictions (16) hold, then from Table 3, the four conditions \( C_1, C_2, C_3, \) and \( C_4 \) are identical; that is, \( \epsilon_h = -\Delta h \). Looking now at the conditions for the wife's reaction functions, we have to distinguish two cases according to the sign of \( \Delta a_w \). Suppose first that \( \Delta a_w < 0 \). Then it is readily seen from Tables 1 and 2 that the pairs (1,1), (1,0), (0,1), and (0,0) occur under the following conditions:

\[
\begin{align*}
(1,1) & \quad \text{if and only if } \Delta u_h + \epsilon_h \geq 0 \text{ and } \Delta u_w + \epsilon_w \geq 0, \\
(1,0) & \quad \text{if and only if } \Delta u_h + \epsilon_h \geq 0 \text{ and } \Delta u_w + \epsilon_w < 0, \\
(0,1) & \quad \text{if and only if } \Delta u_h + \epsilon_h < 0 \text{ and } \Delta u_w + \epsilon_w \geq 0, \\
(0,0) & \quad \text{if and only if } \Delta u_h + \epsilon_h < 0 \text{ and } \Delta u_w + \epsilon_w < 0.
\end{align*}
\]

It suffices now to note that these conditions are exactly identical to the ones that are obtained from the recursive probability model (14)-(15) with the usual dichotomization where \( \Lambda_h = \Delta u_h \) and \( \Lambda_w = \Delta u_w \). The case \( \Delta a_w < 0 \) is similarly studied, and gives the same conditions as above on the errors \( \epsilon_h \) and \( \epsilon_w \). We have therefore established the following proposition.

**Proposition 3:** If the restrictions \( a_h^1 = a_h^0 = 0 \) hold, then the usual recursive probability model using the dichotomization rule is identical to our model in which the observed outcomes are generated as Stackelberg equilibria.

The import of Proposition 3 is that it gives a structural interpretation to the usual recursive probability model in terms of a Stackelberg game. In addition, since the restrictions (16) on the parameters of our model must hold in order for the result in Proposition 3 to hold, it follows that the usual recursive probability model is nested in our proposed model. As an empirical consequence, it is then possible to test the specification of the usual recursive model by testing \( a_h^1 = a_h^0 = 0 \). Finally, given the above interpretation of these restrictions, it can be seen that these restrictions are unrealistic since they impose that the utilities of the husband (from working or not working) do not depend on whether the wife is working. Thus the usual recursive formulation is inappropriate since it implicitly assumes that the leader is indifferent to the follower's action. Let us also note that although the husband is moving first
and in principle should take into account his wife's conditional action when making his decision, the restrictions (16) when imposed lead the husband to ignore his wife's action.

4. IDENTIFICATION AND ESTIMATION

Given the previous expressions for the probabilities $\Pr(i,j)$ of the observed dichotomous variables $Y_h$ and $Y_w$, the log-likelihood function under random sampling is written as:

$$L = \sum_t \log \Pr_t(Y_{ht}, Y_{wt})$$

$$= \sum_t \left( Y_{ht} Y_{wt} \log \Pr_t(1,1) + (1 - Y_{ht}) \log \Pr_t(1,0) + (1 - Y_{ht}) Y_{wt} \log \Pr_t(0,1) + (1 - Y_{ht})(1 - Y_{wt}) \log \Pr_t(0,0) \right)$$

where the subscript $t$ indexes the observations. The probabilities are subscripted by $t$ since $\Delta U_h$ and $\Delta U_w$ are in general functions of explanatory variables. We assume:

$$\Delta U_{ht} = x_{ht}'Y_h \quad \text{and} \quad \Delta U_{wt} = x_{wt}'Y_w$$

where $x_{ht}$ may include characteristics of the $t$-th household and characteristics of the husband. A similar remark applies to $x_{wt}$. We now turn to the conditions under which the parameters $(p, \Delta a_w, c_h, a_{h}, b_{ht}, Y_{ht}, Y_{wt})$ of our model are identified.

In order to discuss identification, we first need to introduce some notation. Define the following partitioned matrix $\mathbf{X}$ as

$$\mathbf{X} = \begin{bmatrix} \mathbf{D}_p \mathbf{X}_p \mid \mathbf{D}_h \mathbf{X}_h \mid \mathbf{D}_w \mathbf{X}_w \end{bmatrix}$$

where $\mathbf{D}_p$, $\mathbf{D}_h$ and $\mathbf{D}_w$ are block diagonal matrices of order $3T$, the $t$-th blocks given as follows:

if $\Delta a_w > 0$

$$\begin{bmatrix} r_t^1 & 0 & 0 \\ 0 & r_t^2 & r_t^3 \\ 0 & 0 & r_t^4 \end{bmatrix}$$

if $\Delta a_w < 0$

$$\begin{bmatrix} r_t^{-1} & 0 & 0 \\ 0 & r_t^2 & 0 \\ 0 & r_t^3 & r_t^4 \end{bmatrix}$$

The elements of the above matrices are described in part (e) of the Appendix. The matrices $\mathbf{X}_h$ and $\mathbf{X}_w$ are of dimension $3T$ by $K_h + 2$ and $3T$ by $K_w + 1$, the $t$-th blocks given respectively as:

$$\begin{bmatrix} -1 & 0 & x_{ht}' \\ 0 & 1 & x_{ht}' \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 1 & x_{ht}' \\ 0 & 0 \end{bmatrix}$$

In addition, $\mathbf{X}_p$ is a unit vector of dimension $3T$.

PROPOSITION 4: The parameters $(p, \Delta a_w, c_h, a_{h}, b_{ht}, Y_{ht}, Y_{wt})$ of the model are identified if and only if $\mathbf{X}$ has full column rank.
As seen in part (e) of the Appendix, the elements of the matrices $D_P$, $D_h$, and $D_w$ are all nonzero. Moreover, these matrices are nonsingular in both cases since they are either triangular matrices or can be made triangular by suitable permutations of rows and columns. By examining matrix $A$ above, it is clear that if $A$ does not have full column rank, it will occur only extremely rarely for some specific values of the parameters as an artifact of certain explanatory variables. We have, although, the following necessary condition for identification.

**COROLLARY 1:** If $\Delta a_w = a_w^1 - a_w^0 = 0$, the model is not identified.

As a practical implication of the corollary for estimation, it must be the case that the initial values chosen for $a_w^1$ and $a_w^0$ not be the same. Otherwise, the information matrix will be nonsingular at the first iteration, and the optimization cannot be carried out. We now turn to estimation.

The estimation routine we employ is a version of the iterative procedure suggested by Berndt, Hall, Hall, and Hausman (1974), where we provide various initial values for $(\Delta a_w, a_h^1, a_h^0, Y_h, Y_w)$ with a grid search over possible values of $\rho$ and iterate until convergence.

5. AN EMPIRICAL EXAMPLE

A. THE MODEL

The following four equations will be used to describe the joint behavior of a representative married couple:

\[
\begin{align*}
W &= \omega_w^1 - \omega_w^0 + \omega_w^0 + \eta_w^0 \\
W &= \omega_h^1 - \omega_h^0 + \omega_h^0 + \eta_h^0 \\
W &= \omega_w^1 - \omega_w^0 + \omega_w^0 + \eta_w^0 \\
W &= \omega_h^1 - \omega_h^0 + \omega_h^0 + \eta_h^0
\end{align*}
\]

Equations (19) and (20) describe the reservation wages, or equivalently, the shadow price of time for the husband and wife, respectively. Note that the wife's decision whether or not to work, given by the dichotomous variable $Y_w$, affects the husband's reservation wage in (19). Analogously, the husband's decision whether or not to work, given by $Y_h$, affects the wife's reservation wage in (20). Equations (21) and (22) describe the market wages for the husband and the wife, respectively. Note that we allow the possibility that one of the determinants of the husband's market wage is whether or not he has a working wife; we make a similar allowance for the wife.

Moreover, let the husband's (wife's) reservation wage play the role of the payoff he (she) derives from not working. Therefore we have $W_h = U_h(0,Y_h)$ and $W_w = U_w(0,Y_w)$. Similarly, let the husband's (wife's) market wage play the role of the payoff he (she) derives from working. We thus have $W_h = U_h(1,Y_h)$ and $W_w = U_w(1,Y_w)$.

From equations (1) through (4) we see that $\Delta U_h = X_hY_h - Z_hY_h$ and $\Delta U_w = X_wY_w - Z_wY_w$. Moreover, note that in specifying the husband's reservation wage and market wage equations, given by (19) and (21) respectively, it may be the case that certain explanatory
variables appear in both equations, implying that the associated coefficient in $\Delta U_h$ will be measuring the difference between the market and reservation wage coefficients. A similar remark holds for the wife. In addition, note that the assumptions on error terms are also satisfied, namely $\epsilon_h = \frac{1}{\eta_h - \eta_h}$ and $\epsilon_w = \frac{1}{\eta_w - \eta_w}$.

We must now specify the set of explanatory variables for the market wage equations and the reservation wage equations for the husband and wife. Market wages for the husband and wife are specified in (23) and (24) respectively. Reservation wages for the husband and wife are specified in (25) and (26) respectively.

\[ U_h(Y_h) = \gamma_h + \gamma_{AGEH} Y_{AGEH} + \gamma_{EDUCH} Y_{EDUCH} + \gamma_{UNEM} Y_{UNEM} + \gamma_{RACE} Y_{RACE} + \alpha_{Yh} Y_h + \eta_h \]

\[ U_h(Y_w) = \gamma_h + \gamma_{AGEW} Y_{AGEW} + \gamma_{EDUCW} Y_{EDUCW} + \gamma_{RACE} Y_{RACE} + \alpha_{Yw} Y_w + \eta_w \]

\[ U_i(Y_h) = \gamma_i + \gamma_{AGEH} Y_{AGEH} + \gamma_{EDUCH} Y_{EDUCH} + \gamma_{UNEM} Y_{UNEM} + \gamma_{RACE} Y_{RACE} + \alpha_{Yh} Y_h + \eta_h \]

\[ U_i(Y_w) = \gamma_i + \gamma_{AGEW} Y_{AGEW} + \gamma_{EDUCW} Y_{EDUCW} + \gamma_{RACE} Y_{RACE} + \alpha_{Yw} Y_w + \eta_w \]

where

- $AGEH$ Age of husband
- $AGEW$ Age of wife
- $AGEW**2$ Squared age of wife
- $EDUCH$ Number of years of formal schooling of husband
- $EDUCW$ Number of years of formal schooling of wife
- $UNEM$ Local unemployment rate
- $RACE$ 1 = Black or Hispanic, 0 otherwise
- $ASSETS$ Family's annual income other than from wages or salaries
- $KIDS1-2$ Number of children in family unit ages 1 and 2.
- $KIDS3-5$ Number of children between ages 3 and 5.
- $KIDS6-13$ Number of children between 6 and 13.
- $KIDS13$ Number of children 13 years or younger
- $KIDS14$ Number of children 14 years or older

The plus and minus signs under the explanatory variables in Equations (23)–(26) indicate the expected impact of each variable in the respective equation. From equations (23) through (26), we have the following expressions for $\Delta U_h$ and $\Delta U_w$:

\[ \Delta U_h = (\gamma_h - \gamma_h) + (\gamma_{AGEH} - \gamma_{AGEH}) Y_{AGEH} + (\gamma_{EDUCH} - \gamma_{EDUCH}) Y_{EDUCH} + (\gamma_{UNEM} - \gamma_{UNEM}) Y_{UNEM} + (\gamma_{RACE} - \gamma_{RACE}) Y_{RACE} + (\alpha_{Yh} - \alpha_{Yh}) Y_h + (\eta_h - \eta_h) \]

\[ \Delta U_w = (\gamma_h - \gamma_w) + (\gamma_{AGEW} - \gamma_{AGEW}) Y_{AGEW} + (\gamma_{EDUCW} - \gamma_{EDUCW}) Y_{EDUCW} + (\gamma_{RACE} - \gamma_{RACE}) Y_{RACE} + (\alpha_{Yw} - \alpha_{Yw}) Y_w + (\eta_w - \eta_w) \]
The data we will use in this study on married couples is from the 1982 wave of the University of Michigan Survey Research Center's Panel Study on Income Dynamics, 1968-1982. The data is restricted to 2012 records for married couples living in the U.S., where both the husband and the wife were able-bodied, neither older than 64 years of age with no nonrelative living with the family (see Bjorn and Vuong (1984)).

B. EMPIRICAL RESULTS

From equations (19) and (21) it will be recalled that not only does the model allow for the possibility that one of the determinants of the husband's reservation wage is whether or not his wife chooses to work, the model also allows for the possibility that the husband's market wage is affected by his wife's decision. Although economic theory suggests that only the former effect should be meaningful, we can test that hypothesis in our model by allowing for the presence of both effects; that is, both \( a^0_w \) and \( a^1_w \) are included. The maximum likelihood estimates of the parameters of the full model are presented in Table 4.

From the t-statistic associated with \( a^1_w \), it follows that \( a^1_w = 0 \) cannot be rejected at any reasonable level of significance, as theory suggests. Although we see from Table 4 that most of the explanatory variables, especially for the wife, have the a priori correct sign and are highly significant, we therefore reestimate the model without \( a^1_w \). These latter results are presented in Table 5.

The value of \( \rho \) that maximizes the log-likelihood function is \(-0.45\). Although it may at first appear surprising that this maximizing value of \( \rho \) is not positive, it must be remembered that \( \rho \) is not simply the correlation between omitted variables in the husband's and wife's equations, but arises from a more complicated relationship between the disturbance terms \( \varepsilon_h \) and \( \varepsilon_w \), viz., \( \varepsilon_h = \eta^1_h - \eta^0_h \) and \( \varepsilon_w = \eta^1_w - \eta^0_w \) as seen in Section 2. From the table we see that both \( A^w_w \) and \( a^0_w \) are significantly different from zero, providing evidence that the wife's decision whether or not to work depends on the husband's decision and vice versa. Although it will be recalled from Section 4 that only the difference \( A^w_w = a^1_w - a^0_w \) can be identified in our model, economic theory again suggests that \( a^0_w \) should be a priori zero. Therefore the estimate \(-1.12\) of \( A^w_w \) is actually an estimate of \(-a^0_w \). With this in mind then, we see from equation (20) that if the husband works, the wife's reservation wage increases as expected since \( a^0_w \) is positive.

It should also be noticed from Table 5 that we can provide a test of Proposition 3. Since \( a^1_h \) is restricted to be a priori zero and \( a^0_h \) is significantly different from zero at the 5 percent level, we can reject the hypothesis that the data are generated by the usual recursive probability model using a dichotomization rule. In other words, we must accept the hypothesis that the husband takes his wife's conditional action into account when making his decision whether or not to work.
# Table 4

$p = .40$

<table>
<thead>
<tr>
<th>Variable</th>
<th>Husband</th>
<th>Wife</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Coefficient</td>
<td>Estimate</td>
</tr>
<tr>
<td>CONSTANT</td>
<td>$\alpha^0_h$</td>
<td>-1.98</td>
</tr>
<tr>
<td></td>
<td>$\alpha^1_h$</td>
<td>-0.256</td>
</tr>
<tr>
<td>$\gamma^0_h - \gamma^0_w$</td>
<td>-0.076</td>
<td>-0.60</td>
</tr>
<tr>
<td>$\gamma^1_h - \gamma^1_w$</td>
<td>0.014</td>
<td>1.75*</td>
</tr>
<tr>
<td>$\gamma^2_h - \gamma^2_w$</td>
<td>0.071</td>
<td>1.81*</td>
</tr>
<tr>
<td>EDUCW</td>
<td>$\gamma^3_h - \gamma^3_w$</td>
<td>0.410</td>
</tr>
<tr>
<td>UNEM</td>
<td>$\gamma^4_h - \gamma^4_w$</td>
<td>-0.040</td>
</tr>
<tr>
<td>RACE</td>
<td>$\gamma^5_h - \gamma^5_w$</td>
<td>-0.330</td>
</tr>
<tr>
<td>ASSET</td>
<td>$\gamma^6_h - \gamma^6_w$</td>
<td>0.410</td>
</tr>
<tr>
<td>KIDS1-2</td>
<td>$\gamma^7_w$</td>
<td>-0.685</td>
</tr>
<tr>
<td>KIDS3-5</td>
<td>$\gamma^8_w$</td>
<td>-0.444</td>
</tr>
<tr>
<td>KIDS6-13</td>
<td>$\gamma^9_w$</td>
<td>-0.212</td>
</tr>
<tr>
<td>KIDS &lt; 13</td>
<td>$\gamma^{10}_h$</td>
<td>0.021</td>
</tr>
<tr>
<td>KIDS &gt; 14</td>
<td>$\gamma^{10}_w$</td>
<td>0.104</td>
</tr>
</tbody>
</table>

log-likelihood value = -1514.93

* significant at the 10% level
** significant at the 5% level
**TABLE 5**

\( \rho = 0.45 \)

<table>
<thead>
<tr>
<th>Coefficient</th>
<th>Husband</th>
<th>Estimate</th>
<th>( t )-Statistic</th>
<th>Wife</th>
<th>Estimate</th>
<th>( t )-Statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>( a^0_h )</td>
<td>-1.82</td>
<td>-2.16**</td>
<td></td>
<td>( a^0_w )</td>
<td>-1.12</td>
<td>-1.92*</td>
</tr>
<tr>
<td>CONSTANT</td>
<td>( \gamma^0_h - \gamma^0_w )</td>
<td>-0.784</td>
<td>-0.63</td>
<td>( \gamma^0_w - \gamma^0_w )</td>
<td>0.31</td>
<td>0.40</td>
</tr>
<tr>
<td>AGEH</td>
<td>( \gamma^1_h - \gamma^1_w )</td>
<td>0.013</td>
<td>1.68*</td>
<td>( \gamma^1_w - \gamma^1_w )</td>
<td>0.083</td>
<td>3.47**</td>
</tr>
<tr>
<td>AGEW**2</td>
<td>( \gamma^2_h - \gamma^2_w )</td>
<td>0.069</td>
<td>1.79*</td>
<td>( \gamma^2_w - \gamma^2_w )</td>
<td>-0.130</td>
<td>-4.31**</td>
</tr>
<tr>
<td>EDUCW</td>
<td>( \gamma^3_h - \gamma^3_w )</td>
<td>-0.038</td>
<td>-1.97**</td>
<td>( \gamma^3_w - \gamma^3_w )</td>
<td>-0.014</td>
<td>-1.50</td>
</tr>
<tr>
<td>UNEM</td>
<td>( \gamma^3_h - \gamma^3_w )</td>
<td>-0.038</td>
<td>-1.97**</td>
<td>( \gamma^3_w - \gamma^3_w )</td>
<td>-0.014</td>
<td>-1.50</td>
</tr>
<tr>
<td>RACE</td>
<td>( \gamma^4_h - \gamma^4_w )</td>
<td>-0.335</td>
<td>-2.64**</td>
<td>( \gamma^4_w - \gamma^4_w )</td>
<td>0.442</td>
<td>5.69**</td>
</tr>
<tr>
<td>ASSET</td>
<td>( \gamma^5_h )</td>
<td>0.400</td>
<td>1.36</td>
<td>( \gamma^5_w )</td>
<td>-0.012</td>
<td>-2.13**</td>
</tr>
<tr>
<td>KIDS1-2</td>
<td>( \gamma^7_h )</td>
<td>0.400</td>
<td>1.36</td>
<td>( \gamma^7_w )</td>
<td>-0.012</td>
<td>-2.13**</td>
</tr>
<tr>
<td>KIDS3-5</td>
<td>( \gamma^8_h )</td>
<td>0.400</td>
<td>1.36</td>
<td>( \gamma^8_w )</td>
<td>-0.012</td>
<td>-2.13**</td>
</tr>
<tr>
<td>KIDS6-13</td>
<td>( \gamma^9_h )</td>
<td>0.400</td>
<td>1.36</td>
<td>( \gamma^9_w )</td>
<td>-0.012</td>
<td>-2.13**</td>
</tr>
<tr>
<td>KIDS &lt; 13</td>
<td>( \gamma^{10}_h )</td>
<td>0.034</td>
<td>0.63</td>
<td>( \gamma^{10}_w )</td>
<td>0.109</td>
<td>0.99</td>
</tr>
<tr>
<td>KIDS &gt; 14</td>
<td>( \gamma^{11}_h )</td>
<td>0.034</td>
<td>0.63</td>
<td>( \gamma^{11}_w )</td>
<td>0.109</td>
<td>0.99</td>
</tr>
</tbody>
</table>

log-likelihood value = -1514.99

* significant at the 10% level
** significant at the 5% level
A priori, we would expect the estimate of \( \alpha^0 \) to be positive; that is, we expect that the wife's decision to work should lower the husband's reservation wage. In contrast, we find that the estimate of \( \alpha^0 \) is negative and significant at the 5 percent level. One possible explanation for this result is that no husband wishes to bear the embarrassment of staying at home when his wife chooses to work; that is, the husband chooses to lower his reservation wage when his wife is working.

Looking again at Table 5, we see that most of the coefficients explaining the wife's decision whether or not to work are in agreement with our expectations and are highly significant. (In reading Table 5, it should be noted that all estimated coefficients represent either differences between market and reservation wages or minus the reservation wage coefficient, as seen in Equations (27) and (28).)

For example, family income from sources other than wages and salaries (ASSET) has the expected effect of increasing the wage at which the wife is willing to accept work outside the home \( (\beta^6 = +0.012) \).

Concerning children, one would certainly expect that mothers would be least likely to leave the home when children are very young and be more inclined to seek outside employment as children become older and more self-sufficient. That is, younger children should have the effect of increasing the mother's reservation wage more than do older children. Indeed, this is what we see from Table 5. Children between the ages of one and two (KIDS1-2) raise the mother's reservation wage more than do children between three and five (KIDS3-5); her reservation wage is higher for children between three and five than for children six to thirteen (KIDS6-13); finally, the mother is more likely to stay at home when her children are between six and thirteen than when they are fourteen years or older (KIDS14). The coefficient on the wife's education (EDUCW) is also consistent with our priori expectation; although an increase in education should increase the wife's market wage, it should also increase her reservation wage. The estimated positive coefficient on the female race dummy (RACE) seems to suggest that women of racial minorities, on average, can command a higher market wage than the wage necessary to entice them into the labor market; that is, minority women are on average worth more in the marketplace than they think they are worth. Turning finally to the effect of age on a wife's decision whether or not to work, a life-cycle model of employment would suggest that women are more likely to work during middle age than either early or late in their life times. That is, the probability that our individual will work exhibits a concave shape. As can be seen from Table 5, the combined effects of a positive linear term on age (AGEW) and a negative quadratic term (AGE**2) does indeed impart an increasing then a decreasing shape with respect to age with a peak at about 32 years of age.

Turning next to the variables used to explain the husband's decision whether or not to work, we see that while some of the coefficients are insignificant, many of the variables to which we attached strong priors are indeed significant. For example, the coefficients attached to the husband's age (AGEH), his education
(EDUCH) and the local unemployment rate (UNEM) are each significant. Since each of these three coefficients measure the difference between the husband’s market wage and his reservation wage, it is not surprising that they all should be close to zero if the husband is behaving rationally; for example, the effect of an increase in education should not only increase an individual’s market wage but should also increase his reservation wage. Finally, we see that the effects of racial discrimination on minorities has the effect of lowering their market wages relative to those of nonminorities.

6. CONCLUSION

In this paper, we presented an alternative approach for formulating simultaneous equations models for qualitative endogenous variables which integrates results in game theory and discrete choice modeling. In this game theoretic formulation, we assume the two individuals play a Stackelberg game in which each player maximizes his own utility; the model was made stochastic by adopting the random utility framework.

A distinctive feature of our model is that it generalizes the recursive models for discrete endogenous variables that have been proposed up to now in the literature; that is, the usual recursive model is nested in our game theoretic model. Although recursive models have been used in the formulation of many econometric problems in which sequential decision making is a distinct feature, these models implicitly assume that the leader is indifferent from the follower’s action. If this is not the case, then the usual recursive models are misspecified since they ignore the optimizing behavior of the leader who is taking into account the conditional action of the second agent when choosing his action. As such, the usual recursive model of a sequential decision making problem is inadequate in many problems. In contrast, our formulation in terms of a Stackelberg model allows for optimizing behavior on the part of both agents.

As an empirical application, we studied the joint decision of a husband and wife whether or not to participate in the labor market. Here it was assumed that the husband was the Stackelberg leader and his wife was the follower where both were fully optimizing; that is, we assumed that the husband knew what action his wife would take and he thus optimized accordingly. Since the usual recursive model is nested in our model, we were able to reject the recursive specification for the problem we studied; that is, we were able to reject the hypothesis that the husband did not take his wife’s conditional action into account when making his decision whether or not to work. In addition, most of the coefficients for which we held strong priors had the correct signs and significant t-statistics.
APPENDIX

a. Conditions for Wife's Reaction Functions

Using Figure 2, reaction function \( \bar{w}_1 \) is characterized by the following two conditions: \( \bar{U}_h(1,0) \leq \bar{U}_h(0,0) \) and \( \bar{U}_h(1,1) \leq \bar{U}_h(0,1) \). Using (3) and (4) from the text, these conditions are equivalent to \( \epsilon_w \geq -(U^1_w - U^0_w + a^1_w - a^0_w) \) and \( \epsilon_w \geq -(U^1_w - U^0_w - \min(0, a^1_w - a^0_w)) \), respectively, which can be combined to give \( \epsilon_w \geq -(U^1_w - U^0_w - \min(0, a^1_w - a^0_w)) \).

Reaction function \( \bar{w}_2 \) is characterized by \( \bar{U}_h(1,0) \leq \bar{U}_h(0,0) \) and \( \bar{U}_h(1,1) \geq \bar{U}_h(0,1) \), which are equivalent to \( \epsilon_w \leq -(U^1_w - U^0_w) \) and \( \epsilon_w > -(U^1_w - U^0_w + a^1_w - a^0_w) \), respectively. When combined, they give the result in the text.

Reaction function \( \bar{w}_3 \) is characterized by \( \bar{U}_h(1,0) < \bar{U}_h(0,0) \) and \( \bar{U}_h(1,1) \leq \bar{U}_h(0,1) \). Using (3) and (4) from the text, these conditions are equivalent to \( \epsilon_w \geq -(U^1_w - U^0_w) \) and \( \epsilon_w \leq -(U^1_w - U^0_w + a^1_w - a^0_w) \), respectively. When combined, we get the result in the text.

Reaction function \( \bar{w}_4 \) is characterized by \( \bar{U}_h(1,0) \leq \bar{U}_h(0,0) \) and \( \bar{U}_h(1,1) < \bar{U}_h(0,1) \). Using (3) and (4) from the text, these conditions are equivalent to \( \epsilon_w \leq -(U^1_w - U^0_w) \) and \( \epsilon_w < -(U^1_w - U^0_w + a^1_w - a^0_w) \), respectively, which when combined give \( -(U^1_w - U^0_w) < \epsilon_w < -(U^1_w - U^0_w + a^1_w - a^0_w) \) if \( a^1_w - a^0_w < 0 \); otherwise \( \bar{w}_4 \) cannot occur.

b. Conditions for Husband's Utility Comparisons

Using Figure 1 in the text, when the wife follows reaction function \( \bar{w}_1 \), the husband compares \( \bar{U}_h(1,1) \) and \( \bar{U}_h(0,1) \). If \( \bar{U}_h(1,1) \geq \bar{U}_h(0,1) \), then from (1) and (2) we have \( \epsilon_h \geq -(U^1_h - U^0_h + a^1_h - a^0_h) \).

When the wife follows reaction function \( \bar{w}_2 \), the husband compares \( \bar{U}_h(1,1) \) and \( \bar{U}_h(0,0) \). When \( \bar{U}_h(1,1) \geq \bar{U}_h(0,0) \), we have \( \epsilon_h \geq -(U^1_h - U^0_h) \).

When reaction function \( \bar{w}_3 \) is used, the husband compares \( \bar{U}_h(1,1) \) and \( \bar{U}_h(0,1) \). When \( \bar{U}_h(1,1) \geq \bar{U}_h(0,1) \), we have \( \epsilon_h \geq - (U^1_h - U^0_h) \).

Finally, Figure 1 shows that when the wife uses \( \bar{w}_4 \), the husband makes a comparison between \( \bar{U}_h(1,0) \) and \( \bar{U}_h(0,1) \). If \( \bar{U}_h(1,0) \geq \bar{U}_h(0,1) \), we have from (1) and (2) that \( \epsilon_h \geq -(U^1_h - U^0_h - a^1_h) \).

c. PROOF OF PROPOSITION 1:

From Table 2, it is clear that reaction function \( \bar{w}_4 \) for the wife cannot occur when \( (a^1_w - a^0_w) \geq 0 \), while reaction function \( \bar{w}_2 \) cannot occur when \( (a^1_w - a^0_w) < 0 \). Thus when \( (a^1_w - a^0_w) \geq 0 \) it follows from equations (5) - (8) that

\[
Pr(0,0) = Pr(\bar{w}_4 \notin C_2) + Pr(\bar{w}_3 \notin C_3),
Pr(1,0) = Pr(\bar{w}_3 \notin C_3),
Pr(0,1) = Pr(W_1 \notin C_1),
Pr(1,1) = Pr(W_1 \notin C_1) + Pr(W_2 \notin C_2).
\]

Similarly, when \( (a^1_w - a^0_w) < 0 \), we have

\[
Pr(0,0) = Pr(\bar{w}_3 \notin C_3),
Pr(1,0) = Pr(\bar{w}_3 \notin C_3) + Pr(\bar{w}_4 \notin C_4),
Pr(0,1) = Pr(W_1 \notin C_1) + Pr(W_4 \notin C_4),
Pr(1,1) = Pr(W_1 \notin C_1).
\]
Now, using the conditions on $\varepsilon_w$ and $\varepsilon_h$ given in Tables 2 and 3, respectively, we can derive the needed comparisons between particular $W_i$, $C_i$, and $\tilde{C}_i$, $i = 1, \ldots, 4$. For the cases $\Delta a_w = (a_1^w - a_0^w) > 0$ and $\Delta a_w = (a_1^w - a_0^w) < 0$, figures 2a and 2b respectively show the areas over the bivariate normal density for $(\varepsilon_h, \varepsilon_w)$ which must be integrated to obtain the four probabilities $Pr(0,0)$, $Pr(1,0)$, $Pr(0,1)$, and $Pr(1,1)$. Without loss of generality, figures 2a and 2b are drawn for the case $a_0^w < a_1^w$. It can be seen from figures 2a and 2b that $I^{A}_a, I^{B}_a, I^{A}_w$ and $I^{B}_w$ correspond to the areas over the bivariate normal density given by (13) in the text. It follows that the probabilities $Pr(0,0)$, $Pr(1,0)$, $Pr(0,1)$, and $Pr(1,1)$ are given by equations (9) - (13) in Proposition 1.

d. First partial derivatives of the Probabilities $Pr(i,j)$: Let $\Phi$ be the univariate normal c.d.f. and let $\varphi$ be the corresponding p.d.f. We then use the relations $\frac{\partial F(x,y,p)}{\partial x} = \varphi(x)\Phi(y - px^*)$, $\frac{\partial F(x,y,p)}{\partial y} = \Phi(x)\varphi(y - py^*)$, and $\frac{\partial F(x,y,p)}{\partial p} = f(x,y,p)$ where a quantity with a "*" means that quantity is divided by the square root of $(1 - p^2)$. In addition, let $f(x,y,p)$ be the p.d.f. corresponding to the bivariate normal c.d.f. $F(x,y,p)$. Then from equations (9) - (13), the first partial derivatives of the probabilities $Pr(i,j)$ use the following:

$$\frac{\partial F(-A_h, -A_w, p)}{\partial a_h} = 0,$$
$$\frac{\partial F(-A_h, -A_w, p)}{\partial a_w} = 0,$$
$$\frac{\partial F(-A_h, -A_w, p)}{\partial p} = f(-A_h, -A_w, p);$$
\[\frac{\partial F}{\partial h} = \Phi(\Delta U_h + a_h)^{*} - \rho(\Delta U_h + a_h) - A_h^* \],
\[\frac{\partial F}{\partial w} = \Phi(\Delta U_w + a_h)^{*} - \rho(\Delta U_w + a_h) - A_w^* \],
\[\frac{\partial F}{\partial a} = \Phi(\Delta U_a + a_h)^{*} - \rho(\Delta U_a + a_h) - A_a^* \],
\[\frac{\partial F}{\partial p} = \Phi(\Delta U_p + a_h)^{*} - \rho(\Delta U_p + a_h) - A_p^* \].

And

\[\frac{\partial A}{\partial h} = \Phi(\Delta U_h + a_h)^{*} - \rho(\Delta U_h + a_h) - A_h^* \],
\[\frac{\partial A}{\partial w} = \Phi(\Delta U_w + a_h)^{*} - \rho(\Delta U_w + a_h) - A_w^* \],
\[\frac{\partial A}{\partial a} = \Phi(\Delta U_a + a_h)^{*} - \rho(\Delta U_a + a_h) - A_a^* \],
\[\frac{\partial A}{\partial p} = \Phi(\Delta U_p + a_h)^{*} - \rho(\Delta U_p + a_h) - A_p^* \].
e. Elements of the matrices $D_{pt}$, $D_{ht}$, and $D_{wt}$

For simplicity, we drop the subscript $t$ in the following expressions.

\[
\begin{align*}
1 + r_s f(A_{Uh}, A_{Uw} + A_{aw}, p) \\
2 + r_s (A_{Uh} + A_{aw}, p) \\
3 + r_s f(A_{Uh} + A_{aw}, p) \\
4 + r_s (A_{uh} + A_{ah}, p) \\
a_h = t(A_{uh} + A_{aw} + pA_{uh}^*) \\
b_h = -t(A_{uh} + A_{aw} + pA_{uh}^*) \\
c_h = t(A_{uh} + A_{aw} + pA_{uh}^*) \\
d_h = -t(A_{uh} + A_{aw} + pA_{uh}^*) \\
a_w = t(A_{uh} + A_{aw} + pA_{uh}^*) \\
b_w = -t(A_{uh} + A_{aw} + pA_{uh}^*) \\
c_w = t(A_{uh} + A_{aw} + pA_{uh}^*) \\
d_w = -t(A_{uh} + A_{aw} + pA_{uh}^*) \\
1 - t(A_{uh}, A_{Uw} + A_{aw}, p) \\
2 - t(A_{uh} + A_{aw}, p) \\
3 - t(A_{uh} + A_{aw}, p) \\
4 - t(A_{uh} + A_{aw}, p) \\
a_h = t(A_{uh} - A_{aw} + pA_{uh}^*) \\
b_h = -t(A_{uh} - A_{aw} + pA_{uh}^*) \\
c_h = t(A_{uh} - A_{aw} + pA_{uh}^*) \\
d_h = -t(A_{uh} - A_{aw} + pA_{uh}^*) \\
a_w = t(A_{uh} + A_{aw} + pA_{uh}^*) \\
b_w = -t(A_{uh} + A_{aw} + pA_{uh}^*) \\
c_w = t(A_{uh} + A_{aw} + pA_{uh}^*) \\
d_w = -t(A_{uh} + A_{aw} + pA_{uh}^*)
\end{align*}
\]

\[c_{w} = -t(A_{Uw} - A_{aw} + pA_{uh}^*) \]
\[d_{w} = t(A_{Uw} - A_{aw} + pA_{uh}^*)\]

f. PROOF OF PROPOSITION 2

Easily established by using either the areas defining the probabilities $Pr(i,j)$, as found in part (c) of this Appendix, or differentiating the probabilities found in Proposition 1.

<table>
<thead>
<tr>
<th>Case 1: $\Delta a_w = (a_{w} - a_{o}) &gt; 0$</th>
<th>Pr(0,0)</th>
<th>Pr(1,0)</th>
<th>Pr(0,1)</th>
<th>Pr(1,1)</th>
<th>Pr(1,0)</th>
<th>Pr(0,1)</th>
<th>Pr(0,1)</th>
<th>Pr(0,1)</th>
<th>Pr(0,1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$a_h$</td>
<td>-</td>
<td>no</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>no</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$a_o$</td>
<td>no</td>
<td>change</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>no</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\Delta a_w$</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$U_h$</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$U_w$</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Case 2: $\Delta a_w = (a_{w} - a_{o}) &lt; 0$</th>
<th>Pr(0,0)</th>
<th>Pr(1,0)</th>
<th>Pr(0,1)</th>
<th>Pr(1,1)</th>
<th>Pr(1,0)</th>
<th>Pr(0,1)</th>
<th>Pr(0,1)</th>
<th>Pr(0,1)</th>
<th>Pr(0,1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$a_h$</td>
<td>no</td>
<td>change</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>no</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$a_o$</td>
<td>no</td>
<td>change</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\Delta a_w$</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$U_h$</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$U_w$</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
g. PROOF OF PROPOSITION 4:

Let $z_t = (Y_{ht}, Y_{wt}, X_{ht}, X_{wt})$ and $\theta = (p, A_{aw}, a_h, a_w, Y_h, Y_w)$. Define:

$$B = E \left[ \sum_{t=1}^{T} \frac{\partial \log f(z_t, \theta)}{\partial \theta} \cdot \frac{\partial \log f(z_t, \theta)}{\partial \theta'} \right] = E \left[ \sum_{t=1}^{T} B_t \right].$$

From Section 4, we have, omitting the subscript $t$, that

$$\frac{\partial \log f(Z, \theta)}{\partial \theta} = \frac{Y_w Y_h}{\text{Pr}(1,1)} \frac{\partial \text{Pr}(1,1)}{\partial \theta} + \frac{Y_h(1 - Y_w)}{\text{Pr}(1,0)} \frac{\partial \text{Pr}(1,0)}{\partial \theta} + \frac{(1 - Y_h) Y_w}{\text{Pr}(0,1)} \frac{\partial \text{Pr}(0,1)}{\partial \theta} + \frac{(1 - Y_h)(1 - Y_w)}{\text{Pr}(0,0)} \frac{\partial \text{Pr}(0,0)}{\partial \theta}$$

Then, \( \frac{\partial \log f}{\partial \theta} \) is given by

$$\begin{bmatrix}
\frac{Y_w Y_h}{\text{Pr}(1,1)} \frac{\partial \text{Pr}(1,1)}{\partial \theta} & \frac{Y_h(1 - Y_w)}{\text{Pr}(1,0)} \frac{\partial \text{Pr}(1,0)}{\partial \theta} \\
\frac{(1 - Y_h) Y_w}{\text{Pr}(0,1)} \frac{\partial \text{Pr}(0,1)}{\partial \theta} & \frac{(1 - Y_h)(1 - Y_w)}{\text{Pr}(0,0)} \frac{\partial \text{Pr}(0,0)}{\partial \theta}
\end{bmatrix}^2$$

where we have used the fact that $Y_h$ and $Y_w$ take on only the values zero or one. Since $Y_h$ and $Y_w$ are random variables where $Y_h = 1, Y_w = j$ with probability $\text{Pr}(i,j), i,j \in \{0,1\}$, we have that

$$\frac{E \left[ \frac{\partial \log f}{\partial \theta} \cdot \frac{\partial \log f}{\partial \theta'} \right]}{\text{Pr}(1,1)} = \frac{1}{\text{Pr}(1,1)} \left[ \frac{\partial \text{Pr}(1,1)}{\partial \theta} \right]^2 + \frac{1}{\text{Pr}(1,0)} \left[ \frac{\partial \text{Pr}(1,0)}{\partial \theta} \right]^2 + \frac{1}{\text{Pr}(0,1)} \left[ \frac{\partial \text{Pr}(0,1)}{\partial \theta} \right]^2 + \frac{1}{\text{Pr}(0,0)} \left[ \frac{\partial \text{Pr}(0,0)}{\partial \theta} \right]^2$$

Proceeding analogously, the remaining terms in $B$ are given by:

$$F \left[ \frac{\partial \log f}{\partial \theta} \cdot \frac{\partial \log f}{\partial \theta'} \right] = \frac{1}{\text{Pr}(1,1)} \left[ \frac{\partial \text{Pr}(1,1)}{\partial \theta} \right]^2 + \frac{1}{\text{Pr}(1,0)} \left[ \frac{\partial \text{Pr}(1,0)}{\partial \theta} \right]^2 + \frac{1}{\text{Pr}(0,1)} \left[ \frac{\partial \text{Pr}(0,1)}{\partial \theta} \right]^2 + \frac{1}{\text{Pr}(0,0)} \left[ \frac{\partial \text{Pr}(0,0)}{\partial \theta} \right]^2$$

Notice that $B$ can be decomposed into $B = A'DA$ where $A$ is of dimension $4T \times K$, $K = K_h + K_w + 4$, that has as its $t$-th block $A_t$ defined as:

$$\begin{bmatrix}
\frac{\partial \text{Pr}(1,1)}{\partial \theta} & \frac{\partial \text{Pr}(1,1)}{\partial \theta} & \frac{\partial \text{Pr}(1,1)}{\partial \theta} & \frac{\partial \text{Pr}(1,1)}{\partial \theta} \\
\frac{\partial \text{Pr}(1,0)}{\partial \theta} & \frac{\partial \text{Pr}(1,0)}{\partial \theta} & \frac{\partial \text{Pr}(1,0)}{\partial \theta} & \frac{\partial \text{Pr}(1,0)}{\partial \theta} \\
\frac{\partial \text{Pr}(0,1)}{\partial \theta} & \frac{\partial \text{Pr}(0,1)}{\partial \theta} & \frac{\partial \text{Pr}(0,1)}{\partial \theta} & \frac{\partial \text{Pr}(0,1)}{\partial \theta} \\
\frac{\partial \text{Pr}(0,0)}{\partial \theta} & \frac{\partial \text{Pr}(0,0)}{\partial \theta} & \frac{\partial \text{Pr}(0,0)}{\partial \theta} & \frac{\partial \text{Pr}(0,0)}{\partial \theta}
\end{bmatrix}$$

and $D$ is a block diagonal matrix of order $4T$, the $t$-th block given by

$$\begin{bmatrix}
\text{Pr}(1,1) & 0 & 0 & 0 \\
0 & \text{Pr}(1,0) & 0 & 0 \\
0 & 0 & \text{Pr}(0,1) & 0 \\
0 & 0 & 0 & \text{Pr}(0,0)
\end{bmatrix}^{-1}$$

The model will be identified if and only if $B$ is nonsingular (see, e.g., Rothenberg (1971)). Since $D$ is of full rank and $4T > K$, a necessary and sufficient condition is that $A$ have full column rank.

From part (d) of the Appendix, it is seen that the partial derivatives
of $Pr_t(i,j)$ with respect to the vector $\theta$ depend on the sign of $\Delta a_w$; we must therefore check that matrix $A$ is nonsingular for both cases.

Case 1: $\Delta a_w > 0$

Substituting into $A_t$ the partial derivatives, using the notation $a_t^{i+}, b_t^{i+}, c_t^{i+}, d_t^{i+}, i = h,w,$ and $r_t^{j+}, j = 1,2,3,4,$ found in the Appendix, we perform the following matrix algebra

(i) add rows $(2+3+4)$ to row 1
(ii) add row 2 to row 4
(iii) add column 3 to column 4
(iv) multiply columns 1, 2 and 6 by $-1$
(v) Switch rows 3 and 4

Rearranging columns and omitting row 1 since it is identically null, we have

$$\bar{A}_t = \begin{bmatrix}
  r_t^{1+} & 0 & b_t^{h} x_h & a_t^{w} x_w & a_t^{w+} x_w' \\
  (r_t^{2+} + r_t^{3+}) & 0 & (b_t^{h} + c_t^{h+}) (c_t^{h+} + d_t^{h+}) x_h & c_t^{w} x_w & (b_t^{w} + c_t^{w+}) x_w' \\
  r_t^{4+} & d_t^{h} & 0 & -d_t^{h+} x_h' & 0 & d_t^{w} x_w \\
\end{bmatrix}$$

We now decompose the resulting matrix $\bar{A}$ into a partitioned matrix

$$\bar{A} = \begin{bmatrix}
  D_p^{1+} & | & D_h^{1+} & | & D_w^{1+} \\
  D_p^{2+} & | & D_h^{2+} & | & D_w^{2+} \\
  D_p^{3+} & | & D_h^{3+} & | & D_w^{3+} \\
\end{bmatrix}$$

where $D_p$, $D_h$, and $D_w$ are each block diagonal matrices of order $3T$, the $t$-th blocks being $D_p^{ht}$, $D_h^{ht}$, and $D_w^{ht}$ respectively, as given in the text.

Case 2: $\Delta a_w < 0$

Substitute into $A_t$ the partial derivatives found in the Appendix, again using $a_t^{i-}, b_t^{i-}, c_t^{i-}, d_t^{i-}, i = h,w,$ and $r_t^{j-}, j = 1,2,3,4.$ Now perform the following matrix algebra on matrix $A$

(i) add rows $(1+2+4)$ to row 3
(ii) add row 4 to row 2
(iii) add column 4 to column 3
(iv) multiply column 6 by $-1$
(v) switch rows 2 and 4

Rearranging columns and omitting row 3 since it identically null, we have

$$\tilde{A}_t = \begin{bmatrix}
  r_t^{1-} & 0 & 0 & 0 & b_t^{h} x_h & b_t^{w} x_w \\
  r_t^{2-} & 0 & b_t^{h-} & b_t^{h-1} & b_t^{w-} & b_t^{w-1} \\
  (r_t^{3-} + r_t^{4-}) & c_t^{h-} & 0 & -c_t^{h-} + d_t^{h-} x_h & c_t^{w-} & (c_t^{w-} + d_t^{w-}) x_w \\
\end{bmatrix}$$

which can be written as $\tilde{A}$.

h. PROOF OF COROLLARY 1:

When $\Delta a_w = 0$, it is seen from Section 4 of the text that $b_t^{h+} + c_t^{h+} = 0$. Therefore matrix $D_{ht}$ is singular for all $t$ which implies that matrix $\bar{A}$ no longer has full column rank.

Q.E.D.
Footnotes

1. Where an individual is indifferent, we arbitrarily assume that he or she will take action 1.

2. Let us note that the husband is fully informed about the utility function of the wife; that is, he not only knows the deterministic components in the utilities (3)-(4) given below, but also the random components. An interesting generalization, which will be pursued in future work, arises when the husband knows only the deterministic components, in which case one has a Stackelberg game under uncertainty (see also Vuong (1982)).

3. Let us note that we allow the utilities $\bar{U}_h(1,Y_w)$ and $\bar{U}_w(1,Y_h)$ to depend on $Y_w$ and $Y_h$ respectively. This contrasts with the formulation adopted in Bjorn and Vuong (1984, Equation (21)-(22)).

4. If $a < c$, $I(a,b,c,d,p)$ is by convention the negative of the integral of the bivariate density over the range $[a,c] \times [d,b]$. A similar remark applies if $b < d$. If both $a < c$ and $b < d$, then $I(a,b,c,d,p)$ is by convention the integral of the bivariate density over $[a,c] \times [b,d]$.

5. We use a common set of explanatory variables (see, e.g., Ashenfelter and Heckman (1974), Gronau (1973), and Heckman (1974)). The same specification was also used in Bjorn and Vuong (1984).
REFERENCES


