Chapter 3

WAGES, INCOME, AND HOURS OF WORK IN THE U.S. LABOR FORCE

Robert E. Hall

I. INTRODUCTION

This paper presents the results of an empirical study of hours of work in the labor force of the United States. The main goal of this research is to obtain better knowledge of the pattern of work effort by wage and income classes within broad demographic groups in the labor force. In brief, our results can be summarized as follows:

1. Husbands of both races in the central age group, 20 through 59, tend to work roughly full time, on the average, and have weak wage and

---

Robert E. Hall is an associate professor of economics at the Massachusetts Institute of Technology. This is a report on research supported by the Office of Economic Opportunity through a grant to the Brookings Institution. The author is grateful to Michael Boskin and Michael Hurd for advice and assistance, to Thomas Moore for computer programming, and to Orley Ashenfelter, Michael Barth, Ralph Beals, David Gordon, Robert A. Gordon, Ronald Grieson, and Sherwin Rosen for comments on an early version.

income effects. Even the relatively unusual individual with a low wage and high outside income tends to work almost full time.

2. Black husbands tend to work somewhat less than white husbands with the same wage and income. The difference is most pronounced in the lowest income groups.

3. Husbands of retirement age, 60 and over, show substantial variation in hours of work, related systematically to wages and income in the expected way.

4. Wives in all age groups are quite sensitive to wages and income. Black wives work substantially more than white wives, after adjustment for wages and income.

5. Single individuals do not have a systematic tendency to work longer hours with higher wages. Some groups show evidence of backward-bending labor-supply curves.

6. Adult sons and daughters and other relatives do not seem to respond to the incomes of the families in which they reside. Their wage response is roughly the same as that of single individuals.

7. Teen-agers who are not in school, work remarkably little and do not have a strong positive response to wages.

8. Race and sex differences are conspicuous for husbands and wives and are almost absent for other groups.

Our main emphasis is on the proper measurement of the economic quantities relevant to the study of labor supply, rather than on the fitting of supply equations derived from an underlying parametric specification of preferences for consumption and leisure. In fact, an important intermediate step in this work is simply the cross-tabulation of average hours of work by the characteristics of individuals and their families and by their wages and incomes. Even without further restriction, the resulting tables provide useful information for some major groups in the labor force. For teen-agers and other smaller groups, averaging methods are used to reduce the influence of random fluctuations. The general approach of the research seems to be successful because of the size and richness of the body of data on which it rests.

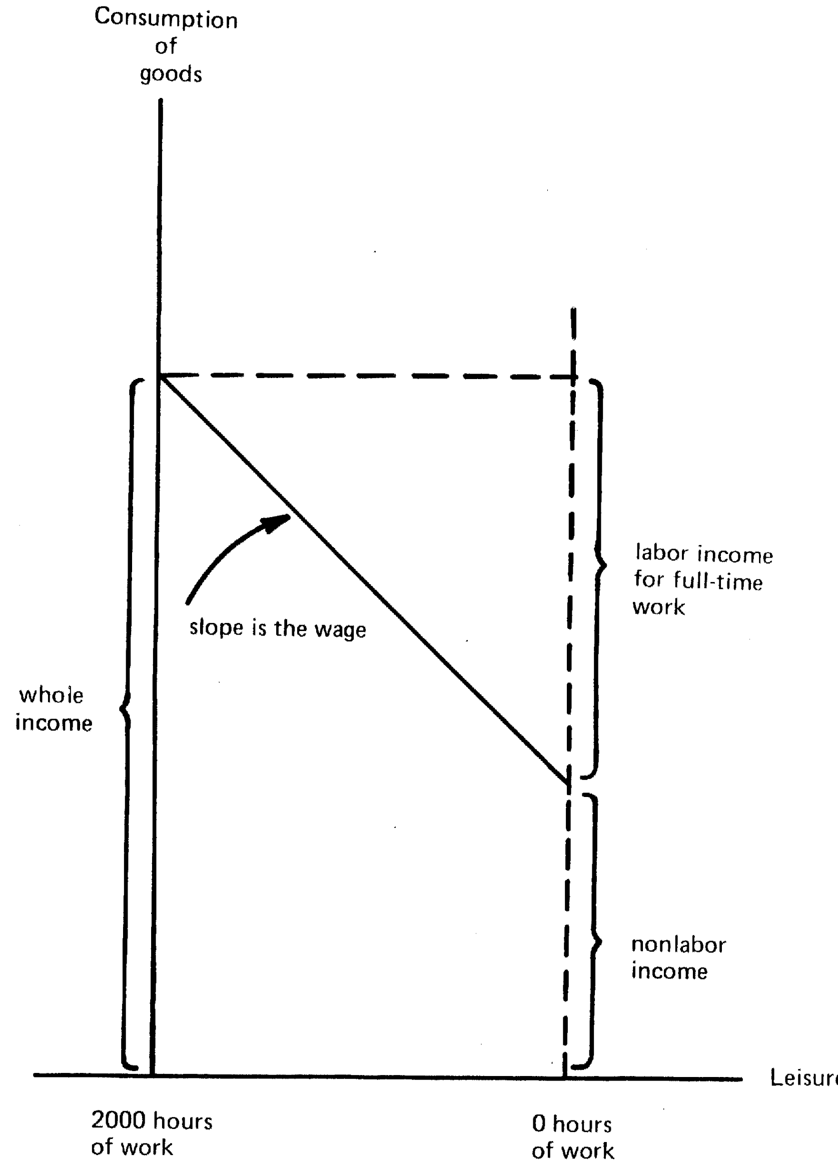
The data are taken from the Survey of Economic Opportunity (SEO) for 1967, a file of data on individuals collected by the Bureau of the Census and compiled by the Brookings Institution and the Office of Economic Opportunity. The SEO is an augmented version of the Current Population Survey (CPS). The augmentations are crucial, however, for this kind of study. First, data on hours and wages were collected from most respondents for the week before the survey in March 1967. One of the main obstacles to the use of data from the CPS and from the decennial Census is the lack of a reliable measure of wages. The availability of information on wages also made it possible to construct estimates of hours of work in 1966, by dividing wage income by the wage. Again, there is no reliable measure of annual hours of work in the CPS, and most investigators have adopted the rather unsatisfactory assumption that hours of work in the week before the survey were typical of all the weeks worked in the previous year. Second, one-half of the SEO sample is drawn from specially selected nonwhite poverty areas. As a result, whites and blacks (who are distinguished from other nonwhites) are approximately equally represented, and comparisons between races are greatly facilitated. Third, extra data on income and assets are available, so that by a series of imputations a reasonably comprehensive measure of income can be constructed.

The study embodies a somewhat unconventional approach to the definition of the economic determinants of labor supply. The usual labor-supply function for an individual is written in terms of his wage and his income from sources other than his own labor. With this convention, the response to an increase in the wage has two conflicting components—a substitution effect tending to increase hours of work and an income effect tending to reduce hours. The pure substitution effect can be inferred from these responses by an appropriate Hicksian income compensation, but the formulation is still somewhat awkward. A more general view is that the labor-supply function of an individual can be written in terms of any two variables that uniquely define his budget constraint.[1] The variables we have chosen are the slope of the budget constraint (the wage) and its intercept with a vertical line corresponding to full-time work. The latter quantity is what we will call *whole income*. It is the amount the individual can spend on goods if he works full time.[2] (See Figure 3.1.) The advantage of writing the labor-supply function this way is that for full-time workers, the pure substitution effect is exactly the effect of changing the wage while holding whole income constant. The Hicksian income compensation is built into the labor-supply function when it is written as a function of whole income. For individuals who work less than full time, the income compensation of our labor-supply function overstates the Hicksian compensation. As long as leisure is not an inferior good, the substitution and income

[1] For simplicity assume that the individual consumes only leisure and one good. As long as the relative prices of the various consumption goods are the same for all people in the sample, this treatment is rigorous.

[2] This measure of income should be distinguished from Gary Becker's notion of *full* income; see his article, "A Theory of the Allocation of Time," *The Economic Journal* 75 (September 1965), which defines full income as the amount an individual could spend on goods if he consumed no leisure at all. The two measures differ by the value of the leisure consumed by a person who works full time.

FIGURE 3.1. PARAMETERS OF AN INDIVIDUAL'S BUDGET CONSTRAINT



effects of wage changes have the same sign in our supply function. Holding whole income constant, we should observe increasing (or, at least, nondecreasing) hours of work as we increase the wage for any individual, as long as he does not work more than full time.

So far we have discussed the case of an isolated individual. In fact, most of the people in the sample live in families, where decisions about hours of work are made jointly by the members of the family. For a family with two adult members, say a husband and wife, we identify three items that enter family preferences: goods, leisure of the husband, and leisure of the wife. The family budget constraint now requires three parameters to describe it completely: two wages to describe its slope and one income to locate its position away from the origin. Once again we use a measure of whole income, defined as the amount of goods the family could consume if both its members worked full time. Again, the response of either member to an increase in his or her wage, with whole income held constant, is at least fully compensated for the income effect and should always be positive.

An additional complication arises in the case of a family: the wife's wage enters the husband's labor-supply function, and vice versa. In a conventional model of family labor supply where the income variable is nonlabor income, the wife's wage has two influences on her husband's supply of labor: a substitution effect of uncertain sign and a presumably negative income effect. The presence of the wife's wage in the husband's supply function is mandatory, even if the substitution effect is zero, owing to the potential strength of the income effect. With our approach, on the other hand, the income effect is exactly compensated by the use of whole income, provided the wife actually works full time. To impose the hypothesis that the cross-substitution effect is zero—that is, that the leisure of the husband and leisure of the wife are neither substitutes nor complements—we simply exclude the wife's wage from the husband's supply function, and vice versa. The result is a saving in parameters—quite important in a study of this sort that uses a very unrestrictive functional form. The substantive advantage of using whole income rather than nonlabor income is that it permits the exclusion of the wages of other family members from the supply equation of each member.

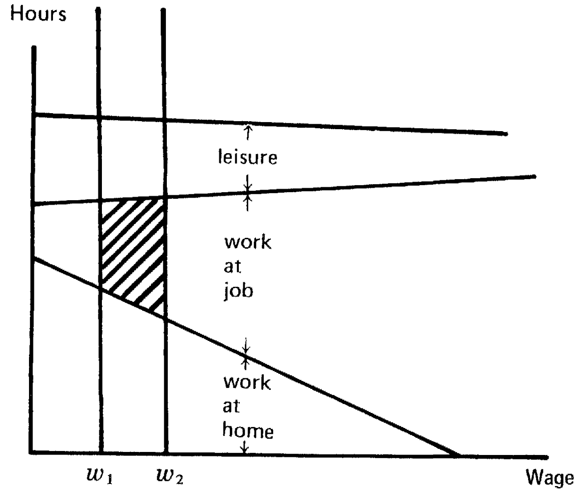
One apparent difficulty with this approach is that wives, in general, do not work full time. Our procedure seems to overcompensate for the income effect of an increase in the wife's wage by assuming that she works full time when, in fact, she may work half time or not at all. The problem here is to interpret the notion of work correctly. In our simple theory, hours of work are the hours of the year not spent enjoying leisure. For wives, this means that hours of work include hours spent caring for children and keeping house. With work measured to include work at home, most wives do work full time. Should work at home be included in calculating the income effect of a change in wages? For an individual who suddenly faces a wage increase but is otherwise unchanged, it is clear that the income effect arises only from the change in the value of the labor services actually sold on the market. In a cross section, however, individuals earning higher wages may also be more productive at home—many of the same talents that are useful to employers are also useful in managing a household. An extreme version of this hypothesis holds that the marginal-productivity schedule for work at home of an individual with a high wage is higher than that of one with a low wage in exact proportion to their wages. Then the income effect should be calculated by multiplying the change in the wage by total hours of work, on the job and at home. The two cases are shown in Figure 3.2. On the left, the marginal-productivity curve for work at home is assumed to be the same for individuals with wages w_1 and w_2. The difference in whole income (measuring the income effect) is just the shaded area corresponding to work on the job. On the right, the marginal-productivity curve has shifted to the right for the individual with wage w_2, and the difference in whole income includes the shaded area corresponding to this shift.

In the results presented in this paper, the extreme hypothesis of a fully compensating shift in marginal productivity of work at home is assumed, and whole income is always calculated on the basis of 2,000 hours of work per year. Because the truth doubtless lies part way between the two polar cases of Figure 3.2, we have probably overcompensated for the income effect of a wage change, so that our wage effects overstate the true substitution effects. The only substantive defect of our procedure is that it probably overstates the income effect of the wife's wage on the husband's hours of work.

An important limitation of the analysis just presented and of the interpretation of the empirical results given later in the paper, is the assumption that the average wage is the same as the marginal wage for each worker. Individuals are assumed to have free choice about their hours of work without paying any penalty in their hourly wages for working less than full time. Research in progress will attempt to measure marginal wages without assuming that they equal average wages.

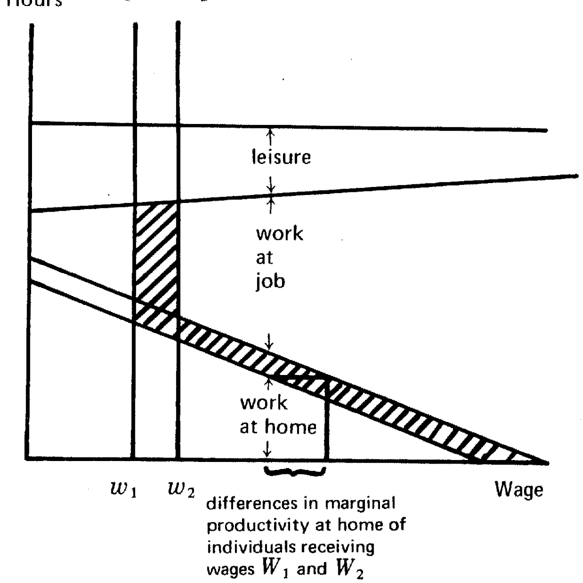
The practical measurement of the variables discussed above—wages and family whole income—is the topic of the bulk of this paper. The problem of measurement cannot be divorced from the problem of the choice of estimation method, so that the discussion must necessarily deal with

FIGURE 3.2. INCOME EFFECTS UNDER ALTERNATIVE ASSUMPTIONS ABOUT THE RELATION BETWEEN THE WAGE AND PRODUCTIVITY AT HOME



case 1

Productivities at home of individuals receiving wages at W_1 and W_2 are the same.



case 2

Productivities are in proportion to wage differences.

some technical econometric issues. Abstracting from income effects, we will begin by considering the simplified cross-section labor-supply equation,

$$L_i = \beta_0 + \beta_1 w_i{}^* + u_i, \tag{3.1}$$

where L_i is hours of work for individual i, $w_i{}^*$ is his wage, and u_i is a random disturbance. Two complications arise in estimating the parameters of this equation.

First, as in many cross sections, the observed wage, w_i, may differ from the true wage, $w_i{}^*$, by a random error of measurement or transitory disturbance, say v_i:

$$w_i = w_i{}^* + v_i. \tag{3.2}$$

We assume that v_i is uncorrelated with $w_i{}^*$.[3] The relation between the measured wage and hours of work is then

$$\begin{aligned} L_i &= \beta_0 + \beta_1 w_i + u_i - \beta_1 v_i \\ &= \beta_0 + \beta_1 w_i + \epsilon_i. \end{aligned} \tag{3.3}$$

Because v_i appears positively in w_i and negatively in ϵ_i, w_i and ϵ_i must necessarily be negatively correlated and the least-squares regression of L_i on w_i must understate the true responsiveness of hours of work to changes in the wage rate. The estimation problem is formally analogous to that of estimating the consumption function from cross-section data.

The second difficulty is peculiar to the study of labor supply: the wage, w_i, is not observed for individuals who decide not to enter the labor force. Omitting the data for these individuals would probably cause a downward bias in the estimated wage response, because the omitted observations are likely to have negative disturbances.

The natural solution to the first of these problems is an instrumental variables estimator. As we shall see, an estimator of this sort is available as a by-product of our proposed solution to the second problem. Suppose we add a second equation to the labor-supply system expressing the hypothesis that the wage measured for an individual depends on certain observable personal characteristics—age, sex, education, experience, and so forth—plus a random disturbance (composed of the error in measurement

[3] This assumption is far from innocuous. It rules out any systematic errors in measuring the wage. For example, it is inconsistent with the hypothesis that the wage should be measured to take into account the amount of unemployment associated with a job. In the absence of adequate data on unemployment, little can be done about this problem. See section V.

or transitory component mentioned above, v_i, and a second error, η_i, corresponding to unobserved personal characteristics). Thus:

$$\begin{aligned} w_i &= \alpha_0 + \alpha_1 x_{i1} + \cdots + \alpha_N x_{iN} + \eta_i + v_i \\ &= \hat{w}_i + \eta_i + v_i, \end{aligned} \tag{3.4}$$

where $x_{i1} \cdots x_{iN}$ are measurements of the characteristics of individual, i; some or all of them may be dummy variables.[4] The application of ordinary least squares to this yields an equation that can be used to calculate an imputed wage for individuals who are not working and for whom a direct wage measurement is not available.[5] The imputed wage is

$$\hat{w}_i = \hat{\alpha}_0 + \hat{\alpha}_1 x_{i1} + \cdots + \hat{\alpha}_N x_{iN}, \tag{3.5}$$

where $\hat{\alpha}_0, \ldots, \hat{\alpha}_N$ are the least-squares estimates.

If there were no errors in observing the wage (if v_i were always zero), then ordinary least squares would be the appropriate estimator for the structural equation, except for the problem that data on wages are missing for some individuals. In this case, the imputed wage could be used in place of the actual wage in a least-squares regression. For nonworkers, the structural equation would become:[6]

$$\begin{aligned} L_i &= \beta_0 + \beta_1(\hat{w}_i + \eta_i) + u_i \\ &= \beta_0 + \beta_1 \hat{w}_i + z_i, \end{aligned} \tag{3.6}$$

where z_i is defined as $u_i + \beta_1 \eta_i$. Because we have implicitly assumed that u_i and v_i are uncorrelated, the variance of z_i is greater than the variance of u_i, and the appropriate estimator is weighted least squares, with lower weights for observations incorporating the imputed wage.

Under our assumption that there are errors in observing the wage w_i, it can no longer be included in a least-squares regression, even when it is available, without giving rise to bias. Consistent estimates are available, however, by applying least squares to the regression obtained by substituting the imputed wage, $\hat{w}_i$, in place of the unobserved true wage, w_i^*,

[4] The actual work here uses log w_i as the left-hand variable, but the point is the same. This equation can be thought of as the demand equation in the supply and demand system for the labor services of the individual.

[5] Much the same approach is used by Edward D. Kalachek and Fredric Q. Raines, "Labor Supply of Income Workers and Negative Income Tax," *Technical Studies,* President's Commission on Income Maintenance Programs (Washington, D.C.: U.S. Government Printing Office, 1970), pp. 159–185.

[6] In this discussion we assume that $\hat{w}_i$ is calculated using the true parameters $\alpha_0, \ldots, \alpha_N$ rather than $\hat{\alpha}_0, \ldots, \hat{\alpha}_N$. In the practical case where only the estimates are available, the situation is more complicated, but all of the conclusions are valid in the limit as the number of observations becomes large.

for all observations in the labor-supply equation. This involves applying least squares to equation (3.4) for all observations, and amounts to the use of an instrumental variables estimator, with the various personal characteristics as instruments. This estimator is used for all results presented in this study.

For the estimator for the labor-supply equation just mentioned, the regression of wages on personal characteristics is a necessary first stage. Section II of this paper is devoted to discussion of an empirical investigation along these lines. Although the basic motivation for this work is to prepare to estimate labor-supply equations, the results are not without interest in themselves.

Measurement of family whole income (section III) proceeds by adding nonlabor income to the sum of the contribution of each family member to the labor component of whole income. The latter is measured as 2,000 hours (or fewer for certain individuals) times the wage rate imputed by the equation of section II. The logic of this method of calculating whole income is essentially the same as that for using the imputed wage in place of the actual wage.

Section IV of the paper discusses the adjustment of wages and whole income to take account of the Federal income tax. Section V discusses the measurement of hours of work and presents a statistical justification for the use of labor income divided by the imputed wage as a measure of hours. Section VI presents the empirical results in summary and in detail, and gives the results of certain hypotheses. Finally, section VII compares the results of this study to those of similar studies, including those of the present volume.

II. THE WAGE EQUATION

Previous studies of the relation between wages and the characteristics of individuals have focused on the estimation of an earnings function rather than a wage function. That is, the left-hand variable has been annual earnings rather than the hourly wage.[7] An earnings function is, in effect, a reduced form of the labor-supply system. Results from earnings functions

[7] Research in this area has been reviewed recently by Zvi Griliches, "Notes on the Role of Education in Production Functions and Growth Accounting," *Education, Income, and Human Capital, Studies in Income and Wealth* 35 (1970); and Jacob Mincer, "Distribution of Labor Incomes: A Survey with Special Reference to the Human Capital Approach," *Journal of Economic Literature* 7 (March 1970).

are not directly relevant for our purposes.[8] Previous investigators have been hampered by the lack of data on hourly earnings and have been forced to adapt their analysis to data on annual earnings, with only fragmentary data on annual hours of work.

The body of data used in the present study is richer in this respect. The SEO reports hourly wage rates (calculated as an average over one week) and a variety of personal characteristics for individuals who worked during the survey week in 1967. Included in our wage study are 8,970 individuals age 14 or over, living in one of the 12 largest Standard Metropolitan Statistical Areas (SMSAs), and employed at a wage or salary earning job. The following characteristics were selected for study: sex, race, years of education, residence at age 16, union membership, and health. The composition of the sample is given in Appendix 3B.

The choice of functional form in a study of this kind is a difficult one. As a tentative choice, we have adopted the following analysis of variance-regression model:

$$\log w_{i,j,k,l,m,n,q,r} = \kappa_{i,j} + \delta_{i,j,k} + \theta_{i,j,l} + \lambda_{i,j,m} + \phi_{i,j,n} + \eta_{i,j,q} + \mu_{i,j,r}, \qquad (3.7)$$

where

$i = 1, 2$ for white and black
$j = 1, 2$ for male and female
$k = 1, \ldots, 9$ for age groups
$l = 1, \ldots, 9$ for years of education groups
$m = 1, \ldots, 12$ for SMSAs
$n = 1, 2$ for U.S. and foreign residence at age 16
$q = 1, 2$ for nonmember or member of a union
$r = 1, 2$ for no health effect on work or some effect.

A full set of interactions is permitted between race, sex, and each of the other characteristics. Within each race-sex group, the effects of the characteristics are assumed to be independent—the age pattern of wages does not vary over education levels, for example. This is an unduly strong restriction, and future work will attempt to relax it within the limitations of the data. Computationally, the present procedure involves separate regressions for each race-sex group, with dummy variables for each of the other characteristics.

The regression results are given in Table 3.1. The coefficients are the

[8] Nor, for that matter, are they appropriate for some of the uses to which they are put. For example, in measuring the return to education, it is whole income (the hourly wage times a standard number of hours) that should be studied, not labor income.

TABLE 3.1. REGRESSION RESULTS FOR THE WAGE EQUATION

	Sex-race Group			
	Male		*Female*	
Characteristic	*White*	*Black*	*White*	*Black*
Constant	1.152	.897	.820	.710
	(.025)	(.027)	(.037)	(.028)
Age				
14 to 15	–.972	–.722	–.429	.028
	(.091)	(.099)	(.121)	(.254)
16 to 17	–.762	–.517	–.264	–.164
	(.056)	(.060)	(.094)	(.063)
18 to 19	–.532	–.316	–.315	–.156
	(.046)	(.042)	(.054)	(.045)
20 to 24	–.263	–.092	–.094	–.066
	(.030)	(.030)	(.041)	(.030)
25 to 34	.000	.000	.000	.000
	—	—	—	—
35 to 44	.118	.075	.026	.013
	(.023)	(.022)	(.038)	(.025)
45 to 54	.162	.080	.036	–.012
	(.024)	(.023)	(.037)	(.026)
55 to 64	.143	–.022	.015	–.105
	(.027)	(.029)	(.043)	(.032)
65 or above	–.075	–.097	–.201	–.205
	(.058)	(.066)	(.082)	(.065)
Years of Education				
0–3	–.380	–.187	–.252	–.358
	(.060)	(.041)	(.106)	(.060)
4–6	–.281	–.152	–.305	–.245
	(.040)	(.030)	(.070)	(.037)
7–9	–.190	–.122	–.235	–.240
	(.024)	(.023)	(.037)	(.024)
10–11	–.092	–.093	–.131	–.177
	(.026)	(.023)	(.038)	(.024)
12	.000	.000	.000	.000
	—	—	—	—
13–14	.098	.106	.117	.179
	(.028)	(.032)	(.038)	(.031)
15	.132	.176	.154	.289
	(.050)	(.065)	(.076)	(.086)
16	.385	.253	.314	.541
	(.032)	(.051)	(.046)	(.047)
17–20	.320	.600	.429	.780
	(.033)	(.062)	(.058)	(.055)

Note: Numbers in parentheses are standard errors.

TABLE 3.1 (cont.)

	Male		Female	
	White	Black	White	Black
SMSA				
Baltimore	–.094	–.079	–.047	–.224
	(.045)	(.032)	(.067)	(.034)
Chicago	–.025	.019	.014	.029
	(.029)	(.031)	(.041)	(.034)
Cleveland	–.138	.020	–.169	–.034
	(.046)	(.044)	(.076)	(.049)
Detroit	.086	.085	–.006	–.035
	(.034)	(.033)	(.052)	(.036)
Houston	–.004	–.237	–.147	–.406
	(.049)	(.037)	(.075)	(.039)
Los Angeles	.021	.096	.042	.031
	(.024)	(.031)	(.035)	(.033)
New York	.000	.000	.000	.000
	—	—	—	—
Philadelphia	–.004	–.021	–.048	–.078
	(.032)	(.034)	(.050)	(.036)
Pittsburgh	–.084	–.055	–.015	–.335
	(.042)	(.074)	(.066)	(.085)
St. Louis	–.005	–.191	–.076	–.209
	(.051)	(.049)	(.070)	(.049)
San Francisco	.099	.183	.043	.085
	(.030)	(.039)	(.042)	(.042)
Washington, D.C.	.053	.004	.082	–.046
	(.037)	(.028)	(.052)	(.029)
Residence at age 16				
U.S.	.000	.000	.000	.000
	—	—	—	—
Foreign	–.146	–.032	–.094	.004
	(.029)	(.073)	(.043)	(.074)
Union membership				
Nonmember	.000	.000	.000	.000
	—	—	—	—
Member	.082	.157	.133	.068
	(.019)	(.017)	(.033)	(.023)
Health				
No effect on work	.000	.000	.000	.000
	—	—	—	—
Some effect	–.105	–.108	–.079	–.053
	(.031)	(.029)	(.048)	(.029)
Standard error of the regression	.435	.360	.487	.356
Sum of squared residuals	576.049	280.065	427.397	231.776

TABLE 3.2. HOURLY WAGES BY AGE

	Male		Female	
	White	Black	White	Black
Age				
14 to 15	1.20	1.19	1.48	2.09
16 to 17	1.48	1.46	1.74	1.73
18 to 19	1.86	1.79	1.66	1.74
20 to 24	2.43	2.24	2.07	1.90
25 to 34	3.16	2.45	2.27	2.03
35 to 44	3.56	2.64	2.33	2.06
45 to 54	3.72	2.65	2.35	2.01
55 to 64	3.65	2.40	2.31	1.83
65 or above	2.93	2.22	1.86	1.66

Note: Estimated wages in New York for individuals with 12 years of education, resident in the United States at age 16, not union members, and in good health. Calculated from Table 3.1.

logs of the multiplicative effects of the associated characteristics. For each characteristic, one group was selected as the reference group and its log coefficient constrained to be zero. The other effects are measured relative to the reference group.

The implications of these results are more easily seen by converting the log coefficients to actual wage levels. In Table 3.2, we give hourly wages for various age groups, holding other characteristics constant. These are estimates of the pure age effect on wages. They are stated in terms of the reference group of New York residents with 12 years of education, but exactly the same pattern of wage variation over age would appear if the results were stated in terms of the wages of any other SMSA-education group. This is a consequence of the assumption of independence.

The results in Table 3.2 show striking variations in the age pattern of wages in the different sex-race groups. Whatever the validity of the assumption of independence of the effects of other characteristics, it is clear that the effects of sex, race, and age are far from independent. The use of single dummies for sex and race would give a seriously distorted view of the differentials in wages by sex and race. For men, the disadvantage suffered by black workers first becomes apparent in the 20 to 24 age group, and becomes much larger from age 25 to 64 in the groups dominated by heads of families. The differential by race (which might be loosely described as a measure of the direct and indirect effects of racial discrimination) is least serious for young workers and most serious for older workers. It should be noted that the differential could be more serious in every age group if a different SMSA were chosen for reference. For example, in

Houston, wages of black workers are almost 25 percent lower than those of white workers, relative to the situation in New York.

Except for teen-agers from age 14 to 17, the differential between white females and white males is larger than the differential by race among men. The striking characteristic of the age pattern of wages for women of both races is the failure of wages to rise with age after the early twenties. This is especially pronounced for black females. The proportional differential between white and black females is substantially smaller than that between white and black males, indicating that black females do not suffer fully from the combined effects of being black (as measured by the differential for males) and of being female (as measured by the differential for whites).

In Table 3.3, we present a similar calculation of wages by years of education, adjusted for other characteristics. For whites of both sexes, the results show the expected upward trend with increasing years of education, except for the male group with graduate education, which is probably heavily weighted with school teachers. The return to completing college is remarkably high for white males. It should be recognized that to the extent that unmeasured personal characteristics are positively correlated with years of education, these results overstate the actual return to additional education. These results confirm the findings of other authors that only a small increase in wages is associated with increased education for black males.

Finally, three remaining characteristics are included at the end of Table 3.1. The first is residence at age 16; foreign residence is associated with 15 percent lower wages for white males and 9 percent lower wages for white females. For blacks, the effect is negligible and statistically insignificant, in accordance with the expectation that the main cost of foreign residence is difficulty with English and the probability that the small number of blacks of foreign origin came from English-speaking countries. The second characteristic is union membership, which has a substantial positive effect on wages, especially for black males and white females. The union effect is not nearly so large as that found by previous investigators using similar data from the 1960 census. This may be a result of the greater disaggregation of the present study (especially by geographical area, not possible with the 1960 data) or because of the tendency for the union differential to shrink during expansionary periods like 1967. Finally, the third characteristic is personal health. Men who report that problems with their health interfered with their work receive wages 10 percent lower than otherwise; the similar effect for women is between 4 and 7 percent. Of course, the main effect of poor health is probably not so much a reduction in wages as a reduction in hours of work, in many cases to zero.

TABLE 3.3. HOURLY WAGES BY YEARS OF EDUCATION

Years of Education	Male		Female	
	White	*Black*	*White*	*Black*
0–3	2.16	2.03	1.77	1.42
4–6	2.39	2.11	1.67	1.59
7–9	2.62	2.17	1.80	1.60
10–11	2.89	2.23	1.99	1.70
12	3.16	2.45	2.27	2.03
13–14	3.49	2.73	2.55	2.43
15	3.61	2.92	2.65	2.72
16	4.65	3.16	3.11	3.50
17–20	4.36	4.47	3.49	4.44

Note: Estimated wage in New York for individuals aged 25 to 34 years, resident in the United States at age 16, not members of unions, and in good health. Calculated from Table 3.1.

Although the coefficients of the first-stage regression just presented are almost without exception entirely reasonable, there is still a great deal of variance around the regression model. The standard errors of the four regressions are all between .35 and .50, indicating that the average error in imputing wages on the basis of personal characteristics is between 35 and 50 percent.[9]

III. THE CALCULATION OF WHOLE INCOME

The whole income of a family is defined as its total nonwage income plus the dollar value of the time of each of its members. In this section, we discuss the measurement of these two components of whole income from the SEO data. The following section discusses modifications of these figures to take into account the Federal income tax.

The SEO presents data on family income according to the definitions used in the Current Population Survey. For our purposes we use only the category of unearned income—comprising rental income, interest and dividends, pensions, social security, and other nonwage income. Several adjustments must be made to the reported total of these for our purposes: (1) the imputed value of durable goods must be added; (2) the treatment of

[9] Because the left-hand variable is in log form, the standard error has the dimension of a percentage error.

interest receipts and expenditures must be put on a consistent basis; (3) the interest component of business and farm income must be added.

(1) The value of three kinds of durables are reported in the SEO: owner-occupied homes, other real estate, and automobiles. For most families, these probably account for the greater part of the total value of durable goods, but for poorer families, the omission of the value of clothing and furniture is significant. Imputed income from durables was calculated as 6 percent of the value of real estate plus 12 percent of the value of automobiles, less rental income.

(2) The CPS definitions treat interest receipts as income but interest payments as part of expenditure. We converted to a net interest income basis by subtracting an estimate of interest payments, calculated as 6 percent of the value of mortgages plus 12 percent of the value of automobile loans plus 15 percent of the value of installment and other credit. For real estate and automobiles, this has the effect of reducing the previous imputation to one on the owner's equity, rather than on the total value.

(3) The interest component of business and farm income was estimated as 33 percent of total business and farm income for each family. Because families with substantial amounts of income from this source were excluded from the study, refinement of this calculation did not seem warranted.

The annual value of each individual's time was calculated as the product of his hourly wage, imputed by the method of section II, and the number of hours available for work. For most adults a full work year of 2,000 hours was assumed. Individuals in school were assumed to have 500 hours available. Individuals reporting physical disabilities that prevented work or limited their amount of work were assigned potential hours of work between 0 and 2,000 hours according to a formula that took into account the nature and length of the disability.

IV. THE FEDERAL INCOME TAX

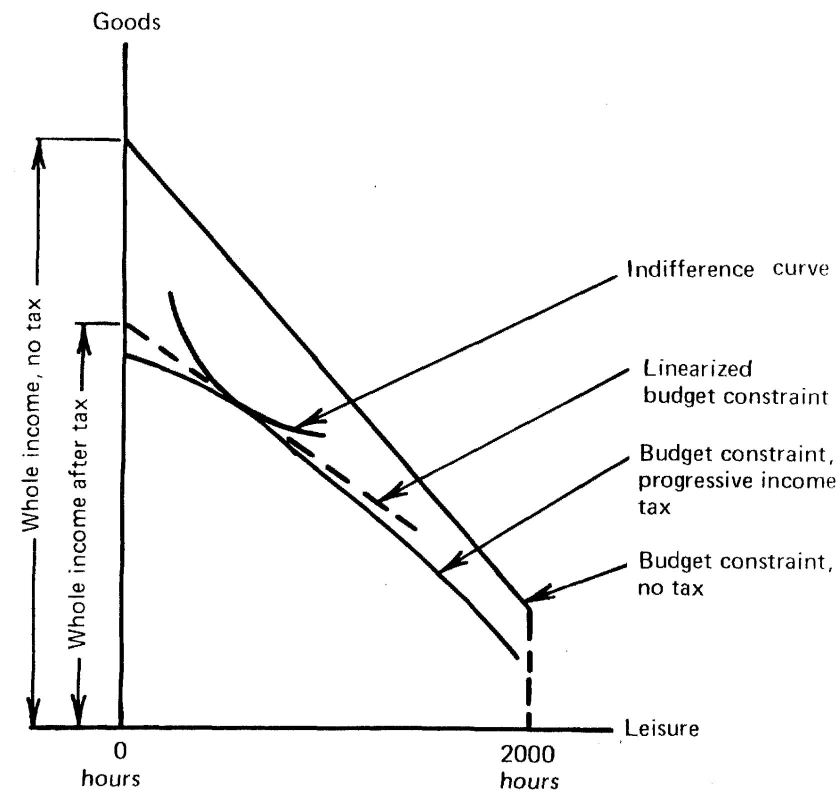
In principle, we need to take into account all taxes imposed on individual or family income, especially those having rates that vary according to income. The most important such tax, however, is the Federal personal income tax; this is the only tax explicitly incorporated in the present study.

The logic of our approach to the income tax can best be seen in the case of a single individual. We have discussed the behavior of an individual in terms of two parameters of his budget constraint—its slope, as measured

by his wage, and its vertical intercept, whole income. Our plan is to develop a method for treating an individual facing an income tax in terms of the same parameters, adjusted for the effect of the tax. The new slope will be the wage after tax, and the new vertical intercept will be whole income after tax. The only obstacle to this treatment is that, under a progressive tax, the slope of the new budget constraint (the wage after tax) is not constant, but declines with increasing hours of work. The after-tax constraint is not a straight line, but is a curve with its concave side toward the origin. Previously we have taken advantage of the fact that any *straight-line budget constraint can be described fully in terms of its slope* and its vertical intercept. Curving budget constraints do not seem to lend themselves to such an easy characterization.

Our approach, shown in Figure 3.3, is to replace the true, curving budget constraint facing an individual subject to a progressive income tax

FIGURE 3.3. ACTUAL AND LINEARIZED BUDGET CONSTRAINTS UNDER A PROGRESSIVE INCOME TAX



with a straight line that is tangent to the true constraint at the point of his actual consumption of leisure and goods. We define the individual's whole income after tax as the vertical intercept of this tangent and his wage after tax as the slope of the tangent; the latter is the wage before tax multiplied by one minus the marginal income tax rate at the individual's actual taxable income. The justification for these definitions is the following: Any individual whose indifference curves have the normal curvature (convex to the origin) will behave in the same way whether he faces the curving budget constraint of a progressive tax or a straight-line constraint, provided the latter is tangent to the former at his point of actual consumption under the former. There is no element of approximation in this procedure.

The tangent budget constraint can be thought of as the result of replacing the progressive income tax with a proportional tax on wage income plus a lump-sum tax. In fact, this is the way that the wage rate and whole income are adjusted for the effect of the tax in our empirical work. First, the actual amount of tax and the marginal tax rate are estimated for each individual by a method described below. Second, each individual's imputed wage is adjusted downward by multiplying by one minus the marginal tax rate. Third, the lum-sum component of the tax is calculated as the difference between the actual amount of tax paid and the amount that would have been paid if only wage earnings had been taxed but the tax were proportional at the marginal rate. Finally, family whole income after tax is calculated as the sum of potential labor income for each family member (the product of the wage after tax and potential hours) plus family nonwage income less the sum of the lump-sum components of the income tax for each member. This procedure automatically accounts for the fact that some (but by no means all) components of nonwage income are subject to the income tax.

The data in the SEO are adequate for a rough calculation of income tax liability for each individual.[10] The only important component of taxable income omitted altogether is income from capital gains. Information about deductible expenditures, however, is generally lacking, as is complete information on support necessary to assign dependents correctly.

The following assumptions were made in calculating tax liability and tax rates:

1. All married couples living together file jointly;
2. Any person with dependents but not filing jointly files as a head of household;

[10] The author is indebted to Benjamin Okner for advice in this part of the work.

3. All taxpayers use the standard deduction formula;
4. Anyone with taxable income over \$600 is self-supporting; all other individuals are dependents of their parents or of the head of the family;
5. All taxable nonwage income is income of the head of the family, except taxable pension income, which is distributed equally among all family members age 65 or over.

With these assumptions, we calculated net taxable income for each individual and looked up his tax liability and marginal tax rate in the tables for the 1966 Federal income tax.

V. THE MEASUREMENT OF HOURS OF WORK

A problem encountered in almost any study of annual hours of work is that a single survey cannot measure individual hours of work over a period as long as a year. No person can recall with any usable accuracy how many hours he has worked in the past year unless he has an unusually regular schedule. As a result, we need to find an indirect approach that makes use of the limited data available to estimate hours of work. In this section we discuss the statistical properties of the estimates obtained by dividing annual wage income by the imputed or observed wage rates. We then go on to discuss the difficulties taking into account unemployment and time spent searching for work in the measurement of total hours of work effort.

The following system of equations describes the estimation problem:

$$L_i = \beta_0 + \beta_1 w_i^* + u_i \qquad (3.8)$$
$$Y_i = w_i^* L_i \qquad (3.9)$$
$$w_i^* = \hat{w}_i + \eta_i \qquad (3.10)$$
$$w_i = \hat{w}_i + \eta_i + v_i. \qquad (3.11)$$

The first equation is the structural equation for hours worked. The second equation is an identity linking wage income, Y_i, the actual wage, w_i^*, and hours of work. The third equation gives the relation between the actual wage and the imputed wage, $\hat{w}_i$, and the fourth equation gives the relation between the observed and imputed wages.

We study the properties of two estimates of hours of work: wage income divided by the imputed wage, $(Y_i/\hat{w}_i)$, and wage income divided by the observed wage, (Y_i/w_i). Neither is a perfect estimate of true hours, (Y_i/w_i^*); $(Y_i/\hat{w}_i)$ is contaminated by the error in imputing the wage, η_i, and (Y_i/w_i) by the error in observing the wage, v_i. We will argue, however, that there are good reasons to believe, on statistical grounds, that $(Y_i/\hat{w}_i)$ will give better results for our purposes.

Because the actual wage and hours are unobserved, we ask the following question: What can we learn from studying the observable relation between estimated hours and the imputed wage, $\hat{w}_i$?[11] More precisely, what is the regression relation or conditional expectation of estimated hours, given $\hat{w}_i$? Now, for the first estimate, $(Y_i/\hat{w}_i)$,

$$\begin{aligned} E(Y_i/\hat{w}_i|\hat{w}_i) &= (1/\hat{w}_i)E(Y_i|\hat{w}_i) \\ &= (1/\hat{w}_i)E(w_i^*L_i|\hat{w}_i) \\ &= \frac{1}{\hat{w}_i}E\{(\hat{w}_i+\eta_i)[\beta_0+\beta_1(\hat{w}_i+\eta_i)+\eta_i]|\hat{w}_i\} \qquad (3.12) \\ &= \beta_0+\beta_1\hat{w}_i+\frac{1}{\hat{w}_i}\beta_1\sigma_\eta^2, \end{aligned}$$

where σ_η^2 is the variance of η_i. Except for the third term, the conditional expectation has the same form as the structural equation. Note that the error, $(1/\hat{w}_i)\beta_1\sigma_\eta^2$, is proportional to the slope, β_1. In particular, if hours are unresponsive to the wage $(\beta_1 = 0)$, then the regression of $(Y_i/\hat{w}_i)$ on $\hat{w}_i$ will give unbiased results. If β_1 is positive, estimates of it will be biased slightly downward, because the error decreases as $\hat{w}_i$ increases. Where the true slope of the labor-supply function is β_1, the slope of the conditional expectation, equation (3.12), is $(1 - \sigma_\eta^2/\hat{w}_i^2)\beta_1$. The value of σ_η^2, the variance of the error in imputing the wage, is not known, but may be about 0.75. Thus the relative error in the slope is

$$0.75/2.00^2 = 0.19 \text{ at } \hat{w}_i = \$2.00, \text{ and } 0.75/4.00^2 = 0.05 \text{ at } \hat{w}_i = \$4.00.$$

Statistical analysis of the second estimate of L_i, (Y_i/w_i), is not as straightforward. We have

$$\begin{aligned} E(Y_i/w_i|\hat{w}_i) &= E\left(\frac{w_i^*L_i}{w_i}|\hat{w}_i\right) \\ &= E\left\{\frac{(\hat{w}_i+\eta_i)[\beta_0+\beta_1(\hat{w}_i+\eta_i)+u_i]}{\hat{w}_i+\eta_i+v_i}\right\}. \end{aligned} \qquad (3.13)$$

This expectation involves a nonquadratic function of the two random variables η_i and v_i, so it cannot be evaluated without further assumptions

[11] The author is grateful to Harold Watts for pointing out an important error in a previous version of what follows.

about their distributions. We can get a rough impression, however, by taking a quadratic approximation to equation (3.13):[12]

$$\begin{aligned} E(Y_i/w_i|\hat{w}_i) &\doteq E(\beta_0+\beta_1\hat{w}_i+u_i)-\frac{1}{\hat{w}_i}(\beta_0+\beta_1\hat{w}_i+u_i)v_i \\ &\quad +\beta_1\eta_i+\frac{1}{\hat{w}_i^2}(\beta_0+\beta_1\hat{w}_i+u_i)v_i^2 \\ &\quad +\frac{\beta_0+u_i}{\hat{w}_i^2}(\eta_iv_i) \qquad (3.14) \\ &\doteq \beta_0+\beta_1\hat{w}_i+\frac{1}{\hat{w}_i^2}(\beta_0+\beta_1\hat{w}_i)\sigma_v^2. \end{aligned}$$

The slope of this relation is $[1 - (2\beta_0/\beta_1\hat{w}_i^3 + 1/\hat{w}_i^2)\sigma_v^2]\beta_1$. If σ_v^2 has the same value as σ_η^2, 0.75, $\beta_0 = 1500$ and $\beta_1 = 200$, then the relative error in the slope is $2 \cdot 1500/200 \cdot 2.00^3 + 1/2.00^2 \cdot 0.75 = 1.59$ at a wage of \$2.00 (the slope is actually reversed), and $2 \cdot 1500/200 \cdot 4.00^3 + 1/4.00^2 \cdot 0.75 = 0.22$ at a wage of \$4.00. Thus, even when the error in imputing the wage is the same size on the average as the error in measuring the wage, it is far better to use the imputed wage in the denominator of the estimate of hours of work.

By no means are all the obstacles to the satisfactory measurement of hours of work purely statistical. One of the most serious difficulties is in the treatment of time spent searching for work. Because our imputed wage, $\hat{w}$, is estimated on the basis of the wages received by individuals for their hours of actual work, neither it nor the measure of imputed annual hours of work derived from it take into account the time required to find a job. In labor markets that are substantially out of equilibrium on the side of excess supply, this could result in a serious underestimation of the total hours of work (including job search) for groups in the labor force experiencing high rates of unemployment. If hours spent looking for work could be measured directly, this figure could be added to our measure of hours at work to get a more comprehensive measure of labor supply. But even with data vastly more detailed than those available it would be almost impossible to separate hours spent looking for work from those spent enjoying leisure. The work reported here does not include any adjustment for periods of unemployment in measuring the amount of labor supplied by an individual.[13] Fortunately, the year we study, 1966, was one

[12] If the distributions of η_i and v_i are symmetric, this is also a cubic approximation because the terms involving third moments vanish.

[13] For an empirical study of unemployment in much the same framework as the one used here, see Robert Hall, "Why is the Unemployment Rate So High at Full Employment?" *Brookings Papers on Economic Activity* 3 (December 1970).

of extraordinarily high employment,[14] so the amount of excessive search time induced by excess supply is probably fairly small.

VI. RESULTS

Individuals meeting the following criteria were included in the sample:

1. Resident in one of the 12 large metropolitan areas identified in the SEO;
2. Not in school in 1966;
3. No disability in 1966 that limited the amount of work the individual could perform;
4. Not in a family with total self-employment income over $1,000 in 1966;
5. Not in a family receiving public assistance;
6. Not a male head of family without a wife;
7. Either white or black;
8. Age 14 years or older.

All but the last three of these restrictions are substantive. The first limits the sample to a relatively homogeneous urban population for whom precise geographical information is available. The second eliminates the difficulty that hours spent attending school voluntarily should be treated in the same way as hours spent working, but hours in school cannot be measured. It also eliminates 14-, 15-, and 16-year-olds who are subject to a variety of restrictions on their hours of work because of compulsory school attendance. The third restriction is necessary because of the tremendous variety of physical and mental disabilities reported in the data. Hours of work of disabled individuals is properly the subject of a separate study. The fourth restriction is a consequence of the difficulty in separating the capital and labor components of proprietary income, and in allocating the labor component among members of the family. The fifth restriction is in many ways the most serious; it resulted in the exclusion of about 900 families. Again, the study of hours of work of members of families receiving public assistance is a separate project in itself.

The variables used in the analysis are defined as follows:

(1) *Annual Hours of Work.* This was estimated by dividing annual earnings by the imputed wage. See section V.

(2) *Position in Family.* The following seven categories were used: husband, wife present; wife, husband present; female head of family; son or other male relative, not head of family; daughter or other female relative, not head of family; single man; single woman.

[14] The unemployment rate averaged 3.8 percent over the year.

(3) *Race.* This was classified as either white or black.

(4) *Age.* The following three categories were used: 14 through 19 years; 20 through 59 years; 60 years and older.

(5) *Number of Adults (individuals 14 years or older) in Family.* This was classified according to the following four categories: 1 adult; 2 adults; 3 or 4 adults; 5 or more adults.

(6) *Children in Family.* The following four categories were used: no children; children of preschool age only (6 years or younger in March 1967); children of school age (7 through 13 years) only; children of both ages.

(7) *Whole Income per Adult.* Whole income after tax was calculated as described in sections III and IV, divided by the number of adults in the family and deflated by the price index given in Appendix 3B. The following five categories were defined: less than $3,000 per year; $3,000 or more, but less than $3,750; $3,750 to $4,500; $4,500 to $5,500; $5,500 or more.

(8) *Hourly Wage.* The imputed hourly wage was calculated from the regression equation of section II, adjusted for the Federal income tax as described in section IV, and also deflated by the price index of Appendix 3C. The following six categories were defined: less than $1.50 per hour; $1.50 or more, but less than $1.75; $1.75 to $2.00; $2.00 to $2.50; $2.50 to $3.00; $3.00 or more.

The reduction of income and wages to categorical variables permits the use of unrestrictive analysis of variance functional specifications that are nonetheless linear in their parameters. For example, by classifying the wage into six categories, we approximate the wage effect by a step function with six steps, each determined by a separate parameter. We avoid the unduly restrictive linear specification implied by the use of the wage itself in a linear regression. This is particularly important in specifying a regression where the left-hand variable, hours of work, is subject to a constraint on its variation—it cannot become negative.

The argument given in section I in favor of the use of imputed rather than actual wages was presented in terms of the linear regression function that we have just ruled out. In Appendix 3A, we discuss the problem of applying the instrumental-variables estimator to a structural equation containing a step-function specification. Our results show that a slight blurring of the estimated coefficients will take place in general, but that there will not be any systematic bias in the overall estimates of the wage or income effects. The blurring will be least serious if the coefficients change smoothly from one step to the next.

The first step in the study of hours of work was the preparation of an exhaustive cross-tabulation of average hours of work by all seven characteristics. This is the least restrictive regression model possible—it permits the wage and income effects to depend on each other and on all five demographic characteristics. The result is a set of several hundred tables, one for each group defined by position in the family, age, race, number of adults in the family, and age of children. In each table, there is a row for each income class and a column for each wage class. Reading across a row, we find the effects of variations in the wage rate on the hours of work of individuals in families in a single whole income class. Because the wages of the individuals themselves enter the calculation of the whole incomes of their families, individuals in the high wage classes toward the right of the tables live in families with less income from other sources than the families of the individuals in the lower wage groups toward the left. This is the Hicksian income compensation discussed in the introduction. It permits us to read the pure substitution effects of wage changes directly from the tables.

Because there are more than ten times as many cells in these tables as there are individuals in the sample, most of the cells in most of the tables are empty. The tables for very small and very large families and other smaller groups are so sparsely filled that they give very little information. For other groups, the unrestricted tabulations are of some interest. Some of the better-populated tables are reproduced in Table 3.4. In each cell, we give the average annual hours of work, the standard error of the average,[15] the participation rate (defined as the proportion of the individuals in the cell who worked 40 or more hours in 1966), and the number of individuals, N, in the cell.

The table for white husbands, age 20 to 59, with wives and preschool children, illustrates some of the strengths and weaknesses of this kind of study. Classifying by demographic characteristics greatly reduces

[15] Defined as

$$\sqrt{\frac{1}{N(N-1)}\sum_{i=1}^{N}(H_i-\bar{H})^2},$$

where N is the number of observations in the cell, $H_1, \ldots, H_N$ are the various observations on hours, and

$$\bar{H}=\frac{1}{N}\sum_{i=1}^{N}H_i.$$

TABLE 3.4. UNRESTRICTED CROSS-TABULATIONS, ANNUAL HOURS OF WORK BY INCOME AND WAGE

A. Husbands, White, Age 20 to 59, 2 Adults and Preschool Children in Family

Annual Whole Income Per Adult		Hourly Wage (dollars) 0–1.50	1.50–1.75	1.75–2.00	2.00–2.50	2.50–3.00	3.00+
$0–$3000	Mean	1705 (1148)	3446 (312)	2949 (729)	2067 (714)	3040 (406)	
	Participation rate	1.00	1.00	1.00	1.00	1.00	
	Number	3	3	3	2	5	0
$3000–$3750	Mean			2299 (154)	2291 (174)	2619 (640)	2073 (463)
	Participation rate			1.00	1.00	1.00	1.00
	Number	0	0	19	28	2	2
$3750–$4500	Mean			1643 (711)	2220 (102)	2223 (78)	1993 (282)
	Participation rate			1.00	1.00	1.00	1.00
	Number	0	1	3	62	74	10
$4500–$5500	Mean				1030 (267)	2052 (58)	2210 (70)
	Participation rate				1.00	1.00	1.00
	Number	0	0	0	6	89	97
$5500+	Mean					1780 (290)	1880 (50)
	Participation rate					1.00	.99
	Number	0	0	0	0	5	92

B. Single Men, Black, Age 20 to 59

Annual Whole Income Per Adult		Hourly Wage (dollars) 0–1.50	1.50–1.75	1.75–2.00	2.00–2.50	2.50–3.00	3.00+
$0–$3000	Mean						
	Participation rate						
	Number	0	0	1	0	0	0
$3000–$3750	Mean	2475 (45)	1486 (186)				
	Participation rate	1.00	1.00				
	Number	2	13	1	0	0	0
$3750–$4500	Mean			1623 (78)	1212 (153)		
	Participation rate			1.00	.95		
	Number	0	1	37	19	0	0
$4500–$5500	Mean			1203 (314)	1547 (72)	926 (153)	
	Participation rate			.83	.97	1.00	
	Number	0	0	6	98	10	0

Note: Numbers in parentheses are standard errors

TABLE 3.4 (cont.)

B. Single Men, Black, Age 20 to 59

Annual Whole Income Per Adult		Hourly Wage (dollars) 0–1.50	1.50–1.75	1.75–2.00	2.00–2.50	2.50–3.00	3.00+
$5500+	Mean				1824 (192)	1655 (102)	1405 (186)
	Participation rate				1.00	1.00	.94
	Number	0	0	0	18	50	16

C. Wives, White, Age 20 to 59, School-Age Children Only

Annual Whole Income Per Adult		Hourly Wage (dollars) 0–1.50	1.50–1.75	1.75–2.00	2.00–2.50	2.50–3.00	3.00+
$0–$3000	Mean						
	Participation rate						
	Number	1	0	0	0	0	0
$3000–$3750	Mean	172 (121)	3 (3)	264 (264)			
	Participation rate	.15	.00	.33			
	Number	13	5	3	1	0	0
$3750–$4500	Mean	105 (57)	91 (66)	543 (251)			
	Participation rate	.25	.11	.42			
	Number	12	18	12	1	0	0
$4500–$5500	Mean	0 (0)	114 (41)	138 (42)	398 (158)		
	Participation rate	.00	.25	.22	.57		
	Number	3	36	76	14	1	0
$5500+	Mean	0 (0)	3 (3)	113 (49)	229 (98)	256 (212)	683 (544)
	Participation rate	.00	.00	.14	.19	.22	.67
	Number	2	13	51	31	9	3

D. Husbands, White, Age 60 and above, 2 Adults and No Children in Family

Annual Whole Income Per Adult		Hourly Wage (dollars) 0–1.50	1.50–1.75	1.75–2.00	2.00–2.50	2.50–3.00	3.00+
$0–$3000	Mean		2998 (502)	2310 (648)		1955 (94)	
	Participation rate		1.00	1.00		1.00	
	Number	1	3	4	1	4	0
$3000–$3750	Mean	2948 (1052)	1204 (952)	1611 (507)	1727 (425)	2844 (678)	
	Participation rate	1.00	.50	.60	.73	1.00	
	Number	2	4	10	11	5	0

TABLE 3.4 (cont.)

D. Husbands, White, Age 60 and above, 2 Adults and No Children in Family

Annual Whole Income Per Adult		Hourly Wage (dollars) 0–1.50	1.50–1.75	1.75–2.00	2.00–2.50	2.50–3.00	3.00+
$3750–$4500	Mean	593 (593)	718 (373)	1691 (310)	1060 (213)	1892 (139)	1970 (357)
	Participation rate	.50	.42	.71	.58	1.00	1.00
	Number	2	12	17	24	13	5
$4500–$5500	Mean		675 (675)	221 (122)	367 (95)	1585 (207)	1931 (185)
	Participation rate		.20	.14	.40	.84	.95
	Number	1	5	29	45	25	19
$5500+	Mean	0 (0)	170 (170)	716 (347)	133 (64)	463 (169)	1121 (138)
	Participation rate	.00	.13	.24	.17	.30	.74
	Number	2	8	17	48	33	46

E. Sons and Other Male Relatives, White, Age 20 to 59, 3 or 4 Adults and No Children in Family

Annual Whole Income Per Adult		Hourly Wage (dollars) 0–1.50	1.50–1.75	1.75–2.00	2.00–2.50	2.50–3.00	3.00+
$0–$3000	Mean			1933 (18)	1584 (376)	1651 (195)	1547 (104)
	Participation rate			1.00	1.00	1.00	1.00
	Number	0	1	2	3	3	2
$3000–$3750	Mean		725 (335)	1948 (171)	1287 (249)	1365 (126)	
	Participation rate		.67	1.00	1.00	1.00	
	Number	0	6	6	4	4	0
$3750–$4500	Mean		1052 (137)		1033 (267)	1227 (274)	511 (295)
	Participation rate		1.00		.90	.88	.67
	Number	0	2	1	10	8	3
$4500–$5500	Mean			1723 (350)	1540 (183)	1511 (136)	1049 (195)
	Participation rate			1.00	1.00	1.00	1.00
	Number	0	0	3	20	20	5
$5500+	Mean					1927 (309)	1010 (152)
	Participation rate					1.00	.86
	Number	0	0	0	1	6	14

the variation in income and wages.[16] Most members of this highly favored group are in the top three wage and income groups. Further, there is a strong tendency for the observations to fall mainly in the cells along the diagonal of the table—husbands with high wages have families with large whole incomes. The principal explanation is simply that the husband's wage income is the dominant component of family whole income. The tendency is accentuated by the fact that well-paid men tend to have wives with higher imputed wages, and also tend to have more property income. These are the remaining important components of whole income. In spite of this difficulty, a great deal can be learned from the comparison of adjacent cells when both have reasonably large numbers of observations. The three such comparisons that can be made in this table in the horizontal direction suggest that the substitution effect of wage changes cannot be very strong for this group. In the second income group, hours decline by 8 per year between the third and fourth wage groups. For the third income group, there is an increase of 3 hours per year between the fourth and fifth wage groups. Only in the fourth income group is there evidence of a noticeable positive effect—between the fifth and sixth wage groups, hours rise by 158 per year or approximately one month at 40 hours per week. The evidence in favor of a negative income effect is considerably stronger. In the fourth wage group, hours drop by 71 between the second and third income groups; in the fifth wage group by 171 hours between the third and fourth income groups; and in the sixth wage group by 330 hours between the fourth and fifth income groups.

The results for single people suffer even more from the close association between wages and whole income. Table 3.4B gives the results for black single men. The wage and income effects that can be discerned from the comparison of adjacent cells are generally of the wrong sign. Reading down the diagonal from the cell in the second wage and second income group, we can see that the perverse wage and income effects cancel out, and individuals tend to work 1,500 to 1,600 hours per year whatever their wage. There is very little evidence of a positive response of hours of work to higher wages.

The results for wives with school-age children, shown in Table 3.4C,

[16] If we classified by all of the characteristics used in the wage regression of section II (used to calculate imputed wages here), there would be no variation in wages and almost none in whole income within each table. This illustrates the order condition for identification—there must be some variables in the first-stage regression that are not in the second stage. Wage and income effects can be identified only if there is something causing them to vary that does not have an independent effect on labor supply. In this study, age, education, location, residence at age 16, union membership, and health are sources of such variation.

suffer much less from correlation between wages and income. The cells below the diagonal are well-populated because many women live in families with a higher whole income per adult than they would have if all members of the family had the same wage as the wife. The table shows that wives generally work rather little, but those with higher wages work substantially more than those with lower wages in the same whole income group. Within a wage group, those with higher incomes work much less than those with lower incomes. These results seem to confirm the general belief that wives are quite sensitive to economic variables in their decisions about working.

Husbands of retirement age are similarly responsive to wages and income, as shown in Table 3.4D. Here, large numbers of individuals appear below the diagonal because they receive income from sources other than work, including income from property, pensions, and so forth.[17]

Table 3.4E presents results for the rather heterogeneous group of adult sons and other male relatives living in families of which they are not the head. These individuals seem to work substantially less than full time, but it is difficult to discern any systematic differences by wage or income groups. In the more detailed discussion that follows, we will suggest that part of the difficulty may arise from the fact that whole income per adult in the complete family may not be the appropriate measure of income for individuals who are not well integrated in the family.

Most of the tables produced in the first phase of the study are subject to random fluctuations that make them difficult to interpret. The logical way to overcome this problem is to use a procedure for averaging the results for related groups. For example, in the case of husbands, we would like to calculate a set of wage effects that represent the average over husbands with children of various ages and with different numbers of relatives living with them. The natural way to carry out this kind of averaging is by estimating the parameters of a regression function in which the effects of some variables are independent of those of others. That is, by excluding interactions we can interpret the regression coefficients as averages for the corresponding effects. The advantage of this method over the more direct method of calculating marginal effects by summing the rows and columns of the tables is that it takes proper account of the unequal numbers of individuals in the cells.

Our next step is to present detailed regression results for various groups in the labor force. We allow a full set of interactions between sex, race, and family position, and all the economic and other demographic characteristics. That is, separate regressions are presented for each group defined by sex, race, and family position, except in the case of teen-agers,

[17] We treat Social Security benefits as nonlabor income, but do not take into account the implicit tax on wages imposed by the system. Work in progress will attempt to incorporate the tax.

who are separated by race and sex only. For each group, we have chosen a specification that permitted certain interactions (for example, between the age and wage effects) and excluded other interactions. In particular, we have assumed that the income and wage effects are independent in every group. Because this assumption is, at best, an approximation valid over a restricted range of incomes and wages, we have eliminated all families whose whole incomes exceed $5,500 per adult per year and all individuals whose imputed wage exceeds $3.00 per hour. The latter exclusion is particularly important for white husbands.

Results for the detailed regressions appear in Table 3.5. Each box contains a regression coefficient and its standard error. Coefficients that are normalized at zero have no standard error reported, and coefficients that could not be estimated because of lack of data are replaced by a dash. In discussing the results, we will refer occasionally to Table 3.6, which presents results for statistical tests comparing blacks and whites of the same sex and family position, and to Table 3.7, which presents the results of tests of the null hypothesis that there are no wage effects.

Our first results are for husbands. The specification for this group permits interactions between income and age effects, between wage and age effects, and between the effects of income and the number of adults in the family. It also has independent effects for the presence and age of children. Results for white husbands are given in Table 3.5A. These results for all members of the group confirm the impression given by the cross-tabulation of the subset of the group of age 20 to 59 having two adults and children of preschool age only, given previously in Table 3.4A—income and wage effects are present and have the expected sign, but they are not very strong. The hypothesis that wage effects are absent for the 20 to 59 age group cannot be rejected. The evidence on this point is fairly good, in that the standard errors for the wage effects in the third and fifth wage groups are small (very few white husbands appear in the first two groups). For husbands of retirement age, the income and wage effects are much stronger. A man of 60 or over, earning $2.00 to $2.50 per hour after taxes, with a wife but no children living with him, works 1,613 hours per year (essentially full time) if the whole income of his family is $6,000 to $7,500 per year, but only 431 hours per year if the whole income is $9,000 to $11,000. The wage effects are equally strong; the hypothesis of their absence is overwhelmingly rejected, with an F-statistic of 19.9.

The presence of additional adults in the family reduces the hours of work of the husband in all but the highest income group. Our discussion of this observation anticipates the results of Table 3.5B showing exactly the opposite effect for black husbands. We have suggested previously that our use of whole income per adult overstates the true income correction

TABLE 3.5. DETAILED REGRESSIONS FOR ANNUAL HOURS OF WORK

A. White Husbands

	Whole Income (dollars)				
	0–3000	*3000–3750*	*3750–4500*	*4500–5500*	
Age					
20 to 59	447 (158)	131 (93)	0	–212 (74)	
60 or above	–502 (212)	–411 (162)	–887 (135)	–1593 (122)	
No. adults					
2	0	0	0	0	
3–4	–418 (168)	– 32 (107)	– 84 (97)	78 (129)	
5 or more	–781 (215)	–281 (203)	–100 (307)	112 (495)	
	Wage (dollars per hour)				
	0–1.50	*1.50–1.75*	*1.75–2.00*	*2.00–2.50*	*2.50–3.00*
Age					
20 to 59	209 (403)	– 54 (279)	83 (127)	0	62 (65)
60 or above	–177 (284)	–234 (174)	18 (126)	0	967 (127)
	Presence of Children				
	None	*Preschool Only*	*School-age Only*	*Both*	
	0	122 (67)	232 (79)	262 (76)	

Standard Error: 852 hrs. per yr.
N: 1397
Constant: 2024 (74)

B. Black Husbands

	Whole Income (dollars)				
	0–3000	*3000–3750*	*3750–4500*	*4500–5500*	
Age					
20 to 59	171 (79)	60 (55)	0	–105 (74)	
60 or above	55 (132)	57 (134)	–600 (130)	–812 (157)	
No. adults					
2	0	0	0	0	
3–4	64 (81)	95 (70)	64 (89)	– 22 (173)	
5 or more	42 (104)	570 (163)	211 (475)		
	Wage (dollars per hour)				
	0–1.50	*1.50–1.75*	*1.75–2.00*	*2.00–2.50*	*2.50–3.00*
Age					
20 to 59	611 (390)	–121 (119)	– 55 (60)	0	–162 (44)
60 or above	–1093 (152)	–928 (136)	–773 (132)	0	93 (254)
	Presence of Children				
	None	*Preschool Only*	*School-age Only*	*Both*	
	0	72 (50)	179 (54)	224 (50)	

Standard Error: 668 hrs. per yr.
N: 1504
Constant: 1775 (49)

Note: Numbers in parentheses are standard errors. There is no part *O* to this table.

TABLE 3.5. (cont.)

	C. White Wives				D. Black Wives			
	Whole Income (dollars)				Whole Income (dollars)			
	0–3000	3000–3750	3750–4500	4500–5500	0–3000	3000–3750	3750–4500	4500–5500
Age								
20 to 59	266	336	0	−181	299	253	0	−387
	(106)	(69)		(49)	(94)	(69)		(87)
60 or above	−159	−209	−509	−521	−989	−945	−1402	−1572
	(204)	(218)	(170)	(160)	(907)	(908)	(879)	(914)
No. adults								
2	0	0	0	0	0	0	0	0
3–4	−253	−425	−245	−202	−154	− 71	–	264
	(115)	(74)	(58)	(61)	(100)	(90)		(205)
5 or more	−372	−609	−459	−223	−319	−423	629	–
	(143)	(121)	(168)	(270)	(130)	(168)	(397)	

	C. White Wives: Wage (dollars per hour)					D. Black Wives: Wage (dollars per hour)				
	0–1.50	1.50–1.75	1.75–2.00	2.00–2.50	2.50–3.00	0–1.50	1.50–1.75	1.75–2.00	2.00–2.50	2.50–3.00
Age										
20 to 59	−647	−303	0	515	504	−630	−310	0	−189	416
	(77)	(64)		(69)	(125)	(126)	(138)		(241)	(265)
60 or above	−473	−170	0	616	470	217	387	0	265	−250
	(155)	(191)		(266)	(522)	(886)	(912)		(1019)	(1252)
Presence of children										
None	0	0	0	0	0	0	0	0	0	0
Preschool only	−243	−455	−594	−754	11	−447	−418	−366	−347	−773
	(101)	(74)	(73)	(116)	(422)	(92)	(112)	(182)	(291)	(458)
School-age only	13	−118	−467	−594	− 27	−177	−235	−252	158	−223
	(121)	(77)	(75)	(93)	(251)	(93)	(132)	(209)	(297)	(345)
Both	−286	−486	−663	−939	−534	−322	−556	−489	−734	−1695
	(119)	(85)	(77)	(132)	(371)	(84)	(114)	(189)	(279)	(915)

C. White Wives — Standard Error: 701 hrs. per yr.
N: 2251
Constant: 1050
(61)

D. Black Wives — Standard Error: 871 hrs. per yr.
N: 1560
Constant: 1402
(121)

E. White Female Heads					F. Black Female Heads				
Whole Income (dollars)					Whole Income (dollars)				
0–3000	3000–3750	3750–4500	4500–5500		0–3000	3000–3750	3750–4500	4500–5500	
295	0	−123	−25		233	0	−183	511	
(270)		(237)	(274)		(130)		(190)	(320)	
Wage (dollars per hour)					Wage (dollars per hour)				
0–1.50	1.50–1.75	1.75–2.00	2.00–2.50	2.50–3.00	0–1.50	1.50–1.75	1.75–2.00	2.00–2.50	2.50–3.00
442	−451	0	124	835	138	−259	0	32	289
(513)	(278)		(210)	(542)	(169)	(170)		(27)	(324)
No. Adults					No. Adults				
1	2	3–4	5 or more		1	2	3–4	5 or more	
–	0	−189	−500		−1295	0	9	20	
		(193)	(522)		(892)		(117)	(257)	
Presence of Children					Presence of Children				
None	Preschool Only	School-age Only	Both		None	Preschool Only	School-age Only	Both	
0	−363	−316	−817		0	−623	−316	−662	
	(435)	(233)	(415)			(230)	(138)	(189)	

E. White Female Heads — Standard Error: 972 hrs. per yr.
N: 123
Constant: 1540
(221)

F. Black Female Heads — Standard Error: 860 hrs. per yr.
N: 259
Constant: 1547
(192)

TABLE 3.5. (cont.)

G. Single Men: White (Income and Wage Effects)

Age	*Whole Income (dollars)* 0–3000	3000–3750	3750–4500	4500–5500
20 to 59	1059 (855)	0	0	938 (255)
60 or above	—	173 (705)	–1303 (559)	–1765 (322)

Age	*Wage (dollars per hour)* 0–1.50	1.50–1.75	1.75–2.00	2.00–2.50	2.50–3.00
20 to 59	–2168 (940)	–648 (405)	0	–1007 (271)	–2096 (412)
60 or above	2094 (912)	– 42 (402)	0	764 (333)	396 (790)

Standard Error: 755 hrs. per yr.
N: 87
Constant: 1975 (224)

H. Single Men: White (Income Effects for 60 or above Only)

Age	*Whole Income (dollars)* 0–3000	3000–3750	3750–4500	4500–5500
	—	116 (753)	–159 (594)	–2053 (336)

Age	*Wage (dollars per hour)* 0–1.50	1.50–1.75	1.75–2.00	2.00–2.50	2.50–3.00
20 to 59	–1397 (464)	–936 (427)	0	–671 (274)	–1430 (380)
60 or above	2094 (980)	– 42 (432)	0	764 (358)	396 (849)

Standard Error: 811 hrs. per yr.
N: 87
Constant: 2263 (225)

I. Single Men: Black (Income and Wage Effects)

Age	*Whole Income (dollars)* 0–3000	3000–3750	3750–4500	4500–5500
20 to 59	–622 (723)	–206 (518)	0	151 (156)
60 or above	—	729 (337)	–322 (232)	–551 (245)

Age	*Wage (dollars per hour)* 0–1.50	1.50–1.75	1.75–2.00	2.00–2.50	2.50–3.00
20 to 59	1140 (723)	142 (518)	0	–175 (167)	–766 (284)
60 or above	–881 (329)	–945 (248)	0	523 (270)	—

Standard Error: 715 hrs. per yr.
N: 253
Constant: 1541 (111)

J. Single Men: Black (Income Effects for 60 or above Only)

Age	*Whole Income (dollars)* 0–3000	3000–3750	3750–4500	4500–5500
	—	727 (335)	–324 (230)	–553 (243)

Age	*Wage (dollars per hour)* 0–1.50	1.50–1.75	1.75–2.00	2.00–2.50	2.50–3.00
20 to 59	932 (515)	– 51 (218)	0	– 50 (125)	–617 (249)
60 or above	–881 (328)	–945 (247)	0	523 (269)	—

Standard Error: 713 hrs. per yr.
N: 253
Constant: 1543 (106)

TABLE 3.5. (cont.)

K. Single Women: White (Income Effects for 60 or above Only)

Whole Income (dollars)			
0–3000	*3000–3750*	*3750–4500*	*4500–5500*
−1193 (306)	−1202 (237)	−1270 (211)	−1305 (212)

Wage (dollars per hour)	*0–1.50*	*1.50–1.75*	*1.75–2.00*	*2.00–2.50*	*2.50–3.00*
Age 20 to 59	−480 (309)	−354 (183)	0	−370 (138)	−894 (292)
60 or above	−299 (183)	51 (211)	0	677 (359)	–

Standard Error: 765 hrs. per yr.
N: 411
Constant: 1858 (109)

L. Single Women: Black (Income Effects for 60 or above Only)

Whole Income (dollars)			
0–3000	*3000–3750*	*3750–4500*	*4500–5500*
297 (533)	−210 (512)	−513 (516)	−807 (524)

Wage (dollars per hour)	*0–1.50*	*1.50–1.75*	*1.75–2.00*	*2.00–2.50*	*2.50–3.00*
Age 20 to 59	212 (158)	104 (181)	0	163 (242)	–
60 or above	−525 (487)	−714 (556)	0	−318 (960)	–

Standard Error: 745 hrs. per yr.
N: 312
Constant: 1373 (140)

M. Sons and Other Male Relatives: White

Whole Income (dollars)	*0–3000*	*3000–3750*	*3750–4500*	*4500–5500*
Age 20 to 59	75 (201)	31 (171)	0	67 (152)
60 or above	−1161 (449)	−2152 (642)	−1568 (342)	−1722 (813)

Wage (dollars per hour)	*0–1.50*	*1.50–1.75*	*1.75–2.00*	*2.00–2.50*	*2.50–3.00*
Age 20 to 59	423 (575)	−645 (236)	0	−429 (182)	−372 (180)
60 or above	−205 (792)	644 (472)	0	224 (304)	640 (461)

No. Adults		
2	*3–4*	*5 or more*
0	−140 (161)	−119 (205)

Presence of Children			
None	*Preschool Only*	*School-age Only*	*Both*
0	631 (387)	−265 (169)	16 (245)

Standard Error: 740 hrs. per yr.
N: 204
Constant: 1927 (223)

N. Sons and Other Male Relatives: Black

Whole Income (dollars)	*0–3000*	*3000–3750*	*3750–4500*	*4500–5500*
Age 20 to 59	−82 (144)	−148 (141)	0	− 11 (218)
60 or above	−601 (698)	−878 (392)	−607 (418)	−1422 (588)

Wage (dollars per hour)	*0–1.50*	*1.50–1.75*	*1.75–2.00*	*2.00–2.50*	*2.50–3.00*
Age 20 to 59	286 (513)	−272 (211)	0	−178 (122)	−567 (176)
60 or above	–	−667 (548)	0	102 (453)	–

No. Adults		
2	*3–4*	*5 or more*
0	301 (140)	127 (166)

Presence of Children			
None	*Preschool Only*	*School-age Only*	*Both*
0	20 (232)	−327 (162)	−97 (195)

Standard Error: 705 hrs. per yr.
N: 213
Constant: 1454 (173)

TABLE 3.5. (cont.)

P. Daughters and Other Female Relatives: White

	Whole Income (dollars) 0–3000	3000–3750	3750–4500	4500–5500
Age 20 to 59	−215 (197)	95 (140)	0	19 (133)
60 or above	−501 (423)	−88 (371)	−453 (343)	−467 (354)

	Wage (dollars per hour) 0–1.50	1.50–1.75	1.75–2.00	2.00–2.50	2.50–3.00
Age 20 to 59	−571 (232)	− 90 (161)	0	−85 (119)	44 (301)
60 or above	−992 (328)	−586 (367)	0	197 (504)	–

No. Adults 2	3–4	5 or more
0	85 (130)	82 (163)

Presence of Children: None	Preschool Only	School-age Only	Both
0	−593 (218)	−199 (134)	−333 (198)

Standard Error: 750 hrs. per yr.
N: 331
Constant: 1418 (163)

Q. Daughters and Other Female Relatives: Black

	Whole Income (dollars) 0–3000	3000–3750	3750–4500	4500–5500
Age 20 to 59	−133 (171)	− 57 (166)	0	− 152 (287)
60 or above	−899 (690)	−857 (715)	−940 (731)	−1241 (687)

	Wage (dollars per hour) 0–1.50	1.50–1.75	1.75–2.00	2.00–2.50	2.50–3.00
Age 20 to 59	279 (159)	393 (166)	0	122 (227)	779 (451)
60 or above	9 (673)	246 (765)	0	–	171 (1076)

No. Adults 2	3–4	5 or more
0	−128 (127)	−73 (161)

Presence of Children: None	Preschool Only	School-age Only	Both
0	−251 (193)	−61 (139)	−293 (172)

Standard Error: 845 hrs. per yr.
N: 323
Constant: 1199 (187)

R. Male Teen-agers: White

Whole Income (dollars) 0–3000	3000–3750	3750–4500	4500–5500
−4 (236)	6 (206)	0	131 (268)

Wage (dollars per hour) 0–1.50	1.50–1.75	1.75–2.00	2.00–2.50	2.50–3.00
0	303 (198)	−74 (268)	220 (302)	–

Position in Family	Presence of Children: None	Preschool Only	School-age Only	Both
Husband	1080 (630)	1326 (878)	–	–
Single	253 (335)	–	–	–
Relative	0	−255 (292)	−251 (196)	−143 (239)

Standard Error: 839 hrs. per yr.
N: 128
Constant: 926 (219)

S. Male Teen-agers: Black

Whole Income (dollars) 0–3000	3000–3750	3750–4500	4500–5500
−373 (171)	−86 (178)	0	57 (229)

Wage (dollars per hour) 0–1.50	1.50–1.75	1.75–2.00	2.00–2.50	2.50–3.00
0	−59 (142)	130 (190)	245 (235)	–

Position in Family	Presence of Children: None	Preschool Only	School-age Only	Both
Husband	–	1190 (367)	–	–
Single	891 (335)	–	–	–
Relative	0	365 (212)	123 (156)	220 (164)

Standard Error: 774 hrs. per yr.
N: 174
Constant: 941 (179)

TABLE 3.5. (cont.)

T. Female Teen-agers: White

Whole Income (dollars)	0–3000	3000–3750	3750–4500	4500–5500	
	-80 (242)	-138 (209)	0	-248 (349)	
Wage (dollars per hour)	0–1.50	1.50–1.75	1.75–2.00	2.00–2.50	2.50–3.00
	0	360 (194)	918 (409)	45 (577)	—

Position in Family	Presence of Children: None	Preschool Only	School-age Only	Both
Wife	-29 (269)	-600 (246)	—	—
Single	1076 (623)	—	—	—
Relative	0	-436 (316)	317 (233)	-167 (302)

Standard Error: 834 hrs. per yr.
N: 119
Constant: 969 (212)

U. Female Teen-agers: Black

Whole Income (dollars)	0–3000	3000–3750	3750–4500	4500–5500	
	35 (196)	125 (190)	0	-319 (289)	
Wage (dollars per hour)	0–1.50	1.50–1.75	1.75–2.00	2.00–2.50	2.50–3.00
	0	-157 (159)	407 (279)	335 (369)	-138 (841)

Position in Family	Presence of Children: None	Preschool Only	School-age Only	Both
Wife	-388 (346)	-309 (191)	—	-688 (790)
Single	203 (411)	—	—	—
Relative	0	295 (198)	285 (169)	140 (208)

Standard Error: 773 hrs. per yr.
N: 179
Constant: 543 (202)

TABLE 3.6. *F*-STATISTICS FOR THE HYPOTHESIS OF NO DIFFERENCE BETWEEN RACES

Family Position	F-*Statistic*	*Degrees of Freedom* *Numerator*	*Denominator*	*Critical* F, *5 Percent Level*
Husbands	7.27	26	2848	1.50
Wives	5.53	37	3735	1.36
Female heads of families	0.55	13	355	1.73
Single men	2.99	14	312	1.70
Single women	1.67	14	694	1.70
Male relatives	1.40	19	377	1.59
Female relatives	1.36	19	614	1.59
Male teen-agers	1.06	12	277	1.75
Female teen-agers	2.25	13	270	1.73

Note: These statistics refer to Table 3.5.

for extra adults if there are increasing returns to scale in running a family. The main difficulty with this explanation is that it does nothing to rationalize our finding of opposite effects for the two races. Part of the differential by race might be explained as follows: There is evidence in the results for relatives that they are not really included in the family decision-making process. If so, the inclusion of the income of low-paid relatives would tend to cause the husband to be classified in too low an income group. In accord with our results, this effect would be strongest in the lowest income group. Moreover, white husbands are more likely than black husbands to be much better paid than their relatives, as our results in section II demonstrate. The bias on this account would be substantially stronger for whites than for blacks, suggesting at least part of the explanation of the different effects by race.

For both races, the presence of school-age children has a larger stimulus to their fathers' hours of work than does the presence of preschool-age children. This suggests that older children are more expensive but require less of their fathers' time.

Results for black husbands are presented in Table 3.5B. The striking feature of these results is that black husbands seem to work substantially fewer hours per year than do white husbands. In the reference group, blacks work 249 hours per year less than whites, more than six weeks less at 40 hours per week. This difference is not just a statistical fluctuation—the hypothesis that blacks and whites have the same coefficients in every cell is rejected decisively. The higher unemployment rate suffered by blacks explains only a fraction of the difference. As we will show later, black wives work longer hours than do white wives and, in fact, almost exactly counterbalance the shorter hours of their husbands.

TABLE 3.7. *F*-STATISTICS FOR THE HYPOTHESIS OF NO WAGE EFFECTS

	White				*Black*			
Family Position, Age	F-*Statistic*	*d.f. n.*	*d.f. d.*	*Critical* F, *5 Percent Level*	F-*Statistic*	*d.f. n.*	*d.f. d.*	*Critical* F, *5 Percent Level*
Husbands, 20 to 59	0.41	4	1370	2.37	4.17	4	1478	2.37
Husbands, 60 or above	19.93	4	1370	2.37	20.94	4	1478	2.37
Wives, 20 to 59	21.93	16	2212	1.65	5.63	15	1523	1.67
Female heads of families, 20 to 59	1.88	4	110	2.45	2.74	4	245	2.37
Men, single, 20 to 59	9.07	3	73	2.76	2.78	4	239	2.37
Men, single, 60 or above	3.03	4	73	2.53	9.90	4	239	2.37
Women, single, 20 to 59	5.76	4	396	2.37	0.39	3	298	2.60
Women, single, 60 or above	5.14	3	396	2.60	0.57	3	298	2.60
Male relatives, 20 to 59	2.77	4	183	2.37	2.87	4	194	2.37
Male relatives, 60 or above	0.78	4	183	2.37	0.98	2	194	3.00
Female relatives, 20 to 59	1.58	4	311	2.37	1.95	4	303	2.37
Female relatives, 60 or above	6.30	3	311	2.60	0.09	3	303	2.60
Male teen-agers	1.00	3	115	2.68	0.68	3	162	2.60
Female teen-agers	2.42	3	106	2.68	1.21	4	164	2.37

Note: d.f.n. means numerator degrees of freedom; d.f.d. means denominator degrees of freedom. These statistics refer to Table 3.5.

Income effects are present for black husbands in the 20 to 59 age group but are only about half as strong as are those for white husbands. As a result, the difference between the hours of work of whites and blacks is largest in the lowest income group and smallest in the highest group. The wage effects for the 20 to 59 age group are relatively small and of the expected sign in the second, third, and fourth wage groups, but the effect in the highest group is negative with a sufficiently small standard error to rule out the possibility that it is a random fluctuation (the perverse effect in the first group, on the other hand, is probably random). There is no explanation for this peculiar finding.

Black husbands of retirement age show a sensitivity to income and wages that is comparable to that of white husbands, although they generally work several hundred hours per year less than whites in similar circumstances.

Results for wives are presented in Table 3.5C and D. The specification for this group is the same as for husbands except that interactions are permitted between the wage effects and the effects of the presence and age of children. White wives in the reference group without children work almost exactly half time (1,050 hours per year). Black wives in the same income-wage group work almost 9 weeks per year more (1,402 hours per year)—not a great deal less than their husbands work. The income effects for white wives are roughly the same as for their husbands and for black wives are considerably stronger than for their husbands.

Black and white wives show approximately the same negative response in hours of work to the presence of additional adults in the family. For wives with children, one might expect that extra adults would stimulate hours of work by helping with the care of the children. However, in families with both children and extra adults, the latter are predominantly teenagers who are still in school and are unavailable most of the day.

Our discussion of the effects of wages and the presence of children will be carried out in terms of the figures in Table 3.8, which were calculated by adding the constant, the wage effects for the 20 to 59 age group, and the wage effects for each age-of-children group, all taken from Table 3.5C and D. Thus, Table 3.8 gives the estimated hours of work for wives age 20 to 59 in families with no extra adults and with whole incomes of $7,500 to $9,000 per year. Black wives work longer hours in almost every cell, and the null hypothesis that wives of the two races have the same coefficients is rejected very strongly. The presence of children reduces hours of work in almost every race-wage group. There do not seem to be important differences in the wage effects by age of children or by race, although, of course, the general level is higher if all the children are of school age. For both races, the overall wage effects are strongly positive, and the hypotheses of no wage effects are clearly rejected in both cases.[18]

[18] The null hypothesis is that all of the wage-age coefficients are zero for the 20 to 59 age group, and all of the wage effects are the same within each age-of-children group.

TABLE 3.8. HOURS OF WORK FOR WIVES BY WAGE GROUP AND AGE OF CHILDREN

	Wage (dollars per hour)				
Children	*0–1.50*	*1.50–1.75*	*1.75–2.00*	*2.00–2.50*	*2.50–3.00*
Black					
None	772	1092	1402	1313	1818
Pre-school only	325	674	1036	966	1045
School-age only	595	857	1150	1471	1595
Both	450	536	913	579	123
White					
None	403	986	1050	1565	1554
Pre-school only	160	531	456	811	1565
School-age only	416	868	583	971	1527
Both	117	500	387	626	1020

Our next results, shown in Table 3.5E and F, are for women who are heads of their families. These results are seriously incomplete because of the exclusion of families receiving public assistance; very few women with young children and low incomes remain in the sample after this exclusion. The remarkable feature of the results is the similarity between the hours of work of female heads and of wives, after taking into account differences in incomes. White female heads in the reference group with one extra adult and no children work 1,540 hours per year, compared to 1,386 hours per year for wives in the same income group. For blacks, female heads work slightly less (1,547 hours per year), than do comparable wives (1,655 hours per year). For both races the pattern of negative income effects and positive wage effects for wives is repeated, although the effects are subject to a great deal more sampling variation, so that the hypothesis of no wage effects is barely rejected for blacks and falls short of rejection for whites. In contrast to our finding for wives, there is no apparent tendency for blacks to work longer hours than whites; the null hypothesis that female heads of families of both races have the same coefficients is nowhere near rejected.

Our results for single individuals, presented in Table 3.5G through L, suffer in most cases from an inability to separate income and wage

effects in the 20 to 59 age group where almost all income is from earnings. Our original specification had independent effects for income and wages separately for the two age groups. Results for this specification are given for men in Table 3.5G and I. These results are sufficiently implausible (income effects are positive and wage effects negative) to force a retreat to a specification that excludes income effects for the younger age group. This specification, used for both men and women, requires a rather different interpretation of the wage effects—they include the income as well as the substitution effects of wage changes. The wage effects trace out the conventional labor-supply curve. For men in the 20 to 59 age group, Table 3.5H and J suggest rather strongly that the labor-supply curves for both races bend backward, at least for higher wage groups. White single men in the third wage group work slightly more than full time (2,263 hours per year), but in the highest wage group they work less than half time (833 hours per year). Whites in the lowest wage group also work less than half time. Black single men work substantially less in the third wage group (1,543 hours per year), slightly more in the highest wage group (926 hours per year), and substantially more in the lowest wage group (2,475 hours per year). The hypothesis that the two races have the same coefficients is rejected. In all four age-race groups for single men, the hypothesis of no wage effects is rejected. Income and wage effects for men of retirement age are similar to those found for husbands.

Results for single women are presented in Table 3.5K and L. White single women in the 20 to 59 age group are remarkably similar to white single men, showing the same positive response to wages in the lower wage groups and a negative response in the higher groups. In all wage groups except the third, white women work roughly the same number of hours as men. The hypothesis of no wage effects for white single women is clearly rejected. Black single women, on the other hand, show very little variation in hours of work by wage groups, and the hypothesis of no wage effects cannot be rejected. They tend to work about three-quarters time in every wage group.

Our next results, in Table 3.5M, N, P, and Q, are for the heterogeneous group of adult relatives who are not heads of the families in which they live. This group includes grown sons and daughters living with their parents; parents living with their sons or daughters; and brothers, sisters, aunts, uncles, cousins, grandparents; and grandsons and granddaughters of the heads of the families. In this group, the wages of individuals are not closely associated with the incomes of the families in which they live, so there is no econometric obstacle to separating the wage and income effects. What we find in all four sex-race groups, however, are very weak income effects and wage effects that resemble those for single individuals in that they are frequently negative rather than positive. We have already mentioned a conjecture that would explain this peculiar finding: relatives may not be sufficiently integrated in the families with whom they reside to show much sensitivity to the families' incomes. If so, and the role of relatives is more like that of paying boarders, then the proper measure of income for these regressions is the whole income of the individual, not that of the family. Exclusion of the proper measure of income has the same effect as in the case of single individuals—the wage effects include both the substitution effects and the competing income effects, and may be negative rather than positive. This conjecture could be tested by including the whole income of individuals in the regressions, but unfortunately data on nonwage income are collected at the level of the family and cannot be allocated reliably among its members.

The results for relatives are sufficiently similar to those for single individuals that detailed discussion is not required. Comparison of all four race-sex groups in all of the wage categories suggests that there are no substantial differences between them in the overall level of hours of work; most individuals work about 1,500 hours per year. The test for differences between races does not contradict this conclusion. The presence of children in the family is associated with a substantial reduction in the hours of work of female relatives, presumably because they help care for the children without a formal arrangement for receiving wages.

Our last results are for teen-agers, Table 3.5R, S, T, and U. They are grouped together by age rather than by family position because teen-agers seem to have characteristics in the labor market very different from similar individuals only a few years older. Our results for teen-agers are necessarily fragmentary because of the exclusion of all those who attended school at any time in 1966.

For teen-agers, as for relatives, income effects are absent or of the wrong sign. The estimates of the wage effects are rather irregular, but only white women show an unmistakably positive response to higher wages. The striking feature of the results is how little teen-agers work when they are living with their parents, as most are. White males work less than half time (926 hours per year) if there are no children 13 years or younger in their families and even less if there are children. Black males living with their parents work slightly longer (941 hours per year) without children and even more if there are children. White females also work just under half time (969 hours per year), but they work considerably less if their parents also have children of preschool age. Black females work only about one-quarter time (543 hours per year) when no children are present, but work longer than white females when there are preschool children.

Emancipated teen-agers living as single individuals tend to work more than those living with their parents, especially black males and white females, who work essentially full time. For men, marriage is associated with at least full-time work, while for women, it is associated with a drop in hours of work.

VII. COMPARISON WITH RESULTS OF OTHER RECENT STUDIES OF LABOR SUPPLY

Four studies have appeared recently with many of the same purposes as the present study: those of Cohen, Rea, and Lerman;[19] Greenberg and Kosters;[20] Hill;[21] and Kalachek and Raines.[22] Table 3.9 presents a brief comparison of various pertinent aspects of the designs of these studies. All five are based on large bodies of data on individuals collected by the Bureau of the Census, either in the regular Current Population Survey (Cohen, Rea, and Lerman; and Kalachek and Raines) or in the Survey of Economic Opportunity. The latter is probably better suited to the estimation of labor-supply functions because it reports weekly earnings as well as weekly hours of work and also reports extensive data on wealth. Surprisingly, neither of the other studies based on the SEO took advantage of the opportunity to use the data on weekly earnings to improve the measure of annual hours of work. One of them, that of Hill, did not even use the obvious measure of wages obtained by dividing weekly earnings by weekly hours but used instead the less satisfactory measure based on annual data.

Investigations based on the CPS are prevented by the shortcomings of the data from including imputed income in the measure of income appearing in the labor-supply function. Because home ownership, the most important source of imputed income, increases with income, the omission of imputed income gives rise to an error in measuring income that is nega-

[19] Malcolm S. Cohen, Samuel A. Rea, Jr., and Robert I. Lerman, "A Micro Model of Labor Supply," U.S. Department of Labor, Bureau of Labor Statistics, Staff Paper No. 4 (Washington, D.C.: U.S. Government Printing Office, 1970).

[20] David H. Greenberg and Marvin Kosters, "Income Guarantees and the Urban Poor: The Effect of Income Maintenance Programs on the Hours of Work of Male Family Heads," in this volume, p. 14.

[21] C. Russell Hill, "The Determinants of Labor Supply for the Working Urban Poor," in this volume, p. 182.

[22] Kalachek and Raines, "Labor Supply of Lower Income Workers and the Negative Income Tax."

TABLE 3.9. COMPARISON OF RECENT STUDIES OF LABOR SUPPLY

Authors	*Data*	*Individuals Included*	*Measure of Hours*	*Measure of Wage*	*Measure of Income*	*Functional Form*	*Method of Estimation*
Cohen, Rea, and Lerman, "A Micro Model of Labor Supply."	CPS-1967	Participants only	(weeks of work + weeks looking) × weekly hours last year	Wage income/measure of hours	CPS income less own wage income, no imputed income	Step functions with interactions	OLS*
Greenberg and Kosters, essay in this volume.	SEO-1967	Participants only	(weeks of work + weeks looking) × hours last week	Wage income last week/hours last week	Family nonemployment income, including imputed income	Linear in wage and income	OLS
Hall, essay in this volume.	SEO-1967	Whole population	Wage income last year/imputed wage	Imputed, based on data on average wage in survey week, adjusted for income tax	Nonemployment income (including imputed income) + value of family members' time	Step functions	TSLS†
Hill, essay in this volume.	SEO-1967	Participants only	(weeks of work + weeks looking) × weekly hours last year	Wage income/measure of hours	Property income, not including imputed income; transfer income (separate variable)	Linear in wage and income with multiplicative interaction	OLS
Kalachek and Raines, "Labor Supply of Lower Income Workers and the Negative Income Tax."	CPS-1966	Whole population	Participation (0 or 1) × weeks of work × weekly hours	Imputed, based on data on average annual wages	Nonemployment income excluding transfers; no imputed income	Separate equations for participation weeks, and hours; each is quadratic in wage and income	TSLS

Sources: CPS = Current Population Survey; SEO = Survey of Economic Opportunity; also, see Figure 3.5.
*OLS = ordinary least squares.
†TSLS = two-stage least squares. See Figure 3.5.

tively correlated with true income. This causes biases in their regression results, although it is not possible to state the directions of these biases a priori.

The five studies fall into two groups according to the method of selecting the data and also, by no coincidence, into the same two groups according to the statistical method of estimation employed. The first groups, consisting of the studies by Cohen, Rea, and Lerman; Greenberg and Kosters; and Hill, included only participants in the labor force, and estimated the resulting conditional labor-supply function by ordinary least squares. The second group, comprising the studies by Kalachek and Raines and by the present author, included nonparticipants (who were able to work) as well as participants, and estimated the unconditional labor-supply function by one or another variant of two-stage least squares. Both studies use a preliminary regression in which the wage is the left-hand variable to impute values of the wage for participants as well as nonparticipants.

The exclusion of nonparticipants requires that the results of the first group of studies be interpreted with some care. A labor-supply function, $L = F(w,y)$, can always be written in the form

$$F(w,y) = G(w,y)P(w,y) \tag{3.15}$$

where $P(w,y)$ is the labor-force participation rate for individuals with wage, w, and income, y, and $G(w,y)$ is the average hours of work of those who work a positive number of hours. Studies limited to participants yield estimates of $G(w,y)$, while studies of the whole population provide estimates of $F(w,y)$. For most purposes estimates of the overall labor-supply function $F(w,y)$ are required—for example, for evaluating the effects of programs for income maintenance—so the first group of studies are incomplete in failing to estimate the participation function, $P(w,y)$. Because the labor-force participation rate is always less than one, $G(w,y)$ will always exceed $F(w,y)$. Perhaps more importantly, if $P(w,y)$ is at all sensitive to w or y, the response of $G(w,y)$ to a change in w or y will be rather different from the response of $F(w,y)$ to the same changes. The two are compared as functions of income in Figure 3.4.

Kalachek and Raines obtain estimates of the overall labor-supply function, $F(w,y)$, by multiplying separate estimates of the conditional labor-supply function, $G(w,y)$, and the participation function, $P(w,y)$. Their procedure suffers from a statistical defect whose importance is hard to evaluate. The labor supply of an individual is

$$H_i = [G(w,y) + \phi_i][P(w,y) + \psi_i], \tag{3.16}$$

FIGURE 3.4. COMPARISON OF HOURS OF WORK OF PARTICIPANTS, $[G(w,y)]$, AND HOURS OF WORK OF THE ENTIRE POPULATION, $[F(w,y)]$

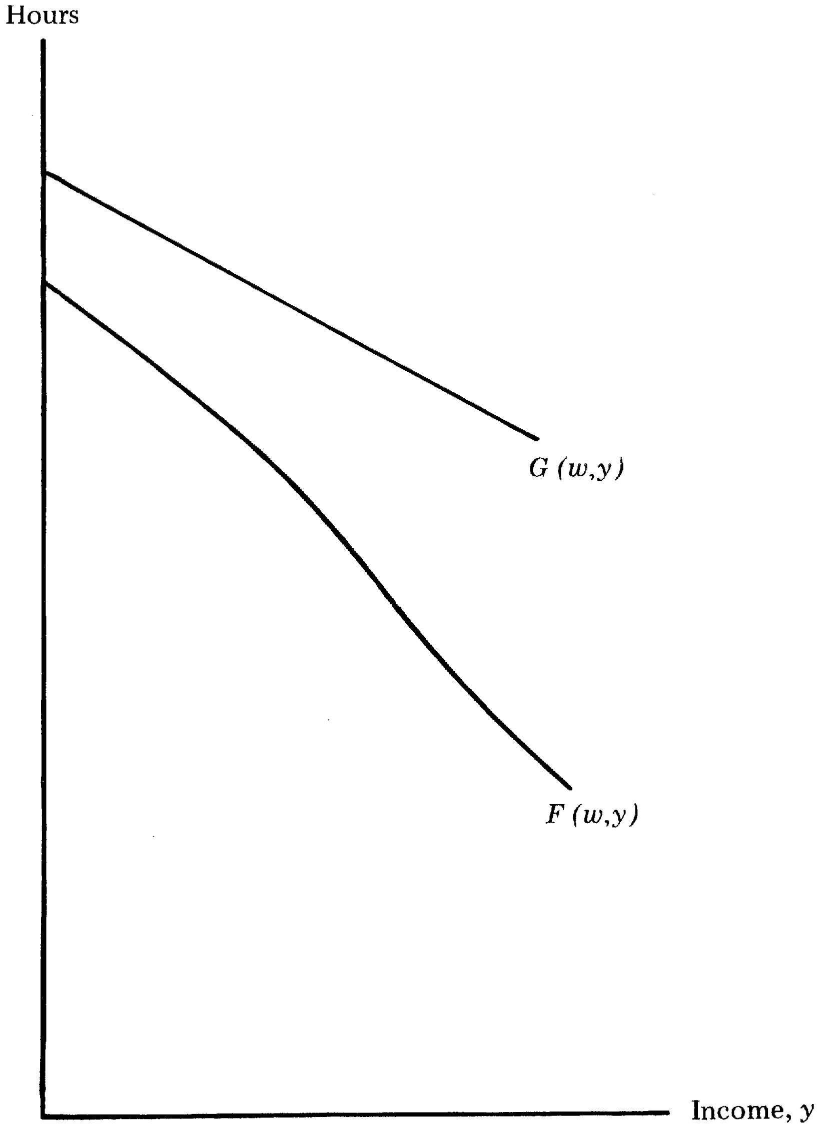
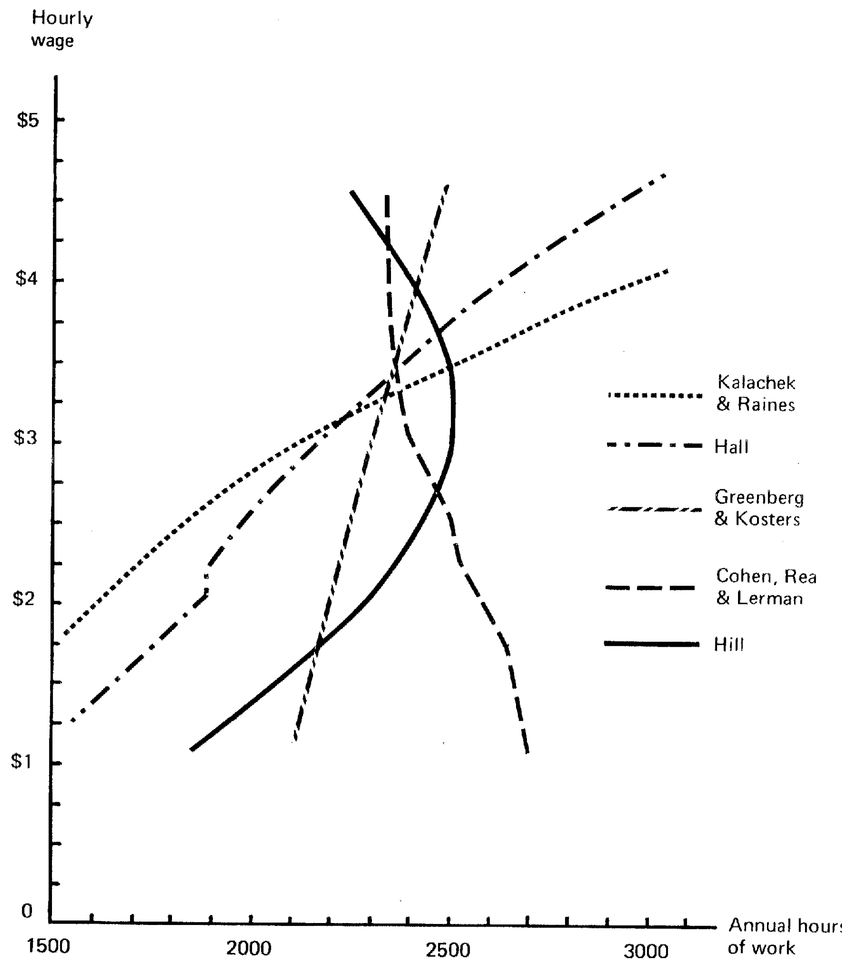


FIGURE 3.5. COMPENSATED LABOR-SUPPLY FUNCTIONS



Sources: 1. C. Russell Hill, "The Determinants of Labor Supply for the Working Urban Poor," in this volume. Results are taken from Table 5.4, p. 201, for white nonpoor male family heads. The estimated curve refers to the following group: no transfer income, 12 years of education for both husband and wife, both in good health, and 2 children.

2. Malcolm S. Cohen, Samuel A. Rea, Jr., and Robert I. Lerman, "Micro Model of Labor Supply," U.S. Department of Labor (Washington, D.C.: U.S. Government Printing Office, 1970), Bureau of Labor Statistics Staff Paper No. 4. Regression results are taken from Table F-3, p. 206. Data are for men age 22 to 54, neither ill nor in school, who worked one or more weeks in 1966. The estimated curve refers to the following group as defined by the right-hand variables in the regression: 12 years of school completed, mar-

where ϕ_i and ψ_i are random disturbances with mean zero. The overall labor-supply function is

$$F(w,y) = E(H_i|w,y) = G(w,y)P(w,y) + Cov\ (\phi_i,\psi_i|w,y). \quad (3.17)$$

There is every reason to believe that the conditional covariance, $(\phi_i,\psi_i|w,y)$, differs from zero and, moreover, that it depends on w and y. In other words, nothing in Kalachek and Raines's procedure guarantees that the residuals from their overall labor-supply equation are uncorrelated with the right-hand variables. The present study avoids this problem altogether by estimating the overall labor-supply function, $F(w,y)$, directly.

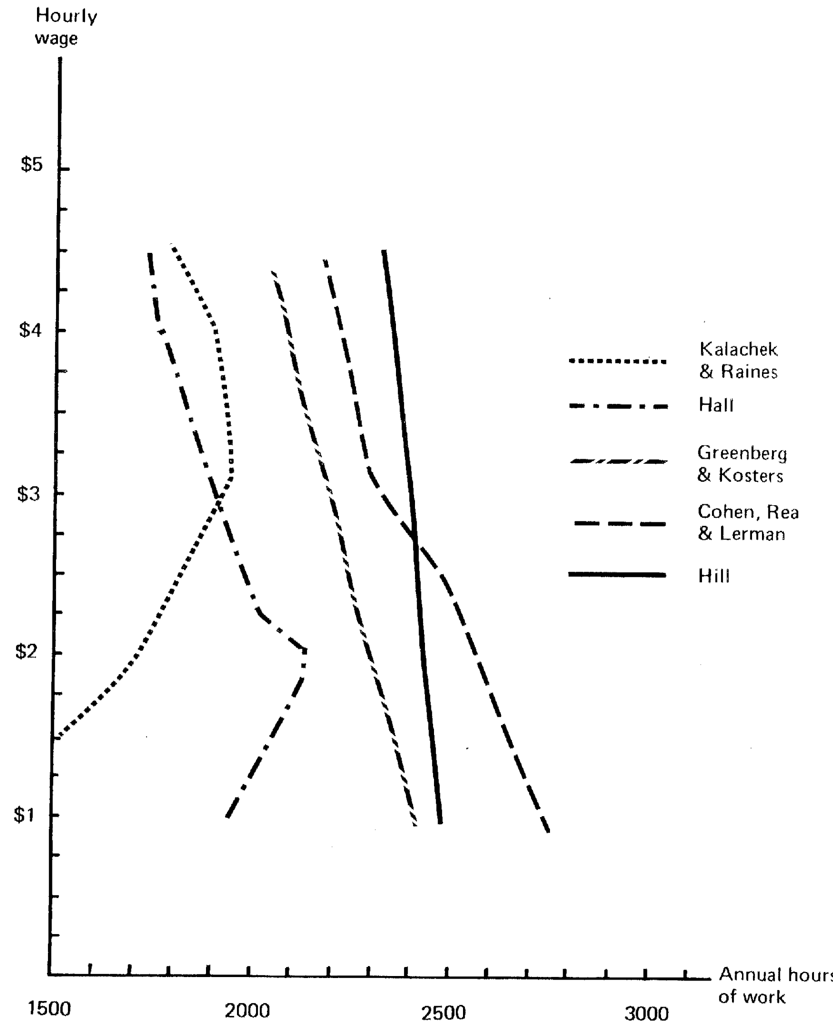
The differences among the five studies are shown graphically in Figures 3.5 and 3.6. Both present hours of work as functions of the wage. In Figure 3.5, the level of economic well-being is held constant as the wage changes, through the Slutsky-Hicks income compensation. The curves shown were obtained by numerical integration of the Slutsky equation; their slopes measure the pure substitution effects of changes in wages. The theory of the consumer suggests that the slope of the compensated labor-supply function should always be positive—an individual whose wage is

ried, spouse present, 2 children under 18, no resident in a poverty tract or in the south, white, and not self-employed. The step function in wages and income was smoothed before the calculation of the Hicksian labor-supply function. All interaction terms were included in the calculations.

3. David H. Greenberg and Marvin Kosters, "Income Guarantees and the Urban Poor: The Effect of Income Maintenance Programs on the Hours of Work of Male Family Heads" (Santa Monica, Calif.: The Rand Corporation, 1970). Regression results are taken from Table B-3a, p. 108, for demographic group 6: male heads of families, age 25 to 55, excluding those who did not work at their usual jobs in the week before the survey. The estimated curve refers to the following group: observed assets equal to expected assets (their "preference variable" equal to zero), no earnings of other family members, 2 dependents other than wife, white, 12 years of school, in good health, and 30 years old. The constant for this regression is reported as 154.486, but calculations suggested that the proper figure was 1540, which was used.

4. Edward D. Kalachek and Fredric Q. Raines, "Labor Supply of Income Workers and Negative Income Tax," mimeographed, 1970. Results are taken from Table 1, p. 171, for males in poor families, from their "reduced form model." The estimated curve refers to the following group: 24 to 35 years old, fathers, wife's wages equal to $1.50 per hour, no other wage-earning relatives, no public assistance, resident of suburban ring of an SMSA in the middle Atlantic region, occupation in clerical and retail sales, married with wife present, 4 persons in family, 1 child less than 6 years old and 1 between 6 and 13 years old, no family member unemployed. The labor-supply function is the product of the three equations presented for participation, weeks, and weekly hours.

FIGURE 3.6. ORDINARY LABOR-SUPPLY FUNCTIONS



Source: See Figure 3.5.
Note: Nonlabor income is held constant at $2000 per year.

increased but whose nonlabor income is decreased correspondingly should work more than before, because the price of his leisure has risen while his total income to divide between goods and leisure has remained the same.

Three of the studies show the expected positive slope in Figure 3.5: the present study, and those by Greenberg and Kosters and by Kalachek and Raines. As one would expect, the conditional labor-supply function of Greenberg and Kosters shows substantially less sensitivity to changes in wages than do the unconditional labor-supply functions of the other two studies. The very large substitution effects of Kalachek and Raines are somewhat suspicious in the light of our earlier criticism of their statistical method.

The other two studies have negative substitution effects for some or all wage levels. In Hill's study, the compensated labor-supply curve bends backward above \$3.00 per hour. This is caused by the rather large positive coefficient of the multiplicative interaction between wages and income—as income becomes negative along the upper part of the compensated labor-supply curve, the wage effect becomes increasingly negative. This shows the danger of too restrictive a parametric specification for the labor-supply function. On the other hand, in Cohen, Rea, and Lerman's study, the substitution effect is negative at every wage level. There is no obvious explanation for this unexpected finding.

As a final comparison of the five studies, we present in Figure 3.6 the ordinary labor-supply curves showing hours of work as a function of the wage with nonlabor income held constant. Three of the studies—those of Cohen, Rea, and Lerman; Greenberg and Kosters; and Hill—have negative slopes for all wage levels. These are the three studies that estimate hours of work conditional upon labor-force participation. The remaining two studies show a positive slope at low-wage levels (up to \$2.00 for the present study and up to \$3.25 for Kalachek and Raines) and a backward-bending relation at higher wage levels.

Appendix 3A

APPLICATION OF THE INSTRUMENTAL-VARIABLES ESTIMATOR TO EQUATIONS THAT ARE NOT LINEAR IN THE RIGHT-HAND VARIABLES

Ordinary least squares equips us to estimate the parameters, $\beta_1, \ldots, \beta_N$, of the regression function,

$$E(y|x) = \sum_{j=1}^{N} \beta_j f_j(x), \qquad (3.18)$$

where $E(y|x)$ is the expected value of some left-hand variable, y, conditional on one or more right-hand variables, x, and $f_j(\cdot)$ are known but not necessarily linear functions. Our problem is to estimate the parameters of the labor-supply equation,

$$L = \sum_{j=1}^{N} \beta_j f_j(w^*) + u. \qquad (3.19)$$

We have fitted values from a first-stage wage regression, $\hat{w}$, related to the unobserved w^* by

$$w^* = \hat{w} + \eta, \qquad (3.20)$$

where the errors, η, and the structural disturbances, u, have the properties $E(\eta|,\hat{w}) = 0$ and $E(u|\hat{w}) = 0$. We ask: what is the relation between the regression function $E(L|\hat{w})$ and the deterministic part of the structural equation? A basic result of econometric theory holds that they are identical if the functions $f_j(\cdot)$ are constant or linear. We are concerned with the extension to the case of nonlinear functions, and in particular with the case of step-function approximations of an arbitrary function.

Written in terms of $\hat{w}$, the structural equation is

$$L = \sum_{j=1}^{N} \beta_j f_j(\hat{w} + \eta) + u, \qquad (3.21)$$

so the regression function is

$$E(L|\hat{w}) = \sum_{j=1}^{N} \beta_j E[f_j(\hat{w} + \eta)|\hat{w}] + E(u|\hat{w}), \qquad (3.22)$$
$$= \sum_{j=1}^{N} \beta_j g_j(\hat{w}),$$

where

$$g_j(\hat{w}) = \int_{-\infty}^{\infty} f_j(\hat{w} + \eta)p(\eta)\, d\eta, \qquad (3.23)$$

and $p(\eta)$ is the probability density function of η. That is, the regression function is a linear combination of functions of $\hat{w}$, each of which is a weighted average of the corresponding function of $\hat{w}$ in the structural equation, with the average taken over values of $\hat{w}$ centered at $\hat{w}$ and the weights given by the density of η. The standard result is a special case of this; if $f_j(\hat{w}) = \hat{w}$,

$$g_j(\hat{w}) = \hat{w} \int_{-\infty}^{\infty} p(\eta)\, d\eta - \int_{-\infty}^{\infty} \eta p(\eta)\, d\eta = \hat{w} = f_j(\hat{w}). \qquad (3.24)$$

If f_j is not linear, however, g_j is not identical to it. Simple extension of the instrumental variables estimator to the nonlinear case is impossible.

Because the functions $f_j(\cdot)$ are known and the density function $p(\eta)$ can be estimated in the first-stage regression, there is no obstacle in principle to calculating the functions $g_j(\hat{w})$ and estimating the parameters β_j by ordinary least squares. However, this extra complication may be unnecessary in some cases. The case of particular interest to us is the step function specification,

$$\left.\begin{aligned} f_j(w) &= 1 \text{ if } w_{j-1} < w \leq w_j, \\ &= 0 \text{ if } w \leq w_{j-1} \text{ or } w > w_j, \\ w_j &= \bar{w}j. \end{aligned}\right\} j = -\infty, \ldots, 0, \ldots, +\infty \qquad (3.25)$$

(We let j range over all negative and positive integers to avoid the complications caused by end effects.) Under some circumstances, the uncorrected regression function,

$$\sum_{j=-\infty}^{\infty} \beta_j f_j(\hat{w}),$$

is a good approximation to the corrected function,

$$\sum_{j=-\infty}^{\infty} \beta_j g_j(\hat{w}).$$

These circumstances are: (1) the density $p(\eta)$ is symmetric about zero, and (2) the β_j coefficients lie along a straight line (i.e., $\beta_j = k_0 + k_1 w_j$). In fact, as we will show, they are equal for a value of $\hat{w}$ at the center of one of the intervals:

$$\hat{w} = \frac{w_{j^*-1} + w_{j^*}}{2}. \qquad (3.26)$$

First,

$$\Sigma\beta_j f_j(\hat{w}) = k_0 \Sigma f_j(\hat{w}) + k_1 \Sigma w_j f_j(\hat{w}) \qquad (3.27)$$
$$= k_0 + k_1 w_{j^*}.$$

On the other hand,

$$\Sigma\beta_j g_j(\hat{w}) = k_0 \Sigma g_j(\hat{w}) = k\Sigma w_j g_j(\hat{w}) \qquad (3.28)$$
$$= k_0 + k_1 \bar{w} \Sigma j p_j,$$

where

$$p_j = P(\hat{w} - w_{j-1}) - P(\hat{w} - w_j), \qquad (3.29)$$

and P is the cdf of η. Now p_j is the distribution of an integer-valued random variable distributed symmetrically about j^*, so its expectation is j^*. Substituting j^* for $\Sigma j p_j$ in (3.28), we can see that it equals (3.27), as asserted.

We conclude from this calculation that if end effects are not too serious, if the values of $\hat{w}$ are distributed more or less evenly in each interval, and if the coefficients lie approximately along a straight line within the range of variation of the error, η, then the use of the uncorrected rather than the corrected regression equation will provide a satisfactory approximation.

Appendix 3B

TABLE 3B.1. WAGE EARNERS IN THE SEO SAMPLE

	Number of Individuals			
	Male		Female	
Characteristics	White	Black	White	Black
Total	3082	2194	1836	1858
Age				
14 to 15	25	14	19	2
16 to 17	74	40	33	36
18 to 19	105	85	115	77
20 to 24	288	204	284	210
25 to 34	739	563	303	412
35 to 44	664	526	357	456
45 to 54	669	462	417	410
55 to 64	431	260	246	210
65 or above	87	40	62	45
Years of education				
0-3	58	102	23	43
4-6	153	234	58	133
7-9	578	560	289	405
10-11	452	446	248	351
12	932	556	744	622
13-14	343	168	213	176
15	84	33	44	18
16	249	57	135	64
17-20	233	38	82	46
SMSA				
Baltimore	105	210	59	185
Chicago	297	241	191	185
Cleveland	102	88	45	64
Detroit	204	213	107	153
Houston	88	140	47	119
Los Angeles	563	249	322	191
New York	838	329	533	314
Philadelphia	239	181	116	154
Pittsburgh	123	26	62	19
St. Louis	80	67	54	65
San Francisco	274	119	191	96
Washington, D.C.	169	331	109	313
Residence at age 16				
U.S.	2796	2168	1671	1833
Foreign	286	26	165	25
Weeks of work in 1966				
27 or more	2856	2029	1537	1566
26 or less	226	165	299	292
Union membership				
Nonmember	2123	1369	1548	1564
Member	959	825	288	294
Health				
No effect on work	2773	1986	1658	1632
Some effect	309	208	178	226

TABLE 3B.2. COMPOSITION OF THE SAMPLE FOR THE STUDY OF HOURS OF WORK

Characteristic	Range or Value	Number of Individuals
Total	–	12937
Position in family	Husband	4499
	Wife	4512
	Female head	553
	Single man	703
	Single woman	949
	Son or other male relative	786
	Daughter or other female relative	935
Race	White	7709
	Black	5228
Age	14 to 19	640
	20 to 59	10068
	60 or above	2229
Number of adults	1	1817
	2	6361
	3 or 4	3927
	5 or more	832
Children	None	7365
	Preschool age only	1980
	School-age only	1851
	Both	1741
Whole income per adult (dollars)	0-3000	2079
	3000-3750	2818
	3750-4500	3186
	4500-5500	2888
	5500+	1966
Imputed hourly wage (dollars)	0-1.50	2731
	1.50-1.75	1971
	1.75-2.00	2075
	2.00-2.50	2633
	2.50-3.00	1848
	3.00+	1679

Appendix 3C

TABLE 3C.1. PRICE LEVEL IN THE 12 METROPOLITAN AREAS IN THE SEO

Metropolitan Area	*Price Level*
Baltimore	94
Chicago	104
Cleveland	103
Detroit	99
Houston	93
Los Angeles	102
New York	109
Philadelphia	100
Pittsburgh	97
St. Louis	101
San Francisco	107
Washington, D.C.	101

Source: Bureau of Labor Statistics, *Three Standards of Living, Spring, 1967,* Bulletin 1570-5, Table 3, p. 35, 4th column.