Wages, Alcohol Use, and Smoking: Simultaneous Estimates

M. Christopher Auld[1]

University of Calgary
and
Institute of Pharmacoeconomics.

August 1998

Abstract

This paper estimates a simultaneous model of moderate and problem drinking, smoking, and wages using a random sample of employed Canadian men. The results indicate that sample selection into alcohol and tobacco use is not negligible. With all else in the system held constant, moderate and heavy drinking are both associated with considerably higher wages than abstention from drinking, whereas smoking is associated with lower wages. Allowing for feedback from wages to substance use is important: if wages are excluded from the substance use equations, the premium to heavy drinking disappears, the premium for moderate drinking rises, and the penalty to smoking is diminished.

JEL Classification: I12
Keywords: alcohol, tobacco, simultaneous equations, maximum simulated likelihood

[1]I thank Mingshan Lu, Herb Emery, Ken McKenzie, and Whipple Smith for helpful discussions. The usual disclaimer applies.

Wages, Alcohol Use, and Smoking: Simultaneous Estimates

M. Christopher Auld

August 1998

1 Introduction.

Analysing the relationships between health, health–affecting behaviours, and indicators of well–being such as labour market outcomes has occupied a considerable portion health economics research. One particular health–affecting behaviour, substance use and abuse, has received much attention in the past decade, partially due to the frequent but controversial and unexpected finding that alcohol and illicit drug use are associated with higher wages, all else equal. This paper estimates a simultaneous structural model of moderate and heavy drinking, smoking, and wages using a random sample of employed Canadian men in order to shed more light on the relationship between alcohol use, tobacco use, and income by controlling for potentially endogenous use decisions, including feedback from wages to consumption.

Alcohol is by a wide margin the most commonly used drug in Canada and the U.S., deleteriously affecting health when used heavily (NIAAA 1993), causing traffic (Ruhm 1995) and other accidents, violence (Lenke 1982, 1990), and suicide (CDC 1990). The largest social costs are often attributed to lost productivity (Rice *et al.* 1990). However, the prevailing view that alcohol use reduces wages was challenged by Berger and Leigh (1988), Cook (1991), and French and Zarkin (1995) who find a positive effect of alcohol consumption on earnings. This result is consistent with the medical literature which finds that moderate alcohol consumption may increase health (Turner, Bennett and Hernandez 1981, Shaper 1988), but the opposite result is found by Mullahy and Sindelar (1989, 1991, 1993) and Hamilton and Hamilton (1997). "Thus," Mullahy and Sindelar (1996) comment, "controversy remains even as to the direction of the effects of alcohol on productivity, much less their magnitude." Levine, Gustafon, and Velenchik (1995) find that use of another common drug, tobacco, reduces wages.

The literature on substance use and labour market outcomes emphasizes the endogenous nature of the decision to use alcohol or tobacco. However, previous research nonetheless often treats substance use as exogenous when investigating its impact on labour market outcomes or uses two–stage methods which perform poorly when the instruments are weak. Methods treating use as exogenous can neither control for sample selectivity nor distinguish a causal effect of substance use on wages from a non–zero income elasticity of demand for substances. In particular, if alcohol is a normal good and tobacco an inferior good, then it will appear in models treating substance use as exogenous that alcohol has a positive effect on wages and smoking a negative effect. A related and apparently overlooked point is the converse: studies of demand for alcohol and tobacco which treat income as exogenous will not consistently estimate the income elasticity of demand for these goods if there does exist a causal effect of substance use on wages: a positive (negative) effect will bias income elasticities up (down).

Studies which treat substance abuse as exogenous come to significantly different conclusions from those which do not. Zarkin *et al.* (1998) do not report estimates from an IV approach to their problem since the estimated effects of alcohol use are on the order of 50 to 200%, and the overidentification restrictions reject the hypothesis that the excluded instruments are orthogonal to the residuals. However, a review of previous work indicates that such results are not anomolous, rather, they are consistent with other methods which incorporate endogenous use decisions. Berger and Leigh's (1988) wage decompositions based on switching regressions modeling the decision to use alcohol indicate positive returns to alcohol use of 45.0% for males, whereas the effect is 8.1% when drinking is treated as exogenous. Hamilton and Hamilton's (1997) approach based on switching regressions with a first–stage multinomial logit model over drinking status yields positive returns to moderate over heavy drinking of 75.8%, but the effect is -6.6% (and insignificant) when drinking in treated as exogenous. In a related context, Kaestner's (1991) results show positive returns to illicit drug use of up to 146% [1] when drug use is instrumented out, and generally (comparing Table 4 to Table 5) that the effect of drug use on wages is much higher when treated as being simultaneously determined with wages. The implication of these results is that sample selection into substance use is important and estimates treating use as if it were striking a random subset of population may be misleading.

The effect of drinking is usually estimated without controlling for smoking status and vice versa. Since smoking and drinking are highly correlated, it is possible that failing to control for alcohol use when estimating the impact of alcohol use on wages confounds the effect of alcohol use with the effect of tobacco use, and vice versa.[2] An exception is Kenkel and Wang (1998), who find that controlling for smoking reduces the estimated impact of drinking on receiving major fringe benefits.

This paper presents results of a model in which moderate and heavy drinking, smoking, and wages are treated as endogenous. Since evaluating

[1] The coefficient on cocaine use in the last 30 days for males, Table 5, column 2, is 0.9024, which implies a percentage difference of $\exp(0.9024) - 1 = 1.46$.

[2] This paper is subject to a similar criticism in that I do not control for illicit drug use, which is also highly correlated with drinking and smoking. However, the population proportion who use alcohol (even heavily) or tobacco is much higher than the proportion using illicit drugs, so I not expect this is a serious caveat.

the likelihood is a numerically intractable problem, a method of maximum simulated likelihood (MSL) is used in order to obtain estimates which are equivalent to the full information maximum likelihood (FIML) estimates of the system as accuracy of the simulation method is allowed to increase without bound. I can then explicitly control for feedback from wages to the propensity to use substances, sample selectivity into drinking and smoking, and the joint decision to drink or smoke. I present simultaneous estimates with and without feedback from wages to substance use decisions in order to determine whether failing to allow for such feedback significantly affects results, and also contrast both simultaneous specifications with the single–equation results.

The findings suggest that the cross–equation correlations have dramatic effects on the estimates. After controlling for endogeneity, both drinking and heavy drinking are associated with much higher wages than abstention from alcohol use, whereas smoking is associated with much lower wages than not smoking. Disallowing feedback from wages to use increases the return to moderate drinking, almost eliminates the premium to heavy drinking, and substantially reduces the penalty to smoking. These significant changes in results suggest that controlling for feedback from wages to substance use should not be ignored. I conclude that treating substance use and abuse as exogenous is decidedly rejected by the data. Even if one finds the results presented here and in previous work treating use as endogenous as implausibly large when interpreted as causal effects, there is no reason to prefer models treating use as exogenous.

The paper is organized as follows. Section 2 presents a description of the dataset and summary statistics. Section 3 outlines the econometric specification (an appendix to this section describes the method used to simulate the likelihood function). Results are presented in section 4. Concluding remarks close the paper in section 5.

2 Data

I use the 1991 Canadian General Social Survey (GSS), the focus of which is health. The survey is a random sample of Canadians aged 15 and over conducted from January through December, 1991. I extract information on full–time employed or self–employed males aged 20 through 59. After removing observations with missing data, the final sample consists of 2,430 males. Summary statistics are presented in Table 1.

I define an individual as being full–time employed if they reported working in excess of 1000 hours in the previous year, which excludes some 18% of the male sample in the relevant age range from the analysis. Since I exclude men who are not employed, the coefficients in the wage equation should be interpreted as reflecting variation in wages conditional on employment rather than the determinants of the more fundamental relationship governing both rejected and accepted wage offers.

Since the GSS does not directly ask about hourly wages, I estimate the wage rate by dividing personal income by annual hours (weeks worked in the last year multiplied by usual number of hours worked per week). The wage estimates have an inherent upward bias since I observe all personal income rather than just salary income or income directly related to self–employment. In order to minimize this bias, individuals reporting any pension income (72 observations) are excluded from the sample. Since most income for prime–age males is generated in the labour market and pension income is excluded, wage estimates based on personal income rather than labour market income is not expected to pose a significant problem.

Unfortunately, the GSS reports income only in categories: $5,000 brackets from $0 to $20,000, $10,000 brackets from $20,000 to $40,000, $20,000 brackets from $40,000 to $80,000, and an over $80,000 group. I do not then observe actual wages even under the strong assumption that annual hours are reported without error, I observe an upper and lower bound on wages, produced, respectively, by diving the upper and lower income bound for the respondent's income category by hours worked. The lower bound for average hourly earnings (computed by assuming that each observation is actually at the lowest income in the bracket) in the sample is \$12.06, whereas the upper bound for average earnings is \$24.06.

Price indexes for alcohol and for tobacco and related products from Statistics Canada are added to these data according to province of residence. The indexes are based on retail prices in major cities in each province as of September 1991. Since I do not include regional dummy variables, the coefficients on price will reflect both response to price and geographic variation, and should therefore not be interpreted as demand curve slopes.

I follow Hamilton and Hamilton (1997) in constructing indicators for moderate drinking, heavy drinking, and abstention from alcohol use. Following these authors definitions, an abstainer (18% of the sample) has not had a drink at least once per month during the last year. A heavy drinker (9.6% of the sample) drinks at least once a week and had eight or more drinks in one sitting on at least one occasion in the last week. Everyone else (that is, those who drink at least once per month, but do not meet the criteria for heavy drinking) are moderate drinkers. The heavy drinking measure requires both frequent alcohol use and an episode of binge drinking, which is a strong predictor of problem drinking (Knupfer 1984). The indicator also classifies a similar portion of the sample as having a potential alcohol problem as Mullahy and Sindelar's (1989, 1991, 1993) diagnosis of alcoholism measure in the Epidemiological Catchment Area Survey, which applies to about 10% of the male sample in those studies, and with estimated prevalences of (U.S.) alcohol dependence in Stinson *et al.* Unfortunately, the GSS does not contain retrospective information on drinking which prevents use of any measure of alcohol's "addiction capital" (in the Becker and Murphy (1988) sense).

A smoker is defined as someone who reports smoking cigarettes on a daily basis. 30% of the sample meet this definition. Occasional smokers are excluded from smoker status, but the number of people who smoke infrequently is small due to the highly addictive nature of nicotine.[3] I also choose not to consider intensity of smoking, conditional on smoking daily, as an additional endogenous outcome. Using number of cigarettes smoked in the last week in a Tobit regression version of the single–equation probit for smokers described in Section 4 did not appreciably change the qualitative results, and Levine *et al.* (1995) do not find that number of cigarettes smoked has an effect on wages. A dummy indicating the respondent tried smoking at or below the age of 14 is used to proxy both "addiction capital" and of unobservable tastes for smoking. Current smokers are 2.2 times more likely to report having smoked before age 14 as current non–smokers.

Religious status is included in the substance use equations both as a proxy for moral sentiment towards substance use and because most religions

[3]In the GSS, for instance, there are over six times as many daily smokers as occasional smokers.

actively discourage drinking and smoking, which may increase the perceived costs of use. A respondent is defined as religious if they report attending church (aside from a special occasions) at least once per month. I construct a further indicator for Roman Catholicism, including it in the drinking equations but not the smoking equation, as preliminary analysis showed no significant difference for Catholics in smoking behavior and the exclusion improves identification. 24% of the sample are classified as religious, of whom more than half (14% of the sample) are Roman Catholic.

The GSS reports age in five year intervals rather than as a continuous measure. I use the midpoints of these categories as the respondent's age rather than adding dummies for each (but one) category, trading a small amount of measurement error for parsimony in specification. Education is measured by indicators for never completing high school (23.6% of the sample), for high school graduation but no subsequent education (16.7%), for education beyond elementary or secondary school but no bachelor's degree or higher (40.3%), and for the completion of a bachelor's or greater degree (19.4%). Occupation is categorized as administrative (4.6%), service (16.4%), primary sector (48.9%) , or managerial (14.4%) based on 1980 Standard Occupation Code classifications. Health status is proxied by indicators based on self–assessed state of health. A respondent reporting "excellent" health status (27.0% of the sample) forms one category, "very good" or "good"

(64.9%) a second, and those reporting "fair" or "poor" the final category (8.1%).[4]

3 Econometric Specification

The econometric model is motivated by the theoretical work of Grossman (1972) and Becker and Murphy (1988). Consider an individual who faces a dynamic optimization problem to allocate time and money between production of health, production of a composite commodity produced from leisure time and money, and production of "relaxation" (Siegal 1989) using alcohol, tobacco, and time. Maximization subject to resource constraints yields optimal use of the addictive goods, labor market productivity, and human capital accumulation in each period. These outcomes are affected by ability, labour market opportunities and, conditionally, on past use of the addictive goods and stock of health capital. This framework suggests that, in any period, use of addictive goods and labor market outcomes are simultaneously determined. In a dynamic framework, health capital and other human capital characteristics would also be endogenously determined. See Kenkel and Ribar (1994) for evidence that alcoholism affects marriage probability, Kenkel and Wang (1998) for evidence that alcoholics sort into different occupations

[4]Bound (1991) finds that such self–reported measures of health are flawed and potentially endogenous, but fare no worse than "objective" measures of health status as predictors of labor market outcomes.

(and possibly self–employment) than others, and Cooke and Moore (1993) and Mullahy and Sindelar (1989, 1991, 1993) for evidence that alcoholism affects human capital acquisition. Since I use cross–section data, I condition on contemporaneous health and human capital variables. The results should then be interpreted with caution as the education and occupation variables are potentially endogenous in a life–cycle context.

Recall that income and consequently wages are only observed in categories. The rule mapping actual income into categories is the same across individuals, but hours worked is reported continuously and varies across individuals, therefore the upper and lower bounds on wages are individual specific. If individual i's income is reported as being between L and U, then that individual's log wage is between $w_i^L \equiv \log(L/h_i)$ and $w_i^U \equiv \log(U/h_i)$, where h_i denotes hours worked. The other observable endogenous outcomes are: A_i, an indicator for being a moderate or heavy drinker, HD_i is an indicator for heavy drinking, and C_i, an indicator for being a smoker. Let a superscript $*$ denote a latent outcome and X_{ji} a vector of exogenous regressors for individual i in equation j.

The latent wage equation is specified:

$$w_i^* = X_{1i}\beta_1 + \phi_0 A_i + \phi_1 HD_i + \phi_2 C_i + \epsilon_{1i}, \tag{1}$$

but since only the bounds on the wage are observed,

$$w_i^L < w_i^* < w_i^H, \tag{2}$$

where $w_i^H = \infty$ if the person reported earning more than \$80,000. Berger and Leigh (1988) and Hamilton and Hamilton (1997) suggest that much of the difference between the earnings of users and non–users can be attributed to differential returns to human capital. Since I do not allow returns to human capital to vary with substance use decisions, the dummy variables representing use will reflect both direct effects of use and effects use has on returns to schooling and occupational choice. However, as noted above, prior differences in schooling and occupational choice between substance users and non–users are conditioned away in this specification.

Drinking status is determined in two stages.[5] In the first stage, a binary decision is made over whether to abstain from alcohol or not. The drinking abstention equation is

$$A_i^* = X_{2i}\beta_2 + \phi_3 w_i + \epsilon_{2i} \tag{3}$$

$$A_i = \begin{cases} 1 & \text{if } A_i^* > 0 \\ 0 & \text{otherwise,} \end{cases} \tag{4}$$

where X_{3i} contains exogenous variables and w_i is the midpoint between w_i^L

[5]I choose to model drinking as two separate decisions rather than an ordered probit because preliminary analysis indicated that constraining the coefficients determining moderate drinking over abstention to equal those determining heavy over moderate drinking to be invalid.

and w_i^U.[6] The wage rate is then measured with error in this and the following equations, however, the approach still allows for wages to feedback to substance use. Integrating over all wages consistent with observed wage category in the substance use equations would add considerable computational burden to solve a relatively small specification problem.

If the person chooses not to abstain, a second choice over whether to be a heavy or moderate drinker is made:

$$HD_i^* = X_{2i}\beta_2 + \phi_4 w_i + \epsilon_{3i} \qquad (5)$$

$$HD_i = \begin{cases} 1 & \text{if } HD_i^* > 0 \\ 0 & \text{otherwise,} \end{cases} \qquad (6)$$

If the person chooses to abstain, this decision is irrelevant and HD_i^* can take on any value without affecting the observed outcome. I have chosen to model drinking decisions in this manner since it allows the two–dimensional space of possible outcomes for the two errors related to a three–outcome discrete choice problem to be divided into *rectangular* areas mapping draws to outcomes. A multinomial probit (MNP) specification, by contrast, either requires different random variables depending on observed choice, which complicates the construction of the covariance matrix when the MNP model is embedded in a system, or splitting the two–dimensional error space into non–rectangular regions, which prohibits the use of the simple simulator I employ

[6]Except when the individual reports earning more than \$80,000, in which case the upper bound is arbitrarily set at \$100,000.

here. To improve identification, I also impose the condition that ϵ_2 and ϵ_3 are uncorrelated, that is, that unobservables which tend to cause drinking are unrelated to unobservables which cause heavy drinking conditional on either moderate or heavy drinking.

Finally, the smoking equation is given by

$$C_i^* = X_{4i}\beta_4 + \phi_5 w_i + \epsilon_{4i} \quad (7)$$

$$C_i = \begin{cases} 1 & \text{if } C_i^* > 0 \\ 0 & \text{otherwise.} \end{cases} \quad (8)$$

Note that drinking status is not included in the smoking equation, nor is smoking status included in the drinking equations. It has been suggested that drugs may have strong cross–price elasticities (e.g. Chaloupka and Laixuthai (1997)). Here, net complementarity (substitutability) between drinking and smoking will then appear as positive (negative) correlation between the error terms in the smoking and drinking equations.

Since there are endogenous regressors in each of the four equations, estimating (2) by maximum likelihood treating ϵ_1 as uncorrelated with any of the explanatory variables or (4) through (8) by appropriate binary variable estimation methods (probit, if the errors are assumed to be normally distributed) will produce inconsistent estimates if the errors are actually correlated.

The FIML estimator maximizes the likelihood of observing the vector of dependent variables $\{w, A, HD, C\}$ conditional on the exogenous variables. The FIML estimator has the advantage of being consistent, asymptotically efficient, imposes all cross-equation restrictions, fully incorporates the limited and endogenously censored nature of the endogenous variables, and has an asymptotically correct covariance matrix, as opposed to two-step methods which require often convoluted corrections to the second–step covariance matrix to account for the presence of generated regressors (see, e.g., Baldwin *et al.* 1993). Under the assumption that the error terms are distributed multivariate normal with zero means, the likelihood function can be expressed:

$$L = \prod_i \int_{b_{1i}}^{u_{1i}} \int_{b_{2i}}^{u_{2i}} \int_{b_{3i}}^{u_{3i}} \int_{b_{4i}}^{u_{4i}} f(s_1, s_2, s_3, s_4; \Sigma) ds_4 ds_3 ds_2 ds_1, \tag{9}$$

where $f(\cdot)$ denotes the multivariate normal distribution with covariance matrix Σ (described below). Letting Z_{ji} denote the complete set of endogenous and exogenous explanatory variables in equation j and γ_j the associated parameter vector, the lower (b) and upper (u) bounds of integration for a given observation are

$$(b_{1i}, u_{1i}) = (w_i^L - Z_{1i}\gamma_1, w_i^U - Z_{1i}\gamma_1), \tag{10}$$

which constrains the latent wage variable to be an element of the set defined by the upper and lower bounds of the relevant wage bracket. The bounds for the dimension representing the decision whether to drink are

$$(b_{2i}, u_{2i}) = \begin{cases} (-\infty, -Z_{2i}\gamma_2) & \text{if } A_i = 0 \\ (-Z_{2i}\gamma_2, \infty) & \text{otherwise,} \end{cases} \tag{11}$$

which are the bounds of integration for the Probit model in a single–equation context, and similarly for the decision whether to drink heavily conditional on choosing not to abstain:

$$(b_{3i}, u_{3i}) = \begin{cases} (-\infty, -Z_{3i}\gamma_2) & \text{if } HD_i = 0 \text{ and } A_i = 1 \\ (-Z_{2i}\gamma_3, \infty) & \text{if } HD_i = 1 \text{ and } A_i = 1 \\ (-\infty, \infty) & \text{if } A_i = 0. \end{cases} \quad (12)$$

$$(b_{4i}, u_{4i}) = \begin{cases} (-\infty, -Z_{4i}\gamma_4) & \text{if } C_i = 0 \\ (-Z_{4i}\gamma_4, \infty) & \text{otherwise.} \end{cases} \quad (13)$$

(14)

The covariance matrix takes the form:

$$\Sigma = \begin{pmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} & \sigma_{14} \\ \sigma_{12} & 1.0 & 0.0 & \sigma_{24} \\ \sigma_{13} & 0.0 & 1.0 & \sigma_{34} \\ \sigma_{14} & \sigma_{24} & \sigma_{34} & 1.0 \end{pmatrix} \quad (15)$$

where σ_{ij} denotes the covariance between ϵ_i and ϵ_j and σ_{11} denotes the variance of ϵ_1. The scale of the equations determining drinking and smoking status is not identified so the variances are normalized to unity. The correlation between the drinking abstention and type of drinker equations is restricted to be zero to improve identification.

Since all of the dependent variables are limited and endogenous regressors appear in each equation, evaluating the likelihood function involves integrating over a four dimensional normal density. Since even the best quadrature methods are slow and inaccurate for integration over more than two dimensions (Hajivassiliou and Ruud 1996), the likelihood function is simulated using the Geweke, Hajivassiliou, and Keane (GHK) simulator. The maximum simulated likelihood (MSL) estimates have many advantages, including being continuous in the parameters, numerically easy to compute, and consistent. A disadvantage is inefficiency, since simulation adds noise. The simulation methodology is described in detail in the Appendix.

4 Results

Tables 2 through 5 present the estimated wage, drinking, moderate versus heavy drinking, and smoking equations, respectively. Three specifications are estimated. One ("single") in which the correlation matrix is restricted to be diagonal, which is numerically equivalent to single equation estimates. One ("wage feedback") with a general correlation matrix (subject to the form specified in equation (15)) and in which substance use potentially depends on wages. And finally one ("no wage feedback") with a general correlation matrix but with coefficients on the wage rate in the substance use equations restricted to zero (that is, $\phi_3 = \phi_4 = \phi_5 = 0$). This last specification is equivalent to an endogenous treatment effects model with multiple treatments, where in this context the "treatments" are decisions to use alcohol or tobacco. The estimated correlation matrixes for the simultaneous models are presented in Tables 6 and 7.

Note first that the wage feedback model fits the data significantly better than the other two specifications. The most restricted model (unreported) has both a diagonal correlation matrix and zero restrictions on the wage coefficients. This specification has a loglikelihood value of -7376.2. Freeing the correlation matrix gives a loglikelihood of -7358.2, whereas freeing the wage coefficients but maintaining a diagonal correlation matrix gives a loglikelihood of -7358.4. Thus the most restrictive model can be rejected in favor of either the single–equation with wage effects or simultaneous equation without wage effects specifications, since the LR statistics are both about 36 with critical values at 1% significance of 11.3 and 15.1, respectively. However, the least restrictive specification, allowing both wage effects on substance use and a general correlation matrix, yields a dramatic improvement of the loglikelihood to -6106.4. This result strongly suggests that models which do not allow substance use to depend on wages are inferior to more general specifications.

The wage equation displays typical results for the socioeconomic variables. Across specifications, wages rise at a decreasing rate with age (experience) and there are significant positive returns to education and marriage. In the single and no wage feedback specifications, self–employment yields lower hourly returns than salaried labour. Similarly, in these two specifications being in less than excellent health is associated with lower wages. When substance use is endogenous and there is feedback from wages to use, neither poor health nor self–employment is associated with a wage penalty, and the returns to marriage and occupations fall. These results suggest that occupation and education are jointly determined with substance use: the model can explain, for instance, the raw positive partial correlation between good health and wages by the correlation between the error in the wage equation and the errors in the substance use equations when the latter errors are constrained to be orthogonal to the wage rate. I interpret this result as suggesting further endogeneity problems rather than that human and health capital returns are biased up by failing to account for selection into substance use.

The differences in the coefficients on drinking and heavy drinking across specifications are not subtle. In the single–equation specification, moderate drinking and heavy drinking are associated with 15.4%[7] and 10% higher wages than abstention, respectively. This result is consistent with previous research which finds positive returns to drinking. However, in the wage feedback model, both moderate and heavy drinking are associated with much higher wages than abstention, about 57% and 47% respectively. This suggests negative sample selection into drinking, which is confirmed by examining Table 5. This result is consistent with Berger and Leigh (1988), who find predicted wages for drinkers 45% higher than non–drinkers when sample selection is taken into account but 8.1% when drinking is treated as

[7]All percentage changes in the discussion are computed as $\exp(c) - 1$, where c is a coefficient on a dummy in the wage equation.

exogenous.

When wages are not allowed to feedback to substance use, heavy drinking and abstention are both associated wage penalties of about 77% relative to moderate drinking. The coefficients on wages in the drinking equations in Tables 3 and 4 suggest that higher wages lead to increased propensity to drink, but lower propensity to drink heavily conditional on not abstaining. When wages are not allowed to feedback to drinking decisions, part of the effect of higher wages leading to more moderate drinking is picked up as increased wages due to moderate drinking. Similarly, if higher wages causing less heavy drinking is not accounted for, a spurious wage penalty to heavy drinking is produced. Here, this effect entirely cancels the positive association between heavy drinking and wages. Further, when wages are not allowed to affect drinking, it appears that sample selection into drinking is less important, but sample selection into heavy drinking becomes more important. Too see this, consider an unobservable change which increase the person's wage. If this truly causes increased propensity to drink, but that causal chain is disallowed by the specification, it will result in an increase in *unobserved* propensity to drink, biasing upwards the correlation between the errors in the drinking and wage equations upwards. Here, the correlation with wage feedback is -0.84, without wage feedback it is -0.38.

In the single–equation estimates, smoking is associated with a wage penalty of 8.0%. This is consistent with Levine *et al.'s* (1995) result that smoking reduces wages by about four to eight percent. It further suggests that one of those authors' proposed mechanisms, a reduction in wages due to higher expected health insurance costs, is not the primary explanation, since a similar result holds under the Canadian socialized health insurance system. When endogenous decisions to smoke are controlled for, but those decisions are not allowed to themselves depend on the wage rate, the penalty increases to 20.4%, suggesting positive sample selection into smoking, confirmed by examining Table 6. In the full model, the penalty to smoking rises to 67.5%, even though increases in wages lead to lower propensity to smoke (Table 4). The sample selection term becomes very large: increases in unobservable factors leading to greater probability of smoking are almost perfectly correlated (0.978) with increased wages. These results suggest that the penalty to smoking may be larger than previously suspected and sample selection into smoking a larger factor than sample selection into drinking.

The drinking equation results show that more education (except in the wage feedback model) increases the propensity to drink. The age profiles suggest that age is not a particularly important factor in the decision as to whether to abstain from alcohol use, nor is health status. Being religious significantly decreases the propensity to drink, although the effect is almost exactly nullified in the case of Catholicism. The alcoholic beverage price index has no effect across specifications. This result likely reflects several factors: a low price elasticity on the decision whether to abstain, low variation in prices across provinces, and the price index picking up other regional effects. The impact of increased wages on the decision to drink increases markedly when wages and drinking are simultaneously determined. The model explains the relatively small point estimate of 0.297 in the single–equation estimates by a large estimate of 1.485 combined with a large negative correlation between the error terms in these equations.

Conditional on not abstaining, the propensity to drink heavily falls with educational status and age and is much lower for religious people, although, again, the religion effect is almost negated in the case of Catholicism. Health has no effect on this decision. The alcohol price index is positive and marginally significant in two specifications, which could be spurious, reflect other geographic variation, or possibly endogenous setting of alcohol prices, that is, provinces with more heavy drinkers and therefore more inelastic alcohol demand (Manning *et al.* 1995) may set higher alcohol tax rates. Heavy drinking decreases as wage rates rise, at a somewhat higher rate (-0.276) when drinking and wages are determined jointly than in the single–equation estimates (-0.162).

The propensity to smoke increases at a decreasing rate with age (in this cross–section), is unaffected by provincial variation in cigarette prices, lower for those in better health, and higher for those who smoked before age 14. When smoking and wages are jointly determined and wages can affect smoking decisions, higher education leads to greater propensity to smoke whereas higher wages lead to much lower propensities to smoke. The coefficient on the wage rate is negative and highly significant in both specifications, but almost an order of magnitude larger in magnitude in the simultaneous estimates, consistent with the large positive correlation between the residuals in the wage and smoking equations. These results suggest that studies of tobacco demand and decisions to start or stop smoking should treat wages or incomes as potentially endogenous explanatory variables.

5 Concluding Remarks

The results of this analysis suggest most importantly that treating substance use as if it were a conditional randomly striking members of the population is inappropriate. A review of the literature shows systematically larger wage impacts for substance use when the statistical model allows for endogeneity, and that result is confirmed here. When moderate drinking, heavy drinking, and smoking are treated as exogenous, they are associated with wage differentials of 15%, 10%, and -8% respectively. When they are endogenously determined but wages are not allowed to feedback to substance use, they are associated with differentials of 77%, 1%, and -20%, respectively. Allowing both endogenous selection and feedback from wages to substance demand implies differentials of 57%, 47%, and -67%, respectively. The last specification fits the data significantly better than either of the first two. These results suggest that allowing feedback from wages to substance use decisions improves the fit of the model much more than only allowing endogenous selection into substance use and that such feedback is an important factor in the endogeneity of substance use to wages. Further, the estimated drinking and smoking equations display sensitivity to whether wages are treated as exogenous or not, in particular, the wage effect on smoking is much more negative and the wage effect on drinking is much larger in magnitude when wages are jointly determined with substance use.

One way of intrepreting the results is that deviation from the social norm of moderate drinking to either heavy drinking (10% of this sample) or drinking abstention (12% of this sample) is associated with lower wages. Similarly, smoking has come to be considered socially inappropriate and deviating from the norm of not smoking is associated with lower wages. Since it is implausible that taking a random member of the population and forcing them to change their smoking or drinking and/or smoking habits would actually double or half their wage rate, it is likely that unobservable heterogeneity is driving the findings.

It is important to emphasize that large wage differentials due possibly to factors unobservable by the econometrician does not justify treating substance use as exogenous. Previous estimates and the results here based on cross–section data vary greatly depending on specification. The best–fitting and theoretically preferred specification implies unbelievably large causal effects. Cross–section data, therefore, appears to be of limited use in unravelling the income/substance use puzzle. More use of panel data in future research may be able to shed more light on the puzzle by controlling for both endogeneity and unobservable heterogeneity to a greater extent than approaches possible in cross–sections.

Several important caveats to this analysis should be mentioned. First, I condition on health capital, educational achievement, marital status, and occupational choice, which will tend to bias the effects of substance use towards zero since previous research has shown all of these outcomes to affected by substance use. Second, I do not allow returns to human capital to vary by substance use status. Third, the set of explanatory variables used is somewhat limited (since the full version of the model already involves some 70 parameters to be optimized over). It is possible that including more controls able to proxy heterogeneity driving both differences in wages and substance use decisions would decrease the estimated causal effect of use on wages. Finally, I condition on full–time employment and examine outcomes for men only, but, as previous research has found, employment is likely endogenous to substance use and there may be important differences across men and women.

Appendix: GHK Simulation of the Likelihood Function.

The GHK simulator exploits the fact that the marginal distributions of a multivariate normal distribution are also normal to break a numerically intractable n dimensional problem down into a sequence of n numerically tractable one dimensional problems. Consider the general problem of computing the probability mass under some rectangular region of such a distribution:

$$P = \int_b^u f(s, \Sigma) ds \tag{16}$$

where b and u represent (possibly improper) bounds of integration and $f(\cdot, \Sigma)$ denotes an n dimensional normal density with covariance matrix Σ and (without loss of generality) zero means. Let L denote the lower Cholesky decomposition of Σ (such that $LL' = \Sigma$) and let e represent an n–vector of independently distributed standard normal random variables. Then a vector x distributed $f(x, \Sigma)$ has the same distribution as Le. The probability of x being in the rectangular area defined by b and u is

$$Pr(b \leq x \leq u), \tag{17}$$

where the inequalities are element by element. This is equal to

$$Pr(b \leq Le \leq u) \tag{18}$$

Since L is lower triangular, this probability can be evaluated recursively. Let $\{l_{ij}\}$ denote the elements of L. The probability of the first element falling within bounds is

$$Pr(b_1 \leq l_{11}e_1 \leq u_1) = Pr(b_1/l_{11} \leq e_1 \leq u_1/l_{11}) \tag{19}$$

Conditional on e_1, the probability of the second element falling within bounds is

$$Pr(b_2 \leq l_{21}e_1 + l_{22}e_2 \leq u_2) = Pr(((b_2 - l_{21}e_1)/l_{22}) \leq e_2 \leq ((u2 - l_{21}e_1)/l_{22}). \tag{20}$$

Since e_1 is not observed, it is simulated by drawing a value consistent with the bounds.[8] The simulated value e_1^r on the r^{th} replication is given by (using an application of the integral transform theorem)

$$e_1^r = F^{-1}[(F(u_1/l_{11}) - F(b_1/l_{11}))z^r + F(b_1/l_{11})], \tag{21}$$

where z^r is the r^{th} draw of a uniformly distributed random variable, F represents the cumulative normal distribution and F^{-1} its inverse. Simulation proceeds by simulating values of e_2 through e_{n-1}, computing the conditional probabilities of each draw at each step (there is no need to simulate a value of the last error term). Let Q_{rm} denote the probability the m^{th} draw falls within bounds, conditional on all previous draws. The GHK simulated probability $\hat{P}$ is the average of the probabilities over replications:

$$\hat{P} = \frac{1}{R}\sum_{r=1}^{R}\prod_{m=1}^{M} Q_{rm}, \tag{22}$$

[8]This draw, however, does not utilize information in the bounds to follow, and therefore the GHK simulator does not simulate the joint density. The simulator is nonetheless unbiased for the probability (but not, therefore, the log probability), see Stern (1997).

where R is the number of replications performed.

Maximization of the sum of the logs of the simulated likelihoods with respect to all the slope parameters and the elements of the correlation matrix[9] produces the MSL estimator.[10] I use 20 replications per evaluation of each observation's contribution, consisting of 10 draws of the uniform random variable and the association antithetic variates, the use of which reduce simulation noise considerable.[11] This estimator is consistent (as long as, in theory, $R \to \infty$ as $N \to \infty$), asymptotically normal, continuous in the parameters, and easy to compute, requiring only computation of univariate normal densities, distributions, and inverse distributions. Further, the effort of computation increases linearly in number of observations, number of replications, and the dimension of the problem. See Lee (1992) and Hajivassiliou and Ruud (1994) for proofs and further discussion of the properties of this estimator, and Stern (1997) for more detailed exposition of the principles involved and a survey of uses.

[9]It is irrelevant whether the correlation, covariance, or Cholesky decomposition of the covariance matrix is maximized over, since there is a one to one mapping between these objects.

[10]Maximization was carried out using a combination of Nelder and Meads's simplex algorithm and a Newton–Raphson based algorithm. All code was written in Fortran 90 and executed using Numerical Algorithm Groups's compiler on a Pentium II–266 running Linux 2.0. One evaluation of the likelihood took approximately one–quarter of a second.

[11]Some authors inflate the standard errors to reflect simulation–induced noise. I do not for two reasons. First, unlike the method of simulated moments (MSM), the MSL covariance matrix cannot be written as the sum of the ML covariance matrix and an $O(r^{-1})$ matrix reflecting simulation error, so it is not clear by what factor the covariance matrix for MSL estimates should be increased. Second, the use of antithetic acceleration can reduce noise to a level where no adjustment is necessary even in the MSM case.

6 References

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

Berger, M. and J. Leigh (1988) "The Effect of Alcohol Use on Wages," *Applied Economics* **20**:1343-51.

Bound, J. (1991) "Self–Reported Objective Measures of Health in Retirement Models," *Journal of Human Resources* **26:**203-138.

Centers for Disease Control (1990) "Alcohol and Related Mortality and Years of Potential Life Lost — United States 1987," *Morbidity and Mortality Weekly Report*, **39**, no. 11 (March 23).

Chaloupka, F. (1991) "Rational Addictive Behavior and Cigarette Smoking," *Journal of Political Economy* **99:**722-742.

Chaloupka, F. and A. Laixuthai (1997) "Do Youths Substitute Alcohol and Marijuana? Some Econometric Evidence," *Eastern Economic Journal* **23:**253-276.

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

Cook, P. and M. Moore (1993) "Drinking and Schooling," *Journal of Health Economics* **12:**411-429.

French, M. and G. Zarkin (1995) "Is moderate Alcohol Use Related to Wages? Evidence from Four Worksites." *Journal of Health Economics* **14**: 319-344.

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

Hajivassiliou, V. and P. Ruud (1994) "Classical Estimation Methods for LDV Models Using Simulation," in *Handbook of Econometrics* Eds. R. Engle and D. McFadden, North Holland: Amstredam and New York.

Hajivassiliou, V. and R. Ruud (1996) "Simulation of Multivariate Normal Rectangle Probabilities and their Derivatives: Theoretical and Computational Results," *Journal of Econometrics* 72(2):85-134.

Hamilton, V. and B. Hamilton (1997) "Alcohol and Earnings: Does Drinking Yield a Wage Premium?" *Canadian Journal of Economics* **30**: 135-151.

Kaestner, R. (1991) "The Effect of Illicit Drug Use on the Wages of Young Adults," *Journal of Labor Economics* **9:**381-411.

Kenkel, D. and D. Ribar (1994) "Alcohol Consumption and Young Adults' Socioeconomic Status," *Brookings Papers on Economic Activity: Microeconomics* Washington, D.C.: The Brookings Institution, 119-175.

Kenkel, D. and P. Wang (1998) "Are Alcoholics in Bad Jobs?" National Bureau of Economic Research #6401.

Lee, L.F. (1992) "On Efficiency of Methods of Simulated Moments and Maximum Simulated Likelihood Estimation of Discrete Response Models," *Econometric Theory* **8:**518-552.

Lenke, L. (1982) "Alcohol and Crimes of Violence: a Causal Analysis," *Contemporary Drug Problems*, 355-3??.

Lenke L. (1990) *Alcohol and Criminal Violence: Time Series Analysis in a Comparative Perspective.* Avhandling: Almqvist & Wiksell.

Levine, P., T. Gustafson, and A. Velenchik (1995) "More Bad News for Smokers? The Effects of Cigarette Smoking on Labor Market Outcomes." National Bureau of Economic Research Working Paper #5270.

Manning, W., L. Blumberg, and L. Moulton (1995) "The Demand for Alcohol: The Differential Response to Price," *Journal of Health Economics* **14:**123-149.

Mullahy, J. and J. Sindelar (1989) "Life–Cycle Effects of Alcoholism on Education, Earnings, and Occupation," *Inquiry* **26**:272-282.

Mullahy, J. and J. Sindelar (1991) "Gender Differences in Labor Market Effects of Alcoholism," *American Economic Review* **81**:161-165.

Mullahy, J. and J. Sindelar (1993) "Alcoholism, Work, and Income." *Journal of Labor Economics* **11**:494-520.

Mullahy, J. and J. Sindelar (1994) "Alcoholism and Income: The Role of Indirect Effects," *The Milbank Quarterly* **72:**259-375.

Mullahy, J. and J. Sindelar (1996) "Employment, Unemployment, and Problem Drinking," *Journal of Health Economics* **15**:409-434.

NIAAA (1993) *Eighth Special Report to the U.S. Congress on Alcohol and Health*, Washington, D.C. N.I.H. Publ. 943699.

Rice, D. *et al.* (1990) *The Economic Costs of Alcohol and Abuse and Mental Illness: 1985* Washington, D.C. Supt. of Documents, U.S. Goverment Printing Office, DHHS Pub. No. (ADM)90-1694.

Shaper, A. (1988) "Alcohol and Mortality in British Men: Explaining the U-Shaped Curve." *Lancet* 1267-73.

Siegel, R. (1989) *Intoxication: Life in Pursuit of Artificial Paradise.* New York: Dutton.

Sindelar, J. (1993) "Measurement Issues in Alcohol Survey Data," in G. Bloss and M Hilton, eds., *Economics and the Prevention of Alcohol Related Problems.* Rockville, MD: National Institute on Alcohol Abuse and Alcoholism, 201-228.

Stern, S. (1997) "Simulation–Based Estimation" *Journal of Economic Literature* **35**:2006-2039.

Stinson, F., D. Scar, and R. Steffens (1992) "Prevalence of DSM–III–R Alcohol Abuse and/or Dependence Among Selected Occupations, United States 1988," *Alcohol Health and Research World* **16**:165-172.

Zarkin, G., M. French, T. Mroz, and J. Bray (1998) "Alcohol Use and Wages: New Results from the National Household Survey on Drug Abuse." *Journal of Health Economics* **17**:53-58.

Table 1.
Summary Statistics

Variable	Mean	Std Dev	Wage	Drinking Abstention	Heavy Drinking	Smoking
log wage (lower)	2.492	0.842				
log wage (upper)	3.191	1.757				
drinker	0.780	0.414				
heavy drinker	0.096	0.295				
smoker	0.300	0.458				
age	37.444	10.421	x	x	x	x
age*age/100	15.106	8.463	x	x	x	x
high school grad	0.167	0.373	x	x	x	x
some college or other	0.403	0.491	x	x	x	x
college grad	0.194	0.396	x	x	x	x
Omitted: high school dropout						
managerial	0.144	0.351	x			
professional	0.157	0.364	x			
administrative	0.046	0.211	x			
primary	0.489	0.500	x			
Omitted: service						
good/very good health	0.649	0.477	x	x	x	x
poor/fair health	0.081	0.273	x	x	x	x
Omitted: excellent health						
married	0.668	0.668	x			
self–Employed	0.188	0.390	x			
religious	0.246	0.431		x	x	x
catholic	0.140	0.347		x	x	
alcohol price	99.682	4.752		x	x	
cigarette price	100.851	6.262				x
young smoker	0.227	0.419				x

Table 2.
Wage Equation

	Single ML		Simultaneous MSL			
			Wage Feedback		No Wage Feedback	
	Coef	t-ratio	Coef	t-ratio	Coef	t-ratio
age	0.064	9.472	0.052	7.148	0.059	8.026
age*age/100	-0.065	-7.796	-0.050	-5.706	-0.059	-6.732
high school grad	0.158	4.891	0.148	4.031	0.097	2.473
some college or other	0.150	5.617	0.136	4.796	0.086	2.698
college grad	0.270	7.306	0.191	5.147	0.156	3.485
Omitted: high school dropout						
managerial	0.171	4.673	0.084	3.131	0.173	4.813
professional	0.159	4.153	0.063	2.201	0.164	4.271
administrative	0.067	1.292	0.054	1.374	0.073	1.252
primary	0.086	3.073	0.050	2.498	0.091	3.343
Omitted: service						
poor/fair health	-0.113	-2.860	0.009	0.226	-0.079	-1.845
good health	-0.077	-3.402	0.002	0.091	-0.080	-3.089
Omitted: excellent health						
married	0.079	3.581	0.032	1.998	0.077	3.494
self–Employed	-0.114	-2.860	-0.016	-0.954	-0.134	-6.038
drinker	0.143	5.866	0.453	15.060	0.574	6.184
heavy Drinker	-0.048	-1.377	-0.068	-1.266	-0.562	-8.034
Omitted: drinking abstainer						
smoker	-0.077	-3.499	-0.516	-26.501	-0.186	-2.180
constant	0.993	7.464	1.124	7.6443	0.927	5.743
log-likelihood	-7358.4		-6106.4		-7358.2	

Notes: N=2,430. Single–equation model estimated using maximum likelihood; simultaneous equation models estimated by maximum simulated likelihood using the GHK simulator. Asymptotic t–ratios computed using standard errors estimated from the outer product of the gradient matrix. The “wage feedback” specification includes the wage rate as a regressor in each of the other three equations in the system.

Table 3.
Drinking Equation

	Single ML		Simultaneous MSL			
			Wage Feedback		No Wage Feedback	
	Coef	t-ratio	Coef	t-ratio	Coef	t-ratio
age	-0.017	-0.832	-0.080	-4.074	-0.006	-0.323
age*age/100	0.008	0.032	0.070	2.907	-0.001	-0.031
high school grad	0.423	4.375	0.034	-0.313	0.466	4.845
some college or other	0.244	3.268	-0.173	-2.128	0.282	3.743
college grad	0.405	4.347	-0.191	-1.992	0.490	5.282
Omitted: high school dropout						
Alcohol price	0.004	0.680	-0.001	-0.278	-0.003	-0.550
Religious	-0.738	-8.338	-0.587	-7.073	-0.749	-8.630
Catholic	0.598	5.307	0.378	3.781	0.591	5.294
Poor/fair health	-0.101	-0.905	-0.083	-0.692	-0.136	-1.211
Good health	0.117	1.751	0.031	0.466	0.094	1.373
Omitted: excellent health						
Wage rate	0.297	5.210	1.485	24.581		
Constant	-0.133	-0.181	-1.105	-1.723	1.072	6.631

Notes: N=2,430. Dependent variable is unity when individual is either a moderate or heavy drinker. Single–equation model estimated using maximum likelihood, simultaneous equation models estimated by maximum simulated likelihood using the GHK simulator. Asymptotic t–ratios computed using standard errors estimated from the outer product of the gradient matrix. Wage rate is log–midpoint of upper and lower bounds for the respondent's wage.

Table 4.
Moderate/Heavy Drinking Equation

	Single ML		Simultaneous MSL			
			Wage Feedback		No Wage Feedback	
	Coef	t-ratio	Coef	t-ratio	Coef	t-ratio
age	-0.729	-0.764	-0.078	-2.558	-0.036	-1.404
age*age/100	-0.033	-1.146	0.068	1.696	-0.010	-0.272
high school grad	-0.002	-0.034	-0.002	-0.023	-0.021	0.191
some college or other	-0.210	-2.809	-0.176	-1.703	-0.189	-1.994
college grad	-0.573	-4.639	-0.457	-3.091	-0.529	-4.002
Omitted: high school dropout						
alcohol price	0.013	1.625	-0.001	-0.071	0.014	1.926
religious	-0.897	-3.445	-0.989	-3.405	-0.885	-3.197
catholic	0.656	2.327	0.814	2.585	0.738	2.436
poor/fair health	0.143	0.916	0.119	0.773	0.126	0.827
good health	-0.038	-0.418	-0.035	-0.373	-0.023	-0.261
Omitted: excellent health						
wage rate	-0.162	-2.084	-0.276	-1.958		
constant	-0.729	-0.764	1.644	1.718	1.072	6.631

Notes: N=1,895. Dependent variable is unity when respondant is a heavy drinker, zero when respondant is a moderate drinker. Single–equation model estimated using maximum likelihood, simultaneous equation models estimated by maximum simulated likelihood using the GHK simulator. Asymptotic t–ratios computed using standard errors estimated from the outer product of the gradient matrix. Wage rate is log–midpoint of upper and lower bounds for the respondent's wage.

Table 5.
Smoker Equation

	Single ML		Simultaneous MSL			
			Wage Feedback		No Wage Feedback	
	Coef	t-ratio	Coef	t-ratio	Coef	t-ratio
age	0.025	1.303	0.095	5.766	0.017	0.905
age*age/100	-0.039	-1.641	-0.100	-4.921	-0.032	-1.328
high school grad	-0.099	-1.131	0.281	3.340	-0.126	-1.406
some college or other	-0.288	-3.971	0.187	2.941	-0.310	-4.241
college grad	-0.524	-5.621	0.186	2.395	-0.578	-6.249
Omitted: high school dropout						
tobacco price	-0.001	-0.220	0.001	0.314	-0.001	-0.282
religious	-0.608	-8.272	-0.189	-4.314	-0.583	-7.867
young smoker	0.558	8.618	0.116	3.472	0.565	8.683
poor/fair health	0.361	3.251	0.086	0.889	0.378	3.377
good health	0.239	3.598	0.087	1.574	0.257	3.764
Omitted: excellent health						
wage rate	-0.174	-3.140	-1.432	-27.993		
constant	-0.294	-0.497	1.192	2.853	-0.551	-0.924

Notes: N=2,430. Dependent variable is unity when respondent is a daily smoker. Single–equation model estimated using maximum likelihood, simultaneous equation models estimated by maximum simulated likelihood using the GHK simulator. Asymptotic t–ratios computed using standard errors estimated from the outer product of the gradient matrix. Wage rate is log–midpoint of upper and lower bounds for the respondent's wage.

Table 6.
Correlation Matrix, Wage Feedback Model
(Standard Errors)

	Wage	Drinking	Heavy Drinking	Smoking
Wage	0.524			
	(0.008)			
Drinking	-0.841	1.00		
	(0.023)	—		
Heavy Drinking	0.206	0.00	1.00	
	(0.047)	—	—	
Smoking	0.978	-0.712	0.262	1.00
	(0.004)	(0.035)	(0.054)	—

Table 7.
Correlation Matrix, No Wage Feedback Model
(Standard Errors)

	Wage	Drinking	Heavy Drinking	Smoking
Wage	0.519			
	(0.012)			
Drinking	-0.382	1.00		
	(0.087)	—		
Heavy Drinking	0.583	0.00	1.00	
	(0.053)	—	—	
Smoking	0.152	0.084	0.189	1.00
	(0.104)	(0.038)	(0.041)	—

NOTE: standard errors on diagonals of correlation matrixes.