1 Introduction

Many think of “structural estimation,” a term I don’t especially like, as involving the estimation of exceedingly complex models, often involving many forms of unobserved heterogeneity, and a bewildering array of functional form and distributional assumptions. In this first talk, my aim is to convince you that this is not the case. Structural estimation is an approach to empirical work that emphasizes, in this order,

1. The development of an economic model of the phenomenon being studied
2. The addition of a stochastic structure if the model itself does not possess one
3. The adaptation of an estimation technique given the nature of the model and the data at hand
4. A consideration of the identification of the “primitive” model parameters given the data, model, and estimator employed.
5. Often, given estimates of primitive parameters, empirical comparative statics exercises and/or counterfactual policy experiments.

In other approaches to empirical work in applied economics, steps in this process are reordered or are (often) missing altogether. To contrast methods and the foci of interest, we will discuss some different approaches to the investigation of the same phenomenon. Most of the discussion will be centered on microeconomic (i.e., cross-sectional) applications.

2 Consumer Demand

Let’s say the goal of the exercise is to estimate the demand for good $x$. A standard utility maximization approach would be the following.
Given a utility function
\[ u(x, z), \]
which is increasing in both arguments,

- Prices \( p_x \) and \( p_z \) and income \( Y \)

\[ x^* = \arg \max u(x, z) \]
subject to \( Y = p_x x + p_z z. \)

Let’s begin the discussion in a highly parameterized (Cobb-Douglas) setting, where
\[ u(x, z) = Ax^\alpha z^{1-\alpha}, A > 0, \alpha \in (0, 1). \]

- In this case,
\[ x^* = \alpha Y \]
or
\[ e^* = \alpha Y, \]
where \( e^* \equiv p_x x^* \)

What does it mean to “take this model to the data”? Is this model too restrictive to be taken seriously, or not? Imagine that we have data from the Current Expenditure Survey (CEX), or the equivalent. The unit of measurement is the household, and detailed information on consumption expenditures over some well-defined period of time are ascertained. Say that expenditures on a number of goods comprising the category \( x \) are summed or each household, \( x_i \). In addition, we have access to household income, \( Y_i \), and (observable) household characteristics, \( M_i \). It is assumed that the households are i.i.d. (independent and identically distributed) draws from a fixed population. The sample data are the sequence \( \{e^*_i, Y_i, M_i\}_{i=1}^N \), where there are \( N \) households in the sample.

- The model as formulated is not suitable as a foundation for estimation. In its most basic form, there is one unknown parameter \( \alpha \). Consider sample case 1, for which
\[ e^*_1 = \alpha Y_1, \]
or
\[ \tilde{\alpha}_1 = \frac{e^*_1}{Y_1}. \]
Now for sample case $i$, $i \neq 1$, we have
\[ \tilde{\alpha}_i = \frac{e_i^*}{Y_i}, \ i = 2, ..., N. \]

If we find that
\[ \tilde{\alpha}_1 = \tilde{\alpha}_2 = ... = \tilde{\alpha}_N, \]
then we could conceivably claim that $\hat{\alpha} = \tilde{\alpha}_1$ was a reasonable estimator of $\alpha$ on some criteria, but not, in general, on any of the usual ones we employ. For example, we could not claim that $\hat{\alpha}$ was consistent, since for this to be the case we would require that, if the population were indefinitely large,
\[ \tilde{\alpha}_i = \hat{\alpha}, \ i = N + 1, ... \]

This is not even really an interpretable statement.

### 2.1 Adding Heterogeneity

There are a large number of ways to add randomness to this model. Let’s begin with perhaps the most obvious (to me, at least).

- **Random preferences.** Say that households have preferences that are i.i.d. draws from a distribution $G(\alpha)$. Then the model is characterized solely in terms of $G$, the estimation of which we could treat in a nonparametric or parametric manner.

- **Nonparametric.** Under the assumption that $e^*$ and $Y$ are measured without error, then the preference weight attached to good $x$ by household $i$ is exactly $\tilde{\alpha}_i$. Then these observations are treated as data, essentially, and we have access to the observations $\{\tilde{\alpha}_i\}$, which are i.i.d. draws from $G$. A nonparametric estimator of $G$ then is simply the cumulative distribution function, or
\[ \hat{G}_N(\alpha) = \sum_{i=1}^{N} \chi[\tilde{\alpha}_i \leq \alpha], \ \alpha \in [0, 1], \]

where $\chi$ is the indicator function.

- Note that this is really an extreme version of population heterogeneity, and represents a case in which no testable restrictions are placed on the model. We know that
\[ \text{plim}_{N \to \infty} \hat{G}_N(\alpha) = G(\alpha) \]

at all points of continuity of $G$. Thus, if $G$ is differentiable everywhere on $(0, 1)$, i.e., has a density, then $\hat{G}_N$ is consistent over all $(0, 1)$.
• Parametric (Unconditional). In this case, we continue not to condition on any observed characteristics of the sample unit \((M_i)\). We merely impose a functional form assumption on the distribution \(G\). We will want to choose a probability distribution that has the appropriate range, which is the interval \((0, 1)\). A simple, one-parameter distribution that satisfies this requirement is the \textit{power distribution}, the c.d.f. of which is
\[
G(\alpha; \theta_1) = \alpha^{\theta_1}, \quad \alpha \in (0, 1), \quad \theta_1 > 0,
\]
with p.d.f.
\[
\theta_1 \alpha^{\theta_1-1}.
\]

• As was true for the nonparametric case, the structural parameter of this specification is \(G\). Since \(G\) has been assumed to belong to the power distribution parameter, there is one (scalar) structural parameter to be estimated, which is \(\theta_1\).

• A natural estimator, especially useful for considering identification issues, is maximum likelihood. Now
\[
\tilde{\alpha}_i = \frac{e_i^*}{Y_i}.
\]
Then the c.d.f. of \(e_i^*\) conditional on \(Y_i\) under our distributional assumption is
\[
F(e_i^*|Y_i) = G(\tilde{\alpha}_i; \theta_1) = G\left(\frac{e_i^*}{Y_i}; \theta_1\right),
\]
and the (conditional) density of \(e_i^*\) is then
\[
f(e_i^*|Y_i) = \frac{\partial F(e_i^*|Y_i)}{\partial e_i^*} = \frac{\partial G(e_i^*/Y_i; \theta_1)}{\partial e_i^*} = Y_i^{-1} g\left(\frac{e_i^*}{Y_i}; \theta_1\right) = Y_i^{-1} \theta_1 \left(e_i^*/Y_i\right)^{\theta_1-1}.
\]
Then the likelihood function is given by
\[
L(\theta_1) = \prod_{i=1}^{N} Y_i^{-1} \theta_1 (\tilde{\alpha}_i)^{\theta_1-1},
\]
and the log likelihood function is
\[
\ln L(\theta_1) = \sum_{i=1}^{N} \{-\ln Y_i + \ln \theta_1 + (\theta_1 - 1) \ln \tilde{\alpha}_i\},
\] (1)
and the maximum likelihood estimator of \(\theta_1\) is
\[
0 = \frac{\partial \ln L(\hat{\theta}_1)}{\partial \theta_1}
\Rightarrow 0 = \frac{N}{\theta_1} + \sum_{i=1}^{N} \ln \tilde{\alpha}_i
\Rightarrow \hat{\theta}_1 = -\frac{N}{\sum_{i=1}^{N} \ln \tilde{\alpha}_i}.
\]
It is straightforward to show that this is the unique solution to the first order condition (due to the negative second partial derivative).

– Note that the parameter \(\theta_1\) is clearly identified given the data. Why? Under our i.i.d. sampling assumption, we could nonparametrically identify the distribution \(G\). The power distribution is a restricted version of \(G\), and therefore the fixed set of parameters upon which it depends are identified given these data (in the nonparametric case, we estimate a function \(G\), in the parametric case we merely have to estimate a finite-dimensional parameter vector \(\theta\)).

– Related to this last point, we may perform specific hypothesis tests to determine whether the power distribution adequately describes the data. If not, we may move to the nonparametric estimator, or to a more general parametric estimator. An obvious choice, in this case, is the Beta distribution, which is characterized by two parameters and which nests the power distribution.

### 2.2 Conditional Heterogeneity

We could also consider introducing conditional heterogeneity into the model, that is, heterogeneity partially based on observable characteristics \(M_i\).

- **Nonparametrics.** The only practical way to estimate a distribution \(G\) conditional on \(M_i\) is if the space to which \(M_i\) belongs is discrete. Say that \(M_i \in \Omega_m = \{m_1, m_2, \ldots, m_L\}\), where \(L\) is the number of “types.” Then partition the sample into \(L\) subsamples, with subsample \(S_j\) containing all those sample households for which \(M_i = m_j\). Then the conditional distribution estimator is given by
\[
\hat{G}_N(\alpha|m_j) = \sum_{i \in S_j} \chi[\tilde{\alpha}_i \leq \alpha], \quad \alpha \in [0, 1], \quad j = 1, \ldots, L.
\]
The conditional c.d.f. estimator is consistent for each $j$. No restrictions link the conditional distributions $G(\alpha|m_j)$ and $G(\alpha|m_k)$, $k \neq j$.

- **Parametrics.** The procedure outlined above is fine, in principle, but requires lots of data and a finite distribution of types (usually only a few) to work well in practice. When $\Omega_m$ is continuous, or discrete but with many elements, the nonparametric approach is not practical. In this case, it is common to impose some restrictions linking the parameters, or moments, of a distribution of $\alpha$ to the observed values $M_i$. This makes the distribution of $\alpha$ a conditional (on $M_i$) distribution.

- Of course, there are many ways to parameterize the model. For example, a fairly natural choice would be to set

$$\theta_{1,i} = \exp(M_i \beta),$$

where $M_i$ is considered to be a $1 \times K$ row vector and $\beta$ is a conformable parameter vector (which is $K \times 1$). The problem now becomes one of estimating $K$ parameters, the elements of $\beta$, which is only slightly more onerous than estimating one ($\theta_1$), and is a lot easier than estimating $L$ separate functions $G(\alpha|m_j)$, $j = 1, ..., L$, when $L$ was finite. This model specification still assumes random preferences, of course, but relates the distribution of the random preferences across values of $M$ in a fairly restrictive way. Estimation of the parameter vector $\beta$ via m.l. proceeds in a natural way. Just substitute (2) into (1) to get

$$\ln L(\beta) = \sum_{i=1}^{N} \{ - \ln Y_i + \ln \exp(M_i \beta) + (\exp(M_i \beta) - 1) \ln \tilde{\alpha}_i \}$$

$$= \sum_{i=1}^{N} \{ - \ln Y_i + M_i \beta + (\exp(M_i \beta) - 1) \ln \tilde{\alpha}_i \},$$

so that the m.l. estimator of $\beta$ is given by

$$\frac{\partial \ln L(\hat{\beta})}{\partial \beta} = 0$$

$$\Rightarrow 0 = \sum_{i=1}^{N} \{ M_i' + M_i' \exp(M_i \hat{\beta}) \ln \tilde{\alpha}_i \}$$

$$\Rightarrow 0 = \sum_{i=1}^{N} M_i' \{ 1 + \exp(M_i \hat{\beta}) \ln \tilde{\alpha}_i \}.$$ 

The matrix of second partial derivatives is

$$\frac{\partial \ln L(\beta)^2}{\partial \beta \partial \beta'} = \sum_{i=1}^{N} M_i M_i' \exp(M_i \beta) \ln \tilde{\alpha}_i$$

$$\leq 0.$$
The m.l. estimator $\hat{\beta}$ is uniquely determined in the sample of size $N$ if
\[
\text{rank}\left(\sum_{i=1}^{N} M_i M_i'\right) = K,
\]
which will be generally true except for some rather pathological cases.

### 2.3 Estimation Based on Measurement Error

The meaning of this title is somewhat ambiguous. Our first approach is based on the highly-structured Cobb-Douglas assumption on $u$ we have pursued up to now. At the beginning of the discussion, we confronted the fact that, with perfect measurement and no heterogeneity, we could never expect to find a common value of $\tilde{\alpha}_i$ across all sample members $i = 1, \ldots, N$. We now consider the case of no heterogeneity and imperfect measurement, which is commonly (albeit implicitly) assumed in most empirical analysis in applied economics.

We began with the relationship
\[
e_i^* = \alpha Y_i,
\]
which we know can be used to develop a valid estimator of $\alpha$. But assume that the expenditures on good $x$ are mismeasured, with
\[
E_i = e_i^* + v_i,
\]
where $v_i$ is independently and identically distributed according to the distribution $R(v)$, and $E(v) = \int v dR(v) = 0$. Then we have
\[
E_i = \alpha Y_i + v_i.
\]
The data available to form an estimator is $\{E_i, Y_i\}_{i=1}^{N}$. There is one scalar, structural parameter to estimate in this case - the common value $\alpha$. We also may be interested in estimating the distribution of measurement errors, $R$, but this distribution, in itself, is not part of the structural model.

Under the mean independence assumption, $E(v_i) = 0$ for all $i$, we know that the OLS estimator of $\alpha$ is unbiased and consistent (given a finite variance of $\nu$). In this case, a perfectly fine estimator of the structural parameter $\alpha$ is the OLS estimator
\[
\hat{\alpha} = \frac{\sum_{i=1}^{N} E_i Y_i}{\sum_{i=1}^{N} Y_i^2}.
\]
Thus OLS can be a structural estimation technique, depending on the application and the model!

1. The mean zero assumption is important, as is the independence assumption, but the “identical” assumption can be easily relaxed (to allow for unrestricted heterogeneity, for example.)
It might be natural to suppose that, if the information on expenditures on good \( x \) are mismeasured, household income \( Y_i \) might also be mismeasured. If we only observe

\[
\tilde{Y}_i = Y_i + \varepsilon_i,
\]

where \( \varepsilon_i \) is independently distributed (over \( i \)) with mean 0 and variance \( \sigma_i^2 \), say. We know that OLS applied to the regression equation

\[
E_i = \alpha \tilde{Y}_i + \xi_i,
\]

which produces the estimator

\[
\bar{\alpha} = \frac{\sum_{i=1}^{N} E_i \tilde{Y}_i}{\sum_{i=1}^{N} \tilde{Y}_i^2}
\]

doesn’t work, in the sense that

\[
E \bar{\alpha} < \alpha,
\]

that is, the estimator is biased toward 0.

There are only a few ways around this problem. The most well-known is to assume that, among the set of variables in \( M \), there are one or more that are orthogonal to \( \varepsilon \) but that are not orthogonal to \( Y \). For simplicity, assume that one of the characteristics in \( M_i \) is the education of the head of the household, which is denoted \( s_i \). Then it is reasonable to assume that

\[
E(s - E(s))(Y - E(Y)) > 0,
\]

or at least, is not equal to 0. By the same token, it may be reasonable to assume that the measurement error process has little or nothing to do with the head of household’s schooling level, at least in the sense that

\[
E(s - E(s))\varepsilon = 0.
\]

Now from (3) we can build the following estimator

\[
E_i = \alpha \tilde{Y}_i + (\nu_i - \alpha \varepsilon_i)
\]

\[
\Rightarrow \quad s_i E_i = \alpha s_i \tilde{Y}_i + s_i (\nu_i - \alpha \varepsilon_i)
\]

\[
\Rightarrow \quad \sum_{i=1}^{N} s_i E_i = \alpha \sum_{i=1}^{N} s_i \tilde{Y}_i + \sum_{i=1}^{N} s_i (\nu_i - \alpha \varepsilon_i).
\]

Now since

\[
\text{plim}_{N \to \infty} N^{-1} \sum_{i=1}^{N} s_i E_i = E(sE)
\]

\[
\text{plim}_{N \to \infty} N^{-1} \sum_{i=1}^{N} s_i \tilde{Y}_i = E(s\tilde{Y})
\]

\[
\text{plim}_{N \to \infty} N^{-1} \sum_{i=1}^{N} s_i (\nu_i - \alpha \varepsilon_i) = \text{Cov}(s(\nu - \alpha \varepsilon)) = 0,
\]

8
the IV estimator
\[ \hat{\alpha}_{IV} = \frac{\sum_{i=1}^{N} s_i E_i}{\sum_{i=1}^{N} s_i Y_i} \]
is a consistent estimator of \( \alpha \), the structural parameter characterizing the model. Thus Instrumental Variables can also be a structural estimation technique when used to consistently estimate a structural parameter. It is not a structural estimation technique when used to consistently estimate a parameter that does not have an explicit interpretation in the context of an economic model.

2.4 Approximation Arguments

To build an estimable model or equation, researchers often appeal to approximation arguments. (Even researchers considered to be “structural modelers” often use this approach. That doesn’t mean it shouldn’t be treated with some suspicion, for the following reasons.)

2.5 An Extended Example

We illustrate some of the basic steps in performing a structural estimation exercise, and the types of choices to be made. This is a simple static (household) labor supply model. In subsequent talks we will present more challenging modeling and estimation problems.

The basic steps, once you have identified a problem of interest and a data set that will be useful in estimating an appropriate model, are:

1. Specifying the objective functions of the agents
2. Specifying the set of constraints each agent faces (expenditures, information, etc.)
3. Specifying how the agents interact within one another. This includes the nature of the equilibrium, if the model is set in partial or general equilibrium, the type of game being played, if one is looking at strategic interactions, etc.
4. Specifying the stochastic structure of the model, and simultaneously, the type of estimator to be used. The stochastic structure referred to in this step is to be understand as including randomness over and above what is already included in the model. For example, if the model itself involves learning or features choice under uncertainty, stochastic assumptions are already necessary in (1).

Our example is a static model of household labor supply decisions. We assume that there are a husband and wife, each with a nonlabor income \( Y_i \).
Objective Function of Agents (1) We assume that each agent has a utility function given by
\[ u_i = u_i(l_i, C) = \alpha_i \ln l_i + (1 - \alpha_i) \ln C, \]
where \( \alpha_i \in (0,1) \), \( l_i \) is the leisure of agent \( i \), and \( C \) is a “public good” purchased by the household. The publicness of the good is what makes this a household problem, essentially. Leisure is a private good.

Constraint Set of Agents (2) Each agent makes a leisure and work time decision, where time at work is \( h_i \). The constraint is
\[ h_i + l_i = T \]
\[ h_i \geq 0 \]
\[ l_i \geq 0 \]

Each agent has an endowment of nonlabor income, \( Y_i \), and faces a wage of \( w_i \).

Household Constraints, Interaction between Spouses (3) The household purchases the public good \( C \), which we assume has price 1, using voluntary contributions by its members. Since there are no private consumption goods (i.e., goods that have value only to one of the spouses), each spouse voluntarily contributes all of their income to purchases of \( C \). In particular, each contributes all of their nonlabor income, so that \( Y = Y_1 + Y_2 \) is spent on \( C \), no matter what other conditions pertain.

For this example, we assume that the agents do not cooperate, that is, they make leisure consumption choices without factoring in the impact of their decision on the other agent. We solve for a Nash equilibrium in the reaction functions, which is reasonably straightforward in this case. The reaction function of agent 1 given the labor supply choice of individual 2 is
\[ h_1^*(h_2; S) = \arg \max_h \alpha_1 \ln(T - h_1) + (1 - \alpha_1) \ln(Y + w_1 h_1 + w_2 h_2), \]
the solution of which is given by
\[ h_1^*(h_2; S) = \begin{cases} (1 - \alpha_1)T - \alpha_1 \frac{w_2 h_2 + Y}{w_1} & \text{if } \geq 0 \\ 0 & \text{if not} \end{cases} \]
and similarly for the reaction function for person 2,
\[ h_2^*(h_1; S) = \begin{cases} (1 - \alpha_2)T - \alpha_2 \frac{w_1 h_1 + Y}{w_2} & \text{if } \geq 0 \\ 0 & \text{if not} \end{cases} \]
The Nash equilibrium is given by
\[
\hat{h}_1 = h_1^*(\hat{h}_2; S) \\
\hat{h}_2 = h_2^*(\hat{h}_1; S)
\]

One complication arises due to the presence of corner solutions. That is, given \( S \), one or neither spouse may supply positive hours to the market. This would clearly be the case as \( Y \to \infty \).

We begin by characterizing the critical wage values \( (w^*_1, w^*_2) \) that have the property
\[
w_1 < w^*_1 \text{ and } w_2 < w^*_2 \iff h^N_1 = 0 \text{ and } h^N_2 = 0,
\]
where \( h^N_i \) is the noncooperative equilibrium leisure choice of spouse \( i \). These critical values are determined as follows. If spouse 2 chooses not to supply time to the market then \( h_2 = 0 \). In this case, the reaction function of agent 1 yields a labor supply choice of
\[
\hat{h}_1(0) = (1 - \alpha_1) T_1 - \alpha_1 \frac{Y}{w_1}.
\]
In this case individual 1 will choose to supply no time to the market whenever
\[
(1 - \alpha_1) T_1 - \alpha_1 \frac{Y}{w_1} \leq 0
\]
\[
\Rightarrow w_1 \leq \frac{\alpha_1}{1 - \alpha_1} \frac{Y}{T_1}
\]
\[
w_1 \leq w^*_1.
\]

We can perform the same analysis for individual 2 conditional on a leisure demand of individual 1 equal to \( T_1 \). This case is completely symmetric, and we find the individual 2 doesn’t participate given that agent 1 doesn’t participate when
\[
w_2 \leq w^*_2 \equiv \frac{\alpha_2}{1 - \alpha_2} \frac{Y}{T_2}.
\]

From this result we know that if \( w_1 > w^*_1 \) or \( w_2 > w^*_2 \) then at least one of the two spouses will work. To determine the conditions under which spouse 1 only, spouse 2 only, or both spouses work, we return to the reaction functions of the two agents. Say that only agent 1 is working at the wage \( w_1 \). Then we know that the labor supply of spouse 1 is given by [4], and we define the critical value \( w^*_2(w_1) \) as that value of the wage offer to individual 2 at which they would just be willing to supply time to the market. This critical value is defined as
\[
0 = (1 - \alpha_2) T_2 - \frac{\alpha_2}{w^*_2(w_1)} [Y + w_1 \hat{h}_1(0)]
\]
\[
\Rightarrow w^*_2(w_1) = \frac{\alpha_2(1 - \alpha_1)}{(1 - \alpha_2) T_2} [Y + w_1 T_1]
\]

11
Then we have agent 1 works and agent 2 does not if and only if \( w_1 > w_1^{**} \) and \( w_2 \leq w_2^*(w_1) \). If the wage offers satisfy these conditions, then the labor supply of spouse 1 is given by \( \hat{h}_1(0) \).

The situation is symmetric with respect to the case of spouse 1 not working while spouse 2 supplies time to the market. In this case we define the critical value

\[
w_1^*(w_2) = \frac{\alpha_1(1 - \alpha_2)}{(1 - \alpha_1)T_1} [Y + w_2 T_2],
\]

and in equilibrium spouse 2 will supply time to the market and spouse 1 will not if and only if \( w_1 \leq w_1^*(w_2) \) and \( w_2 > w_2^{**} \).

If both spouses work, the unique noncooperative equilibrium \((h_1^N, h_2^N)\) is given by

\[
\begin{align*}
    h_1^N &= \hat{h}_1(h_2^N) \\
    h_2^N &= \hat{h}_2(h_1^N),
\end{align*}
\]

which under our functional form assumptions implies

\[
\begin{align*}
    h_1^N &= T_1 - \frac{\alpha_1(1 - \alpha_2)FI}{1 - \alpha_1 \alpha_2} \frac{w_1}{w_1} \\
    h_2^N &= T_2 - \frac{\alpha_2(1 - \alpha_1)FI}{1 - \alpha_1 \alpha_2} \frac{w_2}{w_2},
\end{align*}
\]

where \( FI = Y + w_1 T_1 + w_2 T_2 \) denotes full (household) income. For the equilibrium to be of this type, it must be the case that \( w_1 > w_1^*(w_2) \) and \( w_2 > w_2^*(w_1) \).

\[
\begin{tabular}{|c|c|c|c|c|}
\hline
Equilibrium Type & \( w_1 \) & \( w_2 \) & \( h_1^N \) & \( h_2^N \) \\
\hline
I & \( w_1 \leq w_1^{**} \) & \( w_2 \leq w_2^{**} \) & 0 & 0 \\
II & \( w_1 > w_1^{**} \) & \( w_2 \leq w_2^*(w_1) \) & \( T_1 - \frac{\alpha_1}{w_1} [Y + w_1 T_1] \) & 0 \\
III & \( w_1 \leq w_1^*(w_2) \) & \( w_2 > w_2^{**} \) & 0 & \( T_2 - \frac{\alpha_2}{w_2} [Y + w_2 T_2] \) \\
IV & \( w_1 > w_1^*(w_2) \) & \( w_2 > w_2^*(w_1) \) & \( T_1 - \frac{\alpha_1(1 - \alpha_2)FI}{1 - \alpha_1 \alpha_2} \frac{w_1}{w_1} \) & \( T_2 - \frac{\alpha_2(1 - \alpha_1)FI}{1 - \alpha_1 \alpha_2} \frac{w_2}{w_2} \) \\
\hline
\end{tabular}
\]
2.6 Econometrics of the Model

Whether estimating a noncooperative or (costly) cooperative version of the equilibrium household labor supply model, the underlying parameterization of the household’s problem is invariant. We let household preferences be determined by a draw from a joint distribution over \( \alpha \), the c.d.f. of which we denote by \( G(\cdot; \nu) \), where \( \nu \) is a finite dimensional parameter vector. Since \( \alpha \) are parameters of Cobb-Douglas utility functions, we restrict the support of \( G \) to be \([0,1]^2\). For simplicity, in estimating the model we have assumed that preference draws are independent for husbands and wives, and that the distribution of utility for each spouse is a power distribution with

\[
G_i(\alpha) = \nu^i, \ \alpha \in [0,1], \ \nu > 0.
\]

Thus the probability that \( \alpha_1 \leq A_1 \) and \( \alpha_2 \leq A_2 \) is given by \( A_1^{\nu_1} A_2^{\nu_2} \).  

We have assumed that the wage offer distribution is bivariate log normal, with

\[
\left( \begin{array}{c} \ln w_1 \\ \ln w_2 \end{array} \right) \sim N( \left( \begin{array}{c} \mu_1 \\ \mu_2 \end{array} \right), \Sigma).
\]

In the actual estimation, we allow for individual-specific heterogeneity in the parameters \( \mu_i \). In particular, we assume that \( \mu_i = \beta_i X_i \), where \( X_i \) is a \( K \times 1 \) vector of observable characteristics of spouse \( i \) and \( \beta_i \) is a conformable parameter vector. We allow \( \sigma_{12} \neq 0 \), which means that \( \ln w \) offers are not independently distributed (and of course, neither are the \( w = \exp(\ln(w)) \)). In this way we allow for the possibility of assortative mating, although in an admittedly crude manner.

The model is very parsimoniously specified, with the joint distribution of wages and labor supplies determined by the parameters \( \Omega = \{\nu_1, \nu_2, \beta_1, \beta_2, \Sigma\} \) in the case of noncooperation and \( \Omega' = \{\Omega, \zeta\} \) in the costly cooperation case.\(^3\) We first consider the estimation of these parameters under the assumption that the household equilibrium is noncooperative.

2.7 Estimation of the Noncooperative Model

We form maximum likelihood estimators for \( \Omega \). The likelihood function consists of four components, one for each “type” of equilibrium represented in Table 1. We proceed through them sequentially.

The likelihood of neither spouse working (Type I) is computed as follows. Conditional on the preference parameter vector \( \alpha \), the probability that neither works is the joint probability that the wage offer to spouse 1 is less than \( w_1^{**}(\alpha_1, T_1, Y) \) and the wage offer to

\(^2\)We are restricting the preference parameters to be independent mainly for computational simplicity at this point, though our intention is to allow some dependence in these parameters in our future research. For the moment, all assortative mating is captured through dependence in wage offers.

\(^3\)The noncooperative case is actually nested (in a limiting sense) within the noncooperative case, as \( \zeta \to 0 \).
spouse 2 is less than \( w_2^{**}(\alpha_2, T_2, Y) \), or

\[ F_{W_1, W_2}(w_1^{**}(\alpha_1, T_1, Y), w_2^{**}(\alpha_2, T_2, Y)) \]

The marginal probability of this event, which is its likelihood, is obtained by multiplying the above expression by the density of \( \alpha \) and integrating over the interval \([0, 1]^2\), or

\[ P_I(\nu, \mu, \Sigma) = \int_0^1 \int_0^1 F_{W_1, W_2}(w_1^{**}(\alpha_1, T_1, Y), w_2^{**}(\alpha_2, T_2, Y); \mu, \Sigma) dG(\alpha; \nu). \]

The second and third types of equilibrium in which only one spouse works are symmetric, so we will only discuss the case in which spouse 1 works and spouse 2 does not (Type II). In this case the wage offer to the first spouse must satisfy \( w_1 > w_1^{**} \), while the wage offer to spouse 2 must be less than \( w_2^{**}(w_1) \). Now when the second agent does not work, it is straightforward to compute the first agent’s preference parameter value since

\[
h_1 = T_1 - \frac{\alpha_1}{w_1} [Y + w_1 T_1]
\]

\[
\Rightarrow \alpha_1 = \frac{w_1 (T_1 - h_1)}{Y + w_1 T_1},
\]

and all of the terms on the right hand side of [5] are observable. Thus the likelihood of agent 1’s choice of hours and their wage rate is given by

\[
\left[ \frac{w_1}{Y + w_1 T_1} \right] \left[ \nu_1 \left( \frac{w_1 (T_1 - h_1)}{Y + w_1 T_1} \right)^{\nu_1-1} \right] f_{W_1}(w_1; \mu_1, \sigma_{11}),
\]

where the first term in brackets is the absolute value of the Jacobian, which together with the second term in brackets comprises the density of hours of spouse 1 conditional on the wage rate \( w_1 \). The last term is the marginal density of individual 1’s wage, so that the entire product gives the joint density of hours and the wage of spouse 1.

Turning to the information in individual 2’s decision not to participate, we note that

\[
w_2^*(w_1, \alpha_1, \alpha_2) = \frac{\alpha_2 (1 - \alpha_1)}{(1 - \alpha_2) T_2} [Y + w_1 T_1]
\]

gives the critical spouse 2 wage offer such that all \( w_2 \) below this value lead to a corner solution for individual 2. Conditional on \( \alpha_2 \) (in addition to \( w_1 \) and \( \alpha_1 \)), the probability that agent 2 does not have an acceptable offer is given by

\[
p(w_2 < w_2^*(w_1)|\alpha_1, \alpha_2, w_1) = F_{W_2|W_1}(w_2^*(w_1, \alpha_1, \alpha_2)|w_1).
\]

Unconditional on \( \alpha_2 \) (but conditional on \( w_1 \) and \( \alpha_1 \)) then the probability that individual 2 does not work is

\[
p(w_2 < w_2^*(w_1)|\alpha_1, w_1) = \int F_{W_2|W_1}(w_2^*(w_1, \alpha_1, \alpha_2)|w_1) dG_2(\alpha_2).
\]
Finally, multiplying by the likelihood of $\alpha_1$ and $w_1$ we have the likelihood of the Type II outcome

$$P_{II}(h_1, w_1; \nu, \mu, \Sigma) = \int F_{W_2|W_1}(w_2^*(w_1, \alpha_1, \alpha_2)|w_1; \mu, \Sigma) \nu_2 \alpha_2^{\nu_2 - 1} \, d\alpha_2$$

$$\times \left[ \frac{w_1}{Y + w_1 T_1} \right] \left[ \nu_1 \left( \frac{w_1(T_1 - h_1)}{Y + w_1 T_1} \right)^{\nu_1 - 1} \right] f_{W_1}(w_1; \mu_1, \sigma_{11}).$$

The likelihood of a Type III equilibrium, in which spouse 1 does not work and spouse 2, is computed in an exactly symmetric way, so that

$$P_{III}(h_2, w_2; \nu, \mu, \Sigma) = \int F_{W_1|W_2}(w_1^*(w_2, \alpha_1, \alpha_2)|w_2; \mu, \Sigma) \nu_2 \alpha_2^{\nu_2 - 1} \, d\alpha_2$$

$$\times \left[ \frac{w_2}{Y + w_2 T_2} \right] \left[ \nu_2 \left( \frac{w_2(T_2 - h_2)}{Y + w_2 T_2} \right)^{\nu_2 - 1} \right] f_{W_2}(w_2; \mu_2, \sigma_{22}).$$

The likelihood of a Type IV equilibrium with observed values of $h_1, h_2, w_1,$ and $w_2$ is computed as follows. Define $A_i = (Y + w_1 T_1 + w_2 T_2)/w_i$, $i = 1, 2$, and let

$$\tau = \frac{l_1 l_2}{(A_1 - l_1)(A_2 - l_2)},$$

$$\tau_1 = \frac{\tau}{l_1} + \frac{\tau}{A_1 - l_1},$$

$$\tau_2 = \frac{\tau}{l_2} + \frac{\tau}{A_2 - l_2}.$$ 

Then we can write the implied values of the preference parameters as

$$\alpha_1 = \tau + (1 - \tau) \frac{l_1}{A_1},$$

$$\alpha_2 = \tau + (1 - \tau) \frac{l_2}{A_2}.$$ 

The Jacobian is given by

$$J_N = \left| \begin{array}{cc} \tau_1 (1 - \frac{T_1-h_1}{A_1}) + \frac{1-\tau}{A_1} & \tau_2 (1 - \frac{T_1-h_1}{A_1}) + \frac{1-\tau}{A_2} \\ \tau_1 (1 - \frac{T_2-h_2}{A_2}) & \tau_2 (1 - \frac{T_2-h_2}{A_2}) \end{array} \right|.$$ 

Thus, conditional on the wage draws $(w_1, w_2)$, the likelihood the hours choices $(h_1, h_2)$ is

$$p(h_1, h_2|w_1, w_2) = J_N \nu_1 \nu_2 (\tau + (1 - \tau) \frac{l_1}{A_1})^{\nu_1 - 1} (\tau + (1 - \tau) \frac{l_2}{A_2})^{\nu_2 - 1}.$$ 

After multiplying by the likelihood of the two wage draws, we get the total likelihood of the Type IV equilibrium draw as

$$P_{IV}(h_1, h_2, w_1, w_2; \nu, \mu, \Sigma) = J_N \nu_1 \nu_2 (\tau + (1 - \tau) \frac{T_1-h_1}{A_1})^{\nu_1 - 1} (\tau + (1 - \tau) \frac{T_2-h_2}{A_2})^{\nu_2 - 1}$$

$$\times f_{W_1,W_2}(w_1, w_2; \mu, \Sigma).$$