THE VARIATION IN HOURS OF WORK AMONG INDIVIDUALS

Robert E. Hall

Department of Economics

M. I. T.

September 1, 1972

I am grateful to Daniel McFadden for suggestions at an early stage and to Zvi Body and Meir Kohn for assistance. This research was supported by the National Science Foundation under grant GS28608.

1. Introduction

One of the most important choices made by almost every adult in a modern economy is whether to work at a job and, if so, how much to work. There is now a substantial body of empirical research on labor force participation[1] and on hours of work.[2] Almost without exception, previous work has studied the choice made by the typical individual, and has ignored the tremendous variation in behavior that is revealed in any body of data on individuals. The only mention of this variation made by most authors is an apology for the poor statistical fit of labor supply equations estimated in cross-sections.

This paper attempts to deal directly with variation in the behavior of seemingly identical individuals, both in the theory of individual choice and in the statistical model used in the study of a large cross-section of data on individuals. The theory starts from the supposition that there are unobserved differences in tastes among individuals, and that these differences can be characterized by a probability distribution. This distribution of preferences induces a distribution in hours of work among a group of individuals even when they are presented with identical opportunities and rewards for work. A basic point of the paper is that the behavior of individuals of each type can be deduced from the observed distribution of

---

1. For example, Bowen and Finnegan (1969), Cain ( ), Kalachek and Raines (1969).

2. For example, Cohen, Rea and Lerman (1970), Kalachek and Raines (1969), and the papers in Cain and Watts (1973), including that of the present author.

choices even when the types are unobservable, provided that there is only a single dimension of variation among individuals and a monotonic relation between individual types and their hours of work. If these conditions hold, then previous research has been excessively modest in attempting to measure the labor supply function of only the typical individual-it is also possible to measure the labor supply functions of individuals whose preferences cause them to work relatively little or to work a great deal. In this study we examine the median of the distribution of hours of work as the labor supply function of the typical individual (in our framework it is the median, not the mean studied by previous authors, that is the supply of the typical person, but empirically the median and mean are not very different), the first quartile of the distribution as the labor supply of individuals whose preferences are biased against work, and the third quartile as the supply of hard workers.

The statistical section of the paper describes a general approach to the estimation of probability distributions that depend on observed variables. Since the method is intended to be used on large bodies of data, the emphasis is on specifications that can approximate a wide variety of distributions rather than on a tight parametric specification. Thus we do not assume that the distribution of hours of work can be described by any familiar distribution with only a few parameters.

The empirical section of the paper presents the results of a study of the hours of work of several thousand married individuals in the United States in 1966 and 1967. Separate results are given for the four race-sex groups. In addition to wages and income, the supply functions depend on the size and composition of the family and on the spouse's wage.

2. A Probabilistic Model of Individual Behavior

In this section we develop a simple model that deals explicitly with differences among individuals. We start from the assumption that the tastes of individuals can be described by a single index, θ. All individuals of a particular type, θ, have the same tastes and therefore have the same supply function for labor:

$$\text{(2.1)} \qquad \begin{aligned} h &= S(w,y,p,z,\theta) \\ &= S(x,\theta) \end{aligned}$$

where

$$x = \begin{bmatrix} w \\ y \\ p \\ z \end{bmatrix} .$$

and w is the wage, y is income, p is a vector of prices of other goods, and z is a vector of observed characteristics. We suppose that h decreases with θ, so θ might be called a measure of distaste for work.

Now the supply function cannot be estimated directly since it involves the unobserved index θ. We proceed by assuming that there is an underlying distribution of different types of individuals. We define

$$\text{(2.2)} \qquad G(T) = \text{fraction of the population with } \theta \leq T$$

It is convenient to use the language of probability to discuss this distribution, although at this stage there is nothing inherently random about the subject. Thus we refer to G(T) as the cumulative distribution function of individual tastes - the probability that an individual drawn at random from the population will have $\theta \leq T$. Under these conditions we can estimate the cdf of the demand itself:

$$F(H,x) = \text{Prob}\ [h \leq H] \tag{2.3}$$

since it depends only on observable quantities. Now

$$F(H,x) = \text{Prob}\ [S(x,\theta) \leq H] \tag{2.4}$$

$$= \text{Prob}\ [\theta \leq S^{-1}(x,H)]$$

where $S^{-1}(x,H)$ is the inverse of $S(x,\theta)$ in its second argument:

$$S(x,S^{-1}(x,H)) = H \qquad \text{for all } x,H. \tag{2.5}$$

The inverse exists only if demand is monotonically related to the index of tastes, θ. Then we have

$$F(H,x) = G(S^{-1}(x,H)) \tag{2.6}$$

Although the supply function itself can never be estimated as long as θ is unobservable, something very closely related, and just as useful, can be estimated. We define the α-fractile, $m_\alpha(x)$, as the number of hours of work such that a fraction α of the population supply $m_\alpha(x)$ hours or fewer:

$$F(m_\alpha(x),x) = \alpha \tag{2.7}$$

For example, if $\alpha = \frac{1}{2}$, $m_\alpha(x)$ is the median supply. From equation (2.6) we have

$$G(S^{-1}(x,m_\alpha(x)) = \alpha \tag{2.8}$$

Thus

$$S^{-1}(x,m_\alpha(x)) = m_\alpha \tag{2.9}$$

where m_α is the α-fractile of the distribution of θ. Finally,

$$S(x,m_\alpha) = m_\alpha(x) \tag{2.10}$$

In other words, $m_\alpha(x)$ is exactly the supply function at the unobservable value of θ, $\theta = m_\alpha$. We can treat $m_\alpha(x)$ as the supply function of a particular individual; it should obey the usual laws of demand for an individual. No other estimable function of x can be interpreted in this way. In particular, the expected value of supply,

$$E(h|x) = \int_{-\infty}^{\infty} S(x,\theta)dG(\theta)$$

$$= \int_{-\infty}^{\infty} hdF(h,x)$$

cannot generally be interpreted as a demand function for any single individual or group of identical individuals. Previous studies of labor supply have almost invariably interpreted the mean of the labor supply function as the supply function of the typical individual. Under the assumptions of this paper, it is the median, not the mean, that is typical.

3. Estimation of the Distribution of Hours of Work

The previous section has demonstrated the usefulness of studying the distribution of hours of work. In this section we discuss econometric methods for estimating probability distribution that depend not only on unknown parameters but also on observed variables. Except for special cases such as multiple regression, this subject has received little attention in the statistical literature. Our approach is to estimate the cumulative distribution at a set of points $\overline{H}_1, \ldots \overline{H}_N$. We start by defining the probability that h will fall in a certain interval.

$$f_i(x) = F(\overline{H}_i, x) - F(\overline{H}_{i-1}, x) \tag{3.1}$$

(where $F(\overline{H}_o, x) = 0$). The problem then is to estimate $f_1(x), \ldots, f_N(x)$ and from these obtain the cumulative distribution function at the points $\overline{H}_1, \ldots \overline{H}_N$ by addition. If $\overline{H}_1$ is 0 or some small positive number, then $1-f_1(x)$ is the probability of labor force participation of an individual with wage, income, and characteristics, x. There is, of course, a large empirical literature on labor force participation. The novel feature of our work is to estimate not just the probability of working, but the probabilities of working quarter-time, half-time, full time, overtime, and of working at two jobs.

Our next step is to define the random variable j by

$$\overline{H}_{j-1} < h \leq \overline{H}_j \tag{3.2}$$

that is, j is the index of the interval in which h falls.

Then

(3.3) $$\text{Prob}\ [j = i] = f_i(x).$$

Next we must decide upon a suitable parametric specification for $f_i(x)$. It must obey the two constraints of a discrete probability:

(3.4) $$0 \leq f_i(x) \leq 1$$

and

(3.5) $$f_1(x)+\ldots+ f_N(x) = 1.$$

A convenient specification is the multinomial logit function,

(3.6) $$f_i(x) = \frac{e^{g(x)\beta^{(i)}}}{e^{g(x)\beta^{(1)}}+\ldots+ e^{g(x)\beta^{(N)}}} .$$

where $g(x)$ is a known vector-valued function and $\beta^{(1)},\ldots, \beta^{(N-1)}$ are vectors of parameters to be estimated. We use the normalization $\beta^{(N)} = 0$. With a suitable choice of the function $g(x)$ a set of points on any well-behaved cumulative distribution function $F(H,x)$ can be approximated with arbitrary accuracy, as shown in the following.

<u>Approximation Theorem</u>

Suppose the domain of $f,(x),\ldots, f_N(x)$ is bounded, $|x| \leq A$, and the derivatives are bounded, $\left|\frac{\partial f_i(x)}{\partial x}\right| \leq B$ for all i and all x. Then for any $\varepsilon > 0$ there is a vector function g(x) and a set of vectors $\beta^{(1)},\ldots,\beta^{(N-1)}$ such that

$$\frac{e^{g(x)\beta^{(i)}}}{e^{g(x)\beta^{(1)}}+\ldots+e^{g(x)\cdot\beta^{(N-1)}}+1} - f_i(x) \leq \varepsilon \tag{3.7}$$

for all i, x.

The proof appears in the appendix. Three points should be noted: First, the accuracy of the approximation does not depend at all on the number or spacing of the points on the cumulative distribution that are estimated. Second, the requirement that the $f_i(x)$ be differentiable in x does not rule out discontinuities in the cumulative distribution function F(H,x) in H; it only requires that the location of the discontinuities not depend on x. This is important because the cumulative distribution of hours of work usually is discontinuous at zero hours. Third, the choice of the function g(x) need not depend on F(H,x) except on its smoothness as measured by the bound on its derivatives, B. In practice the choice of g(x) balances the need for accurate approximation against the cost of excessive parameters. Our work uses a g(x) of the following sort:

$$g(x) = \begin{bmatrix} g^{(1)}(x) \\ \vdots \\ g^{(Q)}(x) \end{bmatrix} \tag{3.8}$$

where

(3.9) $$g^{(1)}(x) = 1$$ and

(3.10) $$g_k^{(q)}(x) = \frac{x - \bar{x}_{k-1}^{(m)}}{\bar{x}_k^{(m)} - \bar{x}_{k-1}^{(m)}} \quad \text{if } \bar{x}_{k-1}^{(m)} \leq x_m \leq \bar{x}_k^{(m)}$$

$$= \frac{\bar{x}_{k+1}^{(m)} - x}{\bar{x}_{k+1}^{(m)} - \bar{x}_k^{(m)}} \quad \text{if } \bar{x}_k^{(m)} \leq x_m \leq \bar{x}_{k+1}^{(m)}$$

$$= 0 \quad \text{otherwise.}$$

These are piecewise interpolations between points on a grid of values of a selected element of x (called x_m above). The function obtained as the weighted sum of functions of this kind, $g(x)\beta$, is continuous in x. If one of the elements of x, say x_m, is an integer, and the breakpoints, $\bar{x}_1^{(m)}, \bar{x}_2^{(m)}, \ldots$ are consecutive integers, then the functions $g_k^{(q)}(x)$ are just a set of dummy variables.

We turn now to the problem of estimating the parameters $\beta^{(1)},..,\beta^{(N-1)}$ of the model. The probability of a set of observations, $J_1, \ldots, j_k$ on the statistic j defined in equation (3.2) is

(3.11) $$\text{Prob}\ [j_1,\ldots, j_k] = \prod_{k=1}^{K} \frac{e^{g(x)\beta^{(j_k)}}}{e^{g(x)\beta^{(1)}} + \ldots + e^{g(x)\beta^{(N-1)}} + 1}$$

We may interpret this as the likelihood function for $\beta^{(1)},\ldots,\beta^{(N-1)}$ and obtain estimates by finding the values of the parameters that maximize the likelihood. The resulting estimator has been studied by McFadden (1968) and Theil (1971) in the context of the multinomial logit model (both use a somewhat different notation, but equivalent from the point of view of estimation). They show that as long as there are sufficient observations in each category and the elements of g(x) are not linearly dependent, the likelihood function is strictly concave and thus attains a unique relative maximum which is the global maximum. Further they show that the maximum likelihood estimator is best asymptotically normal.

4. Nonlinear budgets

An important characteristic of the choice of hour of work is that the alternative combinations of work and consumption of goods that are available cannot be described by a linear relation. Even if the marginal and average wages are everywhere equal before taxes, the increasing marginal tax rate of the personal income tax causes the marginal wage as seen by the worker to decline with increasing hours of work. This section presents an explicit theoretical justification for the approach taken both here and in the author's earlier paper (Hall, 1973), to accounting for the curvature of the budget constraint induced by a progressive tax. The same method could be used to account for other sources of nonlinearities, especially the diseconomies of part-time work, which probably cause an increasing marginal wage within a certain range of hours.

We start with two definitions:

1. A <u>regular budget</u>, B, is a set of consumption bundles

$$B = \{x | f(x) \leq 0\} \tag{4.1}$$

where f is defined on all non-negative x, and is increasing and differentiable.

2. A <u>regular preference ordering</u> is an ordering on the set of all non-negative consumption bundles with the properties that more is preferred to less ($x' \geq x$ and $x' \neq x$ implies x' is strictly preferred to x) and that the upper contour set,

$$A(x) = \{x' | x \text{ not strictly preferred to } x'\} \tag{4.2}$$

is closed and strictly convex for all x.

In what follows, all consumers are assumed to have regular preferences and to face regular budgets. We note that a regular budget is closed but not necessarily convex, and that a linear budget $\{x | p.x-y = 0\}$ is a regular budget.

The demand function of the consumer is a function of his budget:

$$D(B) = \text{most preferred } x \in B \tag{4.3}$$

Functions of sets are not at all convenient to deal with, nor are they mathematically economical. To see this, consider the restricted demand function D* (p,y) defined only for linear budgets:

(4.4) $$D^*(p,y) = D(\{x|p'x-y = 0\}).$$

From the theory of revealed preference we know that it is possible to discover the whole set of preferences of an individual by observing his demands when presented with alternative linear budgets. Thus given $D^*(p,y)$ we can reconstruct D(B) from its definition, (4.3). This suggests that $D^*(p,y)$ may be an economical way to organize information about preferences even for a consumer who faces nonlinear budgets. If we have data on $x = D(B)$ and B, we need some way to get an equivalent p and y in order to estimate $D^*(p,y)$. If we knew A(x) then the solution is straightforward; the hyperplane separating A(x) and the convex set $L(x) = \{x'|x' < x\}$, say $\{x''|p.x'' = y\}$, provides the appropriate p and y. Obviously this is not empirically useful since A(x) is unknown. On the other hand, the whole budget set, B, is known, and in particular the derivatives of the function, f, defining the budget, are known at the point of the observed demand, x. The usefulness of the derivatives is shown by the following

<u>Theorem</u>: Consider a regular budget, B, and the associated demand, x. The prices

(4.5) $$p_i = \frac{\partial f(x)}{\partial x_i}$$

and the income

(4.6) $$y = p'x$$

satisfy the following condition:

$$D^{*}(p,y) = D(\{x' | p \cdot x' - y = 0\}) = D(B)$$

5. The Data

Data for this study were obtained from the Survey of Economic Opportunity for 1967. Many of the steps in preparing the data were the same as in the author's earlier study (Hall, 1973), which the reader should consult for a more complete description. Briefly, the data were calculated in the following way:

Wage: For each individual, an imputed hourly wage was calculated from data on age, education, geographic area, and various other personal characteristics. The coefficients of the formula for this imputation were estimated in a preliminary regression for those indiviuals for whom an hourly wage was reported. These regression results are similar to those reported in the earlier study and are not presented here, but they may be obtained from the author.

Income: For the purposes of this study family income is defined as non-labor income, including the imputed value of all types of property, plus the value of the time of all adults in the family evaluated at full-time work (2000 hours per year). This definition implies that the derivatives of the labor supply function with respect to the wage is more like a substitution effect than it would be if income were defined as non-labor income alone. The choice between the two definitions is purely arbitrary and does not affect the interpretation of the derivatives with respect to income at all. The methods for calculating imputed property income were the same as those described in the earlier paper.

Hours of Work

Our measure of annual hours is the one used in the earlier study, annual wage earnings divided by the imputed hourly wage.

6. The Specification

We recall from section 3 that we hope to estimate a set of probabilities,

$$f_i(x) = \text{Prob}\ [\overline{H}_{i-1} < h \leqq \overline{H}_i], \tag{6.1}$$

and that a convenient family of specifications has the following form:

$$f_i(x) = \frac{e^{g(x)\beta^{(i)}}}{e^{g(x)\beta^{(1)}} + e^{g(x)\beta^{(N-1)}} + 1} \tag{6.2}$$

Elements of the x-vector in our study are : hourly wage, w, annual family income divided by the number of adults in the family, y, hourly wage of spouse, s, number of adults in the family, and the number and ages of children. The function g we have chosen is:

(6.3) $g_1(x) = 1$ if the wage is 0, and declines to 0 as w approaches $1.50.

$g_2(x) = 1$ if the wage is $1.50 and declines 0 for higher and lower wages.

$g_3(x) =$ as above for $w = \$2.00$

$g_4(x)$ = as above for $w = \$3.00$.

$g_5(x)$ = as above for $w = \$4.00$.

$g_6(x) = 1$ if $w = \$10.00$; declines to zero as w approaches \$4.00.

$g_7(x) = 1$ if income is 0; declines to 0 as y approaches \$3000.

$g_8(x) = 1$ if $y = \$4000$; declines to 0 as y approoahced \$3000 and \$5000.

$g_9(x) = 1$ if $y = \$5000$; declines to 0 for higher and lower y.

$g_{10}(x) = 1$ if $y = \$10,000$; declines to zero as y approaches \$5000.

$g_{11}(x) = 1$ if the family has 3 or 4 adults; 0 otherwise.

$g_{12}(x) = 1$ if the family has 5 or more adults; 0 otherwise.

$g_{13}(x) = 1$ if the family has children of preschool age only; 0 otherwise.

$g_{14}(x) = 1$ if the family has children of school age only; 0 otherwise.

$g_{15}(x) = 1$ if the family has children of both preschool and school age; 0 otherwise.

$g_{16}(x)$, ..., $g_{21}(x)$ same as $g_1(x)$, ..., $g_6(x)$ for spouse's wage.

The categories of hours of work used in the study are: 1. 0 to 100, 2. 100 to 900, 3. 900 to 1800, 4. 1800-2300, 5. 2300-2800, 6. 2800 or more. Thus, for example, the probability that a person will work approximately full time (category 4) if his wage is \$3.00 per hour, his famimly income is \$4000 per year, and there are no extra adults and only children of school age, is

(6.4)
$$\frac{e^{\beta_4^{(4)} + \beta_8^{(4)} + \beta_{14}^{(4)}}}{e^{\beta_4^{(1)} + \beta_8^{(1)} + \beta_{14}^{(1)}} + \ldots + e^{\beta_4^{(5)} + \beta_8^{(5)} + \beta_{14}^{(5)}} + 1}$$

If his wage were, instead, \$2.50, the numerator of this expression would be

(6.5)
$$e^{\frac{1}{2}\beta_3^{(4)} + \frac{1}{2}\beta_4^{(4)} + \beta_8^{(4)} + \beta_{14}^{(4)}}$$

After estimating the parameters of this model we calculated approximate α-fractiles of the implies distribution of hours of work for various values of x by the following procedure:

Let J be defined by

(6.6)
$$\sum_{j=1}^{J-1} f_j(x) \leq \alpha \leq \sum_{j=1}^{J} f_j(x)$$

Then the approximate α-fractile, $\hat{m}_\alpha(x)$, is

$$(6.7) \qquad \hat{m}_\alpha(x) = \bar{H}_{J-1} + (\bar{H}_J - \bar{H}_{J-1}) \frac{\alpha - \sum_{j=1}^{J-1} f_j(x)}{f_J(x)}$$

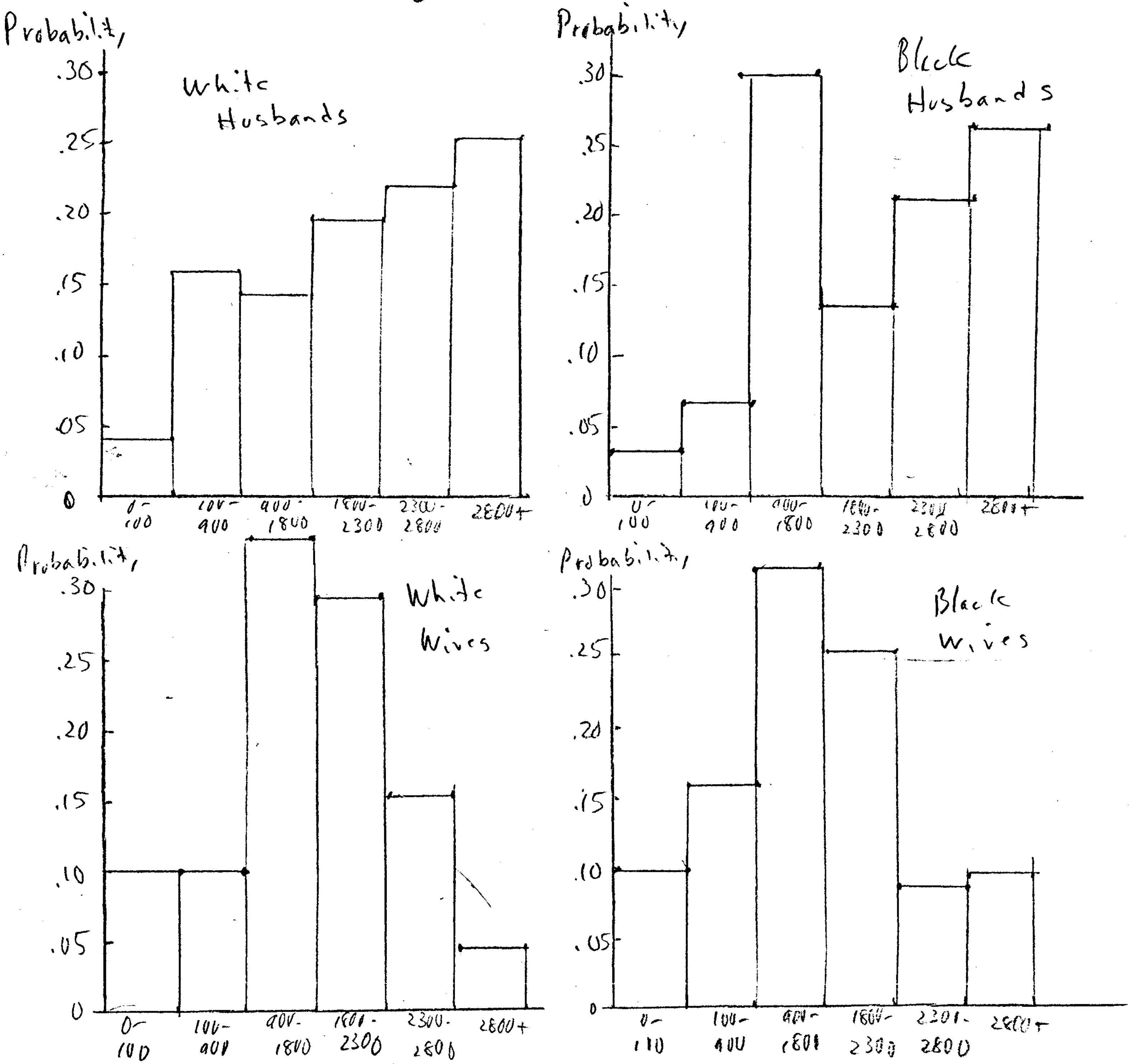
This procedure is exact if the cumulative distribution is linear in h; that is, if the density of h is a step function.

7. Reuslts

It should be apparent that the model of family behavior discussed in section 2 is not precisely applicable to the data discussed in section 5. There are three major sources of discrepancy between the two: First, the data on hours of work contain random errors of measurement. Even if these errors have mean zero, they cause an increase in the dispersion of hours of work beyond that caused by the underlying dispersion of tastes. Second, the model ignores the randomness of the behavior of individuals over time. It is evident that many individuals who work less than full time do so by an unsystematic pattern of full-time work followed by periods of no work. Much, but not all, of this random variation is eliminated by the use of data an hours over a full year. The variation remaining from this source is again attributed to dispersion in individual tastes. Research in progress will attempt to deal explicity with the two sources of variation together. Third, the model rests on the unverifiable hypothesis that the distribution of tastes is independent of individual characteristics. Our interpretation of wage and income effects depends critically on this hypothesis. If individuals with high wages tend also to be hard workers with low values of θ, for example, then the upward shift of the distribution with rising wages cannot be interpreted as measuring the shift of hours of work of a particular individual if his wage were changed. This is a pervasive difficulty in almost all research based on cross sections.

With these problems in mind we present the results of the study. For each of the four color-sex groups there are five vectors of 20 coefficients each, an unwieldy set of results that is postponed to an appendix. Here we discuss various derived results. First is the full distribution for a certain group in the population, presented as a set of histograms in Firgure 1.

Figure 1.
Distributions of Hours of Work



Data for Fig. 0

	0-100	100-900	900-1800	1800-2300	2300-2800	2800+
White husbands	.04	.16	.14	.19	.22	.25
Black husbands	.03	107	.30	.13	.21	.26
White wives	.10	.10	.34	.28	.15	.03
Black wives	.10	.16	.32	.25	.08	.09

Figure 2
Labor Supply Functions

White Husbands

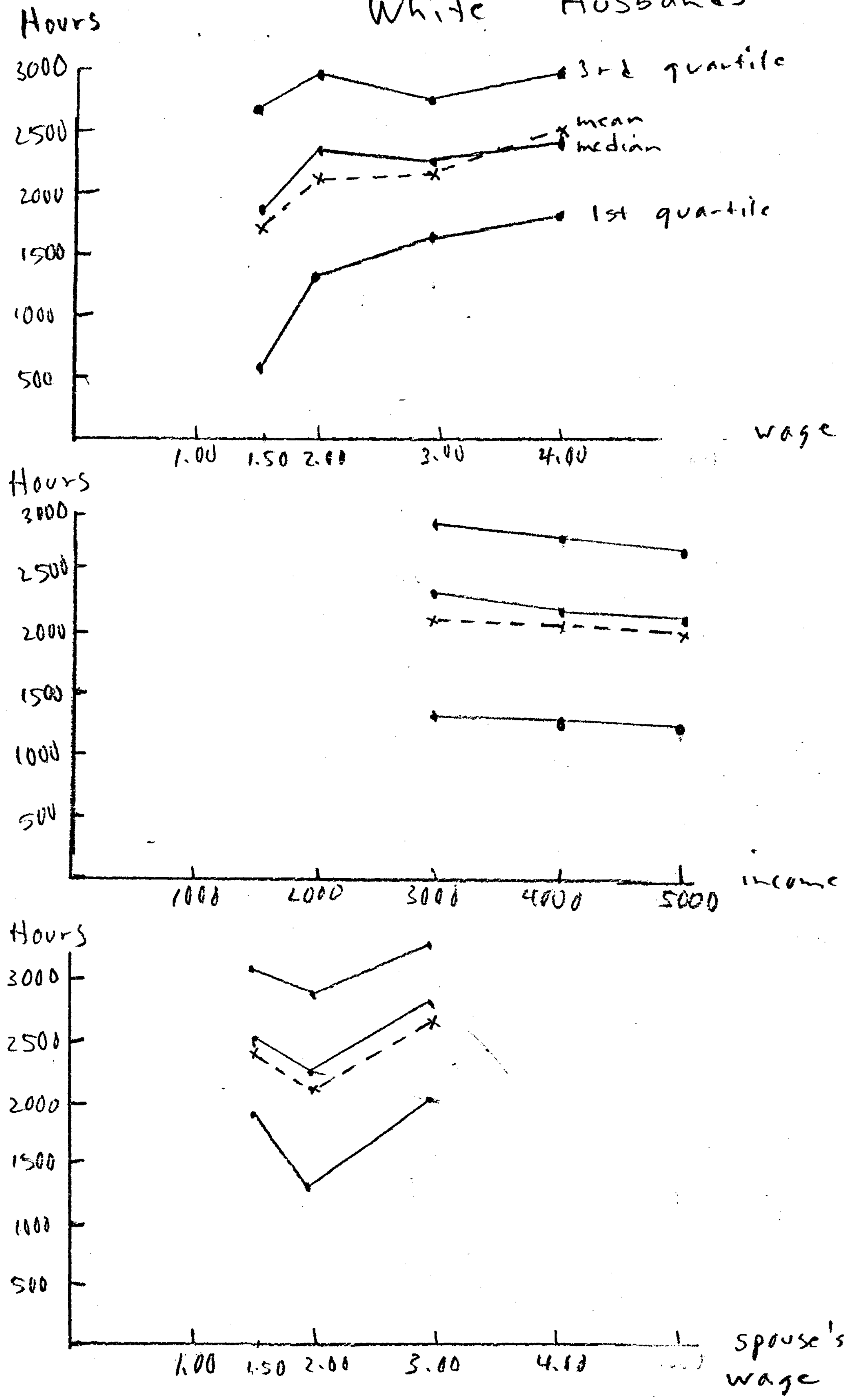
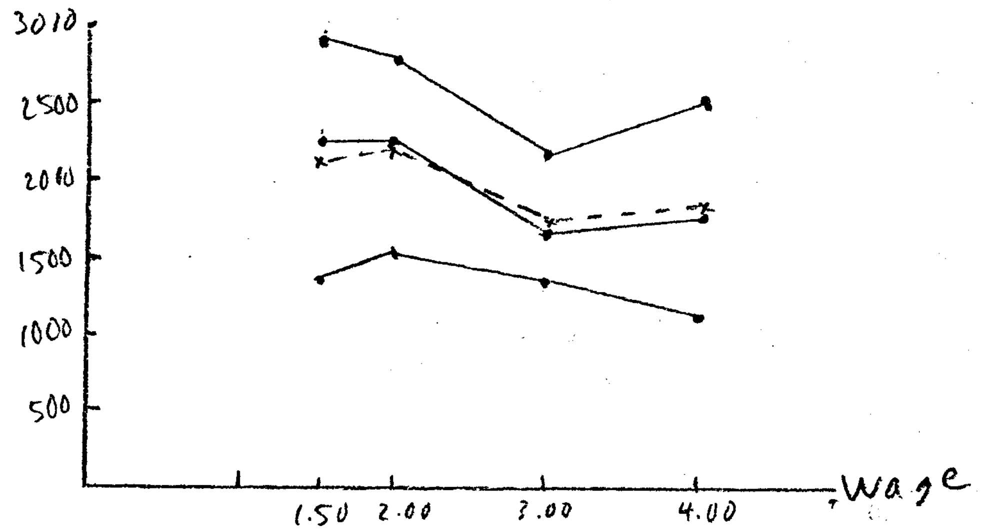
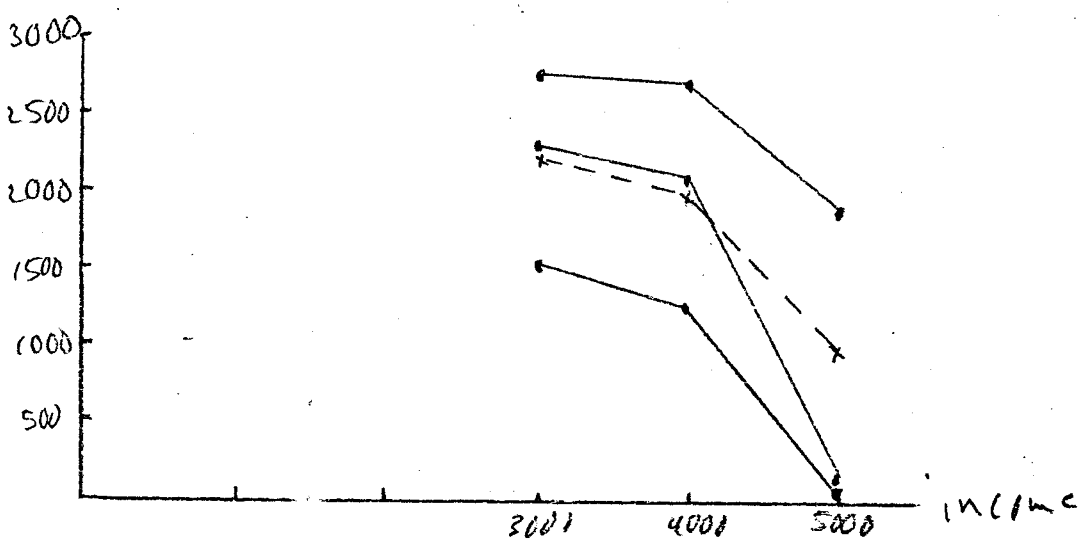


Figure 2 (continued)
Black Husbands





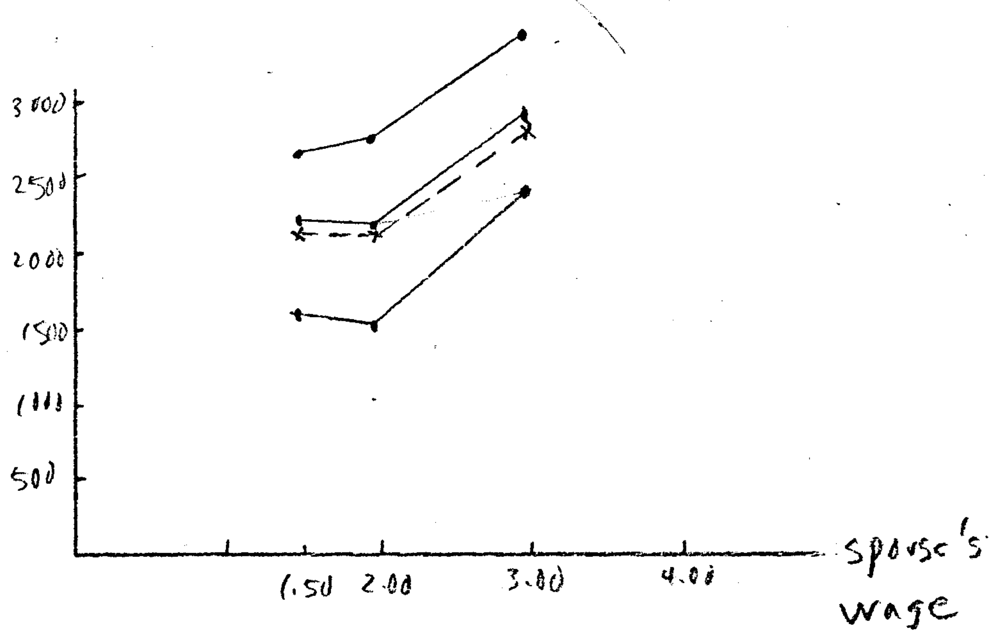


Figure 2 (continued)
White Wives

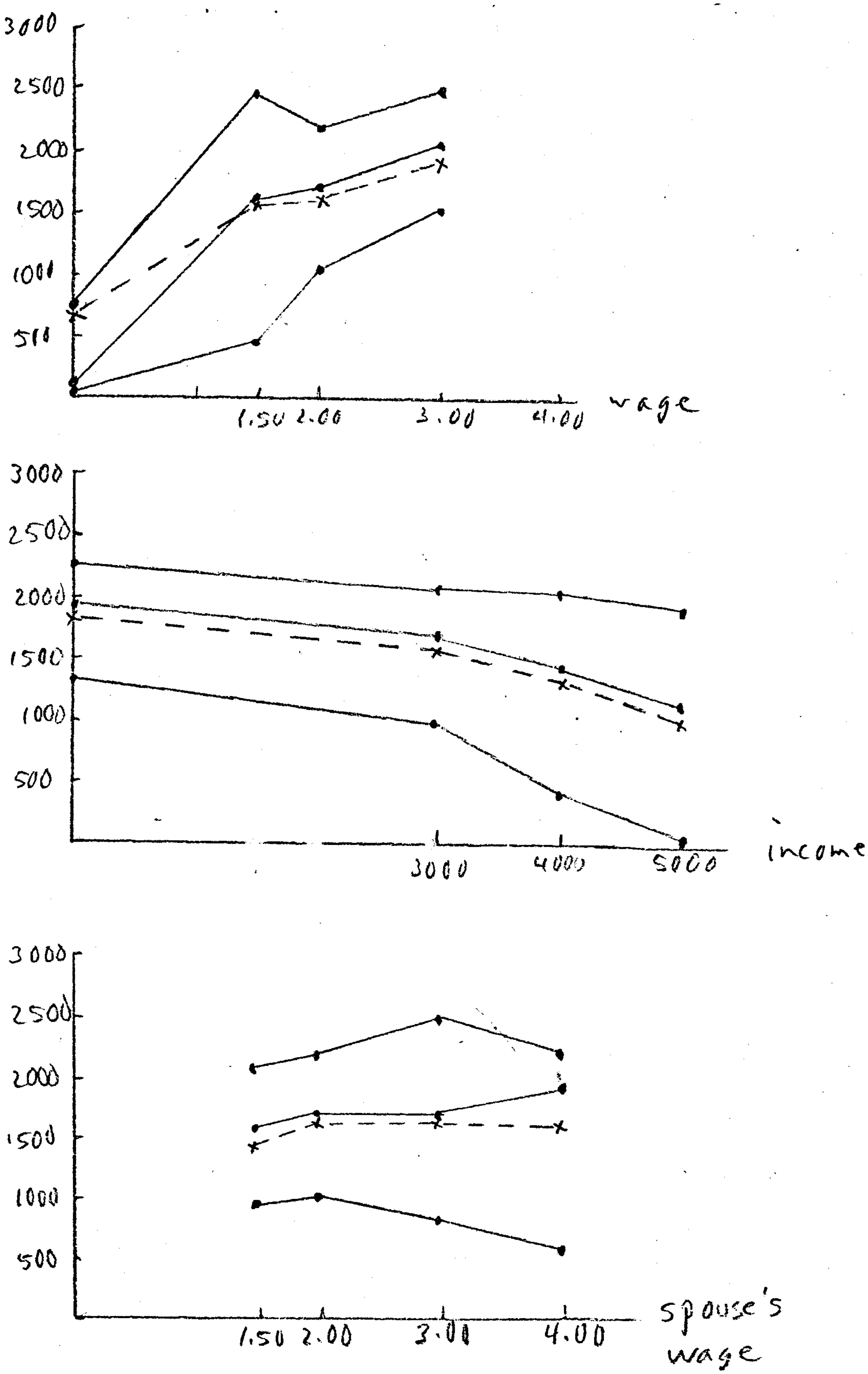
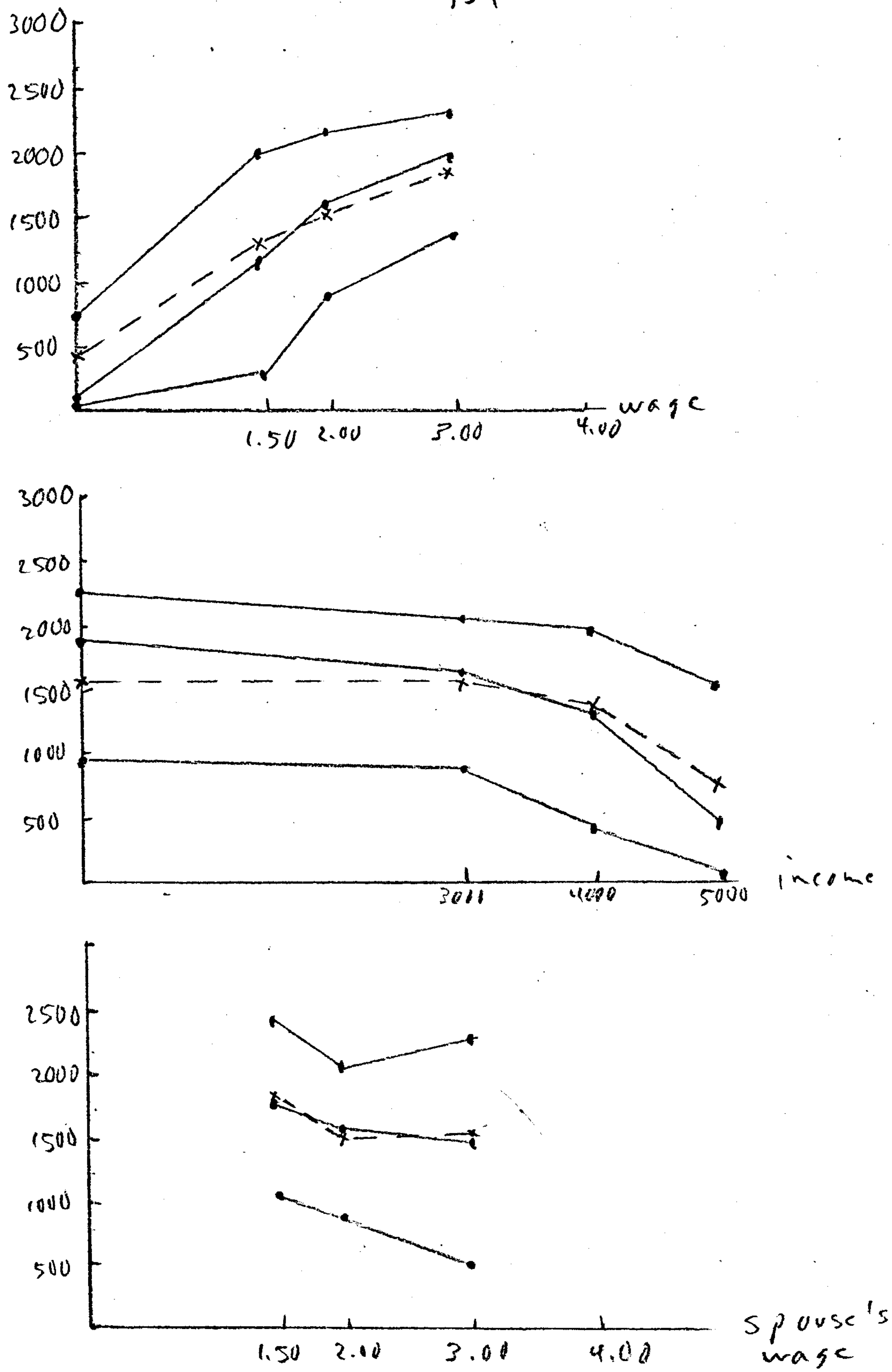


Figure 2 (Continued)
Black wives



Data for Fig. 2

White husbands

Wage	1st q	median	3rd q	mean
1.50	544	1811	2700	1707
2.00	1227	2238	2810	2081
3.00	1522	2121	2581	2027
4.00	1683	2204	2879	2228
Income				
0				
3000	1227	2238	2810	2081
4000	1178	2101	2738	2009
5000	1168	2030	2641	1953
Spouse's wage				
0	1352	1810	2318	1933
1.50	1874	2429	3084	2388
2.00	1227	2238	2810	2081
3.00	2094	2708	3347	2590

Data for Fig. 2

Black husbands

Wage	1st g	median	3rd g	mean
1.50	1364	2214	2863	2143
2.00	1505	2206	2739	2189
3.00	1247	1616	2122	1720
4.00	1098	1693	2437	1722
Income				
3000	1505	2206	2739	2189
4000	1274	1992	2648	1976
5000	48	96	1837	988
Spouse's wage				
1.50	1670	2232	2645	2191
2.00	1505	2206	2739	2189
3.00	2303	2890	3445	2849

Data for Fig. 2

White wives				
wage	1st q	median	3rd q	mean
0	37	74	755	707
1.50	406	1687	2406	1543
2.00	1021	1688	2172	1568
3.00	1395	1957	2321	1873
Income				
0	1339	1930	2261	1849
3000	1021	1688	2172	1568
4000	482	1400	2014	1321
5000	82	1103	1886	1095
Spouse's wage				
1.50	955	1560	2065	1478
2.00	1021	1688	2172	1568
3.00	807	1678	2304	1552
4.00	678	1827	2177	1506

Data for Fig. 2

Black wives

Wage	1st q	Median	3rd q	Mean
0	42	85	723	447
1.50	232	1198	2007	1255
2.00	859	1587	2143	1542
3.00	1365	1898	2215	1800
Income				
0	951	1808	2286	1638
3000	859	1587	2143	1542
4000	465	1300	2004	1334
5000	60	516	1701	888
Spouse's wage				
1.50	1029	1759	2400	1762
2.00	859	1587	2143	1542
3.00	512	1446	2214	1515

Appendix 1.

Approximation Theorem

Consider the discrete probability

$f_i(x)$, $i=1,\ldots, N.$, where x is a vector of length M. Suppose the domain is bounded, $|x| \leq A$ and the derivatives are bounded,

$\left|\frac{\partial f_i(x)}{\partial x}\right| \leq B$, a vector. Then for any $\varepsilon>0$ there is a vector function $g(x)$ and a set of vectors $\beta^{(2)}, \ldots, \beta^{(N)}$ such that

$$\left| \frac{e^{g(x)\beta^{(i)}}}{1+e^{g(x)\beta^{(2)}} + \ldots + e^{g(x)\beta^{(N)}}} \right| \leq \varepsilon$$

for all i, x.

Proof:

Let $N_m = 2\frac{MB_m A_m}{t}$, $m = 1, \ldots. M$,

j_m = greatest integer less than $N_m(\frac{X_m}{A_m} + 1)$

$$j = \sum_{m=1}^{M} \left(\prod_{n=1}^{m-1} N_m \right) j_m$$

and $\bar{x}_m^{(j)} = \frac{j_m A_m}{N_m} - A_m$

Then $\left|x_m - \overline{x}_m^{(j)}\right| \leq \frac{\varepsilon}{M\ B_m}$.

Let $\beta_j^{(i)} = \log f_i(\overline{x}^{(j)}) - \log f_i(\overline{x}^{(j)})$

and $g_k(x) = 1$ if $k = j$

$= 0$ otherwise

Then $$\frac{e^{g(x)\beta^{(i)}}}{1 + e^{g(x)\beta^{(2)}} + \ldots + e^{g(x)\beta^{(N)}}} = f_i(\overline{x}^{(j)})$$

Now $\left|x_m - x_m^{(j)}\right| = \frac{E}{M\ B_m}$, $m = 1, \ldots, M$

So $\left|f_i(x) - f_i(\overline{x}^{(j)})\right| = \varepsilon$ by the Remainder Theorem. QED

Appendix 2

Theorem on Nonlinear Budgets:

Consider a regular budget, B, and the associated demand, x. The prices

$$p_i = \frac{\partial f(x)}{\partial x_i}$$

and the income

$$y = p \cdot x$$

satisfy the following condition:

$$D(\{x' | p.x' - y \leqq 0\}) = D(B)$$

Let $B^* = \{x' | p.x' - y \leqq 0\}$. We need to show that $C = B^* \cap A(x)$ contains only the single point, x. Suppose, on the contrary, that there were another element, x', in C. Then there would be still another element, x", in the interior of C, obeying $x'' < \frac{1}{2}(x + x')$; if not, $\frac{1}{2}(x + x')$ would be on the boundary of A(x), a violation of strict convexity. Then every bundle $\hat{x}$ of the form

$\hat{x} = (1-\theta) x + \theta x''$. $0<\theta<1$, is in the interor of C and thus is strictly prefered to x. Now $f(\hat{x}) = f((1-\theta)x + \theta x'')$ is a differentiable function of θ, so for any $\varepsilon>0$ there is a $\delta>0$ such that

$$\left|f(x) + \theta \sum_{i=1}^{N} (x''_i - x_i)\, f_i(x) - f((1-\theta)x + \theta x'')\right| \leqq \varepsilon\theta$$

for all $\theta \leqq \delta$.

In particular, let ε be the difference between the values of x and x" at the prices p :

$$\varepsilon = p.x - p.x'' \ . \ ;$$

$\varepsilon > 0$ since x" is in the interior of C.

Now $$\varepsilon = - \sum (x''_i - x_i)\, f_i(x)$$

so for $\theta = \delta$,

$$\left|f(x) - E\theta - f((1-\theta)x + \theta x'')\right| \leqq \varepsilon\theta$$

or $$f(x) - f((1-\theta)x + \theta x'') \geqq 0$$

But then $\hat{x} = (1-\theta)x + \theta x''$ is in B and $\hat{x}$ is strictly preferred to x, in which case it is impossible that $x = D(B)$. We conclude that C has only a single element, x.

Appendix 3

Results

White husbands	$\beta^{(1)}$ (0-100)	$\beta^{(2)}$ (100-900)	$\beta^{(3)}$ (900-1800)	$\beta^{(4)}$ (1800-2300)	$\beta^{(5)}$ (2300-2800)
Wage					
0	-115.40 (124.84)	-7.34 (8.22)			
1.50	-1.55 (.84)				
2.00					
3.00					
4.00					
10.00					
Income					
0					
3000					
4000					
5000					
10000					