

COWLES FOUNDATION DISCUSSION PAPER NO. 35

Note: Cowles Foundation Discussion Papers are preliminary materials circulated privately to stimulate private discussion and critical comment. References in publications to Discussion Papers (other than mere acknowledgment by a writer that he has access to such unpublished material) should be cleared with the author to protect the tentative character of these papers.

Working Wives: An Econometric Study

Richard Rosett

June 14, 1957

This discussion paper is a somewhat shortened version of a dissertation presented for the degree of Doctor of Philosophy at Yale University.

Working Wives: An Econometric Study

1. The Secular Trend in the Participation of Women in the Labor Force¹

In the seventeenth Decennial Census Year, 1950, more than eighteen million American women were enumerated as members of the labor force. Of these, more than eight million were married and living with their husbands; five and one-half million of these married women were the mothers of children under eighteen years of age. In the fifty years from 1900 to 1950, while the American labor force tripled in size, the number of women in the labor force quadrupled, and the number of married women in the labor force increased elevenfold. During this period the participation rate² for men over twenty years of age has remained virtually constant both in the aggregate and by age groups. An examination of participation profiles by age for women, however, reveals that the participation rate in every age group moved up steadily until 1940, and that between 1940 and 1950 there was a drastic change in the shape of the profile. This means that

1. The data for this study were made available by the Survey Research Center of the University of Michigan with the permission of the Board of Governors of the Federal Reserve System. I wish to express my gratitude for the opportunity to analyze this body of data. The International Business Machine Corporation made available, through the facilities of the Watson Scientific Computing Laboratory, 175 hours of computing time on the IBM data processing machine. New York University granted an additional eight hours on its own machine. Without the use of the computers and the advice and assistance of the staff of the Watson Laboratory, the statistical analysis undertaken in this study would not have been feasible. The members of my dissertation committee were Professor James Tobin, chairman, Professors William Fellner, and Harold Guthrie. All three, but especially Professor Tobin, contributed generously, in the form of encouragement and advice, to the preparation of this thesis. The responsibility for any errors is mine alone.

2. With respect to aggregate participation, participation rate refers to the proportion of all individuals in the specified category who are gainfully employed or seeking work. With respect to individuals, participation rate means the ratio of the individual's participation to some commonly accepted full time rate.

Table 1
Women Workers, 1870, 1954

Years Aged 10 years and over	Number (in thousands)	<u>Women Workers</u>	
		<u>Percent of all workers</u>	<u>Percent of all women</u>
1870	1,917	15	13
1880	2,647	15	15
1890	4,006	17	17
1900	5,319	18	19
1910	7,445	20	21
1920	8,637	20	21
1930	10,752	22	22
Aged 14 years and over			
1920	8,430	20	23
1930	10,697	22	24
1930 ¹	10,396	22	24
1940 ¹	13,015	24	26
1940 ²	13,840	25	27
1945	19,570	36	37
1947	16,320	28	30
1949	17,167	28	31
1950	18,063	29	32
1951	18,607	30	33
1952	18,798	30	33
1953	18,912	30	32
1954	19,726	31	33

1. Labor force figures for 1930 estimated and for 1940 adjusted by the Census Bureau to make them comparable.

2. Civilian labor force figures for 1940 adjusted by the Census Bureau to make them comparable with later years.

Source: United States Department of Commerce, Bureau of the Census. Data for Decennial Census Years are based on the Censuses. Data for other years are from current population samplings. These data are reproduced here from Bulletin No. 242 of the Women's Bureau of the United States Department of Labor and from Bulletin No. 255.

changes in the participation rate of the population, both in the aggregate and within age groups, are due almost entirely to changes in the participation of women.

Factors which are responsible for changes in the participation rate of women fall logically into three categories: (1) Secular changes in the demand for women workers, and in the attitudes of women and their husbands toward employment of the wife. A detailed discussion of the factors which affect secular trend is beyond the scope of this investigation. (2) The life cycle: such events as marriage, the birth of a child, the child's entry into school, and the death of the husband. (3) The financial position of the family; such factors as changes in the husband's or wife's wage and changes in their assets and debts. This chapter will be devoted to a description of the secular trend and its interaction with the effect of the life cycle.

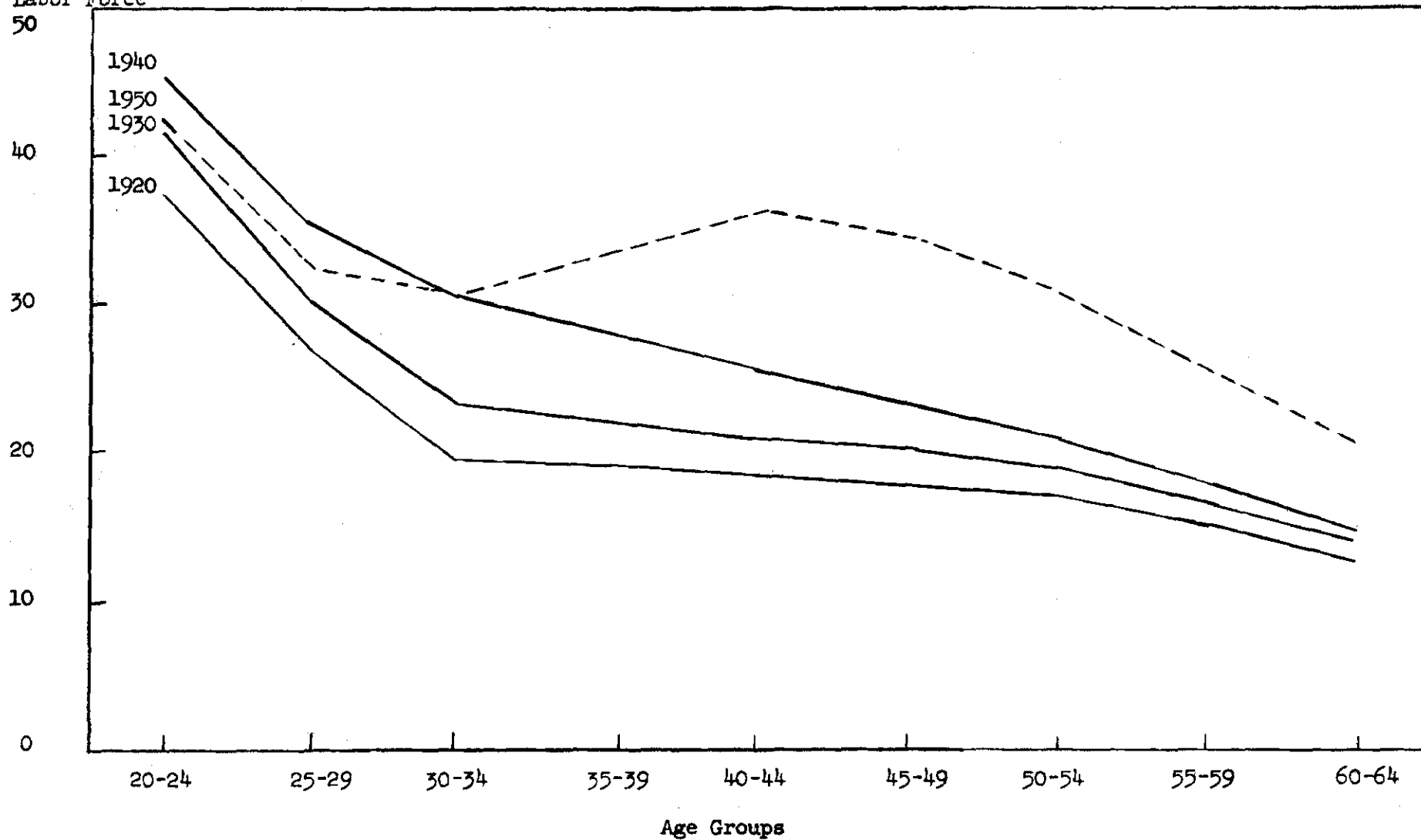
Although the aggregate participation rate of women has been moving steadily upward and shows no sign of leveling off, there is some reason to believe that it will soon become stable and that the striking characteristics of the 1950 profile by age groups will persist. Figure 1 shows profiles for the years 1920, 1930, 1940, and 1950. The data represented in these charts have been adjusted to make them comparable with the 1940 census. Figure 2 compares actual and predicted profiles for 1950 and 1955. The data in Figure 2 are comparable with current population reports. In order to adjust for differences in definition, it is necessary to inflate 1940 census data to make them comparable with current population reports. The predictions shown in Figure 2 were made by John D. Durand.¹

1. John D. Durand, The Labor Force in the United States, 1890-1960, (New York: Social Science Research Council, 1948). Projections were also made by the Bureau of the Census through 1975. Durand did separate analyses of native whites, native non-whites, and non-natives for both men and women. He predicted future participation profiles for each of these quite different groups, made an attempt to account for the effects of World War II, and combined his separate predictions to obtain his aggregate predictions. The Census projections were based on more simple-minded linear extrapolations.

Figure 1 - Participation Rates of Women by Age:

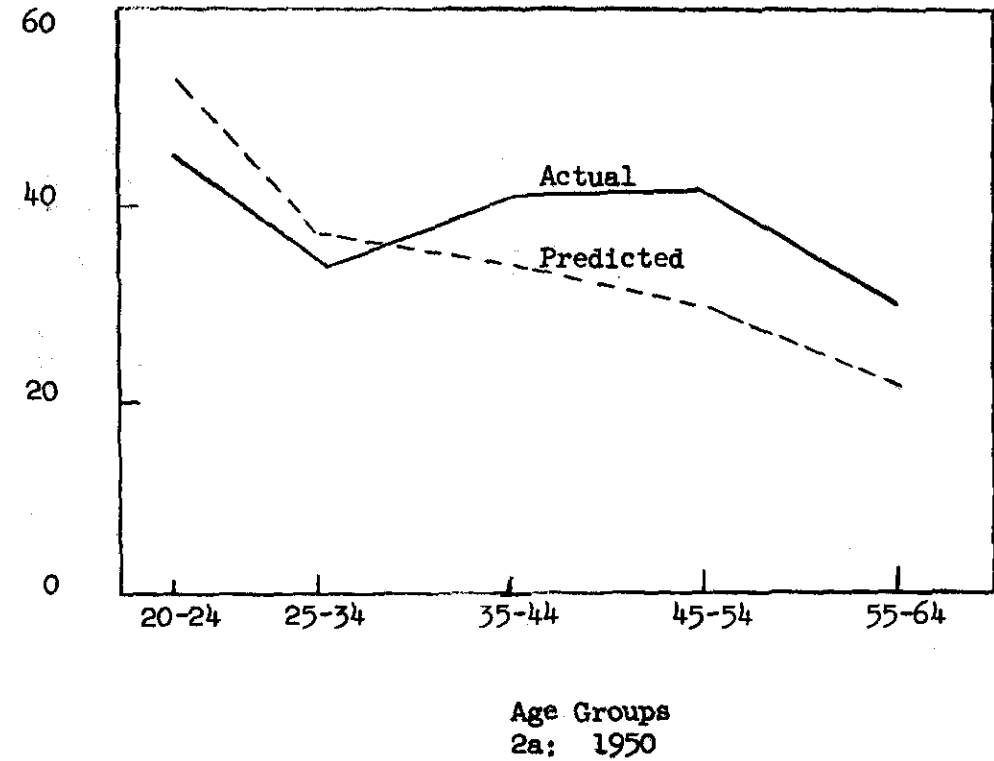
1920 - 1950

Percentage of
Women in the
Labor Force

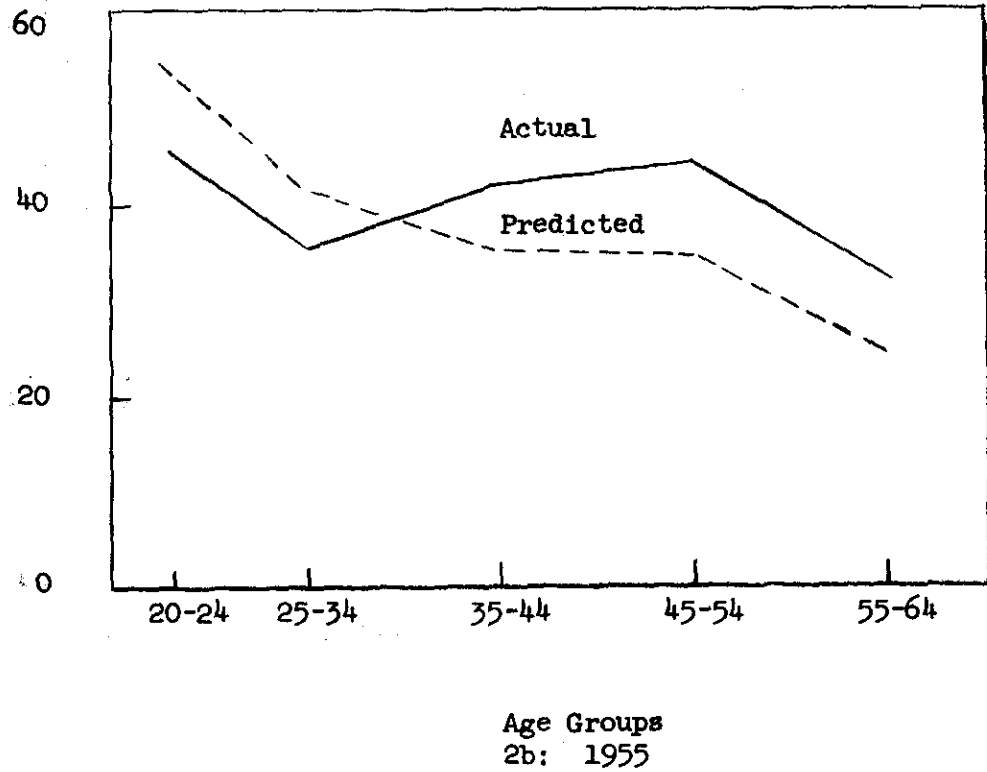


Source: 1920-1940 John D. Durand, *Ibid.*, p. 208 and 1950 Census of Population, Special Report P-E, No. 1-A, Employment and Personal Characteristics, p. 16.

Figure 2 - Actual and Predicted
 Participation Rates for Women
 by Age: 1950 - 1955



Percentage of Women in the Labor Force



Durand made projections on the basis of 1890-1945 data and predicted the participation profiles, by age, for 1950, 1955, and 1960. The errors of prediction made by Durand may themselves cast some light on what happened to change the shape of the profile. Most important is the fact that he overestimated the participation rate among 20-24 year olds by about eight percent. He also overestimated the increase in the 25-34 year old age group. In general, Durand seems to have expected the participation rate in all age groups to continue to increase much as it had in the past.

An explanation of the change in the shape of the profile which is consistent with the nature of Durand's errors is as follows:

(1) The probability that a woman will enter the labor force at any time during her life is determined by whether or not she has been in the labor force before that time. It is higher if she has had a job before.

(2) The pattern of employment for a woman is closely related to her life cycle. If she is working when she marries, she will leave the labor force as soon as her first child is born. As her youngest child grows older, the probability that she will enter or re-enter the labor force increases.

(3) Increasing employment opportunities for women¹ and changes in customs and mores have been responsible for a secular upward trend in the participation rates of all women.

1. For a full discussion of the nature and extent of the change in employment opportunities for women, see National Manpower Council, Womanpower, (New York: Columbia University Press, 1957). An analysis of the change in employment opportunities for women can be divided into two parts: changes in wage rates and changes in the variety of jobs open to women. On the question of changes in wage rates, see W. S. Woytinsky and associates, Employment and Wages in the United States, (New York: Twentieth Century Fund, 1953), p. 505. For data on changes in the variety of women's occupations, see Changes in Women's Occupations.

(4) The secular trend in the participation rates of women in the labor force is leveling off. We have come through a period of transition and are entering a new period of stability in the pattern of employment of women.

If the first three hypotheses are correct, it is possible to infer the fourth from an examination of the data represented in Figure 1. For the purpose of this discussion, define a cohort as a group of individuals who fall within the same five year age group in a decennial year. A cohort will be designated by the year in which it was 20-24 years old. Thus, the 1920 cohort was 20-24 years old in 1920, and in 1950 it was 50-54. A census profile covering the ages 20-65 consists of nine cohorts. A cohort profile of participation would represent the participation rates of a single cohort as it grew older.

If the proposed explanation were correct, one would expect a cohort profile to look something like the 1950 census (cross-section) profile. Women would leave the labor force as they have children, and return as the children grow up. They would then work until age or other circumstances make them retire. Figure 3 shows the participation profiles for a set of cohorts. These are incomplete because of the absence of inter-census data and the lack of adequate data before

Footnote 1, page 6, continued

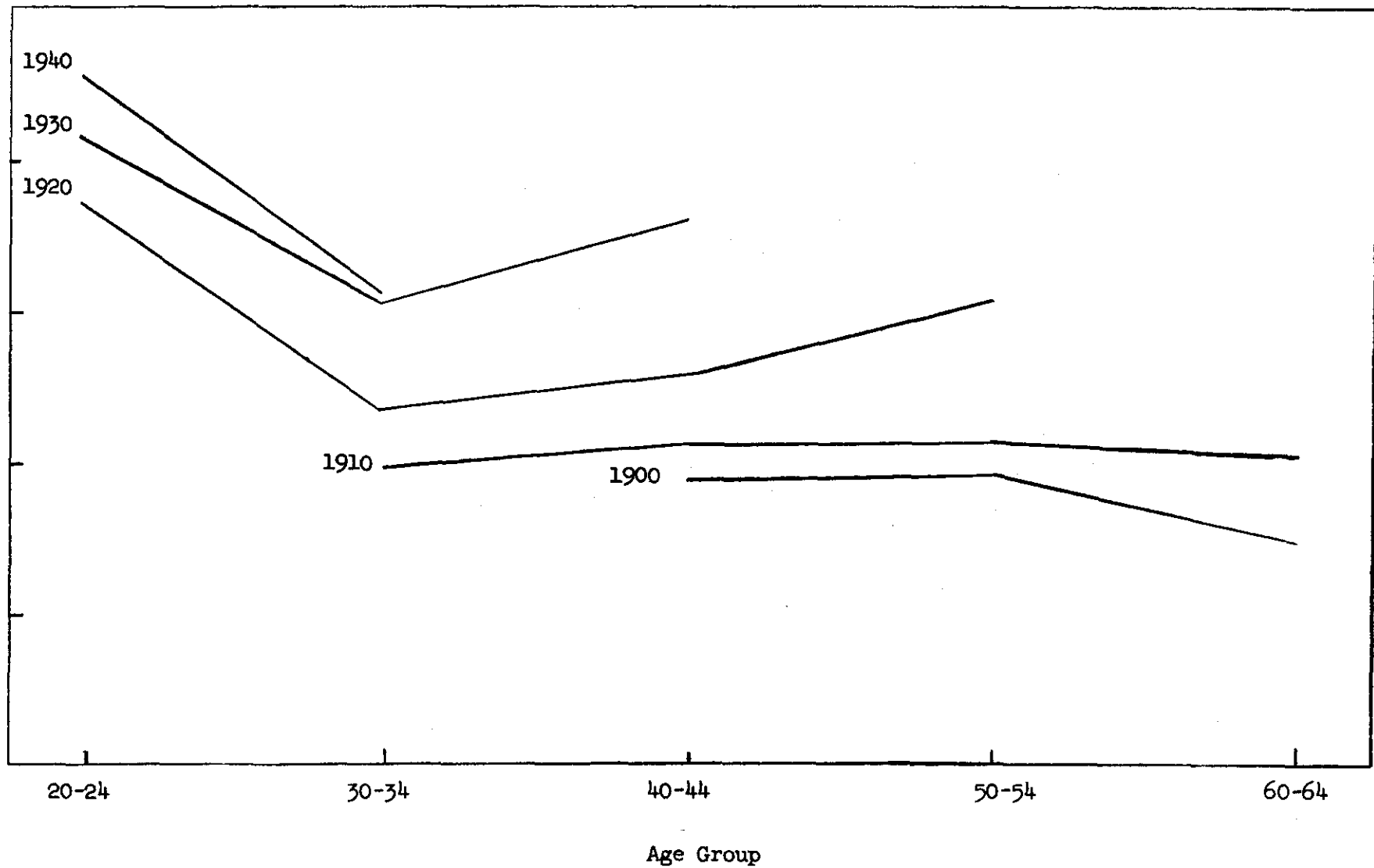
1940-1950, (Woman's Bureau Bulletin No. 253; Washington: United States Department of Labor, 1953). In spite of the title, this bulletin contains data covering the period 1900-1950. For a more detailed discussion of the period 1900-1930, see S. P. Breckinridge, Women in the Twentieth Century, (New York: McGraw-Hill Book Publishing Company, 1933).

All of these books agree on two main points: (1) The median wage rate for women has been, with minor fluctuations, about one-half the median wage rate for men since 1900. (2) While there are many instances of women moving into occupations formerly thought to be exclusively men's occupations, most of the increase in the employment of women has been due to the rapid increase in the relative importance of jobs traditionally held by women. This may partly explain the failure of women's wages to rise relative to men's wages.

1920. Incomplete as they are, they do suggest that the cohort profiles actually reflect the hypothesized life cycle pattern. On the basis of the hypothesized explanation, one would expect that if one cohort participates at a higher rate than another in any age group, it will participate at a higher rate in every age group. This seems to be the case.

If the participation rate among 20-24 year olds is moving up rapidly enough, the shape of the cohort profile will not show up in a census profile. Young cohorts will participate at higher rates than older cohorts, and although women with growing children will be re-entering the labor force, they will not bring their participation rate up to that of any younger cohort, even that of the cohort which is at its own minimum participation rate because of child-bearing. On the other hand, if the participation rate among 20-24 year olds were to become stable, after five years the two youngest cohorts would be similar in cohort participation profile. After ten years, the three youngest cohorts would be similar in this respect, and at the end of forty-five years, a cross section would consist of nine cohorts, all similar in cohort participation profile. In other words, if the cohort profiles remain stable long enough, a cross section will look exactly like a cohort profile. It is interesting to note that in Figure 2, the actual cross-section profiles for 1950 and 1955 are almost identical from the age of 20 to the age of 44. In Figure 1, it will be noted, the participation rate of 20-24 year olds in 1950 was only slightly higher than it was in 1930. The rate for 1940 was higher than in 1930 or 1950, but this might be attributed to the institution of the military draft and the unusual world situation in that year. Taken together and considered within the framework of the hypotheses under discussion, these facts could suggest that the cohort participation profile has been almost constant for more than twenty years. It might be expected that the 1960 census will show little or no increase in the participation rate among

Figure 3 - Cohort Participation Profiles for Women



Source: See source for Figure 1.

younger women. To this argument it is necessary to make two important qualifications. First, if the hypothesized explanation for the change in the shape of the cross section profile is correct, the cohort profiles themselves are influenced by the secular trend. It seems reasonable that if the secular trend were removed, the cohort profile would drop more sharply during child bearing, and rise less steeply during child-rearing. Thus, if the level of the cohort profiles has become stable, cohort profiles (and eventually the cross section profile) will be shaped somewhat differently from the cohort profiles of the past. Second, some of the bulge in the 1950 cross section profile may be due to the effects of World War II. Surely many women who worked for the first time during the war stayed in the labor force permanently, or at least found it easier, thereafter, to re-enter the labor force because of their work experience. This is not enough, however, to explain the unusually small increases in the young age groups.

In the discussion of cohort profiles, it was assumed that the life cycle of a working woman is typically that of a woman who marries and has children. In 1900, sixty-seven percent of the women in the labor force had never been married. This was partly due to the fact that participation rates are highest for every cohort just before marriage, and in 1900 thirty-four percent of the women of working age had never been married. The average age of marriage was higher than it is now. In 1950 only twenty-nine percent of the women in the labor force had never been married, and most of these were young women. In other words, the cohort profile is certainly dominated by the life cycle pattern of married women, and even in 1900, despite the apparent preponderance of single women in the female labor force, married women probably constituted a majority among older working women.¹

1. See Woman's Bureau Bulletin No. 246, Table 3, p. 9.

The relationship between cross-section data and time series data is especially relevant to a study, like this one, in which an attempt is made to infer, from a cross-section, something about the lifetime behavior of individuals. The data used in this study were drawn from the 1954 Survey of Consumer Finances. Even the first casual glance at the data revealed that the most important factors which influenced the decision of the wife to work had to do with the life cycle of the family: the presence or absence of children, their ages (especially the age of the youngest), and the number of years married. All of these things are a function of time. Even the incomes of members of the family and their debt and liquid assets are related to the life cycle and hence to time. A cross section tells us something about a number of cohorts at a point in time. If there are important differences between cohorts, it is impossible to infer anything about the effect of time on a cohort. It would have been incorrect to infer, on the basis of a 1940 cross section, that women become less likely to work as they grow older. The statement that older women worked less than younger women in 1940 is a description of the state of affairs in 1940, but it is not a correct prediction of the behavior of women as they grow older. To the extent that such factors as income, debt, and liquid assets are related to the life cycle, inferences about their effects on the decision of the wife to work could be faulty. The latter difficulty can be remedied somewhat by the use of age as an independent variable.

If the cohort profile is shifting up rapidly, an analysis of a cross section could be almost worthless to someone interested in forecasting the future participation rates of women in the labor force. Such an analysis would reveal nothing about either of the two most important components of change: the secular shift of the cohort profiles, and the shape of the profiles themselves. In such

a case it would be necessary to look at a number of cross sections in order to examine the secular change in the 20-24 year age group, and to follow a number of cohorts through at least parts of their life cycles. On the other hand, if the cohort profile has been reasonably stable for a long time, the need to analyze secular change is obviated, and a single cross section is sufficient for an analysis of the cohort profile.

To summarize: This is to be a study of the factors which affect the participation rates of women in the labor force. Since the study is based entirely on a single cross-section, its predictive value depends entirely on the accuracy with which the cross-section reflects the cohort profile. If the cross-section was taken at a point in time when the cohort profile was changing rapidly, the value of the study is substantially lower than if the cross section was taken at a point in time when the cohort profile was stable. There is some reason to believe that a cross section taken in 1953 does accurately reflect the shape of the cohort profile for women between the ages of 20 and 44. The case, however, is far from solid, and it may very well be that what follows is nothing more than a description of the state of the world in 1953.

2. The Model of Consumer Choice and the Determination of the Wife's Participation Rate

The hypotheses which will be tested in this study fit easily into a simple model of consumer choice. This section will be concerned with a description of the model, and of the extent to which hypotheses were modified to conform to the limits imposed by the adequacy of the data.

Assume that the basic decision-making unit in the consuming sector of the economy consists of a household of individuals who pool their resources and make joint decisions concerning expenditures. Although the actual decision process within the household may be complicated, the household may be assumed to have the usual sort of indifference curves¹ and to respond to changes in the state of the world as if it were, indeed, an individual. Since this is to be an examination of the factors that influence the decision of the wife to work, the discussion which follows will deal only with spending units which consist of at least a husband and wife.

1. For a discussion of whether or not it is possible or reasonable to think of a family or spending unit as having the usual sort of indifference curves, see P. A. Samuelson, "Social Indifference Curves," Quarterly Journal of Economics, LXX (February, 1956), 1-22. See especially the theorem on page 16. The important assumptions which must be made if such indifference curves can be shown to exist are: (1) "as you take equal amounts of a good away from one [member of the spending unit], you must give increasing amounts to another if [spending unit] welfare is to remain constant," and (2) within a spending unit there is "optimal reallocation of income so as to keep each member's dollar expenditure of equal ethical worth..."

For such a spending unit

$$2.1 \quad \begin{aligned} U &= U(Y, P) \\ Y &= H + R + WP \end{aligned}$$

where U is a monotonic increasing function of the spending unit's preference ordering, everywhere differentiable, and having the proper convexity.

Y = total family income

P = wife's participation rate¹

H = husband's full time wage rate

R = property income

W = wife's full time wage rate

It is assumed that the husband's participation rate is equal to one. Given H , R , and W , the household chooses P , and thus Y , so as to maximize U . The alternative to participation will be called leisure, although it is recognized (and is in fact, an essential part of this analysis) that the alternative to participation is most often service in the home.

If an increase in participation will decrease the marginal utility of income and increase the marginal disutility of participation (which is the same as increasing the marginal utility of leisure) then income

1. One unit of participation is defined as a commonly accepted full-time rate (i.e., an eight hour day, five day week, fifty weeks a year).

is a normal good¹ and it can be shown that at a regular constrained maximum

$$2.2 \quad \frac{\partial Y}{\partial H} > 0$$

$$2.3 \quad \frac{\partial Y}{\partial W} > 0$$

If, in addition, an increase in income will increase the marginal disutility of participation and decrease the marginal utility of income, then leisure is a normal good¹ and

$$2.4 \quad \frac{\partial P}{\partial H} < 0$$

$$2.5 \quad \frac{\partial P}{\partial W} > \frac{\partial P}{\partial H}$$

Figure 4 represents an indifference map which has these properties.

1. The conditions stated are sufficient, but not necessary. The necessary and sufficient condition for income to be a normal good is

$$\frac{\frac{U_{yp}}{U_y}}{\frac{U_{pp}}{U_p}} < 1$$

and the necessary and sufficient condition for leisure to be a normal good is

$$\frac{\frac{U_{yy}}{U_y}}{\frac{U_{yp}}{U_y}} < 1$$

As stated in the text, these conditions each have a zero to the right of the inequality sign. By definition of a normal good, income is a normal good if (2) holds, and leisure is a normal good if (4) holds. For an analysis of the cases in which either income or leisure is an inferior good, see Appendix A.

Figure 4 - An Indifference Map
 Combinations of Income and
 Participation

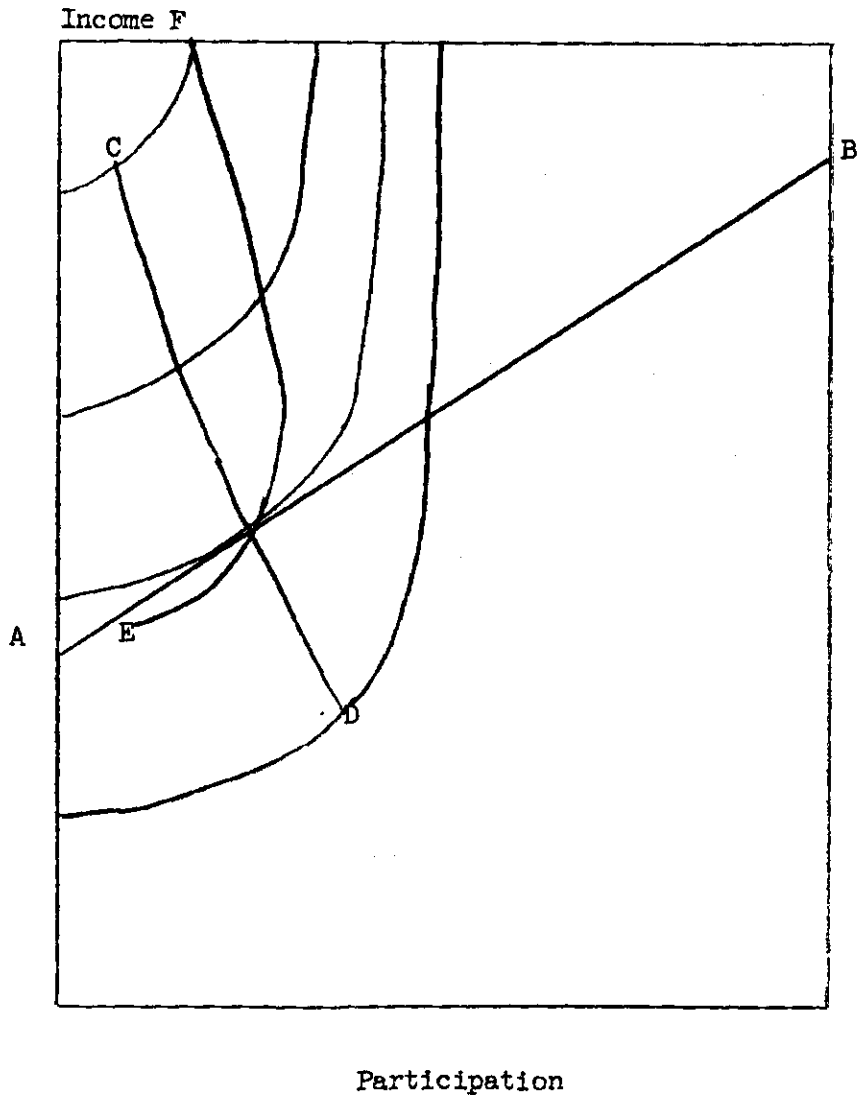
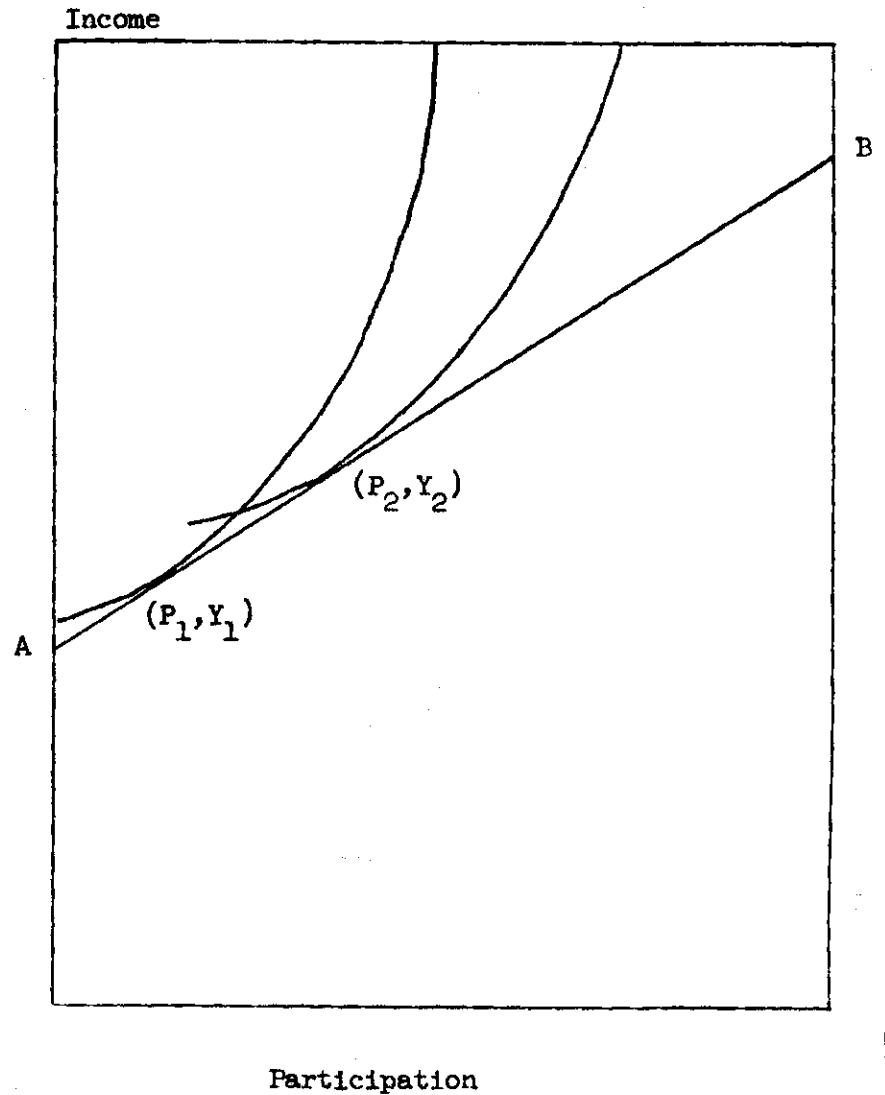


Figure 5 - Effect on P and Y of a
 Shift in Indifference Curves



The line AB represents the budget constraint. CD is the locus of maxima for different values of H and a fixed W. EF is the locus of maxima for a fixed value of H and a variable W.

Four of the hypotheses to be tested are (2.2), (2.3), (2.4), and (2.5). While it would be possible to construct a model explicitly including all of the factors suggested in Section 1 as influencing the decision of the wife to work, it is easier to treat most of them as affecting the slopes of the indifference curves.

The slope of an indifference curve is:

$$2.6 \quad r(Y,P) = - \frac{U_p}{U_y}$$

Assume that two spending units have utility functions U and V and that

$$2.7 \quad r(Y,P) = - \frac{U_p}{U_y}$$

$$s(Y,P) = - \frac{V_p}{V_y}$$

If for every point (Y,P)

$$2.8 \quad r(Y,P) > s(Y,P)$$

then given the same budget constraint the two spending units will maximize utility at points (Y_1, P_1) and (Y_2, P_2) respectively and

$$2.9 \quad Y_1 < Y_2$$

$$P_1 < P_2$$

This can be shown as follows: Assume that U is at a regular constrained maximum at (Y_1, P_1) then from the first order conditions for a maximum

$$2.10 \quad r(Y_1, P_1) = W$$

but from (2.8)

$$2.11 \quad s(Y_1, P_1) < W$$

As P is increased along the budget line,

$$\begin{aligned} 2.12 \quad \frac{dV}{dP} &= WV_y + V_p \\ &= WV_y + s V_y \\ &= (W-s)V_y \end{aligned}$$

Since $s(Y_1, P_1) < W$, at the point (Y_1, P_1)

$$2.13 \quad \frac{dV}{dP} > 0$$

Thus, at the point (Y_1, P_1) , V is increasing with P . If V is a well-behaved utility function, there is some point (Y_2, P_2) on the budget line which will maximize utility and

$$2.14 \quad P_1 < P_2$$

The proof that $Y_1 < Y_2$ follows from the fact that $W > 0$.

Figure 5 is a graphic representation of this proof.

Thus anything which increases (or decreases) the slopes of the indifference curves at every point will unambiguously decrease (or increase)

the participation rate, given a budget constraint. It is now possible to enlarge the set of hypotheses concerning the financial position of the family.

If every spending unit has some desired (positive) net assets position which it attempts to achieve partly through control of its income flow, it should be expected that, at any point on the (Y,P) plane, r will be greater if L (liquid assets) is greater, and r will be smaller if D (debt) is greater. This would follow if the assets position has no effect on the marginal disutility of participation, but does have an effect on the marginal utility of income. Since a spending unit will save or liquidate debt only if doing so will increase U_y , it is hypothesized that

$$2.15 \quad \frac{\partial P}{\partial L} < 0$$

$$2.16 \quad \frac{\partial P}{\partial D} > 0$$

Since the effect of debt is primarily a consequence of the fact that debt represents a claim against current income, it seems reasonable to distinguish between personal and mortgage debt.¹ A spending unit with no mortgage debt is paying rent which is, in some respects, equivalent to paying off a mortgage. There is no such counterpart to the claim against current income represented by personal debt. If the two kinds of debt are

1. It seems even more reasonable to distinguish between the two kinds of debt in the statistical tests reported below because the debt position reported in the survey is the ex post debt position. What was wanted was the ex ante debt position. If wives work to pay off debt, this difference could lead to an identification problem. The problem of identification is less likely to be serious in the case of mortgage debt than in the case of personal debt.

distinguished from one another it is expected that when D_p and D_m represent claims against current income of the same magnitude

$$2.17 \quad \frac{\partial P}{\partial D_p} > \frac{\partial P}{\partial D_m} > 0$$

Similar reasoning can be used to derive hypotheses about the effects of the characteristic events of the life cycle. The effects of marriage itself and of the death of the husband will be ignored. The effect of the birth of a child is somewhat ambiguous because it can be expected to affect both U_y and U_p . Nothing need be said about the increase in the marginal utility of income which results from the presence of a child in a spending unit. A child will increase the marginal disutility of participation for two reasons. First, caring for one's own children has high ethical value in our culture. Second, if a mother does work, it is necessary to pay someone to care for the child. It seems highly likely that the increase in the marginal disutility of participation will be greater than the increase in the marginal utility of income. As the child grows older, it becomes more and more acceptable to entrust it to the care of someone other than the mother. This is the same as saying that the ethical contribution to $-U_p$ decreases. It is not so easy to see the consequences of increasing the number of children. While hiring someone to care for one child may be bad, it is not much worse to hire someone to care for several children (although it may be more expensive). It seems likely, however, that the mother of several children, if she works, will actually be required to do much of the housework herself even if she does hire someone to care for her children while she is away. To the extent that this is true, the marginal disutility of participation will increase as the number of children increases. Since the marginal

disutility of participation increases very fast as participation approaches some physiological limit, and since the mother of several children is likely to spend a substantial part of her apparent leisure doing housework, it is hypothesized that increasing the number of children will increase r and therefore decrease P .

A problem closely related to the effect of the number of children, is the effect of the number of adults. In almost all cases where there are more than two adults in a spending unit, the extra adults are parents of either the husband or wife. If these are retired parents, their presence contributes to the marginal utility of income, and decreases the marginal disutility of participation. A grandparent is a good substitute for the mother in caring for children, and even if there are no children, the presence of an extra adult can be expected to reduce the burden of housework on the wife. Thus, increasing the number of adults in a spending unit will increase the participation of the wife.

To summarize: The set of hypotheses concerning the financial and demographic factors which affect the participation rate can be expressed as follows if the effects are taken to be additive and linear:

$$2.18 \quad P = a_0 + a_1W + a_2H + a_3D_p + a_4D_m + a_5L + a_6C + a_7A \\ + a_8N_c + a_9N_a$$

P = wife's participation rate

W = wife's wage rate

H = husband's wage rate

D_p = personal debt

D_m = mortgage debt

L = liquid assets

2.18 (continued)

$C =$ (0 if there is no child under 6 years of age
 (1 if there is a child under six years of age

$A =$ age of the youngest child if he is under six
 years of age

$N_c =$ number of children

$N_a =$ number of adults

2.19

$$a_1 > a_2$$

$$a_2 < 0$$

$$a_3 > a_4 > 0$$

$$a_5 < 0$$

$$a_6 < 0$$

$$a_7 > 0$$

$$a_8 < 0$$

$$a_9 > 0$$

$$a_2 + 1 > 0 \quad (\text{this is equivalent to } \sqrt{2.27})$$

and for W in the relevant range,¹ the condition equivalent to (2.3) is

$$a_0 + 2Wa_1 + \sum_{i=2}^9 a_i X_i > 0$$

Modifications of the model

The most serious inadequacy of the 1954 Survey of Consumer Finances for

1. W in relevant range means that W must be great enough so that the wife is actually participating.

the purpose of testing the hypotheses proposed above, is the lack of any direct information about P . The data contain no information about the wife at all except her total wages, salaries, and professional income for the year 1953. In the definition of the model, this figure represents WP . It was clear that the data could be used only if it were possible to construct some sort of estimate of W . It was decided that the best estimate of W which could be obtained from the data was H . The dependent variable in (2.18) was replaced with $\frac{WP}{H}$. If men usually marry women with talents, ability and intelligence similar to their own, there should be some positive relationship between the wage rates of men and their wives. In addition to this, husbands and wives are more or less uniformly affected by regional differences in wage rates. A more complete discussion of this assumption will be reserved for Section 5.

If it is true that in a cross-section W is a function of H , then the sign of $\frac{\partial P}{\partial H}$ for a cross-section is no longer unambiguous. Also, since no direct information about W is included in the data, it is difficult to test the hypotheses about the effects of W on P and Y .

Nothing has been said until now about the role of R (property income) in the budget equation of the model. Since H and R play an identical role -- they simply determine the Y -intercept of the budget line -- and since the relationship between R and W is probably much weaker than the relationship between H and W , it is possible to salvage the hypotheses simply by: (1) replacing hypotheses about H with hypotheses about R , and (2) replacing hypotheses about W with hypotheses about the effects of changes in H given that $H + R$ remains constant.

While the introduction of R and the consequent changes in the set of hypotheses make it possible (it is hoped) to circumvent the lack of direct information about P and W , it is important to note that a subtle change has occurred in the interpretation of the hypotheses. The admission that W is a function of H is fundamentally a denial of the ceteris paribus assumption which must underlie the use of cross-section data. If W is a function of H for the reasons cited above, then it is clear that a change in H represents not a ceteris paribus change in wage rates, but a change in wage rates because of a change in the talents and abilities of the wage earners, and therefore a change in the labor markets in which they offer their services. In other words, the participation of a woman whose husband can earn \$10,000 a year is different from the participation of a woman whose husband can earn \$5,000 a year. It is not just the wage rate that is different. The job is different, her intelligence is different, and many of her experiences have been different. In order to correct for this difficulty as much as possible, it was decided to include some variable which would at least indicate the differences in education between spending units. Since the sample includes no information about the education of the wife, the variable chosen was the education of the husband. Since more education makes more interesting work available to the wife, it was hypothesized that education would reduce the marginal disutility of participation. If education has any effect on the marginal utility of income, it seems reasonable to believe that U_y increases as education increases. Thus, increasing education will decrease r and increase P .

Finally, it was decided to add a set of variables which would serve a double purpose: (1) account for changes in the expenditure pattern as a

spending unit grows older, and (2) remove, to some extent, the differences in cohort participation levels. Young spending units are likely to be making initial investments in durable goods and saving toward the down payment on a house. At the same time, the income of the husband is likely to be low relative to what it will be later in the life cycle. If the wife's participation in the labor force is used to reduce the differences in lifetime income and expenditure patterns of the spending unit, it is to be expected, after accounting for the effects of children, that wives in older spending units will participate at a lower rate than wives in younger spending units.

Summary: The Hypotheses to be Tested

The relationship which was actually estimated in this study is:

$$2.20 \quad \frac{FW}{H} = b_0 + b_1H + b_2R + b_3D_p + b_4D_m + b_5L + b_6C \\ + b_7A + b_8N_c + b_9N_a + b_{10}E_1 + b_{11}E_2 \\ + b_{12}E_3 + b_{13}M_1 + b_{14}M_2 + b_{15}M_3 + b_{16}M_4$$

and

$H, R, D_p, D_m, L, C, A, N_c,$ and $N_a,$ have already been defined.

$$(E_1, E_2, E_3) = \begin{array}{l} (0,0,0) = \text{no education} \\ (1,0,0) = \text{1-8 years of education} \\ (0,1,0) = \text{9-12 years of education} \\ (0,0,1) = \text{13 or more years of education} \end{array}$$

M_1 = number of years married if four or less.

$$(M_2, M_3, M_4) = \begin{array}{l} (0,0,0) = \text{married 1-4 years} \\ (1,0,0) = \text{married 5-9 years} \\ (0,1,0) = \text{married 10-19 years} \\ (0,0,1) = \text{married 20 or more years} \end{array}$$

The hypotheses to be tested are as follows:

$$\begin{aligned}
 & b_1 - b_2 > 0 \\
 & b_2 < 0 \\
 & b_3 \quad b_4 > 0 \\
 & b_5 < 0 \\
 & b_6 < 0 \\
 2.21 \quad & b_7 > 0 \\
 & b_8 < 0 \\
 & b_9 > 0 \\
 & 0 < b_{10} < b_{11} < b_{12} \\
 & b_{13} < 0 \\
 & b_{14} < b_{15} < b_{16} < 0 \\
 & b_2 + 1 > 0
 \end{aligned}$$

(equivalent to [2.2])

and on the assumption that $W = c + dH$, and letting $H + R = K$, for W in the relevant range

$$\frac{1}{d}(b_0 + 2 [b_1 - b_2] \left[\frac{W - c}{d} \right] + b_2 K \sum_{i=3}^{16} b_i X_i) > 0$$

(equivalent to [2.3])

3. The Data Employed in this Study

The data used in this study were drawn from the 1954 Survey of Consumer Finances.¹ The survey is conducted annually by the Survey Research Center of the University of Michigan at the direction of the Board of Governors of the Federal Reserve System. The 1954 sample, which was drawn in January, February, and early March, consisted of observations of 3000 spending units, of which 1592 were selected for this study. Spending units were dropped principally because they fell into one or more of the following categories: the head was a farmer, retired, or unemployed; the spending unit did not consist of at least a husband and wife; or one or more of the variables in the relationship was not ascertained.

The definitions of debt, liquid assets, and income employed by the Survey Research Center were satisfactory for the purposes of this study, but it should be kept in mind that the data were inadequate in two respects. Direct information about the wife's participation rate was missing and debt and liquid assets were ex post rather than ex ante.

It was decided that since there is no reason to believe that the weights supplied by the Survey Research Center will reduce heteroscedasticity in the relationship being examined, it would be satisfactory to use unweighted observations.

1. For a discussion of the methods employed in conducting the Surveys of Consumer Finances, see "Methods of the Survey of Consumer Finances," Federal Reserve Bulletin XXXVI, 7 (July, 1950), 795-809. For a discussion of the definitions used in the Surveys, see J. B. Lansing, "Concepts Used in Surveys," Contributions of Survey Methods to Economics, ed. by L. R. Klein, (New York 1954: Columbia University Press), pp. 9-48.

4. The Statistical Model

The statistical model employed in this investigation is of interest because it is applicable whenever a dependent variable is limited by an upper or lower bound, or when the relationship between the dependent and independent variables is nil in some range of the independent variable. It is the very nature of economics that it involve relationships having these characteristics. Saving cannot exceed income, dissaving is limited by the assets and credit resources of the firm or spending unit, purchases by consumers cannot be negative, and participation in the labor force cannot be negative or more than the flesh can bear. Decisions involving transactions may be insensitive to small changes in the state of the world because of transactions costs.

Non-negativity of participation simply means, within the framework of the model presented in Section 2, that the constrained maximum attained by the spending unit might be a corner rather than a regular maximum. The participation rate will be zero whenever $r(H+R,0) \geq W$. For two-thirds of the spending units in the sample, the participation rate is zero.

Although there is no sense in which P can be negative, it is clear that if two women are not in the labor force, one may be closer to entering the labor force than the other. One woman might go to work if her wage rate increased slightly; the other might be disinclined to work even if her wage rate increased substantially. It is possible to construct an index of the disinclination to work simply by re-interpreting the relationship

$$4.1 \quad \frac{WP}{H} = b_0 + \sum_{i=1}^{16} b_i X_i$$

Instead of interpreting the linear combination of independent variables as determining the wife's participation rate, let $b_0 + \sum_i b_i X_i$ be thought of as an index of the wife's desire to work. When the index is positive or zero, its numerical value is equal to the wife's participation rate; when the index is negative, it is a measure of the wife's disinclination to participate in the labor force. If, for given values of X_i , deviations (u) from the index (I) due to factors which have been excluded from the index are random and normally distributed with zero mean and standard deviation σ , the participation of the wife is determined as follows:

$$I = b_0 + \sum_{i=1}^{16} b_i X_i + u$$

$$4.2 \quad \frac{WP}{H} = 0 \quad (I - u < 0)$$

$$\frac{WP}{H} = I - u \quad (I - u \geq 0)$$

The procedure for estimating a relationship of this sort is discussed by Tobin in Estimation of Relationships for Limited Dependent Variables¹ and is outlined in Appendix B of this paper.

There is one important point which must be made here. While the relationship which is hypothesized is linear, the locus of expected values of $\frac{WP}{H}$ is not linear. That $E(I | X_1, \dots, X_{16}) \neq E(\frac{WP}{H} | X_1, \dots, X_{16})$ is easily verified for values of X_i that give a negative expected value for I . If

1. J. Tobin, Estimation of Relationships for Limited Dependent Variables, Cowles Foundation Discussion Paper No. 3, (published in mimeograph form by the Cowles Foundation 1955).

the relationship hypothesized above is correct, and if an ordinary linear regression is fitted, the estimate will approximate the curvilinear locus of expected values of $\frac{WP}{H}$ in the range where observations are concentrated. A relationship of the sort under examination in this paper is interesting precisely because it might throw some light on what would happen if there were a shift (i.e., an increase in the aggregate birth rate) such as to move all spending units to a different range of the independent variables. The linear approximation could be valueless for such purposes. Another aspect of the same point is that apparently significant changes over time in the parameter estimates of linear regressions fitted to data in which the dependent variable is limited, may not represent changes in the underlying structure, but simply changes in the range of the independent variables.

To summarize: Two thirds of the wives in the sample were not participating in the labor force when the 1954 Survey of Consumer Finances was taken. But a woman who is not working can be, in some sense, close to (or far from) participating in the labor force. If this is the case, it is necessary to construct a model which specifies that the concentration of the independent variable at zero is a consequence of the fact that participation cannot be decreased when it is already zero, even if the independent variables change in the direction of decreasing participation.

5. Presentation and Interpretation of the Estimated Relationship

The relationship which was estimated using the technique discussed in section 4 and Appendix B is as follows:

$$\begin{aligned}
 I = & .51908 - .00007056 H - .0001246 R + .0000336 D_p + .0000101 D_m \\
 & (.225) (.00000270) (.0000179) (.0000149) (.00000518) \\
 & + .0000226 L - .96797 C + .11879 A - .06238 N_c + .00238 N_a \\
 & (.0000434) (.0881) (.0205) (.0180) (.0533) \\
 5.1 \quad & + .33101 E_1 + .40513 E_2 + .47158 E_3 \\
 & (.159) (.159) (.161) \\
 & - .13580 M_1 - .69293 M_2 - .75075 M_3 - 1.03989 M_4 \\
 & (.0455) (.127) (.125) (.125) \\
 \hat{\sigma}_u = & .69862 \\
 & (.0346)
 \end{aligned}$$

$$\frac{WP}{H} = 0 \quad (I < 0)$$

$$5.2 \quad \frac{WP}{H} = I \quad (I \geq 0)$$

H = husband's full time wage rate

R = property income

D_p = personal debt

D_m = mortgage debt

L = liquid assets

C = (0 if no child under six years of age
 1 if there is a child under six years of age

A = age of the youngest child if under six years of age,
 otherwise zero

N_c = number of children in spending unit

N_a = number of adults in spending unit

$(E_1, E_2, E_3) =$

- (0,0,0) no education
- (1,0,0) grammar school education
- (0,1,0) high school education
- (0,0,1) college education

M_1 = number of years married if under five

$(M_2, M_3, M_4) =$

- (0,0,0) married under five years
- (1,0,0) married five to ten years
- (0,1,0) married ten to nineteen years
- (0,0,1) married twenty years or more

The numbers in parentheses under the parameter estimates are the estimated standard errors.

There are three uses to which this estimated relationship can be put. It is possible, given values of the independent variables, to compute:

(1) The estimated index of inclination (or disinclination) of the wife to participate. This is simply the numerical value of I .

(2) The estimated expected value of $\frac{WP}{H}$, given I . The formula for computing $E\left(\frac{WP}{H} \mid I\right)$ and a table giving E as a function of I are given in Appendix C. Loosely speaking, I represents the desired participation rate. Actual participation is subject to a lower limit which is often effective, and an upper limit which is seldom effective. This means that the distribution of observed participation rates is skewed upward and that

the mean participation rate will be higher than the desired participation rate. It is obvious, for instance, that even when I is negative the mean value of $\frac{WP}{H}$ cannot be less than zero.

(3) The estimated proportion of spending units for which $\frac{WP}{H}$ will be greater than zero, i.e., the proportion of women who will be in the labor force. The relationship between this proportion and I is tabulated in Appendix C.

The hypotheses formulated in Section 2 can be thought of as hypotheses concerning the effects of the relevant variables on (1) the participation index, (2) the expected value of $\frac{WP}{H}$, or (3) the proportion of women who will participate. This is possible because both (2) and (3) are monotonic increasing functions of (1).

Of the hypotheses which were tested, only two were not supported by the data. Although the effect of increasing the number of adults is to increase the numerical value of the index, the coefficient of N_a is not significantly different from zero. The coefficient of L is statistically insignificant, and the sign of the coefficient is positive. (It was hypothesized to be negative.) The expected effect of L may be included in the coefficient of R . R was introduced into the relationship after a few iterations had been completed. Before R was introduced, the coefficient of L was negative and relatively large. By analogy, it might be hypothesized that D_p and D_m could, with profit, be replaced in the index with a figure representing the debt claims against current income for interest and amortization.

In every other case the coefficients were of the expected sign and significantly different from zero at the .05 level of significance. The appropriate one-tail test was used in each case. It is not enough, however, that the coefficients be statistically significant. With a sample of this size, extremely small coefficients could be statistically significant, but substantively uninteresting because changes within in the relevant range of the independent variables would cause no appreciable change in the dependent variable. In the case of the coefficient of N_a , for instance, even if the standard error had been much smaller so that the coefficient was statistically significant, increasing the number of adults by as much as five will increase I by only .01. This represents an even smaller increase in the value of $E\left(\frac{WP}{H}\right)$.

When this investigation was undertaken, it was feared that, in comparison with demographic variables, economic variables might explain very little of the wife's (or spending unit's) decision as to participation. It is gratifying to be able to report that the economic variables seem to have important effects.

The Effects of Husband's Earned Income, H

Figure 6 shows the relationship between H and I . R , D_p , D_m , L , C , A , and N_c are zero and N_a is two. Relationships were graphed for each of 12 education-years married groups, and the range of the line for each group reflects the relevant range of H . The relevant range was chosen so as to include about ninety-five percent of the observations in each group. In the groups involving M_1 , it is assumed that M_1 is equal to three. The

fact that all groups have the same slope is, of course, a consequence of the simple linear hypothesis.

Figure 7 represents the relationship between H and the expected value of $\frac{WP}{H}$ for nine of the twelve groups. M_3 groups were eliminated from this figure simply so that it would be easier to read. Since this is the relationship which represents observable behavior, it is from this that one should make a judgment as to the importance of H . In most of the groups there is a drop of .2 in the expected value of $\frac{WP}{H}$ over the range of H .

The differences in the over-all participation levels among the twelve groups can leave no doubt as to the importance of education and the number of years married. It should be remembered that the effect of years married should be interpreted with caution. While it is possible that the estimated parameters reveal something about the effects of the life cycle, it is also possible that the differences are largely due to differences in cohort participation rates.

The Effects of Property Income, R

Figures 8a and 8b, respectively, represent the effects of changes in property income on I and $E(\frac{WP}{H})$ for eight education-years married groups. The lowest education groups are not shown because the range of R in these groups is quite narrow. Figure 8 is drawn to the same scale as Figures 6 and 7 so that it is possible to compare the effects of changes in H to the effects of changes in R . For the purpose of drawing Figure 8, all variables were assumed to have the same values as in Figures 6 and 7 and

Figure 6 - Expected Values of I Given H

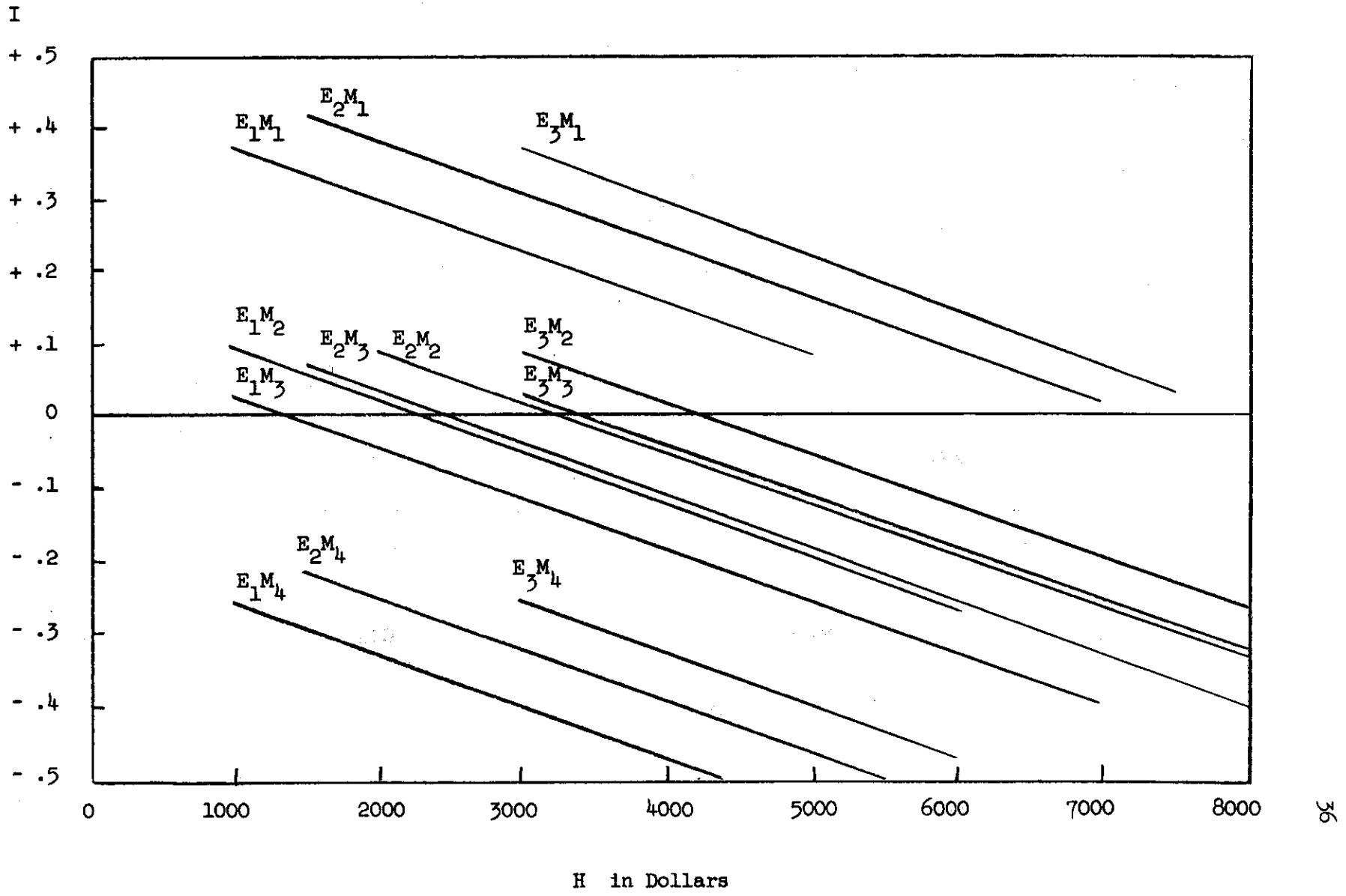
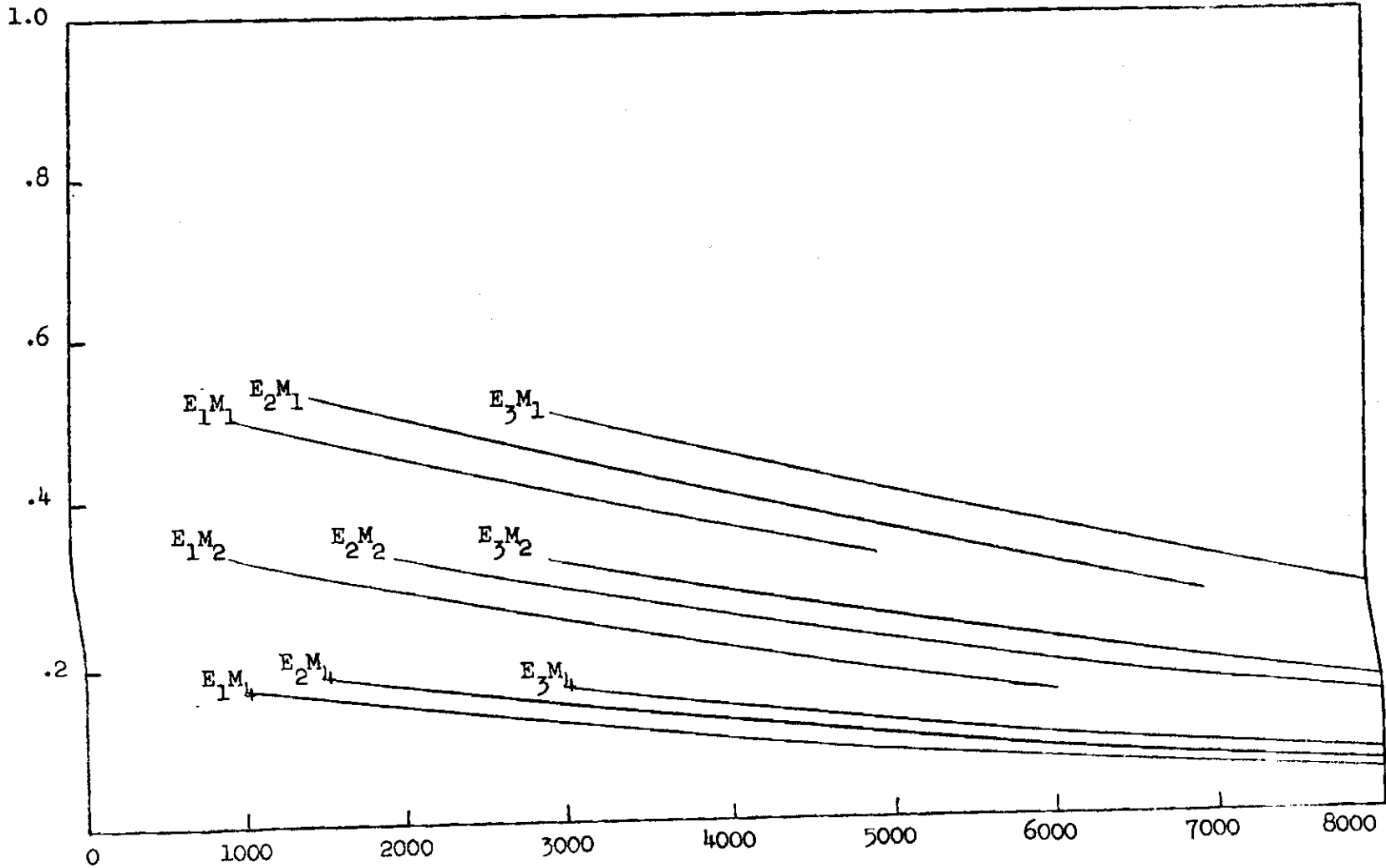


Figure 7 - Expected Values of $\frac{W P}{H}$ Given H

$\frac{W P}{H}$



H in Dollars

H is assumed to be zero. As was hypothesized, the effect of R in decreasing I and $E(\frac{WP}{H})$ is greater than the effect of H. This is because of the correlation between W and H. Since, in the cross-section, increases in H imply increases in W, it is expected that the income effect of increasing H will be somewhat mitigated by the substitution effect of the associated increase in W. An increase in R is equivalent to an increase in H while W is held constant.

The Effects of Personal and Mortgage Debt

Figures 9 and 10 represent the effects of changes in personal and mortgage debt. For the purposes of the two figures $H = \$5000$, R, L, C, A, and N_c are zero, N_a is two, and E_2 and M_2 are both one. In Figure 9, D_m is zero and in Figure 10 D_p is zero. These two variables are the least powerful of the statistically significant explanatory variables. If it were possible to use ex ante personal debt, the estimated coefficient might be larger. A spending unit may have little ex post personal debt just because the wife has been working to pay it off. The coefficient of mortgage debt is more likely to be a good estimate of the effects of ex ante mortgage debt simply because mortgages are usually large relative to income. The coefficient of mortgage debt is suspect for another reason. The size of the mortgage is more or less closely related to the life cycle. Spending units usually buy a house when they are young and pay it off over a long period. To the extent that the coefficient of mortgage debt represents effects which should have been removed by variables accounting for years married, it does not give a correct estimate of the effect of debt itself.

Figure 8a - Expected Values of I Given R

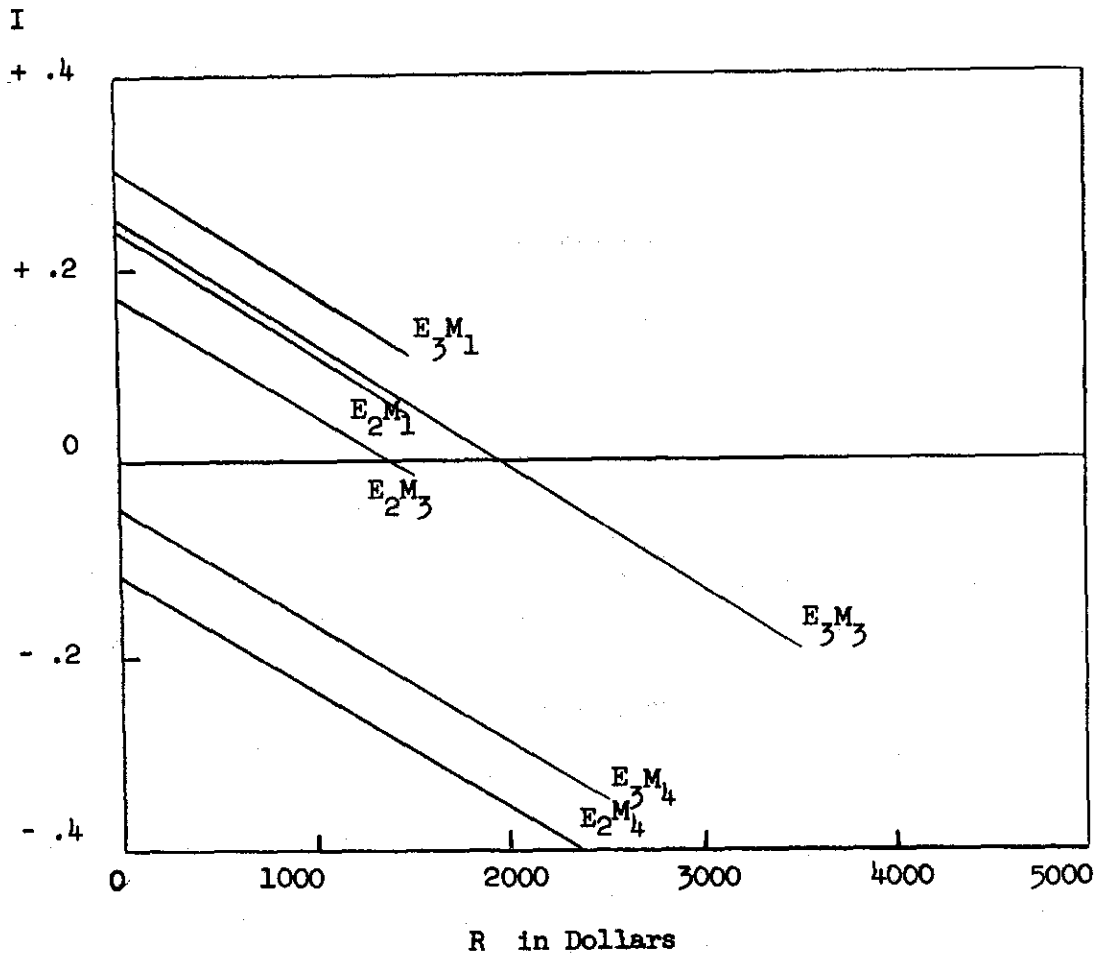
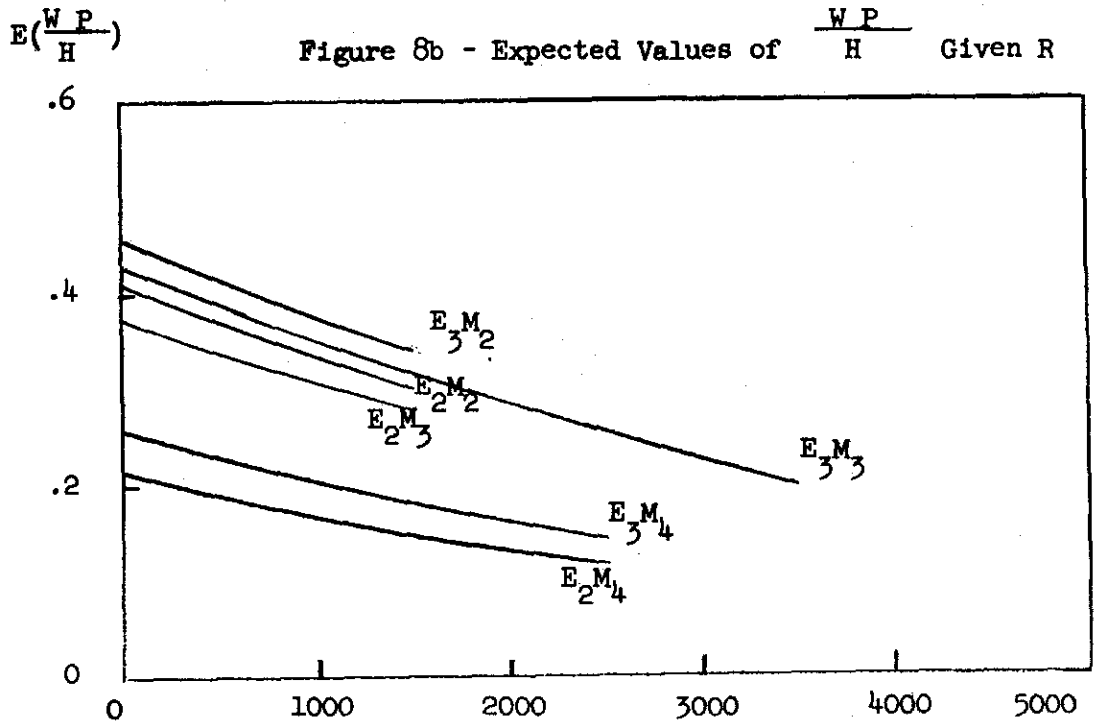
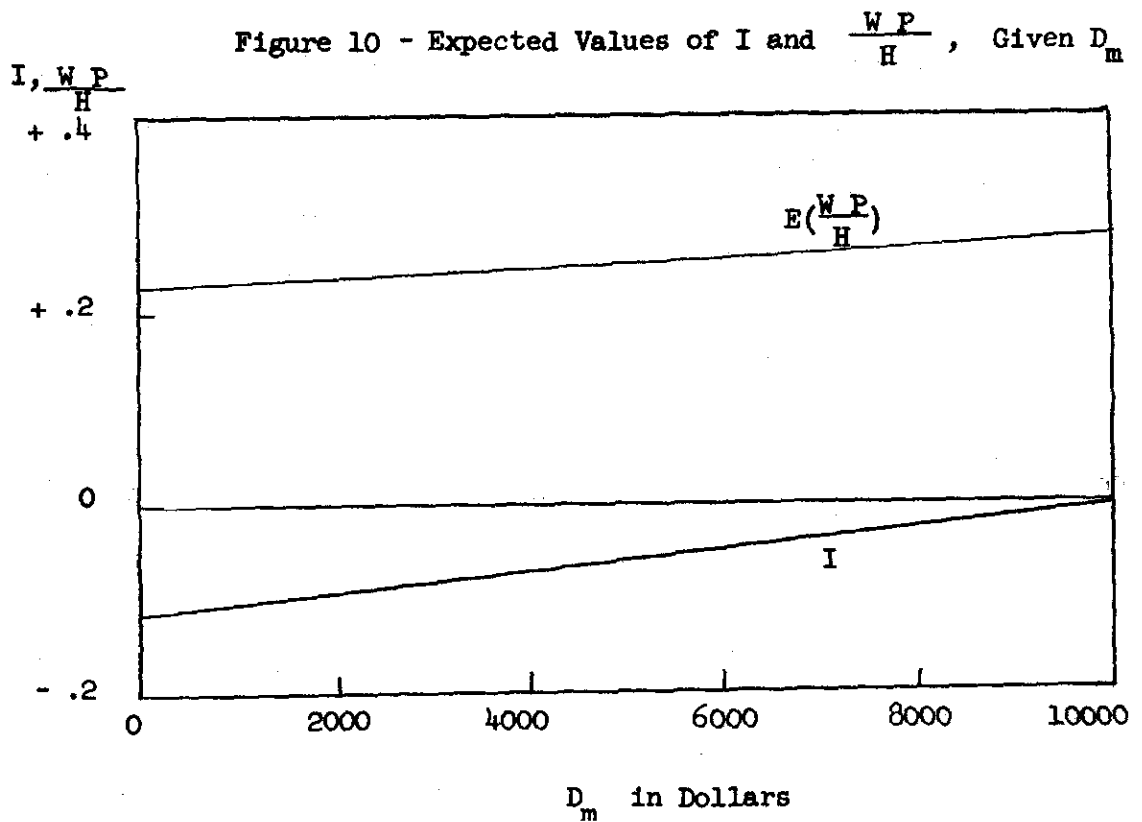
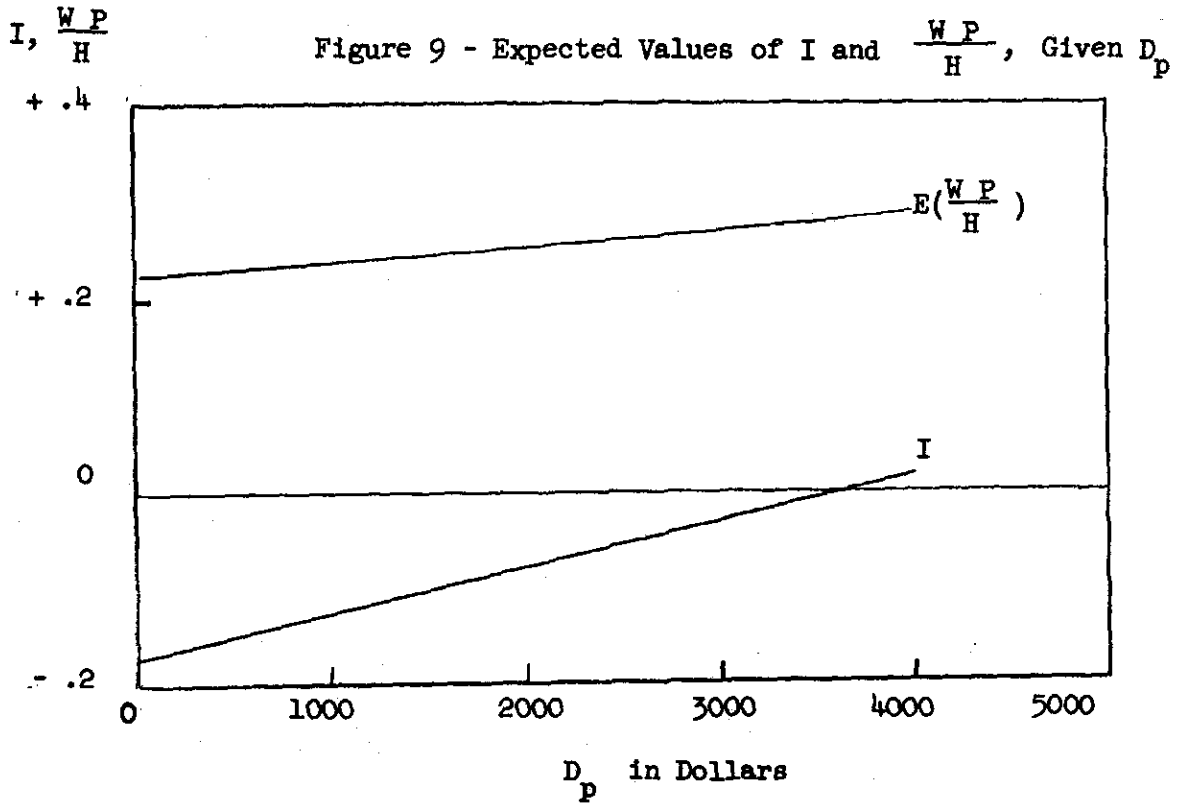


Figure 8b - Expected Values of $\frac{WP}{H}$ Given R





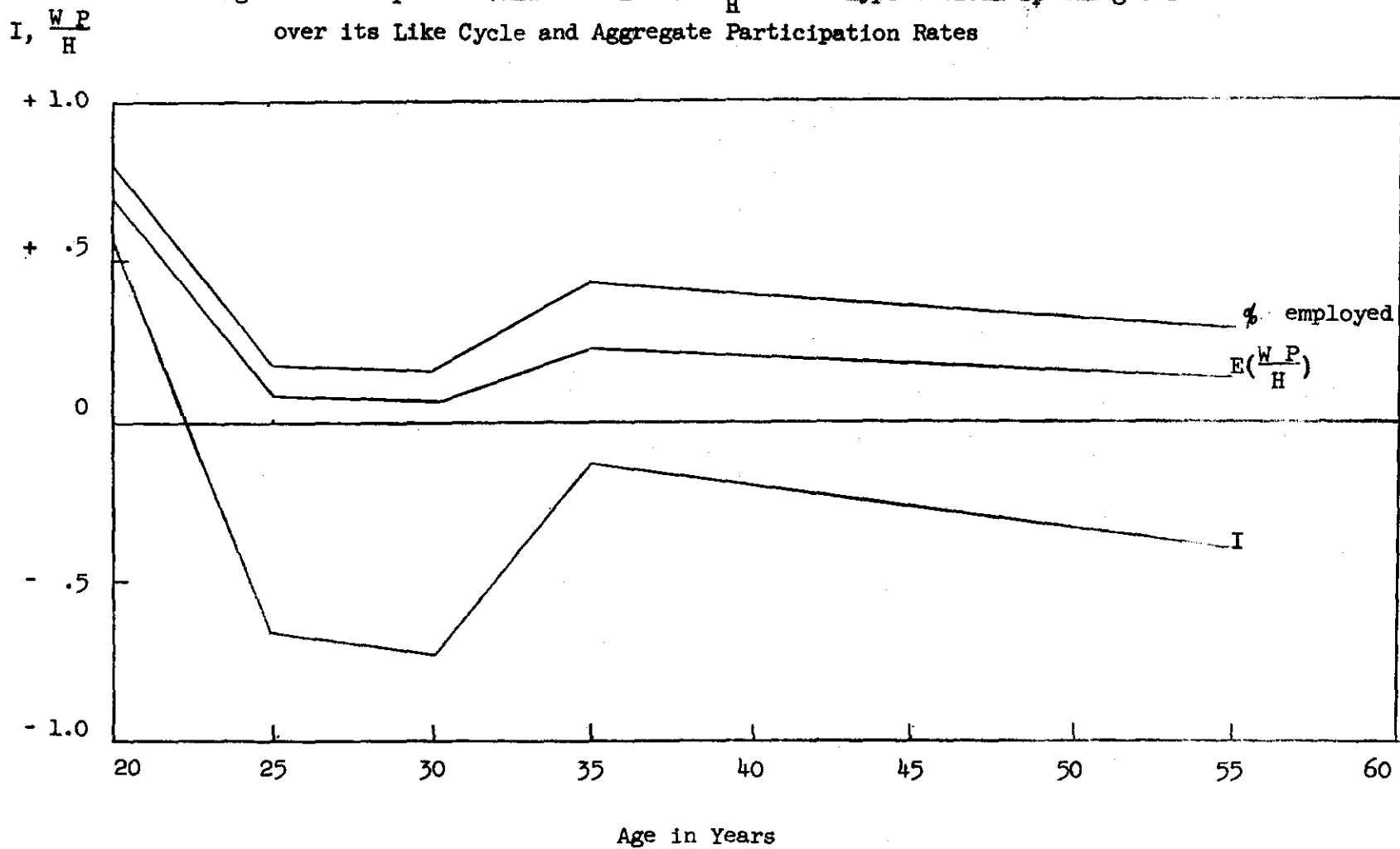
A Hypothetical Life Cycle

Figure 11 represents the life cycle of a hypothetical spending unit. The x-axis measures the age of the wife. She marries at 20, has her first child three years later, and a second child four years after that. Her husband is earning \$3000 when they marry, his income rises to \$4000 by the time she is 30 and remains constant thereafter. Starting at zero, their property income rises at the rate of \$5 per year throughout their lives. They contract \$500 of personal debt in the first year of marriage. Five years later, it rises to \$1500 and remains at the same figure until just after she is 50, when it drops to \$1000. They obtain a \$10,000 mortgage when she is 25 and they pay it off linearly at the rate of \$500 per year. They start with zero liquid assets and accumulate at the rate of \$20 per year. Their parents never come to live with them, and her husband has a high school education.

The points on the graph were computed for ages 20, 25, 30, 35, and 55. The reasons for not computing the points between 35 and 55 are purely aesthetic. Mechanically, effects of having been married over 20 years are felt as soon as the wife reaches the age of 40. This is an artifact of the nature of the data and the consequent form of the estimated relationship. It seems more reasonable to distribute over a longer time span the estimated effect of having been married over twenty years.

The broken line, which represents the proportion of such spending units for which $\frac{WP}{H}$ is greater than zero, is of interest if comparisons are to be made between the results of this investigation, and the aggregate data discussed in Section 1. In order to make a valid comparison it would

Figure 11 - Expected Values of I and $\frac{W.P.}{H}$ of a Hypothetical Spending Unit over its Like Cycle and Aggregate Participation Rates



be necessary to compute a set of life cycles representative of the population and to aggregate. Note, for instance, that about eighty percent of all 20 year old women whose husbands earn \$3000 and who have no children, are expected to be in the labor force. The lack of correspondence between this figure, and the actual percentage of 20 year old women who are in the labor force is partly due to the fact that many married 20 year old women already have children. When this hypothetical woman is 35 years old, she seems to be more representative of the population. The proportion of such women who participate when they are 35 years old is about .4. The aggregate participation rate for women of 35 in 1955 was about 40 percent.

The Correlation Between Husband's Earned Income, H, and Wife's Earned Income, W.

It was difficult to obtain evidence to support the assumption that there is correlation between wage rates of husbands and wives. There are census tabulations of wife's earned income by husband's earned income, but since the earned income of a wife is determined by her participation rate as well as by her wage rate, these tabulations could not be used to support an assumption about the relationship between wage rates. It was, however, possible to obtain some indirect evidence. Table 3 is the result of a tabulation done by the Bureau of the Census on the basis of a current population sample. Define O_h and O_w respectively to be the occupations of the husband and wife. Table 3 gives $\Pr(O_w | O_h)$. It was also possible to obtain median full time wages for both men and women in each of the occupation groups. It is possible, by making some simplifying assumptions, to throw some light on the relationship between the wage rates of husbands and wives.

Assume that

$$5.3 \quad \text{Median } (X \mid O_x) = E(X \mid O_x)$$

and

$$5.4 \quad E(W \mid O_w, O_h, H) = E(W \mid O_w)$$

The first of these assumptions is obviously incorrect, because of the well-known skewness of income distributions. The second assumption, (5.4), while unsupported by any evidence, is probably a conservative assumption for the purpose of demonstrating the relationship between W and H . Assumption (5.4) means that once the wife's occupation is known, information about her husband's occupation or wage rate will add nothing to an explanation of her wage rate. Given (5.3) and (5.4), it is possible to obtain

$$5.5 \quad E(W \mid O_h) = \sum_{O_w} E(W \mid O_w) \cdot \text{Pr}(O_w \mid O_h)$$

for all O_h .

Since $E(H \mid O_h)$ is known, it is possible to obtain a set of points $(E\bar{W} \mid O_h, E\bar{H} \mid O_h)$. A regression was fitted to these points, and the following relationship was obtained.

$$5.6 \quad E(W \mid O_h) = 1470 + .2825 E(H \mid O_h)$$

This can be loosely interpreted as a regression of W on H .

Figure 12 represents the effect of changes in W on I , $E\left(\frac{WP}{H}\right)$ and P when $H + R = \$5000$. The spending unit is assumed to have no debt, liquid assets, or children. E_2 is equal to one and M_1 is equal to three. The

Table 3

OCCUPATION OF EMPLOYED WIVES BY OCCUPATION OF THEIR HUSBANDS: MARCH 1956

Occupation of wife and median full-time, wage	Occupation of the husband and median full-time, wage							
	Professional, technical, and kindred workers \$5668	Managers, officials, & proprietors, except farm \$5477	Clerical and kindred workers \$4248	Sales Workers \$5205	Craftsmen, foremen, & kindred workers \$4766	Operatives and kindred workers \$4117	Service Workers \$3674	Laborers except farm and mine \$3104
Percent	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
Professional, technical & kindred workers \$3559.	32.1	9.5	12.3	13.3	10.4	6.7	5.5	2.5
Managers, officials, and proprietors, except farm \$2851	4.1	17.3	2.9	9.7	3.4	3.0	4.2	2.2
Clerical & kindred workers \$3109	39.7	36.2	43.4	44.9	31.3	21.2	24.1	14.2
Salesworkers \$3099 ⁽¹⁾	8.0	17.0	11.4	14.1	10.1	9.9	4.9	3.3
Craftsmen, foremen, and kindred workers (2)4	1.8	1.1	1.2	2.1	1.4	1.1	1.9
Operatives & Kindred workers \$2532	9.0	8.6	15.3	5.9	24.2	32.7	18.8	29.0
Service workers \$1769	6.4	9.1	13.7	10.6	18.0	24.7	41.4	43.8
Laborers, except farm & mine ⁽²⁾	0.3	0.5	-	0.3	0.5	0.5	-	3.1

(1) Estimated

(2) Not available

SOURCE: Current Population Survey (unpublished tabulation) Bureau of the Census. Median full-time wages, Current Population Report, Series P-60. No. 23 (November, 1956)

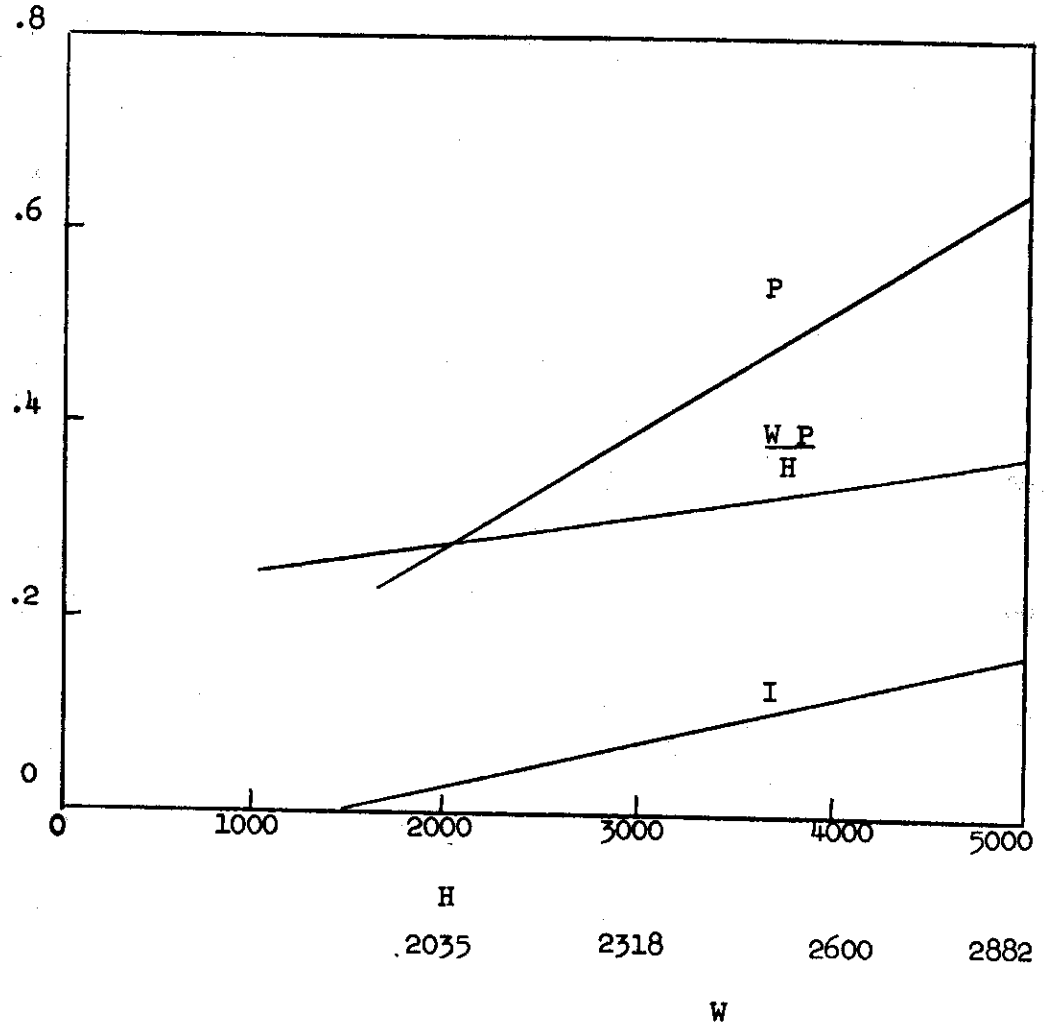
relationship between P and W could be interpreted as a supply curve for the wife's labor. I was computed for combinations of $H + R = \$5000$, $E\left(\frac{WP}{H} \mid I\right)$ was computed and converted into P by using (5.6) to multiply through by $\frac{H}{W}$. The fact that all three curves are positively sloped supports the last hypothesis listed in the summary to Section 2 -- that the derivative of income with respect to the wife's wage rate is positive. Since increases in the wife's wage rate increase participation, it follows from the fact that W is positive that the change in total spending unit income will be positive.

Summary of the Tests of Hypotheses

All of the hypotheses set forth in Section 2 were supported by the data except for the hypothesis that increasing N_a would increase participation and that increases in L would decrease participation. The hypotheses which were supported by the data follow:

- (1) Increases in R will decrease participation more than increases in H . ($b_1 - b_2 > 0$)
- (2) Increases in property income will decrease participation.
($b_2 < 0$)
- (3) Debt will increase participation, and personal debt will increase participation more than mortgage debt. ($b_3 > b_4 > 0$)
- (4) The presence of a child under six years of age will decrease participation. ($b_6 < 0$)
- (5) Participation will increase as the child grows older. ($b_7 > 0$)
- (6) Participation will decrease as the number of children increases.
($b_8 < 0$)

I, $\frac{W}{H}$, P Figure 12 - Effects of W and H on I, $\frac{W}{H}$, and P



(7) Participation will increase as education increases.

$$(0 < b_{10} < b_{11} < b_{12})$$

(8) Participation will decrease as the number of years married increases. $(0 > b_{13} > b_{14} > b_{15} > b_{16})$

(9) Increases in husband's wage will increase spending unit income.

$$(b_2 + 1 > 0)$$

(10) Increasing the wife's wage will increase spending unit income.

This last hypothesis must be interpreted with some care. When the hypothesis was stated in Section 2, it was stipulated that W must be great enough so that the wife is actually participating. If the wife is not participating, the hypothesis has no meaning; changes in her wage rate cannot decrease spending unit income, when she cannot decrease her participation. For the purposes of testing this hypothesis, it is necessary to examine changes in observable behavior, changes in the expected value of $\frac{WP}{H}$ which result from a change in W . Let E be the expected value of $\frac{WP}{H}$. It is possible to demonstrate that the derivative of E with respect to W is positive as follows:

$$(5.7) \quad \frac{\partial E}{\partial W} = \frac{\partial E}{\partial I} \cdot \frac{\partial I}{\partial W}$$

Since the derivative of E with respect to I is positive (see Appendix C) it is necessary only to show that the derivative of I with respect to W is positive. By (5.6) the derivative of I with respect to W is

$$(5.8) \quad \frac{b_1 - b_2}{d} = \frac{.0000540}{.2825} > 0$$

Since W is greater than zero, this is sufficient to show that an increase in W will increase Y provided that $\frac{WP}{H}$ is an appropriate measure of the wife's participation rate.

Evaluation of the Estimated Relationship

The qualifications which must be made in evaluating the usefulness of the estimated relationship are as follows:

- (1) The inadequacies of the dependent variable used make it extremely unlikely that the relationship could accurately predict mean individual participation rates.
- (2) The estimated relationship between W and H is suspect for reasons noted above, and, while illuminating, cannot be used to mitigate the first qualification.
- (3) The relationship between I and the independent variables was hypothesized to be linear and additive. Given better data more sophisticated hypotheses could be formulated and tested so as to improve the predictive value of the relationship.
- (4) The effects of variables closely related to the life cycle should be examined through the use of a set of samples taken in successive years. The possible inadequacy of a single cross-section was discussed in Section 1.
- (5) It would be desirable to distinguish between markets in which the wife might participate and to have information as to the training and prior job experience of the wife.

The value of this study is twofold:

- (1) It is clear that there are sensitive relationships between the decision of the wife to work and a set of demographic and economic variables. The results of this investigation suggest that it would be worthwhile to obtain and analyze a more refined body of data.
- (2) The relationship can be used to estimate the proportion of spending

units for which $\frac{WP}{H}$ is greater than zero. This can be unambiguously interpreted as the aggregate participation rate, defined in Section 1. An example of how such a relationship could be useful is as follows: In the 1930's there was some question as to the extent to which unemployment was increased because of the entrance of women into the labor force when their husbands lost their jobs or suffered pay cuts. It would be useful to know, in such a situation, how many unemployed women would leave the labor force when their husbands begin to make more money. This is exactly the kind of question which could be answered by estimating the aggregate participation rate of wives, given the distribution of wage rates and other relevant independent variables.

Conclusion

The findings of this study strongly support the general hypothesis that the decision of a wife to work is influenced by easily observed economic and demographic characteristics of the spending unit of which she is a member. Although the numerical estimates which resulted from this analysis are of limited usefulness, they at least indicate the direction in which further investigation would be fruitful. There is reason to believe that more intensive analysis of more refined data would yield useful and interesting insights into this aspect of spending unit behavior.

Appendix A

The Model

$$A.1 \quad U = U(Y, P)$$

$$A.2 \quad Y = H + R + WP'$$

The conditions for a constrained maximum are

$$A.3 \quad U_p + WU_y = 0$$

$$A.4 \quad U_{pp} + 2WU_{yp} + W^2U_{yy} < 0$$

Since H and R play an identical role in the model as it stands, R will be assumed equal to zero. Holding W constant and differentiating with respect to H,

$$A.5 \quad \frac{\partial Y}{\partial H} = \frac{WU_{yp} + U_{pp}}{\Delta}$$

$$A.6 \quad \frac{\partial P}{\partial H} = \frac{U_{yp} + WU_{yy}}{\Delta}$$

Holding H constant and differentiating with respect to W,

$$A.7 \quad \frac{\partial Y}{\partial W} = \frac{U_p + P(WU_{yp} + U_{pp})}{\Delta}$$

$$A.8 \quad \frac{\partial P}{\partial W} = \frac{U_y + P(U_{yp} + WU_{yy})}{\Delta}$$

where $\Delta = (A.4)$.

U_{yy} and U_{pp} can both be shown to be negative from (A.4) so that

only the sign of U_{yp} is in doubt. If U_{yp} is positive, there are three possibilities:

$$A.9 \quad WU_{yp} + U_{pp} > 0$$

$$A.10 \quad U_{yp} + WU_{yy} > 0$$

$$A.11 \quad \text{both (A.9) and (A.10)} < 0$$

It can easily be shown that (A.9) and (A.10) cannot hold simultaneously. Multiply (A.10) by W and add (A.9). Since $W > 0$ the sum will be greater than zero. But this contradicts (A.4) and therefore cannot be true.

If (A.9) holds, Y is an inferior good. The slope of an indifference curve is

$$A.12 \quad r(Y,P) = - \frac{U_p}{U_y}$$

Holding Y constant and differentiating with respect to P , at a regular maximum

$$A.13 \quad \frac{\partial r}{\partial P} = - \frac{WU_{yp} + U_{pp}}{U_y}$$

Thus, if (A.9) holds, increasing P will decrease the rate at which increases in P must be compensated for by increases in Y .

If (A.10) holds, leisure is an inferior good.

$$A.14 \quad \frac{\partial r}{\partial Y} = - \frac{U_{yp} + WU_{yy}}{U_y}$$

Thus, if (A.10) holds, increasing Y will decrease the rate at which increases in P must be compensated for by increases in Y . It is now possible to construct the following table of signs:

If

then	Income and leisure are normal goods	Income is an inferior good	Leisure an inferior good
$\frac{\partial Y}{\partial H}$	+	-	+
$\frac{\partial P}{\partial H}$	-	-	+
$\frac{\partial Y}{\partial W}$	+	±	+
$\frac{\partial P}{\partial W}$	±	±	+
$\frac{\partial P}{\partial W} - \frac{\partial P}{\partial H}$	+	+	±

A Modification of the Model

Assume that

$$\text{A.15} \quad W = a + bH$$

then

$$\text{A.16} \quad \frac{\partial Y}{\partial H} = \frac{b(U_p + P(WU_{yp} + U_{pp})) + WU_{yp} + U_{pp}}{\Delta}$$

$$\text{A.17} \quad \frac{\partial P}{\partial H} = \frac{b(U_y + P(U_{yp} + WU_{yy})) + U_{yp} + WU_{yy}}{\Delta}$$

Dropping the assumption that R is equal to zero,

$$\text{A.18} \quad \frac{\partial Y}{\partial R} = \text{(A.5)}$$

$$\text{A.19} \quad \frac{\partial P}{\partial R} = \text{(A.6)}$$

It can now be seen that

$$\text{A.20} \quad \text{(A.15)} - \text{(A.17)} = b\text{(A.7)}$$

$$\text{A.21} \quad \text{(A.16)} - \text{(A.18)} = b\text{(A.8)}$$

Appendix B¹

The Statistical Model

Define $P(x)$ to be the value of the unit normal distribution function, $Q(x) = 1 - P(x)$, and $Z(x)$ is the value of the unit normal density function.

$$B.1 \quad I = \sum_{i=0}^{16} b_i X_i + u \quad (X_0 \text{ is defined to be always equal to } 1)$$

$$B.2 \quad \frac{WP}{H} = 0 \quad (I - u < 0)$$

$$B.3 \quad \frac{WP}{H} = I - u \quad (I - u \geq 0)$$

where u is normally distributed with zero mean and standard deviation σ .

$$B.4 \quad \Pr\left(\frac{WP}{H} = 0 \mid I\right) = \Pr(u > I) = Q(I/\sigma)$$

$$B.5 \quad \Pr\left(\frac{WP}{H} > x \geq 0 \mid I\right) = \Pr(u < I - x) = P([I - x]/\sigma)$$

The distribution function for $\frac{WP}{H}$, given I is

$$B.6 \quad \begin{aligned} F(x; I) &= 0 & (x < 0) \\ F(x; I) &= Q([I - x]/\sigma) & (x \geq 0) \end{aligned}$$

The density function is

$$B.7 \quad f(x; I) = (1/\sigma) Z([I - x]/\sigma) \quad (x > 0)$$

1. This appendix is a presentation, in outline form, of the application of the statistical model, described by J. Tobin in Cowles Foundation Discussion Paper No. 3, to the problem being examined here. It should be obvious to anyone who has seen Discussion Paper No. 3, that the author's debt to Professor Tobin is a heavy one.

For the purpose of simplifying computations, it is convenient to normalize on σ instead of on the coefficient of the dependent variable. Thus (B.1) is rewritten

$$B.8 \quad b'I = I' = \sum_{i=0}^{16} b'_i X_i + u'$$

where $b'_i = b_i/\sigma$ and $b' = 1/\sigma$. The error term u' is unit-normally distributed. The distributed function, F , has a mass point at $x = 0$. It is therefore necessary, in forming the likelihood function, to distinguish between observations at the mass point where f is not defined, and observations for which f is defined. Numbering the 1592 observations so that the first 1067 are limit observations, the likelihood of the sample is

$$B.9 \quad \phi(b', \hat{b}'_0, \dots, \hat{b}'_{16}) = \prod_{i=1}^{1067} F(0; I'_i) \cdot \prod_{i=1068}^{1592} f\left(\frac{W_i P_i}{H_i}; \hat{I}'_i\right)$$

$$= \prod_{i=1}^{1067} Q(\hat{I}'_i) \cdot \prod_{i=1068}^{1592} \hat{b}'_i Z\left(\hat{I}'_i - \hat{b}'_i \frac{W_i P_i}{H_i}\right)$$

The natural logarithm of ϕ is

$$B.10 \quad L(\hat{b}', \hat{b}'_0, \dots, \hat{b}'_{16}) = \sum_{i=1}^{1067} \ln Q(\hat{I}'_i) + 525 \ln \hat{b}' - (525/2) \ln 2\pi$$

$$+ \sum_{i=1068}^{1592} \left(\hat{I}'_i - \hat{b}'_i \frac{W_i P_i}{H_i} \right)$$

Taking the derivatives of L with respect to \hat{b}'_k ($k = 0, \dots, 16$) and \hat{b}' and setting them equal to zero,

$$L_{b_k} = - \sum_{i=1}^{1067} \frac{Z(\hat{I}_i) X_{ki}}{Q(\hat{I}_i)} - \sum_{i=1068}^{1592} (\hat{I}_i - \hat{b}') \frac{W_i P_i}{H_i} X_{ki} = 0$$

B.11

$$(k = 0, 1, \dots, 16)$$

$$L_b = 525/\hat{b}' + \sum_{i=1068}^{1592} (\hat{I}_i - \hat{b}') \frac{W_i P_i}{H_i} \frac{W_i P_i}{H_i} = 0$$

Since this system of equations is non-linear, it is necessary to use some sort of iterative procedure to evaluate $(\hat{b}', \hat{b}'_0, \hat{b}'_1, \dots, \hat{b}'_{16}) = \hat{B}'$

The procedure employed was Newton's method which is to approximate L with a quadratic having the same first and second derivatives as L at some point \hat{B}'_0 . The quadratic is maximized to obtain \hat{B}'_1 , and the first and second derivatives are evaluated at this new point. The procedure is repeated until the changes in the parameters estimates from the n th to the $(n+1)$ st iteration are small enough (i.e., small with respect to the standard errors of the coefficient estimates). An initial estimate was obtained by approximating the function $-Z(x)/Q(x)$ with a line and solving the system of equations (B.11) for \hat{B}'_0 . The second derivatives of L are

$$L_{b_k b_t} = \sum_{i=1}^{1067} X_{ki} X_{ti} \left(\frac{Z(\hat{I}_i)^2}{Q(\hat{I}_i)^2} - (\hat{I}_i) \frac{Z(\hat{I}_i)}{Q(\hat{I}_i)} - \sum_{i=1068}^{1592} X_{ki} X_{ti} \right)$$

$$B.12 \quad L_{b_k b'} = \sum_{i=1068}^{1592} X_{ki} \frac{W_i P_i}{H_i} + \frac{525}{\hat{b}'} \quad (k, t = 0, 1, \dots, 16)$$

$$L_{bb} = - \sum_{i=1068}^{1592} \left(\frac{W_i P_i}{H_i} \right) - 525/\hat{b}'^2$$

Let M be the matrix of second derivatives of L and v the vector of first derivatives. The $(n+1)$ st approximation to \hat{B}' is obtained from the n th by solving the matrix equation

$$B.13 \quad M(B'_{n+1} - B'_n) = -v$$

A computing program for the IBM 650 Magnetic Drum Data Processing Machine was designed which could solve problems of the sort described above as follows:

- (1) The limiting value of the dependent variable can be any number (not just zero).
- (2) The limiting value of the dependent variable can vary from observation to observation within the sample.
- (3) The maximum number of parameters which can be estimated is 20.
- (4) There is no limit to the number of observations which can be processed.

It should be noted, however, that every observation must be processed in each iteration. Since (for 18 parameters) each observation required eleven seconds of computing time per iteration, the computing cost is not negligible.

- (5) The intermediate outputs of the program are

$$M'_n, v'_n, M'^{-1}_n, \hat{B}'_{n-1} - \hat{B}'_n, \text{ and } \phi(\hat{B}'_n).$$

Estimates of the variances and co-variances of the parameter estimates can be obtained from the negative of the inverse matrix of second derivatives evaluated when L is at a maximum. Table 4 gives the initial trial values,

the intermediate approximations to the parameter estimates, the value of ϕ for each iteration, the final estimates of the parameters, and the estimated standard errors of the parameter estimates. In order to obtain the coefficients of the relationship which is being investigated here, it is necessary to divide every parameter estimate by \hat{b}' . The variables C and R were not introduced until the 5th iteration. The computing program was re-scaled after the second iteration in order to obtain more digits in the parameter estimates. Total machine time for the problem was about 175 hours (about 50 hours for programming, and 125 hours for computing). ϕ was evaluated at several points between the old and new \hat{B}' after each iteration, and the point which gave the highest likelihood was used in the next iteration. In several cases, somewhat less than a full step proved to be better than a full step, but in the last three iterations, the full step was taken.

Table 4

INTERMEDIATE AND FINAL ESTIMATES OF \hat{b}_i

	\hat{b}_0	\hat{b}_1	\hat{b}_2	\hat{b}_3	\hat{b}_4	\hat{b}_5	\hat{b}_6	\hat{b}_7
\hat{b}_i	2.0	1.4	2.79	1.8980	1.5128	1.3522	1.3946	1.6421
\hat{b}_0	1.342	1.442	1.442	.82332	.85292	.73844	.76060	.48595
$\hat{b}_1 \times 10^5$	0.0	-3.0	-4.0	-4.42	-4.638	-5.152	-5.633	-5.941
$\hat{b}_2 \times 10^5$	-	-	-	-	-	-2.24	-4.14	-5.56
$\hat{b}_3 \times 10^5$	0.0	0.0	34.0	14.13	8.49	5.93	4.71	3.99
$\hat{b}_4 \times 10^5$	0.0	0.0	24.0	13.39	8.67	5.87	4.26	3.26
$\hat{b}_5 \times 10^5$	0.0	0.0	-7.0	-6.076	-5.384	-4.310	-3.485	-2.841
\hat{b}_6	-	-	-	-	-	-.35942	-.63141	-.71515
\hat{b}_7	-.032	-.048	-.048	-.05948	-.07166	-.00788	.03722	.05913
\hat{b}_8	-.059	-.103	-.128	-.16363	-.13007	-.11298	-.09560	-.08848
\hat{b}_9	.014	-.316	-.316	-.24404	-.17888	-.12670	-.09394	-.06722
\hat{b}_{10}	-.062	.052	.053	.46521	.44579	.42632	.42335	.42130
\hat{b}_{11}	-.077	.045	.045	.52654	.50870	.49014	.49082	.49028
\hat{b}_{12}	.037	.142	.143	.61196	.61404	.59262	.59482	.57882
\hat{b}_{13}	-.164	-.236	-.245	-.20235	-.22495	-.20402	-.20374	-.16788
\hat{b}_{14}	-.708	-1.038	-1.038	-.89444	-.95267	-.90428	-.92895	-.81069
\hat{b}_{15}	-.518	-.788	-.789	-.68667	-.76373	-.81727	-.90771	-.81355
\hat{b}_{16}	-.590	-.925	-.925	-.90498	-1.04268	-1.12012	-1.24069	-1.13321
$\hat{\phi}$	4.159	1.127	2.239	1.127	1.444	3.602	1.396	1.021
\hat{x} 10 to the	-784	-443	-618	-507	-459	-432	-411	-406

(continued)

Table 4 (cont'd)

INTERMEDIATE AND FINAL ESTIMATES OF \hat{B}'

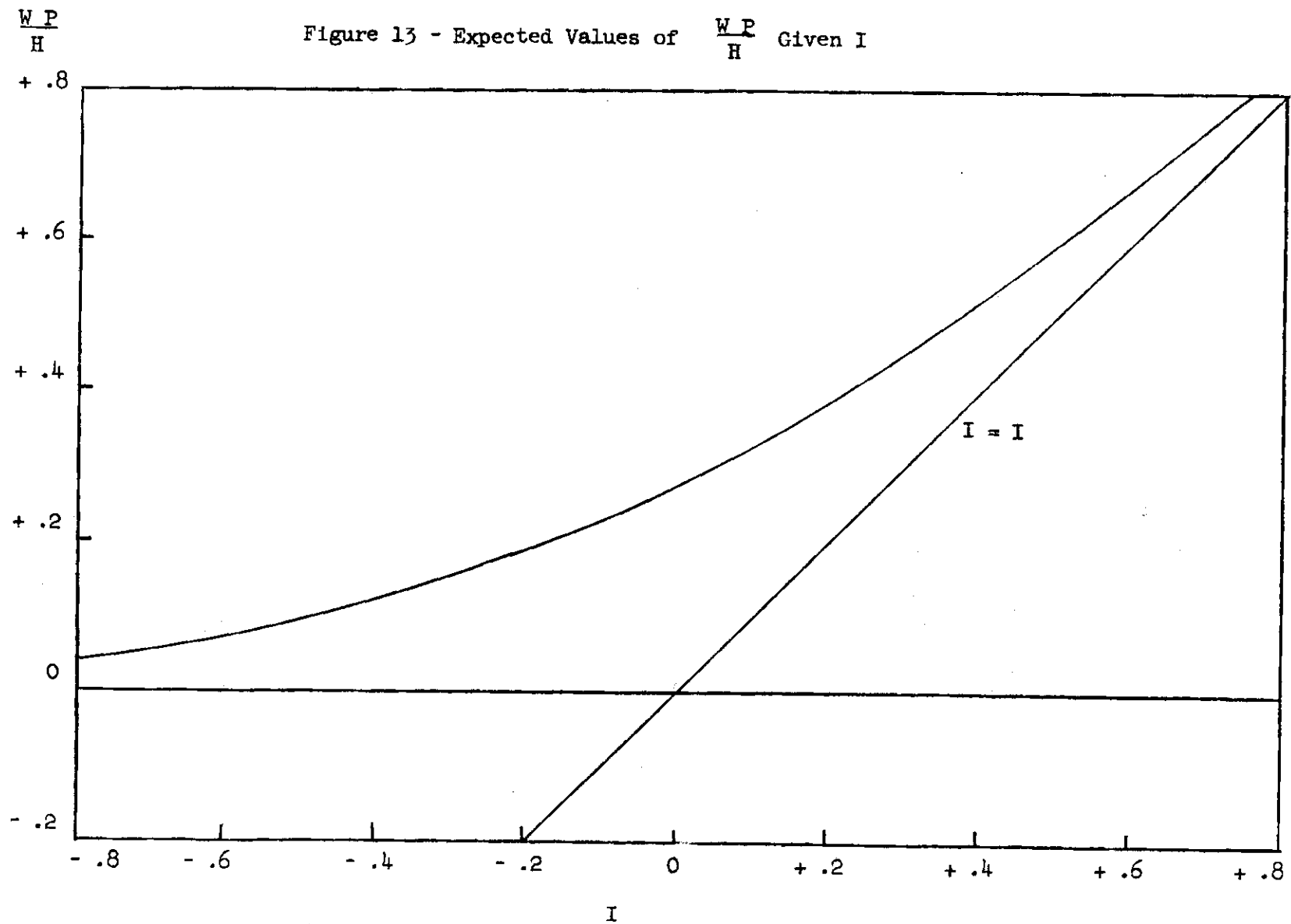
	\hat{B}'_8	\hat{B}'_9	\hat{B}'_{10}	\hat{B}'_{11}	\hat{B}'_{12}	\hat{B}'_{13}	\hat{B}'_{13}
\hat{b}'_1	1.5234	1.4722	1.4398	1.4275	1.4296	1.4314	.0495
\hat{b}'_0	.60798	.67502	.72311	.74589	.74385	.74301	.3218
$\hat{b}'_1 \times 10^5$	-6.602	-7.181	-7.901	-8.803	-9.908	-10.100	.3874
$\hat{b}'_2 \times 10^5$	-6.99	-8.37	-10.23	-13.81	-17.33	-17.83	2.572
$\hat{b}'_3 \times 10^5$	3.80	3.80	3.92	4.02	4.50	4.81	2.140
$\hat{b}'_4 \times 10^5$	2.67	2.27	1.88	1.55	1.41	1.45	.7421
$\hat{b}'_5 \times 10^5$	-2.270	-1.776	-1.175	-.478	-.004	.323	.6214
\hat{b}'_6	-.89395	-1.02364	-1.15462	-1.27814	-1.34516	-1.38555	.1261
\hat{b}'_7	.08574	.10564	.12673	.14793	.16128	.17004	.02940
\hat{b}'_8	-.09204	-.09022	-.08888	-.08842	-.08873	-.08929	.02581
\hat{b}'_9	-.05183	-.03993	-.02673	-.01237	-.00309	-.00348	.07632
\hat{b}'_{10}	.43079	.43833	.44756	.45803	.46704	.47381	.2274
\hat{b}'_{11}	.50311	.51463	.51931	.54550	.56449	.57990	.2269
\hat{b}'_{12}	.59298	.61568	.62747	.64498	.66079	.67502	.2310
\hat{b}'_{13}	-.18484	-.19254	-.19673	-.19690	-.19519	-.19439	.06513
\hat{b}'_{14}	-.89556	-.93945	-.97043	-.98650	-.98879	-.99186	.1822
\hat{b}'_{15}	-.92983	-.99371	-1.04153	-1.06823	-1.07275	-1.07455	.1795
\hat{b}'_{16}	-1.28470	-1.37181	-1.44031	-1.48257	-1.48786	-1.48850	.1787
ϕ	1.266	8.070	1.788	7.326	1.013	1.308	
X							
10 to the	-402	-401	-398	-397	-395	-395	

Table 5

EXPECTED VALUES OF $\frac{WP}{H}$ GIVEN I
 AND THE PROBABILITY OF $\frac{WP}{H} > 0$ GIVEN I

I	$P(\frac{WP}{H} > 0)$	$E(\frac{WP}{H})$
- 1.0	.0762	.0239
- .9	.0989	.0326
- .8	.1261	.0438
- .7	.1582	.0580
- .6	.1952	.0756
- .5	.2370	.0972
- .4	.2833	.1232
- .3	.3340	.1540
- .2	.3873	.1901
- .1	.4431	.2316
0.0	.5000	.2787
.1	.5569	.3316
.2	.6127	.3901
.3	.6660	.4540
.4	.7357	.5232
.5	.7630	.5922
.6	.8048	.6756
.7	.8418	.7580
.8	.8739	.8438
.9	.9011	.9326
1.0	.9238	1.0239

Figure 13 - Expected Values of $\frac{W P}{H}$ Given I



Bibliography

- Breckinridge, S. P. Women in the Twentieth Century. New York: McGraw Hill Book Publishing Co., 1933.
- Ducoff, Louis J. and Hagood, Margaret J. Labor Force Definition and Measurement. New York: Social Science Research Council, 1947.
- Durand, John D. The Labor Force in the United States, 1890-1969. New York: Social Science Research Council, 1948.
- Glick, Paul A. American Families. Census Monograph Series. New York: John Wiley and Sons, Inc., 1957.
- Hill, Joseph A. Women in Gainful Occupations, 1870-1920. Census Monograph IX. Washington: United States Government Printing Office, 1929.
- Klein, Lawrence R. and Morgan, James N. "Results of Alternative Statistical Treatments of Sample Survey Data," Journal of the American Statistical Association, XLVI 256 (December 1951) pp. 442-460.
- Kyrk, Hazel. The Family in the American Economy. Chicago: University of Chicago Press, 1953.
- Lansing, John B. "Concepts Used in Surveys," Contributions of Survey Methods to Economics, edited by L. R. Klein. New York: Columbia University Press, 1954.
- Leven, Maurice. The Income Structure of the United States. Washington D.C.: The Brookings Institution, 1938.
- Long, Clarence D. "The Labor Force under Changing Income and Employment." New York: National Bureau of Economic Research, 1956. (Mimeographed pre-publication manuscript)
- "Methods of the Survey of Consumer Finances," Federal Reserve Bulletin, XXXVI 7 (July, 1950), pp. 795-809.
- Miller, Herman P. Income of the American People. Census Monograph Series. New York: John Wiley and Sons, Inc., 1955.
- National Manpower Council. Womanpower. New York: Columbia University Press, 1957.

Samuelson, Paul A. "Social Indifference Curves," Quarterly Journal of Economics LXX (February, 1956), pp. 1-22.

Tobin, J. "Estimation of Relationships for Limited Dependent Variables," Cowles Foundation Discussion Paper No. 3. New Haven: Cowles Foundation for Research in Economics at Yale University, 1956 (revised 1956). (Mimeographed)

United States Department of Commerce, Bureau of the Census. "Current Population Reports" Series P-50 and P-60. Washington: United States Government Printing Office.

United States Department of Labor. Bulletins of the Women's Bureau. Washington, D. C.: United States Government Printing Office.

The War and Women's Employment. International Labor Office. Montreal: 1946.

Woytinsky, W. S., and Associates. Employment and Wages in the United States. New York: Twentieth Century Fund, 1953.