NBER WORKING PAPERS SERIES

ALCOHOLISM, WORK, AND INCOME OVER THE LIFE CYCLE

John Mullahy

Jody L. Sindelar

Working Paper No. 3909

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
November 1991

NBER Working Paper #3909
November 1991

ALCOHOLISM, WORK, AND INCOME OVER THE LIFE CYCLE

ABSTRACT

We find that alcoholism decreases labor force participation among prime age males, and therefore decreases the income of this group. The effects of alcoholism on the labor force participation of younger and older males and on the wage rates of prime age males are not significantly positive. We also find that alcoholism affects income indirectly through its effects on individual characteristics such as schooling and marital status, as well as directly through labor force participation rates after controlling for these indirect effects.

John Mullahy
Department of Economics
Trinity College
Hartford, CT 06106
and NBER

Jody L. Sindelar
Department of
Epidemiology and Public
Health
Yale School of Medicine
P.O. Box 3333
New Haven, CT 06510

Alcoholism, Work, and Income over the Life Cycle[1]

by John Mullahy[2] and Jody L. Sindelar[3]

...individuals, besides, may sometimes ruin their fortunes by an excessive consumption of fermented liquors...

Adam Smith, The Wealth of Nations

I. Introduction

It is widely believed that alcoholism is a major social problem with potentially important economic consequences.[4] Alcoholism is a prevalent disorder in the U.S., affecting about one in twenty individuals at any one time and one in ten individuals at some point during their lives. Males are three times more likely than females to suffer from alcoholism.

A prevailing view on alcoholism is that it has depressant effects on earnings, income, and wages even after controlling for other important determinants of these labor market outcomes.[5] Despite Adam Smith's early observation and the current popularity of this view, its veracity has been challenged by the hypothesis that increased alcohol consumption -- at least within some range -- might actually be productive.[6] Some of this controversy is resolved immediately by distinguishing "alcohol consumption" and "alcoholism," but this distinction settles only part of the debate.

The main purposes of this paper are to suggest that a more complete answer to the alcoholism-income question requires several other important perspectives that have heretofore been absent from the debate, and to provide some evidence supporting this notion. First, we demonstrate that the relationships between alcoholism and indicators of labor market success may vary over the lifetime. That is, we find that the relationships between alcoholism and contemporaneous measures of labor market success depend strongly on age. Second, we presen evidence that alcoholism's most consequential role in income determination may be more that of a labor supply shifter than a wage shifter. When studying only workers, therefore, small and insignificant effects of alcoholism on income

might tend to be found to the extent that labor supply behavior responds more to alcoholism than do wages. Finally, we show that the magnitude and significance of the effects of alcoholism on earnings depend on the extent to which one controls for alcoholism-related covariates.

Before proceeding, it is useful to review some background information on alcoholism. Throughout the paper we use "alcoholism" as a convenient term to summarize both the "alcohol dependence" and "alcohol abuse" disorders as defined by the American Psychiatric Association (1980,1987). Although much about alcoholism remains unknown, most experts would agree on some general characteristics of alcoholics (see NIAAA (1990)). For instance, it is now widely accepted that alcoholics come from all socioeconomic, demographic, and occupational groups. Moreover, about ten percent of males and three percent of females are actively alcoholic at any point in time. For males, young adulthood is when symptoms of alcoholism are most prevalent; symptoms tend to decline after this peak.[7] In addition, it is commonly observed that alcoholism tends to run in families, with the prevalent current wisdom being that there is a strong genetic dimension to such tendencies (see NIAAA (1990)).[8]

Despite these broad agreements, there remain many lively debates on issues such as the etiology of alcoholism, the effectiveness of alternative treatment methods,[9] and the lifecycle course of alcoholism problems. For instance, there have to date been insufficient longitudinal data to determine much about the lifecycle course of alcoholism problems. More important for our purposes is the debate about whether alcoholism is, in a biomedical sense, a disease, or rather is a complex set of health outcomes that arise from freely chosen behavior. Alcoholism is widely considered in the medical literature to be a disease with a significant genetic basis. However, this "disease model" has not always held sway and even today is not universally embraced, with some preferring to analyze alcoholism not from the disease perspective but rather from that of conscious choice.[10]

This background provides the setting for the remainder of the paper. Section II discusses the data used in the empirical analysis and the characteristics of our sample. Section III presents a model and an econometric strategy for estimating the role of alcoholism in a human capital framework. Section IV provides some evidence concerning the life cycle dimensions of the relationships between alcoholism and labor market outcomes. Section V presents a variety of econometric estimates of the role of alcoholism as a determinant of labor market success. Section VI concludes with a discussion of the findings and of unresolved issues.

II. Data and Sample Characteristics

This analysis is based on Wave I of the New Haven site of the Epidemiological Catchment Area survey conducted under the auspices of the U.S. National Institute of Mental Health (NIMH). The New Haven ECA survey is part of a larger NIMH-funded data collection effort surveying individuals 18 years old and older to assess psychiatric disorders in a population based sample. The ECA surveys were designed primarily to assess the distribution of mental disorders in a community setting.[11] Prior to the availability of the ECA data, there were no large U.S. samples assessing individuals' psychiatric disorders that contained reliable measures of the disorders, including alcoholism. The ECA data are particularly well-suited for study of alcoholism, as they provide medically sophisticated measures of alcoholism and other mental disorders and, of particular importance for our study, information on labor market outcomes as well as socioeconomic and demographic characteristics of individuals and their households.[12,13]

Between 1980 and 1981, Wave I of the New Haven survey was completed, yielding 5,034 observations, a 77.6% completion rate. The New Haven SMSA was (approximately) the catchment area sampled, this area comprising 13 towns with a total adult population of 420,000. Two coordinate groups were sampled in this residential survey: all adults (18+), and individuals 65 and over. From the

5,034 observations in Wave I, we initially restrict our attention to males aged 22 to 64. The focus here is on males, both because they are far more likely to suffer from alcoholism than are females and because of the considerable body of accumulated research regarding the specification of earnings models for males.[14] We select this age group in order to mitigate problems of incomplete education (on the younger end) and retirement (on the older end). Given the oversampling of the elderly at the New Haven site, the upper end restriction reduced our sample size considerably.[15]

Table 1 describes the dependent and independent variables and table 2 displays the descriptive statistics of the sample. With the exception of alcoholism, these tables are largely self-explanatory.[16] Alcoholism is measured in two ways: whether or not the individual ever met the criteria for diagnosis of alcoholism (ALC-EVER=1 or 0); and whether or not the individual who ever met the criteria had symptoms in the past year (ALC-YEAR=1 or 0). Assessment of alcoholism and other mental disorders in the ECA is via a professionally designed survey instrument, the Diagnostic Interview Schedule (DIS). The DIS consists of a battery of questions on symptoms.[17] These symptoms are then used to obtain diagnoses consistent with the American Psychiatric Association's (1980,1987) criteria for diagnoses of mental disorders.[18] Such a diagnostic approach avoids both the potential for self-reporting biases with regard to alcoholism as well as the self-selection problem in which only individuals who seek treatment can be determined to have mental disorders. Although the issue of the accuracy and quality of the DIS-based diagnoses is an open one, it is fortunate for our purposes that the ECA's DIS-based diagnosis of alcoholism has been found to have good correspondence with alternative diagnostic approaches (Anthony et al. (1985)).

III. Earnings, Human Capital, and Alcoholism

The tradition of including measures of individuals' health status as components of their human capital in wage and earnings functions is well

established.[19] The basic framework posits an earnings function

$$y = y(H,K,X) + \varepsilon \tag{1}$$

where y is some measure like log-earnings; H is a vector of measures of the health components of human capital; K is a vector of non-health human capital measures (schooling, experience, etc.); X is a vector of other covariates (age, race, sex, etc.); and ε is a stochastic error, generally assumed to satisfy $E(\varepsilon|H,K,X)=0$.

Following this tradition, we specify H=(A,S), where A is a vector of measures of alcoholism and S represents other health outcomes. The econometric counterpart to (1) is specified to be linear,

$$y_t = z_t\alpha + \varepsilon_t, \tag{2}$$

where y_t is log-income; z_t (1×k) summarizes all exogenous variables, $z_t=(A_t,S_t,K_t,X_t)$; and α (k×1) is a vector of unknown parameters. In addition, we specify a probit equation (Heckman (1979)) to describe the outcome of full-time participation in the labor market. The observed binary participation variable, I_t, is generated by a linear latent variable model whose error term is assumed to be homoskedastic normal,

$$I_t^* = z_t\beta + \eta_t, \tag{3}$$

with β (k×1) a vector of unknown parameters. Thus $I_t=1(I_t^*>0)$, where 1(.) is the 0-1 indicator function. One standard result that will be of use below is that, in general,

$$E(y_t|I_t=1,z_t) = z_t\alpha + \gamma\lambda_t \neq z_t\alpha,$$

where λ_t is the inverse Mills ratio under an assumption of normality for ε_t, and γ is an unknown scalar parameter that is a function of $\text{cov}(\varepsilon,\eta)$. It is assumed throughout that there are available N independent observations on (y_t, I_t, z_t).

That other components of human capital (schooling, experience, marital status, etc.) may be correlated with, and to some degree determined by, the health component(s) is usually ignored in the context of earnings function estimation. To the extent that the health components are structural determinants of the non-health components (e.g. if schooling attainment depends structurally on health), then it is very easy to understate empirically the total productivity of health capital, i.e. dy/dH, since indirect effects operating through the non-health components would not generally be captured. A fundamental point can be made in the context of the human capital framework sketched above (reinforced empirically by results presented below in table 6). Suppose for simplicity that $y=y(A,K)+\varepsilon$ with $E(\varepsilon|A,K)=0$, and consider how $E(y|A,K)$ varies with A:

$$dE(y|A,K)/dA = y_A + y_K dK/dA. \qquad (4)$$

Equation (4) emphasizes what we consider to be a potentially important omission in measuring the productivity effects of alcoholism, or, for that matter, of any disorder of interest. That is, the total effects of A on y are given not simply by the partial derivative y_A that holds all else constant, but rather by the total derivative that allows K to vary in response to variations in A. Accordingly, both direct (y_A) as well as indirect ($y_K dK/dA$) channels of influence must be admitted as possibilities if all the effects of A on y are to be evaluated and well understood in the design of policies targeted to mitigate or prevent alcoholism-related problems. For instance, see Cook and Moore (1990) and Mullahy and Sindelar (1989,1990a,1990b) for recent discussions of the relationships between alcohol use and educational attainment.

It is also important to note that standard static earnings-health models like (1) typically assume that the health variables are econometrically exogenous. In a life cycle context, such an assumption may be tenuous since health and labor market outcomes will be jointly determined in a health production context. Unfortunately, the data at our disposal do not suggest any reasonable instruments that would enable us to control for possible nonzero correlation between A and ε. Even if candidate instruments were available that could reasonably be assumed to be correlated with A but uncorrelated with unobservables, the potential for bias is considerable when such instruments have low correlation with included endogenous variables, particularly when sample sizes, as here, are modest.[20]

Accordingly, mainly out of econometric necessity, we treat alcoholism as a predetermined or exogenous determinant of labor market and other sociodemographic outcomes, much like health status is typically treated in such models. This perspective on alcoholism is consonant with most of the medical literature discussed earlier that considers alcoholism a disease; whether this would be appropriate under a conscious choice perspective is a more tenuous matter. This paper thus implicitly assumes the disease perspective: alcoholism is "predetermined" because of genetics, family background, or whatever other reason. Our empirical results are thus reasonable to the extent that the exogeneity of alcoholism is a valid maintained assumption, but we readily admit that biased estimates of alcoholism's role in labor market success may arise should this assumption be invalid.[21] Moreover, we focus primarily on ALC-EVER as the measure of alcoholism, since whether one has ever had an alcoholism problem is less likely to be correlated with contemporaneous unobservables than whether one is currently drinking to excess.[22]

IV. Life Cycle Dimensions to Alcoholism Problems

This section compares how labor force participation and earnings profiles vary over the life cycle for alcoholics (defined in various ways) and non-

alcoholics. This life cycle perspective provides some insights into why earlier studies have come up with conflicting results about the effect of alcohol consumption and alcoholism on earnings, income, and wages. While some studies have found negative effects (Berry and Boland (1977)), others have found no significant effects (Benham and Benham (1982)), and still others have found positive impacts (Berger and Leigh (1988)).

In our analysis of life cycle profiles we are particularly interested in the youngest and oldest age categories as it is these groups for whom seemingly counterintuitive relationships between alcoholism and labor market success might arise. For the youngest group, alcoholism may tend to increase labor market participation and thus increase earnings. One mechanism consistent with such a finding would be that those with alcohol problems would be more likely to have trouble in school and either drop out of school or work more hours even while attending school part-time. They would thus be working more hours and accumulating more labor market experience; increases in both hours and wage rates would thus tend to result in greater earnings. At a later point in the life cycle, however, the effects of non-alcoholics' greater educational attainment would overtake the experience advantages of the young alcoholics.

The oldest age group could also exhibit seemingly inconsistent working and earning profiles. Over their lives, alcoholics may accumulate less financial capital (pensions, savings, etc.) than would non-alcoholics so that early retirement may not be a reasonable option. Instead these aging alcoholics may continue employment later in life so that they may have higher contemporaneous labor market participation and income (but not necessarily wealth) as compared to their non-alcoholic counterparts who have begun to retire.

Table 3 compares the probability that an individual is employed full time (that is, whether the individual worked for pay all 12 months in the previous year, vacations excepted) for alcoholics and non-alcoholics by age group. The top row in the table suggests that for all ages combined, alcoholism (occurring either within the past twelve months (ALC-YEAR) or at any point in one's

lifetime (ALC-EVER)) has a dampening effect on full-time work propensity, although the difference is statistically significant only for alcoholism symptoms in the past year.

The rest of the table shows how these relationships vary across age categories. The second through fifth rows suggest that there are, in general, important differences in full-time work propensity, and also show how the effects of alcoholism vary over the life cycle. The youngest and the oldest age groups both show a positive labor force participation response to alcoholism measured by ALC-EVER by not ALC-YEAR. Note that the differences are not statistically significant, and that the sample of alcoholics is quite small in the oldest group. Conversely, for what might be considered the prime-aged males, ages 30-59, the effects of alcoholism are generally negative, significant and quite large, as seen in rows 3 and 4. (The exception is for current symptoms (ALC-YEAR) for individuals aged 30-44.) Interestingly, when considering current symptoms only (ALC-YEAR) the differences are far less striking, suggesting the possibility that the damages associated with alcoholism are much more subtle, far-reaching, and indirect than simply whether or not an individual currently is manifesting symptoms. We take up an econometric examination of this issue in Section V.

Table 4 displays the results of a similar analysis, with the focus now on income. The top row of the table demonstrates that for our sample of males, the effects of alcoholism on personal income are negative and fairly large. This result holds whether one considers the full sample (columns 2-5) or restricts attention to the sample of males who are full-time workers (columns 7-10). For full-time workers, income is likely to be a better proxy for earnings and, therefore, for productivity.[23] Again, however, aggregation over the age groups masks considerable heterogeneity in these relationships across the age groups. Statistical significance aside for the moment, alcoholism (measured either way) appears to have little effect on the incomes of the youngest group, a positive effect for the oldest group, but important negative effects on the incomes of

the prime-aged group in the middle.

This simple yet revealing examination of the data suggests that the effects of alcoholism vary in important ways over life cycle.[24] Whether alcoholism has positive, negative, or no effects on income and labor force participation depends on the point in the life cycle. This suggests that the various cross-sectional age groups should be examined separately and should not merely be pooled in analyses.[25] Given the pronounced differences in how alcoholism relates to income and to full-time work propensity between the 30-59 age group and both the younger and the older groups, most of the econometric analysis that follows will focus on this prime-aged group.[26] The overwhelming importance of these life cycle effects is unlikely to be captured fully by using the entire sample and resorting to the use of age polynomials in a regression model.[27]

Although we believe that this glimpse of the data is enlightening, it is insufficiently structured to address many questions one would like answered. Accordingly, we turn in the next section to a more structured econometric approach.

V. Econometric Estimates

In order to assess the relative magnitudes of the direct and indirect effects of alcoholism on labor market success, it is useful to consider a set of baseline estimates in which potentially important correlates of alcoholism have been omitted, thereby allowing the alcoholism coefficient estimate to absorb their impacts to the extent that such correlation is present. These results are presented in table 5, which compares estimates of LOG-INCOME (i.e. eq. 2) and FULLTIME (eq. 3) models for the entire sample 22-64 and for the 30-59 subsample. Consistent with the results in tables 3 and 4, we find that the magnitude of the effects of alcoholism depends on the age composition of the sample. For both LOG-INCOME and FULLTIME, the point estimates of the alcoholism effects are markedly larger and statistically more precise in the 30-59 sample. For this age group, ever having alcoholism has significant and large depressant effects

on both income and labor force participation. Whether these are meaningful estimates of structural effects, or merely the confounding indirect effects of omitted human capital covariates, will be considered below.[28]

The estimates presented in table 6 illustrate what might be thought of as the direct and indirect effects of alcoholism (specifically, ALC-EVER) on LOG-INCOME. In columns 2 through 6 we add to the variables in the column 1 specification additional variables that may be correlated with, and perhaps influenced by, alcoholism, thereby demonstrating how controlling for an increasingly more-inclusive vector of human capital covariates in z_t affects inferences about the effects of alcoholism. (For reference, column 2 of table 6 replicates the results from column 2 of table 5.)

First, note that the coefficient estimates for PHYSICALLY HEALTHY, educational attainment (measured either by SCHOOLING or by the HIGH SCHOOL and COLLEGE dummies), MARRIED, and OTHER INCOME are statistically significant by conventional standards. The net effect on the ALC-EVER coefficient point estimate of including these variables is substantial.[29] Using $\exp(\hat{\alpha}_j)-1$ to estimate the percentage change in $E(y_t|z_t)$ due to turning on the j-th dummy variable, the direct effect of ALC-EVER shrinks from a 31% reduction in mean income (based on the column 1 estimate) to a 17% reduction (based on the column 6 estimate) when we control for all the other human capital covariates.

We interpret the coefficient of ALC-EVER in column 1 as an estimate of the full effect of alcoholism on income. In this case, the coefficient picks up all the effects of alcoholism -- direct and indirect -- that could occur through the omitted human capital variables. Interpreting the difference between the transformed ALC-EVER coefficient in column 1 and that in column 6 as the magnitude of the indirect effects of alcoholism suggests that such indirect effects are substantial, approximately one-half the total effect. This suggests that studies will vary in their estimates of the labor market effects of alcoholism depending on the extent to which they control for covariates correlated with alcoholism, a simple yet critical point that is largely

unrecognized in the pertinent health literature.

We turn now to consider a variety of alternative model specifications and estimation strategies that are designed to determine the robustness of what we view as our central results in table 6. These "sensitivity analyses" are presented in tables 7, 8, and 9.

In table 7, the model estimates displayed in column 1 are used to assess whether there are effects of current alcoholism symptoms on income beyond those attributable to ALC-EVER. Accordingly, both ALC-EVER and ALC-YEAR are included as covariates. We find that little additional information is contained in the ALC-YEAR variable. When only ALC-YEAR is included (column 2), its effects are somewhat stronger than when included with ALC-EVER, but not up to conventional criteria for statistical significance.

In rather different ways, the model estimates presented in columns 3 and 4 of table 7 attempt to control for individuals' "baseline" circumstances. Column 3 includes two dummy variables describing the timing of onset of alcoholism symptoms: symptoms occurring prior to age 19 (ALC-PRE19) and symptoms occurring between ages 19 and 22 (ALC-1922). While both point estimates have the anticipated negative sign, neither is statistically significant.

The model estimated in column 4 takes a somewhat different approach. The sample here is defined so that the educational attainment of the fathers of the ECA survey respondents is exactly 12 years, thus explaining the smaller sample size of N=123. All individuals in this subsample can thus be said to have at least one (but certainly not all) initial conditions in common. Despite the small sample size, the point estimate of the ALC-EVER effect remains statistically significant and is considerably larger in absolute value than in the comparable specification in column 6 of table 6. On the basis of these estimates we are led to conclude that some forms of unobserved heterogeneity in baseline or initial conditions are likely to be correlated with the included health covariates. How and to what extent this is so is an important research issue that demands considerably richer data than that available here.

In column 5 of table 7 we consider the possibility that estimates of the alcoholism coefficients may be corrupted by the omission of measures of other emotional or mental disorders that are correlated with alcoholism. Accordingly, we include in the vector of explanatory variables a dummy variable indicating whether the individual ever suffered from antisocial personality disorder -- a potentially important comorbidity of alcoholism -- as well as a dummy variable indicating whether the individual reports his mental emotional health to be excellent/good versus fair/poor. Neither of these additional mental health covariates is statistically significant, and, more importantly, their inclusion has no material effects on the point estimate of the ALC-EVER coefficient or its significance (comparing column 1 of table 7 to column 5).

The last results in table 7 are an assessment of the sensitivity of our results to the recoding of the income variable. We consider two respecifications of the LOG-INCOME measure, each using different lower and upper censoring cut-offs. For values of income below some L we treat LOG-INCOME as if lower censored, for values of LOG-INCOME between L and some U we use the interval midpoints (as in the preceding analysis), and for values of LOG-INCOME greater than U we treat LOG-INCOME as if upper-censored. The new measures thus have the character of a two-limit Tobit model, and should provide a useful check on the estimates presented above. To attain robustness beyond that offered by ML estimation of such a Tobit model under normality and homoskedasticity assumptions, we instead estimate a doubly-censored version of the symmetrically-censored least squares (SCLS) estimator proposed by Powell (1986).[30] Unlike standard ML Tobit, consistency of Powell's SCLS estimator relies only on the ε_t having conditionally symmetric densities (in addition to some other technical regularity conditions discussed by Powell).

Two sets of SCLS estimates are estimated, and the results are presented in columns 6 and 7 of table 7. The appropriate comparison results in both cases are those in column 6 of table 6. In column 6, we set L=ln(.5) and U=ln(100) in order to account for censoring of just the lower and upper categories in the

data. In column 7 we impose more stringent censoring, L=ln(5) and U=ln(50), given the possibility that the income data within the lower tail below $5,000 and within the upper tail above $50,000 may be most prone to misreporting and because the interval midpoints are less satisfactory proxies in the tails than in the center of the distribution. In the first case (column 6), the results differ only trivially from the results in table 6. When the more stringent censoring is applied to the data (column 7) the differences are somewhat more marked, yet the point estimate of alcoholism's effect is nonetheless in the same ballpark as in the previous set. We are led to conclude that few if any important inferences are being missed because of the censored income variables.

While the results in tables 6 and 7 permit a relatively straightforward interpretation of direct and indirect effects, it is unlikely that models such as these -- linear in alcoholism -- capture all the subtleties of the relationships between alcoholism and labor market success. To this end, we present an additional set of estimates of models in which the ALC-EVER measure is permitted to interact with one or more of the other human capital covariates. While such interactions permit a much enhanced resolution of alcoholism's role in labor market success, they blur to some extent the distinction between what are direct and what are indirect effects.

The results of these interaction models are displayed in table 8. In the first column, ALC-EVER is interacted with the HIGH SCHOOL and COLLEGE dummies. The results here are striking: while the point estimate of the ALC-EVER coefficient remains significant, the significant interactions with both HIGH SCHOOL and COLLEGE suggest that there are some important synergistic effects between schooling attainment and alcoholism. The heteroskedasticity-consistent Wald test statistics for the joint significance of the two interaction terms and for the ALC-EVER coefficient along with the two interaction terms are 8.56 and 9.93, both corresponding to $p<.02$ for χ^2 variates with d.f.=2 and d.f.=3, respectively.

The results for the HIGH SCHOOL variables are quite sensible: non-alcoholics who complete high school are best off, alcoholics who complete high school are next, non-alcoholics who fail to complete high school follow, while worst off are alcoholics who do not complete high school.[31] For the COLLEGE variables, the one possibly curious aberration from this pattern is that the alcoholic college graduates marginally outperform their non-alcoholic counterparts.[32]

Columns 2 and 3 in table 8 present two analogous model estimates, the ALC-EVER interactions now with SCHOOLING (column 2) and with SCHOOLING and AGE (column 3).[33] The positive point estimate on the interaction term in column 2 suggests the possibility that additional years of schooling may mitigate the otherwise adverse effects of alcoholism; a Wald test indicates that the ALC-EVER and ALC-EVER×SCHOOLING coefficient estimates are jointly significant at $p<.05$. Given this schooling effect, however, the results in column 3 suggest that there is little additional effect -- positive or negative -- due to an ALC-EVER×AGE interaction; indeed, the joint significance of the two interaction terms in the column 3 specification is $p>.18$.

Column 4 in table 8 considers an alcoholism interaction with physical health status. While the coefficient estimate on the interaction is not significant by usual standards, the point estimates (as they were for the HIGH SCHOOL variables) accord well with common sense: non-alcoholics in good health are best off, alcoholics in good health are next, non-alcoholics in poor health follow, while worst off are alcoholics in poor health. Wald test statistics for the joint significance of the linear and interaction ALC-EVER terms and for these along with the linear PHYSICALLY HEALTHY term are 5.99 ($p<.05$) and 14.81 ($p<.01$), these corresponding to χ^2 variates with d.f.=2 and d.f.=3, respectively. Given our relatively small sample size, it is no small task to tease out second-order effects. Nonetheless, the results in table 8 are sufficiently strong to suggest -- at least for schooling attainment -- that there is indeed some important interplay in how alcoholism and schooling jointly

determine labor market outcomes.

Finally, our last set of alternative specifications are motivated because individual income as recorded in the ECA represents income from all sources -- labor earnings, non-labor earnings, transfer payments, etc. Accordingly, we attempt to obtain a measure of income that is likely to be closer to a measure of earnings that, we feel, would better measure the productivity effects of alcoholism. To derive such a measure, we first take the subsample of individuals that are full-time workers. Then we identify the subsample of observations reporting that they received no transfer payments in the form of social security, disability, welfare, or unemployment compensation. The intersection of the subsamples that report full-time work and no transfer receipt (FULL-NOTRANS=1 in order for the observation to be in the subsample) is a subsample for which we feel the reported individual income measure better approximates earnings.[34] We model this process econometrically as a two-step Heckman selection process, with a probit model describing the determinants of FULL-NOTRANS, and a λ-corrected LOG-INCOME model estimated on the subsample for which FULL-NOTRANS=1. It should be stressed that because all variables in the FULLTIME model (interpreted as a labor supply model) must of necessity be included in the earnings equation (interpreted as labor supplied times the wage rate), the sample selection process that drives the selection into the population of full-time workers is identified only by the nonlinear functional form of the selection-correction term, λ.

The results of this exercise are presented in table 9. For both the baseline (columns 1 and 2) and the augmented (columns 3 and 4) models, the results suggest that the direct effects of ALC-EVER on income are working largely through participation effects rather than through wage/productivity effects. That is, the ALC-EVER effect is statistically much stronger in the equation determining FULL-NOTRANS than in the equation determining LOG-INCOME conditional on FULL-NOTRANS=1.[35]

The general finding is consistent with the findings from table 2 where the income differences between alcoholics and non-alcoholics are greater for all individuals as compared to the differences conditional on working full-time. Although evidence from both tables 2 and 9 is consistent with this interpretation, there are clearly many alternatives to examine before such a conclusion could be confirmed. Moreover, we share the concern common in applied microeconometrics when the selection model for conditional LOG-INCOME is identified solely by the nonlinearity of the λ term. At a minimum, however, the results are suggestive of some avenues for future research pursuits.

The results that we present are in many respects not definitive. Other interpretations and confounding factors have not yet been eliminated. For example, alcoholism may itself be a symptom of deeper problems that may also result in reduced earnings, so that elimination of alcoholism per se would not necessarily imply that earnings would be of the same magnitude of otherwise similar individuals without alcoholism. Furthermore, as we have stressed repeatedly, the direction of causation cannot be determined; lower earnings may certainly be a factor in the onset of alcoholism symptoms.[36] Of course, the altogether separate possibility that unobserved heterogeneity is ultimately driving all the outcomes must be admitted.

VI. Summary

Studies of the effects of alcohol on earnings, income, and productivity have yielded conflicting results. The popular view that has been confirmed in several studies is that alcohol problems, or, more specifically, alcoholism, have depressant effects on income. However, other studies have found insignificant effects or even positive effects related to alcohol. Part of the confusion owes to the difference between alcohol consumption and alcoholism.[37] The fact that studies have used different measures of income and/or have focused on different populations (e.g. workers only) only compounds the confusion.

Our finding of important life cycle aspects has several implications. One is that the answer one obtains from empirical analysis of the effect of alcoholism (or alcohol) on income may depend on the point in the life cycle one is examining (or the proportion of the sample population in each age category). Another issue emphasized by taking the lifetime perspective is that income alone may not be an accurate measure of well-being. Those alcoholics who earn more in youth, withdraw from school, and work more hours are not necessarily better off. Similarly, the older alcoholic who may have greater income and less leisure time is also not necessarily in a preferred position. Our results also suggest that alcoholism has a more significant impact on the likelihood of working than it does on how much earned when working. Studies may thus vary in their estimates of the impact of alcoholism depending on whether or not they condition on workers only or use a broader sample.

We also have shown that the extent to which one controls for variables correlated with alcoholism (e.g. schooling and marital status) has a large impact on the estimated effects of alcoholism. Estimates of the magnitude of the effect of alcoholism on earnings may thus differ across studies depending on the extent to which one controls for such covariates. The full effect of alcoholism may be estimated by omitting such correlated variables. However, one could equally well be interested in the estimating the effect of alcoholism on earnings after controlling for the indirect effects; the different estimates correspond to different lines of inquiry.

To summarize, our results suggest that, at least for the prime-aged males, alcoholism has negative and significant impacts on labor force participation and income. While this paper has not solved all margins of the alcohol-wage puzzle, we feel it has suggested some important directions for future research. These would include -- but not be limited to -- further examination of alcoholism's life cycle course, direct versus indirect effects on labor market outcomes, interactions in the schooling-alcoholism process, and differential impacts on wages and participation.

Notes

1. An earlier version of this paper was presented at the Alcohol and Public Policy session at the 1990 AEA Meetings in Washington. The authors would like to thank Phil Cook, Michael Grossman, Paul Portney, Chris Ruhm, David Salkever, Paul Schultz, seminar participants at Harvard, Johns Hopkins, and Vanderbilt, and the referee for helpful suggestions and comments on earlier drafts. NIAAA Grant 1R01AA08394 to Yale University provided research support. The usual disclaimer applies.

2. Dept. of Economics, Trinity College, Hartford, CT 06106; Resources for the Future; and NBER. Research supported in part by a University Fellows grant from Resources for the Future.

3. Dept. of Epidemiology and Public Health, Yale University, Box 3333, New Haven, CT 06510; and Institution for Social and Policy Studies, Yale University, Box 16A Yale Station, New Haven, CT 06520.

4. U.S. Dept. of Health and Human Services, National Institute on Alcohol Abuse and Alcoholism (1990) (henceforth: NIAAA (1990)).

5. Berry and Boland (1977) and Rice et al. (1990) are representative of such findings.

6. See Berger and Leigh (1988) for evidence of beneficial labor market effects of alcohol, Cook (1990) for additional discussion of such results, and Shaper et al.

(1988) for a discussion of the findings of a positive health effect and some evidence explaining why the positive effects may be incorrect interpretations.

7. See Vaillant (1983), Chapter 3, for an interesting discussion.

8. It might also be noted that alcoholism is generally estimated to be quite prevalent in the homeless population (Institute of Medicine (1988)) as well as in some areas of the institutionalized population. The prevalence rate in the homeless population is suggested to be 20%-45% (NIAAA (1990)). Since this analysis is based on a residential sample, the interesting and important issue of the effects of alcoholism in the homeless and institutionalized populations will not be treated here. Nonetheless, to the extent that these "nonresidential" populations have greater than average propensities to be alcoholic and lower than average labor market success, our results would likely tend to underestimate the economic impacts of alcoholism in the entire population.

9. Evidence on the effectiveness of treatment is mixed (see, for example, Hayashida et al. (1989) and NIAAA (1990)). Indeed, lacking evidence on the efficacy of inpatient treatment, many third-party payers are no longer covering inpatient treatment. Many individuals seek repeated treatment and try a variety of different approaches, ranging from self-help groups, to inpatient group therapy, to drug treatment. Given remission after such sequences of treatment, it is not clear which, if any, treatment was effective, if the sequence itself mattered, or if individuals self-selecting into treatment were those relatively more likely to succeed.

Moreover, according to a prevalent view, alcoholics who are not currently manifesting symptoms are "recovering" or "in remission," although they never

completely recover; "once an alcoholic, always an alcoholic," even though symptoms may not be manifested currently.

10. See, for example, Fingarette (1988) who maintains that the disease view has been perpetuated by various special interest groups that benefit from alcoholism's classification as a disorder. Perhaps not surprisingly, his perspective is the minority in the medical literature. The "rational addiction" theory of Becker and Murphy (1988) should also be noted in this context.

11. For details on the ECA surveys, see Eaton and Kessler (1985), Reiger et al. (1984), and Robins et al. (1981).

12. We restrict attention in this analysis to the New Haven site because its data on labor market outcomes are much richer than the labor market data available from the other four survey sites: Durham, NC; Baltimore, MD; St. Louis, MO; and Los Angeles, CA.

13. Before the availability of the ECA surveys, studies of the economic and social consequences of alcoholism had of necessity relied on data that were unsatisfactory in one way or another. Weaknesses have included self-diagnoses of alcoholism, data obtained from individuals' visits to medical care facilities, unavailability of important covariates, and others. For instance, in an often-quoted study Berry and Boland (1977) relied on a data set that included only household (not individual) income data and data on alcoholism for only one individual per household. Moreover, until the ECA surveys were conducted, assessments of disorders on the basis of state-of-the-art psychiatric diagnostic criteria were

not available in large community data sets; mental health was typically assessed only by direct self-reporting of a diagnosis in large data sets.

14. See Willis (1986) for a good survey of earnings function estimation, and Mullahy and Sindelar (1990c,1991) for a discussion of gender differences in the effects of alcoholism and other mental health problems.

15. The reduction in sample size from the original 5034 observations to the 555 observations we use in much of the econometric analysis is due to the following set of restrictions:

a. Nonelderly (ages 18-64): 2458 remaining observations;

b. Initial age cutoff (ages 22-64): 2237 remaining observations;

c. Restriction to estimation sample (ages 30-59): 1420 remaining observations;

d. Restriction to males: 604 remaining observations;

e. Miscellaneous missing data: 555 remaining observations.

16. It might be noted, however, that the variables SCHOOLING and INCOME are created using interval midpoints. For SCHOOLING, 17 years was used for the open-ended upper interval "grad school." The 1980-81 survey asks the respondent to report income in the preceding year; income is thus expressed in 1979-80 dollars. INCOME consists of both labor income and other income "brought into" the household by the individual. For this measure, "0.5" was used for the bottom interval "less than $1,000" and "120" was used for the upper open-ended interval "over $100,000." While this approach is admittedly ad hoc, it greatly simplifies the econometrics. We present some evidence in table 7 that explicitly accounting for the censoring of the income measure yields results that differ little from those obtained using the "fill in the upper end" method.

17. Several issues concerning the disorder measures might be noted. First, the data on disorders are not self-reported in the sense that individuals are asked directly, for example, "Do you suffer from alcoholism?" or "Are you schizophrenic?"; rather, an indirect method (the DIS) that involves a battery of diagnostic questions for each disorder is used to ascertain the diagnoses of both current and past disorders. Nonetheless, the input information for the DIS is provided by the respondent. Accordingly, the responses to these questions may involve recall bias. Second, severity is not measured, although we attempt to control for possible severity effects by using the individual's self reported emotional/mental health status in some specifications. In alternative specifications, we control for antisocial personality disorder in analyzing the effects of alcoholism. This is potentially important since antisocial personality is one of the most frequently observed comorbidities of alcoholism for males.

18. Although the DIS criteria conform to the DSM-III definition (American Psychiatric Association (1980)) of alcohol abuse and dependence (which we refer to as "alcoholism"), we revised the diagnosis of alcoholism to conform more closely to the more recent DSM-III(R) criteria (American Psychiatric Association (1987)). In principle, this revision could have been important because the DSM-III(R) definition drops some of the labor market behavior symptoms used in diagnosis, clearly a problem when using the diagnosis as an explanatory variable in labor market outcome models, as we do below. However, for our category of alcohol abuse and dependence, for no observation in the sample was the definition of alcoholism changed with the revision.

19. See, for instance, Grossman (1972), Bartel and Taubman (1979,1986), and Mitchell and Butler (1986).

20. See Nelson and Startz (1990).

21. Even though our data do not allow us to meaningfully address the cause-effect issue in the empirical analysis, we offer the simple yet often overlooked point that a nonzero dK/dA implies that there is some correlation between K and A -- causality may or may not be present. From a policy perspective, it is important to recognize that over the course of the life cycle A and K are likely to be interwoven in an intricate, complicated manner, and that attributing "costs" to alcoholism without cognizance of such interrelationships -- i.e. failing to account for possible indirect as well as the direct effects of alcoholism on earnings -- is likely to lead to underestimates of such costs.

22. The use of the DSM-III(R) measure of alcoholism also helps justify its exogeneity in labor market equations relative to the DSM-III measure which includes symptoms of trouble at work (see note 18).

23. In Section V we assess in greater detail this question about the role of alcoholism as a determinant of income when the focus is on full-time workers only.

24. Interpreting alcoholism as a "hazardous behavior" the finding that there are important life cycle dimensions to the dimensions to the effects of alcoholism should not be surprising. See Ippolito (1981) and Ehrlich and Chuma (1990) for discussion.

25. The above discussion has been cast in a life cycle context. It is possible, of course, that the effects are not age effects as much as they are cohort

effects, but neither our modest sample, nor (to our knowledge) any other contributions in the literature (NIAAA (1990)), have been up to the task of disentangling these elements. Accordingly, we will continue the discussion as if they are age effects, but readily admit that competing hypotheses must be entertained.

26. Using the specification reported in column 6 of table 6, we conduct a Wald test to determine whether it is reasonable to pool the 30-44 and the 45-59 age groups. The test statistic, distributed χ^2 with 9 d.f. under the null, has a value of 10.27. Since the corresponding p-value exceeds .3, pooling would not appear to be unreasonable. In some preliminary estimates of income models for the younger group we find that the alcoholism coefficient estimate is positive (as expected) but insignificant.

27. See Murphy and Welch (1990) for an interesting discussion of the nonlinear structure of age effects in earnings models.

28. Note for both LOG-INCOME and FULLTIME the estimated AGE and AGE SQUARED effects for the 30-59 sample are less significant and somewhat smaller than they are for the full sample. Since the most steeply sloped segment of the age-income profile is likely to be at ages less than 30, this is not a surprising result. See Murphy and Welch (1990) for further discussion.

29. Entering individually each of the omitted human capital variables, schooling attainment as measured by the two school success dummies (HIGH SCHOOL and

COLLEGE) had the largest single impact on the coefficient of ALC-EVER. The age 30-59 sample correlation matrix for the human capital covariates is:

	ALC-EVER	SCHOOLING	PHYSICALLY HEALTHY	HIGH SCHOOL	COLLEGE	MARRIED
SCHOOLING	-0.1449					
PHYS. HLTHY.	-0.1426	0.2596				
HIGH SCHOOL	0.0066	-0.1913	0.0895			
COLLEGE	-0.1291	0.7969	0.1459	-0.6755		
MARRIED	-0.1364	-0.0598	0.0596	0.0824	-0.0653	
OTH.INCOME	0.0124	0.0241	0.0123	0.0737	-0.0274	0.1231

In addition, in related work we have found negative and significant correlations between early onset of alcoholism and educational attainment (Mullahy and Sindelar (1990b)).

30. One way to characterize the SCLS estimator is as the (implicit) solution $\hat{\alpha}$ of

$$\hat{\alpha} = [\Sigma_{t=1}^{T} (\omega_{t1} + \omega_{t2}) z_t' z_t]^{-1} \times [\Sigma_{t=1}^{T} z_t' (\omega_{t1} \min\{y_t, 2z_t\hat{\alpha} - L\} + \omega_{t2} \max\{y_t, 2z_t\hat{\alpha} - U\})],$$

where

$$\omega_{t1} = 1(L < z_t\hat{\alpha} < (U+L)/2)$$

and

$$\omega_{t2} = 1((U+L)/2 < z_t\hat{\alpha} < U).$$

See Sota (1989) for a recent application.

31. It is useful to recall the often overlooked interpretation of coefficients when dummy variables are interacted. That is, with dummies z_1 and z_2, and

$$E(y|z) = \beta_1 z_1 + \beta_2 z_2 + \beta_{12} z_1 \times z_2 + \ldots \text{ terms not involving } z_1, z_2 \ldots ,$$

then the omitted category is $(z_1, z_2)=(0,0)$. The effects relative to $(0,0)$ (e.g. a non-alcoholic, non-high-school-grad) of the three other (z_1, z_2) combinations are:

$(1,0) = \beta_1$ (e.g. alcoholic, non-high-school-grad)

$(0,1) = \beta_2$ (e.g. non-alcoholic, high-school-grad)

and

$(1,1) = \beta_1 + \beta_2 + \beta_{12}$ (e.g. alcoholic, high-school-grad)

32. Cook and Moore (1990) suggest several plausible competing hypotheses under which successful completion of college could be negatively or positively related to students' drinking behavior.

33. We also estimated several alternative versions of the model, where quadratics in SCHOOLING and interactions between SCHOOLING and AGE were included. In none of these specifications, however, were the SCHOOLING^2 or the AGE×SCHOOLING coefficient estimates individually or jointly significant. These results are available on request.

34. It might also be noted that when only full-time workers are sampled, earnings are more likely to proxy for wages since the variation in hours over the year is likely to be reduced considerably.

35. Interestingly, while its significance is quite low, the point estimate on ALC-EVER in the LOG-INCOME model estimated conditional on FULL-NOTRANS=1 differs little from its value when the full sample is used for estimation.

36. Vaillant (1983) quotes Enoch Gordis (now Director of NIAAA): "Changes in personality or mood are now recognized to be largely the consequence of alcoholism, not its cause."

37. For instance, in a study of alcohol consumption (not alcoholism), Berger and Leigh (1988) find a positive effect of consumption on earnings.

References

American Psychiatric Association. 1980. Diagnostic and Statistical Manual of Mental Disorders, Third Edition. Washington: American Psychiatric Association.

_____. 1987. Diagnostic and Statistical Manual of Mental Disorders, Third Edition, Revised. Washington: American Psychiatric Association.

Anthony, J. et al. 1985. "Comparisons of the DIS and a Standardized Psychiatric Diagnosis." Archives of General Psychiatry 42: 667-675.

Bartel, A. and P. Taubman. 1979. "Health and Labor Market Success: The Role of Various Diseases." Review of Economics and Statistics 61: 1-8.

_____ and _____. 1986. "Some Economic and Demographic Consequences of Mental Illness." Journal of Labor Economics 21: 243-256.

Becker, G.S. and K.M. Murphy. 1988. "A Theory of Rational Addiction." Journal of Political Economy 96: 675-700.

Benham, L. and A. Benham. 1982. "Employment, Earnings, and Psychiatric Diagnosis." in V. Fuchs, ed. Economic Aspects of Health. Chicago: University of Chicago Press.

Berger, M.C. and J.P. Leigh. 1988. "The Effect of Alcohol Use on Wages." Applied Economics 20: 1343-1351.

Berry, R. and J. Boland. 1977. The Economic Costs of Alcohol Abuse. New York: Free Press.

Cook, P.J. 1991. "The Social Costs of Drinking." in Expert Meeting on Negative Social Consequences of Alcohol Use. Oslo: Norwegian Ministry of Health and Social Affairs.

_____ and M.J. Moore. 1990. "Drinking and College Completion." Presented at the Annual Meetings of the American Economic Association, Washington, DC.

Eaton, W. and L. Kessler. 1985. Epidemiological Methods in Psychiatry: The NIMH Epidemiologic Catchment Area Program. New York: Academic Press.

Ehrlich, I. and H. Chuma. 1990. "A Model of the Demand for Longevity and the Value of Life Extension." Journal of Political Economy 98: 761-782.

Fingarette, H. 1988. Heavy Drinking. Berkeley: U. of California Press.

Frank, R. and P. Gertler. 1990. "Assessment of the Measurement Error Bias for Estimating the Effect of Mental Distress on Income." Journal of Human Resources (forthcoming).

Grossman, M. 1972. "On the Concept of Health Capital and the Demand for Health." Journal of Political Economy 80: 223-255.

Hayashida, M. et al. 1989. "Comparative Effectiveness and Costs of Inpatient and Outpatient Detoxification of Patients with Mild-to-Moderate Alcohol Withdrawal Syndrome." New England Journal of Medicine 320: 358-365.

Heckman, J.J. 1979. "Sample Selection Bias as Specification Error." Econometrica 47: 153-161.

Institute of Medicine. 1988. Homelessness, Health, and Human Needs. Washington: National Academy Press.

Ippolito, P.M. 1981. "Information and the Life Cycle Consumption of Hazardous Goods." Economic Inquiry 19: 529-558.

Luft, H. 1975. "The Impact of Poor Health on Earnings." Review of Economics and Statistics 57: 43-47.

Mincer, J. 1974. Schooling, Experience, and Earnings. New York: Columbia University Press for NBER.

Mitchell, J.M. and J.S. Butler. 1986. "Arthritis and the Earnings of Men: An Analysis Incorporating Selection Bias." Journal of Health Economics 5: 81-98.

Mullahy, J. and J. Sindelar. 1989. "Lifecycle Effects of Alcoholism on Education, Earnings, and Occupation." Inquiry 26 (Summer): 272-282.

_____ and _____. 1990a. "An Ounce of Prevention: Productive Remedies for Alcoholism." Journal of Policy Analysis and Management 9: 249-253.

_____ and _____. 1990b. "Direct and Indirect Effects of Alcoholism on Human Capital." Mimeo, Yale University.

_____ and _____. 1990c. "Gender Differences in the Effects of Mental Health on Labor Force Participation." in I. Sirageldin, ed. Research in Human Capital and Development: Female Labor Force Participation. Greenwich, CT: JAI Press.

_____ and _____. 1991. "Gender Differences in Labor Market Effects of Alcoholism." American Economic Review 81 (Papers and Proceedings): 161-165.

Murphy, K.M. and F. Welch. 1990. "Empirical Age-Earnings Profiles." Journal of Labor Economics 8: 202-229.

Nelson, C.R. and R. Startz. 1990. "The Distribution of the Instrumental Variables Estimator and Its t-Ratio When the Instrument Is a Poor One." Journal of Business 63: S125-S140.

Powell, J.L. 1986. "Symmetrically Trimmed Least Squares Estimation for Tobit Models." Econometrica 54: 1435-1460.

Reiger, D.A. et al. 1984. "The NIMH Epidemiologic Catchment Area (ECA) Program:

Historical Context, Major Objectives, and Study Population Characteristics." Archives of General Psychiatry 41: 934-941.

Rice, D.P. et al. 1990. The Economic Costs of Alcohol and Drug Abuse and Mental Illness: 1985. Washington: PHS/ADAMHA DHHS Publication No. (ADM) 90-1694.

Robins, L.N. et al. 1981. "National Institute of Mental Health Diagnostic Interview Schedule: Its History, Characteristics, and Validity." Archives of General Psychiatry 38: 381-389.

Shakotko, R.A. and M. Grossman. 1982. "Physical Disabilities and Post-Secondary Educational Choice." in V. Fuchs, ed. Economic Aspects of Health. Chicago: University of Chicago Press.

Shaper, A., G. Wannamethee, and M. Walker. 1988. "Alcohol and Mortality in British Men: Explaining the U-Shaped Curve." The Lancet (Saturday, Dec. 3): 1267-1273.

Sota, M.S. 1989. "Relaxing Distributional Assumptions in the Estimation of Female Labour Supply Models." Oxford Bulletin of Economics and Statistics 51: 429-450.

Sournia, J.-C. 1990. A History of Alcoholism. Oxford: Basil Blackwell.

U.S. Department of Health and Human Services, National Institute on Alcohol Abuse and Alcoholism. 1990. Seventh Special Report to the U.S. Congress on Alcohol and Health. Washington: PHS/ADAMHA/NIAAA DHHS Publication No. (ADM) 90-1656.

Vaillant, G.E. 1983. The Natural History of Alcoholism: Causes, Patterns, and Paths to Recovery. Cambridge, MA: Harvard University Press.

White, H. 1980. "A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity." Econometrica 48: 817-838.

Willis, R.J. 1986. "Wage Determinants: A Survey and Reinterpretation of Human Capital Earnings Functions." in O. Ashenfelter and R. Layard, eds. Handbook of Labor Economics, Vol 1. Amsterdam: North-Holland.

Table 1

Variable Definitions

Variable	Definition
FULLTIME	= 1 if individual worked 12 months for pay in previous year (including paid vacations), = 0 else
TRANSFER RECIPIENT	= 1 if individual reported receiving transfer payments in the form of social security, disability, welfare, or unemployment compensation, = 0 else
FULL-NOTRANS	= 1 if FULLTIME=1 and TRANSFER RECIPIENT=0, = 0 else
INCOME	Of household's total income before taxes for past year, including salaries, wages, social security, welfare, and any other income, how much was earned or brought in by individual (÷1000)
LOG-INCOME	Natural log of INCOME
ALC-EVER	= 1 if ever met criterion of alcoholism, = 0 else
ALC-YEAR	= 1 if symptoms of alcoholism present in past year if ever met the criterion, = 0 else
ALC-PRE19	= 1 if earliest symptoms of alcoholism present at age 18 or earlier, = 0 if earliest symptoms later or never
ALC-1922	= 1 if earliest symptoms of alcoholism present between ages 19 and 22, = 0 if earliest symptoms at other time or never
ANTISOCIAL PERSONALITY	= 1 if ever met criterion for antisocial personality, = 0 else
MENTALLY HEALTHY	= 1 if individual reports emotional/mental health excellent or good, = 0 if reports fair or poor
PHYSICALLY HEALTHY	= 1 if individual reports physical health excellent or good, = 0 if reports if reports fair or poor
AGE	Age in years
AGE SQUARED	AGE squared
WHITE	= 1 if race is white, = 0 if race is nonwhite
SCHOOLING	Years of completed schooling
HIGH SCHOOL	= 1 if $12 \leq$ SCHOOLING ≤ 15, = 0 else
COLLEGE	= 1 if SCHOOLING ≥ 16, = 0 else
MARRIED	= 1 if currently married, = 0 else
OTHER INCOME	Other household income, measured as the greater of zero or household income minus personal income (in thousands)

Table 2

Sample Descriptive Statistics
(N.Obs. = 897 for Full Sample, = 555 for Subsample Ages 30-59)

Variable	Mean		Minimum		Maximum	
Left-Hand Side	Full Sample	Ages 30-59	Full Sample	Ages 30-59	Full Sample	Ages 30-59
FULLTIME	.764	.838	0	0	1	1
FULL-NOTRANS	.737	.809	0	0	1	1
INCOME	20.073	23.423	.500	.500	120.0	120.0
INCOME*	22.704	25.068	.500	.500	120.0	120.0
INCOME**	22.851	25.233	.500	.500	120.0	120.0
LOG-INCOME	2.717	2.912	-.693	-.693	4.787	4.787
LOG-INCOME*	2.959	3.067	-.693	-.693	4.787	4.787
LOG-INCOME**	2.962	3.070	-.693	-.693	4.787	4.787
TRANSFER RECIPIENT	.120	.090	0	0	1	1
Right-Hand Side						
ALC-EVER	.206	.204	0	0	1	1
ALC-YEAR	.106	.101	0	0	1	1
ALC-PRE19	.080	.059	0	0	1	1
ALC-1922	.064	.056	0	0	1	1
PHYSICALLY HEALTHY	.887	.899	0	0	1	1
ANTISOCIAL PERSONALITY	.090	.079	0	0	1	1
MENTALLY HEALTHY	.891	.906	0	0	1	1
AGE	39.689	41.861	22	30	64	59
WHITE	.861	.858	0	0	1	1
SCHOOLING	13.465	13.447	2	2	17	17
HIGH SCHOOL	.465	.436	0	0	1	1
COLLEGE	.359	.371	0	0	1	1
MARRIED	.658	.723	0	0	1	1
OTHER INCOME	5.584	4.532	0.0	0.0	58.5	57.5

* Computed on subsample for which FULLTIME=1, N.Obs.=685,465.
** Computed on subsample for which FULLTIME=1 and TRANSFER=0; N.Obs.=661,449.

Table 3

FULLTIME Workers:
Percentages by Age Group
and ALC-EVER Status
(N. Obs. in Parentheses; Decrease or Increase
Due to ALC=1 Relative to ALC=0 in Brackets)

Ages	All Observations	ALC-EVER Status		ALC-YEAR Status	
		ALC-EVER=0	ALC-EVER=1	ALC-YEAR=0	ALC-YEAR=1
All Ages	.764	.775	.719	.776	.663
	(897)	(712)	(185)	(802)	(95)
		{-.056}		{-.113}**	
Age Subgroups					
22-29	.675	.652	.738	.678	.657
	(243)	(178)	(65)	(208)	(35)
		{+.086}		{-.021}	
30-44	.845	.875	.733	.873	.643
	(348)	(273)	(75)	(306)	(42)
		{-.142}***		{-.230}***	
45-59	.826	.858	.684	.829	.786
	(207)	(169)	(38)	(193)	(14)
		{-.174}**		{-.043}	
60-64	.566	.565	.571	.568	.500
	(99)	(92)	(7)	(95)	(4)
		{+.006}		{-.068}	

For two-tailed tests of differences in proportions between ALC=0 and ALC=1,
* p<.10
** p<.05
*** p<.01

Table 4

INCOME: Means (in Thousands of $), by Age Group,
ALC-EVER Status, ALC-YEAR Status, and FULLTIME Status
(Number of Observations in Parentheses; Percent Decrease or
Increase Due to ALC=1 Relative to ALC=0 in Brackets)

	All Observations					FULLTIME = 1				
		ALC-EVER		ALC-YEAR			ALC-EVER		ALC-YEAR	
Ages	All Obs.	= 0	= 1	= 0	= 1	All Obs.	= 0	= 1	= 0	= 1
All Ages - 22-64	20.1	20.8	17.4	20.5	16.9	22.7	23.2	20.7	22.9	20.7
	(897)	(712)	(185)	(802)	(95)	(685)	(552)	(133)	(622)	(63)
		{-16.3%}***		{-17.6%}**			{-10.8%}		{-9.6%}	
Age Subgroups										
22-29	12.9	12.8	13.2	12.9	12.7	15.6	15.7	15.4	15.6	15.7
	(243)	(178)	(65)	(208)	(35)	(164)	(116)	(48)	(141)	(23)
		{+3.1%}		{-1.6%}			{-1.9%}		{+0.6%}	
30-44	23.5	24.2	21.1	24.3	18.1	25.2	25.2	25.3	25.4	23.3
	(348)	(273)	(75)	(306)	(42)	(294)	(239)	(55)	(267)	(27)
		{-12.8%}		{-25.5%}**			{+0.4%}		{-8.3%}	
45-59	23.3	24.7	16.8	23.3	21.9	24.8	25.7	19.6	25.0	22.3
	(207)	(169)	(38)	(193)	(14)	(171)	(145)	(26)	(160)	(11)
		{-32.0%}***		{-6.0%}			{-23.7%}***		{-10.8%}	
60-64	19.0	18.8	20.6	18.8	22.2	23.8	23.4	29.4	23.4	36.3
	(99)	(92)	(7)	(95)	(4)	(56)	(52)	(4)	(54)	(2)
		{+9.6%}		{+18.1%}			{+25.6%}		{+55.1%}**	

For two-tailed tests of differences in means, allowing for different variances,
* p<.10
** p<.05
*** p<.01

Table 5

Baseline LOG-INCOME and FULLTIME Specifications:
Comparing Results for Alternative Age-Inclusion Criteria
(asymptotic t-stats in parentheses,
heteroskedasticity-consistent for OLS)

	LOG-INCOME (OLS)		FULLTIME (PROBIT)	
Variable	Ages 22-64 (1)	Ages 30-59 (2)	Ages 22-64 (3)	Ages 30-59 (4)
ALC-EVER	-0.173	-0.312	-0.186	-0.465
	(2.43)	(3.33)	(1.58)	(3.01)
AGE	0.165	0.105	0.221	0.176
	(9.83)	(2.50)	(7.38)	(1.99)
AGE SQUARED	-0.002	-0.001	-0.003	-0.002
	(9.18)	(2.34)	(7.33)	(2.00)
WHITE	0.388	0.490	0.269	0.419
	(4.43)	(4.47)	(2.00)	(2.35)
PHYSICALLY HEALTHY	0.493	0.522	0.837	0.948
	(4.65)	(3.33)	(5.81)	(4.90)
CONSTANT	-1.509	-0.232	-4.466	-3.745
	(4.36)	(0.26)	(7.33)	(1.98)
N. Obs.	897	555	897	555

Table 6

OLS Model Estimates with Alternative Human Capital Covariates
(Subsample Ages 30-59; N. Obs.=555)
(heteroskedasticity-consistent t-stats in parentheses)

	Dependent Variable: LOG-INCOME					
Variable	(1)	(2)	(3)	(4)	(5)	(6)
ALC-EVER	-0.369	-0.312	-0.249	-0.236	-0.200	-0.188
	(3.86)	(3.33)	(2.77)	(2.69)	(2.31)	(2.20)
AGE	0.113	0.105	0.125	0.121	0.116	0.109
	(2.61)	(2.50)	(3.11)	(3.02)	(2.92)	(2.77)
AGE SQUARED	-0.001	-0.001	-0.001	-0.001	-0.001	-0.001
	(2.53)	(2.34)	(2.81)	(2.73)	(2.65)	(2.53)
WHITE	0.568	0.490	0.349	0.331	0.280	0.245
	(4.91)	(4.47)	(3.20)	(3.04)	(2.58)	(2.23)
PHYSICALLY HEALTHY	--	0.522	0.415	0.414	0.389	0.392
		(3.33)	(2.59)	(2.57)	(2.51)	(2.49)
SCHOOLING	--	--	0.065	--	0.070	0.072
			(5.67)		(6.39)	(6.69)
HIGH SCHOOL	--	--	--	0.284	--	--
				(3.15)		
COLLEGE	--	--	--	0.570	--	--
				(5.48)		
MARRIED	--	--	--	--	0.319	0.361
					(4.27)	(4.83)
OTHER INCOME	--	--	--	--	--	-0.017
						(3.49)
CONSTANT	0.074	-0.232	-1.446	-0.797	-1.453	-1.234
	(0.08)	(0.26)	(1.68)	(0.94)	(1.72)	(1.50)

Table 7

Alternative Model Specifications and Estimators (Subsample Ages 30-59)

Columns 1-3: Alternative Alcoholism Measures
Column 4: Father's Schooling Exactly 12 Years for All Obs.
Column 5: Inclusion of Possible Psychological Comorbidities
Columns 6-7: Symmetrically-Censored Estimator
(heteroskedasticity-consistent t-stats in parentheses)

	Dependent Variable: LOG-INCOME						
	OLS					SCLS	
Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)
ALC-EVER	-0.199	--	--	-0.330	-0.191	-0.174	-0.120
	(1.63)			(2.43)	(1.98)	(2.23)	(1.73)
ALC-YEAR	0.021	-0.150	--	--	--	--	--
	(0.12)	(1.31)					
ALC-PRE19	--	--	-0.099	--	--	--	--
			(0.57)				
ALC-1922	--	--	-0.175	--	--	--	--
			(1.29)				
ANTISOCIAL PERSONALITY	--	--	--	--	0.010	--	--
					(0.05)		
MENTALLY HEALTHY	--	--	--	--	0.044	--	--
					(0.30)		
AGE	0.109	0.105	0.107	0.114	0.109	0.113	0.102
	(2.75)	(2.64)	(2.72)	(1.65)	(2.77)	(3.29)	(3.41)
AGE SQUARED	-0.001	-0.001	-0.001	-0.001	-0.001	-0.001	-0.001
	(2.51)	(2.39)	(2.47)	(1.44)	(2.53)	(3.02)	(3.17)
WHITE	0.245	0.234	0.238	0.154	0.244	0.203	0.168
	(2.23)	(2.12)	(2.17)	(0.79)	(2.21)	(2.16)	(2.00)
PHYSICALLY HEALTHY	0.391	0.419	0.410	1.324	0.377	0.332	0.242
	(2.48)	(2.66)	(2.53)	(3.26)	(2.37)	(2.38)	(1.95)
SCHOOLING	0.072	0.074	0.075	0.089	0.072	0.072	0.068
	(6.62)	(6.60)	(6.78)	(4.16)	(6.70)	(7.24)	(7.48)
MARRIED	0.362	0.373	0.383	0.156	0.360	0.351	0.318
	(4.84)	(4.94)	(5.02)	(1.16)	(4.78)	(4.95)	(4.70)
OTHER INCOME	-0.017	-0.018	-0.018	-0.009	-0.017	-0.016	-0.014
	(3.49)	(3.50)	(3.54)	(1.50)	(3.46)	(3.58)	(3.16)
CONSTANT	-1.246	-1.223	-1.299	-2.333	-1.256	-1.216	-0.759
	(1.49)	(1.46)	(1.57)	(1.53)	(1.51)	(1.66)	(1.17)
N. Obs.	555	555	555	123	555	555	555

Table 8

Alcoholism Interaction Models
(Subsample Ages 30-59; N. Obs.=555)
(OLS estimates; heteroskedasticity-consistent t-stats in parentheses)

Column 1: Schooling Dummies and Interactions
Columns 2-3: Schooling and Age Interactions
Column 4: Physical Health Status Interaction

	Dependent Variable: LOG-INCOME			
Variable	(1)	(2)	(3)	(4)
ALC-EVER	-0.577	-0.824	-1.036	-0.596
	(2.92)	(2.05)	(1.76)	(2.05)
AGE	0.106	0.109	0.107	0.109
	(2.76)	(2.82)	(2.73)	(2.80)
AGE SQUARED	-0.001	-0.001	-0.001	-0.001
	(2.53)	(2.57)	(2.52)	(2.55)
WHITE	0.239	0.243	0.242	0.247
	(2.21)	(2.22)	(2.21)	(2.26)
PHYSICALLY HEALTHY	0.374	0.370	0.375	0.227
	(2.42)	(2.39)	(2.41)	(1.19)
ALC-EVER × PHYSICALLY HEALTHY	--	--	--	0.481
				(1.59)
HIGH SCHOOL	0.166	--	--	--
	(1.77)			
COLLEGE	0.453	--	--	--
	(4.18)			
ALC-EVER × HIGH SCHOOL	0.480	--	--	--
	(2.11)			
ALC-EVER × COLLEGE	0.658	--	--	--
	(2.92)			
SCHOOLING	--	0.062	0.061	0.072
		(5.50)	(5.35)	(6.66)
ALC-EVER × SCHOOLING	--	0.050	0.053	--
		(1.71)	(1.82)	
ALC-EVER × AGE	--	--	0.004	--
			(0.45)	
MARRIED	0.357	0.366	0.367	0.375
	(4.79)	(4.91)	(4.90)	(5.10)
OTHER INCOME	-0.018	-0.018	-0.018	-0.018
	(3.60)	(3.62)	(3.61)	(3.61)
CONSTANT	-0.398	-1.072	-1.009	-1.102
	(0.49)	(1.33)	(1.23)	(1.33)

Table 9

Estimates for Fulltime Workers Receiving No Transfers:
Probit FULL-NOTRANS and Heckman-Corrected
LOG-INCOME for FULL-NOTRANS=1
(Subsample Ages 30-59)
(asymptotic t-stats in parentheses)

	Reduced Form		Full Model	
	FULL-NOTRANS=1 (Probit)	LOG-INCOME (OLS)	FULL-NOTRANS=1 (Probit)	LOG-INCOME (OLS)
Variable	(1)	(2)	(3)	(4)
ALC-EVER	-0.402	-0.151	-0.286	-0.181
	(2.67)	(0.74)	(1.82)	(1.04)
AGE	0.256	0.104	0.259	0.230
	(3.05)	(0.85)	(2.99)	(2.06)
AGE SQUARED	-0.003	-0.001	-0.003	-0.003
	(3.10)	(0.82)	(3.06)	(2.01)
WHITE	0.497	0.350	0.255	0.283
	(2.92)	(1.34)	(1.37)	(1.45)
PHYSICALLY HEALTHY	0.836	0.096	0.764	0.456
	(4.36)	(0.19)	(3.84)	(1.08)
SCHOOLING	--	--	0.056	0.101
			(2.34)	(3.77)
MARRIED	--	--	0.550	0.542
			(3.74)	(2.31)
OTHER INCOME	--	--	-0.028	-0.033
			(3.12)	(2.55)
λ	--	0.218	--	1.475
		(0.18)		(1.50)
CONSTANT	-5.462	0.410	-6.245	-4.581
	(3.03)	(0.11)	(3.26)	(1.40)
N. Obs.	555	449	555	449