Alcohol use and the wage returns to education and work experience

By: Jeremy W. Bray, Jesse M. Hinde, and Arnie P. Aldridge

This is the peer reviewed version of the following article:

Bray, J.W., Hinde, J.M., Aldridge, A.P. Alcohol use and the wage returns to education and work experience. *Health Economics*. 2018; 27: e87– e100. https://doi.org/10.1002/hec.3565

Abstract:

Despite a widely held belief that alcohol use should negatively impact wages, much of the literature on the topic suggests a positive relationship between nonproblematic alcohol use and wages. Studies on the effect of alcohol use on educational attainment have also failed to find a consistent, negative effect of alcohol use on years of education. Thus, the connections between alcohol use, human capital, and wages remain a topic of debate in the literature. In this study, we use the 1997 cohort of the National Longitudinal Survey of Youth to estimate a theoretical model of wage determination that links alcohol use to wages via human capital. We find that nonbinge drinking is associated with lower wage returns to education whereas binge drinking is associated with increased wage returns to both education and work experience. We interpret these counterintuitive results as evidence that alcohol use affects wages through both the allocative and productive efficiency of human capital formation and that these effects operate in offsetting directions. We suggest that alcohol control policies should be more nuanced to target alcohol consumption in the contexts within which it causes harm.

Keywords: alcohol | human capital | wages

Article:

1 INTRODUCTION

Economists have long assumed that alcohol use negatively impacts wages through adverse effects on human capital (Bray, 2005; Cook & Moore, 2000; Renna, 2007). Studies examining alcohol use and wages have failed to find a consistent, negative relationship between the two, however. Rather, much of the literature suggests a positive relationship between nonproblematic alcohol use and wages, the so-called "drinker's bonus" (Auld, 2005; Böckerman, Hyytinen, &

Maczulskij, 2017; Cook & Moore, 2000). Studies on the effect of alcohol use on educational attainment have also failed to find a consistent, negative effect of alcohol use on years of education (Renna, 2008), although studies examining alcohol use and educational performance suggest that alcohol use may negatively impact academic performance (Powell, Williams, & Wechsler, 2004; Williams, Powell, & Wechsler, 2003). Thus, the effects of alcohol use on wages and the role that human capital plays in those effects, if any, remain uncertain.

Despite the long-running interest in the topic, only one paper, Bray (2005), links alcohol use to wages and human capital using a unified theoretical and empirical model. Bray derived a theoretical model of wage determination linking alcohol use to wages via a human capital production function that includes school enrollment, labor market experience, and alcohol use as inputs. He estimated this model using the 1979 cohort of the National Longitudinal Survey of Youth (NLSY79) and an estimation method that accounted for both self-selection into the wage equation and the endogeneity of alcohol use, employment, and schooling decisions. He found that moderate alcohol use had positive effects on wages via human capital formation but heavy drinking offset that effect partially; however, his conclusions are limited because the NLSY79 does not adequately measure individuals' histories of alcohol use.

In this study, we replicate Bray (2005) using the more recent NLSY97 and make two substantial contributions. First, beyond being more recent, the NLSY97 permits the creation of more complete drinking histories than does the NLSY79 and includes a measure of binge drinking that is more clinically and policy relevant. Second, by using more complete drinking histories, we isolate the education- and work experience-related components of the drinker's bonus from other possible mechanisms. We find that any drinking is associated with reduced wage returns to education whereas binge drinking is associated with increased wage returns to both education and work experience. We interpret these counterintuitive results as evidence that alcohol use affects wages through negative allocative and positive productive efficiency effects on human capital formation. This interpretation is consistent with prior studies that find that alcohol use negatively affects school performance through the choice of study habits (Powell et al., 2004; Williams et al., 2003) and with prior studies that find that alcohol use improves cognitive performance (Elias, Elias, D'agostino, Silbershatz, & Wolf, 1999; Peele & Brodsky, 2000).

2 BACKGROUND

Numerous studies have examined the relationship between alcohol use and wages or earnings (Barrett, 2002; Berger & Leigh, 1988; Bryant, Sumaranayake, & Wilhite, 1992; French & Zarkin, 1995; Heien, 1996; Kenkel & Ribar, 1994; Mullahy & Sindelar, 1993; Zarkin, French, Mroz, & Bray, 1998). Most studies motivate the connection between alcohol use and wages by referencing either immediate, direct productivity effects of alcohol; immediate, indirect effects of alcohol through health; or indirect, delayed effects through health or human capital. Practically all studies recognize the potential for alcohol use to be endogenous in the wage equation, but they vary in the methods used to address this potential endogeneity.

A related literature in economics examines the effects of alcohol use on education. Early studies in this literature found negative effects of alcohol use on educational attainment (Cook & Moore, 1993; Yamada, Kendix, & Yamada, 1996), but these studies did not account for the endogeneity of alcohol use. Later studies in this literature control for the endogeneity of alcohol use and find little effect of alcohol use on educational attainment (Dee & Evans, 2003; Koch & Ribar, 2001). Studies examining school performance find that alcohol use negatively affects school performance through allocative efficiency mechanisms such as study time or other study habits (Powell et al., 2004; Williams et al., 2003).

Despite the central role of educational attainment in wage determination, only Bray (2005) modeled the dynamic, endogenous accumulation of human capital and alcohol use when estimating the effect of alcohol use on wages. Bray developed a theoretical model that explicitly models the effect of alcohol use on wages through the production and accumulation of human capital. Bray used discrete factor approximation to simultaneously model wages and the alcohol use, school enrollment, and labor supply of young men in the NLSY79. Bray found that alcohol consumption increases the wage return to education by an insignificant 3.3% and the wage return to work experience by a statistically significant 3.8%; heavy drinking reduces these gains by an insignificant 1.7% for education and 0.1% for experience, for a net increase of 1.6% in the wage returns to education and 3.7% for experience.

Bray's (2005) conclusions are limited because the NLSY79 does not collect data on alcohol use in every survey wave. Furthermore, the NLSY79 does not ask alcohol use questions that align with current, policy-relevant definitions of binge or heavy drinking. Nonetheless, Bray's analysis remains the only empirical work in the economics literature to link alcohol use to wages via a formal model of human capital formation and wage determination.

3 THEORETICAL MODEL

We summarize the model originally proposed by Bray (2005); more complete derivations of the model can be found in Bray (2005) and Bray (2000). The model begins with an expanded Mincer-type wage equation (Willis, 1992):

$$\ln(w_t) = \beta_0 + \beta_1 X_t + \beta_2 K_t + \beta H_t + \xi_t \tag{1}$$

The natural logarithm of wage in period t, w_t, is a function of several components: X_t is a vector of observable individual demographics; K_t is the stock of human capital at the beginning of the period; H_t is the health stock at the beginning of the period; and ξ_t represents an error term capturing unobserved heterogeneity.

Assume that all human capital is homogeneous, that there is no human capital depreciation, and that individuals can produce additional units of human capital via schooling or work experience. Alcohol use enters the human capital production function because it may affect the cognitive and psychomotor abilities necessary to learn new skills, thus changing the efficiency of learning for an individual. Equation 2 summarizes the relationship between the stock of human capital (K_t), schooling (s_t), labor market experience (l_t), and alcohol use (a_t).

$$K_t = K_{t-1} + k(s_{t-1}, l_{t-1}, a_{t-1}) \tag{2}$$

Assume that the accumulation of the health stock follows a similar process as described by the following equation:

$$H_t = H_{t-1} + h(m_{t-1}, a_{t-1}) \tag{3}$$

where m_t is a composite good reflecting all inputs to the health production function other than alcohol.

Taking a second-order Taylor series expansion of k around the fixed point $k(\bar{s}, \bar{l}, \bar{a})$, a first-order Taylor series expansion of h around $h(\bar{m}, \bar{a})$, recursing both production functions back to the initial period, and substituting the results into Equation 1 yields

$$\begin{aligned} \ln(w) = \beta_0 + \beta_1 X_1 + \beta_2 K_0 + \beta_3 H_0 + \pi_1(t-1) + \pi_2 \sum_{j=1}^{t-1} s_j + \pi_3 \sum_{j=1}^{t-1} l_j + \pi_4 \sum_{j=1}^{t-1} m_j \\ + \pi_5 \sum_{j=1}^{t-1} a + \pi_6 \sum_{j=1}^{t-1} s_j a_j + \pi_7 \sum_{j=1}^{t-1} l_j a_j + \pi_8 \sum_{j=1}^{t-1} s_j l_j + \pi_9 \sum_{j=1}^{t-1} s_j^2 \\ + \pi_{10} \sum_{j=1}^{t-1} l_j^2 + \pi_{11} \sum_{j=1}^{t-1} a_j^2 + \epsilon_t \end{aligned} \tag{4}$$

where

$$\pi_5 = \beta_2\left[k_a(\bar{s}, \bar{l}, \bar{a}) - \bar{l} k_{la}(\bar{s}, \bar{l}, \bar{a}) - \bar{a} k_{aa}(\bar{s}, \bar{l}, \bar{a}) + \bar{s} k_{sa}(\bar{s}, \bar{l}, \bar{a})\right] + \beta_3 h_a(\bar{m}, \bar{a}),$$

$$\pi_6 = \beta_2 k_{sa}(\bar{s}, \bar{l}, \bar{a}),$$

$$\pi_7 = \beta_2 k_{la}(\bar{s}, \bar{l}, \bar{a}),$$

k_z refers to the derivative of k with respect to function argument z, and all other terms are defined in Appendix S1.

In Equation 4, π_5 does not isolate the effect of alcohol on human capital from its effects on health. Instead, it captures the effects of alcohol use on wages as operating through all mechanisms. However, π_6 and π_7 isolate the effect of alcohol use on the marginal human capital product of schooling and experience, respectively. Because the wage return to human capital (β_2 from Equation 1) is assumed to be positive, the signs of π_6 and π_7 are determined by the sign of the cross-partial derivative of the human capital production function. Thus, if alcohol use enters the human capital production function, the return to education or labor market experience in those years with alcohol use will be different than in those years without.

4 DATA AND METHODS

4.1 Data

Data are from the NLSY97 and associated geocode data. Since 1997, the NLSY97 has been administered annually to the same nationally representative sample of 8,984 individuals born between the years 1980 and 1984. The survey collects data on each respondent's family and living environment, their educational and labor market experiences, and their alcohol use. The RTI International institutional review board reviewed this study and deemed it exempt from further review.

The sample for this analysis comprises 51,813 person-year observations from 5,330 unique individuals spanning the years 1997 to 2009. Because we begin modeling alcohol and school enrollment decisions prior to an individual's first year of earning a wage, only 24,129 observations on the same 5,330 individuals contribute wage information. Once we exclude a person-year for a given individual from our analysis sample, we exclude all subsequent person-years for that individual. We exclude 844 persons and 18,380 person-year observations corresponding to ages 12–16 because youth are required to attend school until age 16. Given the look-back period of the NLSY97 survey, the first year of discretionary school enrollment is reported in survey waves starting at age 17 and older. We exclude 2,440 respondents and 13,085 person-year observations because their panel of person-year observations contain all missing or invalid wages. We exclude 302 persons and 3,019 person-years because they have missing alcohol use information. Finally, we exclude 12 persons and 81 person-years on respondents ages 18 or 19 in 1997 because we cannot model the accumulation of their human capital stocks before ages 18 or 19.

We measure hourly wage using the NLSY created hourly pay variable reflecting the hourly rate of pay in the first job listed in the NLSY job roster. Hourly wage is set to missing when a respondent is enrolled in school to focus the analysis on post-schooling wages. We also set wages to missing if the hourly wage was less than $2/hr or greater than $200/hr. Wages are deflated to a base year of 2008 using the consumer price index. School enrollment is defined as self-reported enrollment in school at any time since the date of the last interview.

Hours worked are set to zero if the responded reported being in school to make the study definition of work experience internally consistent with our measure of wages. We divide hours worked in a year by 2000 to scale work experience into a full-time equivalent (FTE) measure.

Detailed alcohol use questions in the NLSY97 are anchored to the past 30 days, so we create categorical drinking variables to proxy for typical alcohol use over the relevant reporting period rather than use semicontinuous measures of alcohol use. Policy makers and clinicians consider a variety of drinking behaviors to reflect problematic use of alcohol, many of which could be measured in the NLSY97. The National Institute on Alcohol Abuse and Alcoholism defines low-risk drinking limits as no more than four drinks per day or 14 drinks per week for men and more than three drinks per day or seven drinks per week for women and defines binge drinking based on blood alcohol concentration levels exceeding 0.08 g/dl (National Institute on Alcohol Abuse and Alcoholism, 2015). The Centers for Disease Control and Prevention (CDC) defines binge drinking as five or more drinks one occasion for men and four or more for women and defines heavy drinking as consuming 15 drinks or more per week for men and 8 or more for women (CDC, 2016). The Substance Abuse and Mental Health Services Administration uses the same definition of binge drinking as the CDC but defines heavy drinking as bingeing on five or more days in the past month (Center for Behavioral Health Statistics and Quality, 2016). Given these definitions, we define drinkers as individuals who reported any past month use and binge drinkers as individuals who reported consuming five or more drinks on a single occasion at least once in the past month.

The stock variables in Equation 4 are measured as the accumulation of the relevant variables previously described: cumulative years of school enrollment, cumulative FTEs worked, cumulative squared FTEs, cumulative years of any alcohol use, cumulative years of binge drinking conditional on any use, and the cumulative interactions of school enrollment and FTEs with both alcohol use variables. In addition to these stock variables, we include a set of degree indicators to capture potential discontinuities in the wage–schooling relationship. Because the decision to obtain a degree is captured by the school enrollment decision, we treat the degree indicators as exogenous conditional on the education stock variable.

Exogenous variables included in the model are as follows: indicators for living in one of four geographical regions; indicators for age cohort; indicators for gender and race/ethnicity; three age indicators, one each for the normal age range for attending high school (<18), college (18–21), and postcollege (>21); a continuous percentile score on the Armed Services Vocational Aptitude Battery to measure innate ability with missing values imputed using mean replacement; and the continuous local area unemployment rate.

4.2 Empirical model

We operationalize Equation 4 as follows:

$$\begin{aligned}\ln(w_{it}) = \gamma_0 &+ \gamma_1 EXP_{it} + \gamma_2 EXP2_{it} + \gamma_3 EXPDRK_{it} + \gamma_4 EXPBNG_{it} + \gamma_5 ED_{it} \\ &+ \gamma_6 EDDRK_{it} + \gamma_7 EDBNG_{it} + \gamma_8 DRK_{it} + \gamma_9 BNG_{it} + \gamma_{10} \boldsymbol{deg_{it}} + \gamma_{11} \boldsymbol{x_{it}} \\ &+ u_{it}\end{aligned} \tag{5}$$

where w_{it} is the hourly wage for person i in period t, $\boldsymbol{x}_{it}$ is a vector of exogenous demographic and labor characteristics described above, and u_{it} is an error term. *EXP* and *EXP*2 refer to cumulative FTEs and the sum of the squared FTEs. *ED* refers to cumulative years of school enrollment. *DRK* refers to cumulative years of any alcohol use; *BNG* refers to cumulative years of binge drinking, conditional on any use. *EXPDRK* and *EXPBNG*, *EDDRK*, and *EDBNG* represent the cumulative interactions of FTEs and enrollment with the alcohol indicators. $\boldsymbol{deg}_{it}$ refers to a series of indicators for the terminal degree earned by the individual as of period t. The squared terms from the Taylor series expansion for enrollment and the alcohol indicators are dropped because the values are either 0 or 1 and therefore exactly replicate the first-order terms.

In this model, γ_3, γ_4, γ_6, and γ_7 are the primary coefficients of interest. They represent the returns to wages from work experience or education for years in which the respondent consumed any alcohol and/or binge drank in that year. In terms of human capital production, these interaction terms correspond to π_6 and π_7 in Equation 4. Within the context of Equation 4, negative coefficients for these interaction terms imply that alcohol consumption reduces the marginal product of education or work experience in forming human capital. Positive coefficients imply that alcohol consumption increases the marginal human capital product of education or work experience. γ_8 and γ_9 correspond to π_5 in Equation 4, which captures the effects of alcohol use on wages as operating through all causal mechanisms. γ_1 and γ_2 correspond to π_3 and π_{10} in Equation 4, which capture the wage return to the marginal human capital product of work experience in which alcohol is not consumed. γ_5 corresponds to π_2 in Equation 4, which captures the wage return to the marginal human capital product of an additional year of schooling in which alcohol is not consumed.

4.3 Estimation

Proper estimation of Equation 5 requires addressing both endogeneity and sample selection within a panel data context. The work experience, education, and alcohol use variables in Equation 5 are endogenous because of the accumulated higher order terms captured in ε_t in Equation 4 and because we do not include covariates to capture the health stock included in Equation 4. Although the NLSY97 includes information on self-reported health status and a variety of health behaviors, we chose to not model health to better focus the analysis on the role of alcohol use in influencing wages through human capital mechanisms and to reduce the number of endogenous variables to be modeled. Importantly, the included endogenous variables are all predetermined stocks and so there is no potential for reverse causality. All endogeneity concerns arise from unobserved heterogeneity. Sample selection arises because we only model wages for those individuals working in a post-education job. Thus, selection into the wage equation is correlated with the decision to stop investing in formal education and to begin a working career.

To explore these estimation issues, we estimate Equation 5 using five estimation methods: ordinary least squares (OLS) regression with cluster robust standard errors, cross-sectional instrumental variables (IV), fixed effects (FE) panel data regression, random effects IV (RE-IV), and a discrete factor approximation method (DFM; Heckman & Singer, 1984; Mroz, 1999). Although potentially biased, OLS provides estimates of the conditional correlations between our variables of interest and wages, which serve as a useful baseline against which to judge the other results. IV ignores the panel structure of the data but corrects for endogeneity using the identifying instruments discussed below. IV does not correct for both endogeneity and sample selection unless the sample selection is completely captured by the modeled endogenous variables. FE uses the panel structure of the data to control for any unobserved heterogeneity that does not vary over the time frame of the panel. FE controls for both endogeneity and sample selection biases, but only to the extent that they are the result of time-invariant, individual-level unobserved heterogeneity. RE-IV includes an individual-level random effect to control for the panel structure of the data and uses the same identifying instruments as the IV model to control for endogeneity. Like IV, RE-IV only corrects endogeneity bias; it does not correct sample selection bias.

Discrete factor approximation method is a semiparametric, quasi-maximum likelihood estimation technique that controls for both sample selection and endogeneity caused by both time-varying and time-invariant unobserved heterogeneity. DFM estimates are identified using the longitudinal variation in the data, the correlation across outcomes within a person-year observation, and the identifying instruments discussed below. To implement the DFM, we jointly estimate a wage equation and one equation for each of the behaviors that contribute to the stocks included in the wage equation: any alcohol use, binge drinking, school enrollment, employment, and FTEs. We model the contemporaneous decisions for alcohol use, enrollment, and employment as binary measures with logistic error terms and FTEs and wages as continuous measures with normally distributed error terms. All equations in the system are correlated via a time-invariant random effect and an idiosyncratic random effect shared across all equations. The distributions of both random effects are approximated using discrete distributions. The DFM likelihood function is presented in Appendix S2.

4.4 Identifying instruments

The 14 identifying instruments used for this analysis include a series of nine family background variables; five price variables consisting of county-level inverse distance-weighted average prices for pizza, movies, beer, and wine; and state-level prices for tuition at 4-year public universities.

Family background variables from the NLSY97 are a set of five indicator variables describing the number and type (biological or not) of parents in the respondent's household at the time of the first interview in 1997 and the highest grade completed by the respondent's mother and father. Two binary variables were included for missing mothers' education and missing fathers' education. Family background characteristics have been used as identifying instruments in previous studies of the wage returns to education (e.g., Ashenfelter & Zimmerman, 1997; Card, 1999) because they are not typically observed by employers and so can only influence wages through some indirect mechanism, such as education. Recently, studies expand beyond education to other mechanisms such as residential choice or parental investments, which are themselves related to human capital accumulation to varying degrees. Regardless of the mechanism, however, most studies assume that family background affects wages only through some form of unobserved heterogeneity, not directly. Thus, we assume that they are excludable from the wage equation conditional on the modeled unobserved heterogeneity.

Price data are from the Council for Community and Economic Research (C2ER) cost of living index (COLI) dataset. We merge C2ER COLI price data to the NLSY97 at the county level and by quarter of survey administration. For each county in which an individual resides, we calculate inverse distance-weighted average prices using the prices from all markets present in the COLI data during that quarter. For each COLI market in the data, we calculated the distance from the geographic centroid of the market to the geographic centroid of every county in the United States. Using this distance, we then create inverse distance weights for each county and quarter. If a COLI market is geographically equivalent to a U.S. county, then the “weighted” prices for that county in that quarter are simply the prices from the COLI market that coincides with that county. Additional details on the price data used are available upon request.

Tuition data are from the Integrated Postsecondary Education Data System maintained by the National Center for Education Statistics. Average tuition at 4-year public universities is merged to the NLSY97 by state and year.

The IV and RE-IV models use the family background variables and the contemporaneous price and tuition variables as identifying instruments (14 identifying instruments for nine endogenous regressors). In the DFM, the alcohol use, enrollment, employment, and labor supply decisions that accumulate to form the endogenous variables in the wage equation are modeled in each period as functions of the family background variables and the contemporaneous price and tuition variables in each period, along with all covariates included in the wage equation. The family background, price, and tuition variables are excluded from the wage equation.

5 RESULTS

Table 1 provides descriptive statistics on the respondents in our analysis data. Time-varying characteristics are summarized as of the last observed period in the data. In the last observation, the average wage was approximately $15. The average years of work experience was about 3 years, and the average years of school enrollment was approximately 4 years, which is consistent with nearly 60% of respondents having a high school diploma or GED as the highest degree attained in their last observation. On average, respondents reported drinking in the last 30 days in approximately 6 years of data. The sample was balanced on gender (49% male) and mostly white (54%). Most respondents were between 13 and 16 years old in 1997, but about 20% were either 12 or 17 years old in 1997. The average respondent was in the 46th percentile of the ASVAB.

Table 1. Descriptive statistics

	Mean	***SD***
Time-varying characteristics in last period		
Hourly wage	15.43	11.49
Log hourly wage	2.58	0.53
Work experience stock		
Years of work experience	3.32	2.72
Years of work experience, squared	3.63	3.59
Work experience and no alcohol use	0.75	1.46
Work experience and any alcohol use	2.58	2.53
Work experience and binge drinking	1.34	1.96
Education stock		
Years of school enrollment	4.44	2.63
School enrollment and no alcohol use	1.31	1.72
School enrollment and any alcohol use	3.13	2.67
School enrollment and binge drinking	1.44	1.98
Alcohol stock		
Years of no alcohol use	2.47	2.87
Years of alcohol use	6.25	3.46
Years of binge drinking	2.99	3.03
Highest degree attained		
Less than high school	0.12	0.33
High school diploma or GED	0.59	0.49
Associate's degree	0.06	0.24
Bachelor's degree	0.20	0.40
Master's degree	0.03	0.16
Professional or doctoral degree	0.006	0.08

	Mean	***SD***
Time-invariant characteristics		
Male	0.49	0.50
Race/ethnicity		
White	0.54	0.50
Black	0.24	0.43
Hispanic	0.21	0.41
Age in 1997		
12	0.12	0.33
13	0.19	0.39
14	0.21	0.41
15	0.22	0.41
16	0.19	0.39
17	0.07	0.26
ASVAB percentile score	46.02	26.65

Note. N = 5,330 individuals.

Figure 1 shows trend lines for enrollment by alcohol consumption category. Through Period 3, enrollment rates across all three drinking categories are similar. Between Periods 3 and 6, the enrollment rate among abstainers falls more quickly than among nonbinge or binge drinkers. The enrollment rate among abstainers falls 34 percentage points, from 63% in Period 3 to 29% in Period 6, whereas the rate for nonbinge or binge drinkers falls only 21 percentage points, from about 66% in Period 3 to about 45% in Period 6. Starting in Period 6, the decline in enrollment among drinkers accelerates whereas the decline among abstainers slows, such that by about Period 9, the enrollment rate across all three categories is approximately equal, with nonbinge drinkers having a slightly higher enrollment rate than abstainers or binge drinkers.

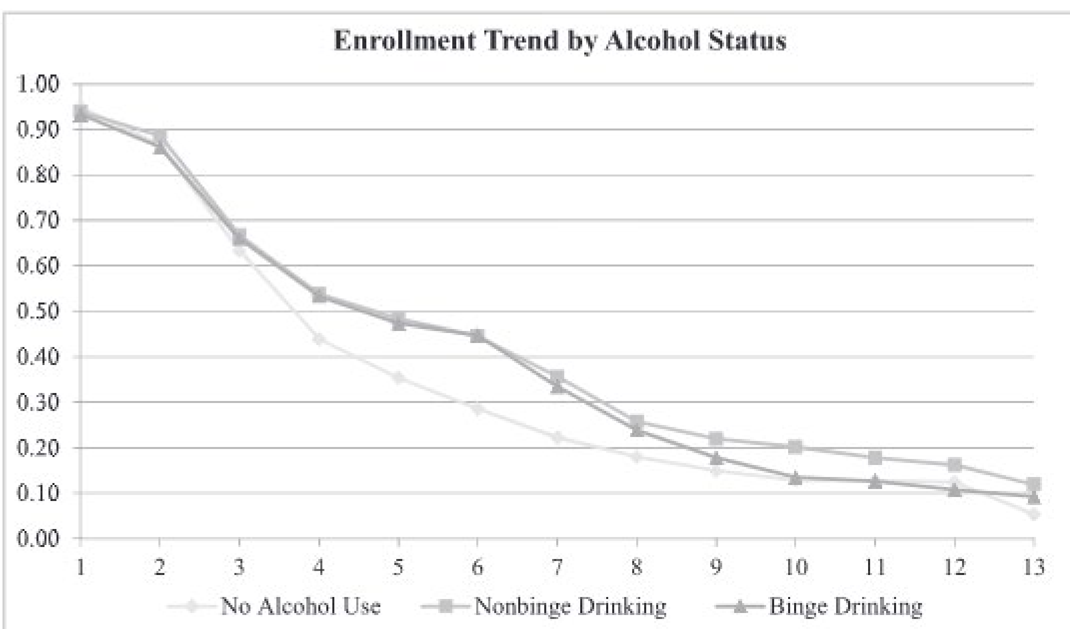

Figure 1. Enrollment trend by alcohol status

Figure 2 presents post-schooling employment trend lines by alcohol consumption category. As expected, employment rates trend upwards over the study time frame. All three drinking types have similar employment trajectories, with two possible exceptions. In Periods 5 and 6, abstainers have slightly higher employment rates than drinkers: Abstainers have employment rates of 52% and 58% in Periods 5 and 6, compared to approximately 46% and 50% for drinkers. Binge drinkers have higher employment rates in Periods 9 through 12 (approximately 80%) compared to abstainers and nonbinge drinkers (approximately 70%).

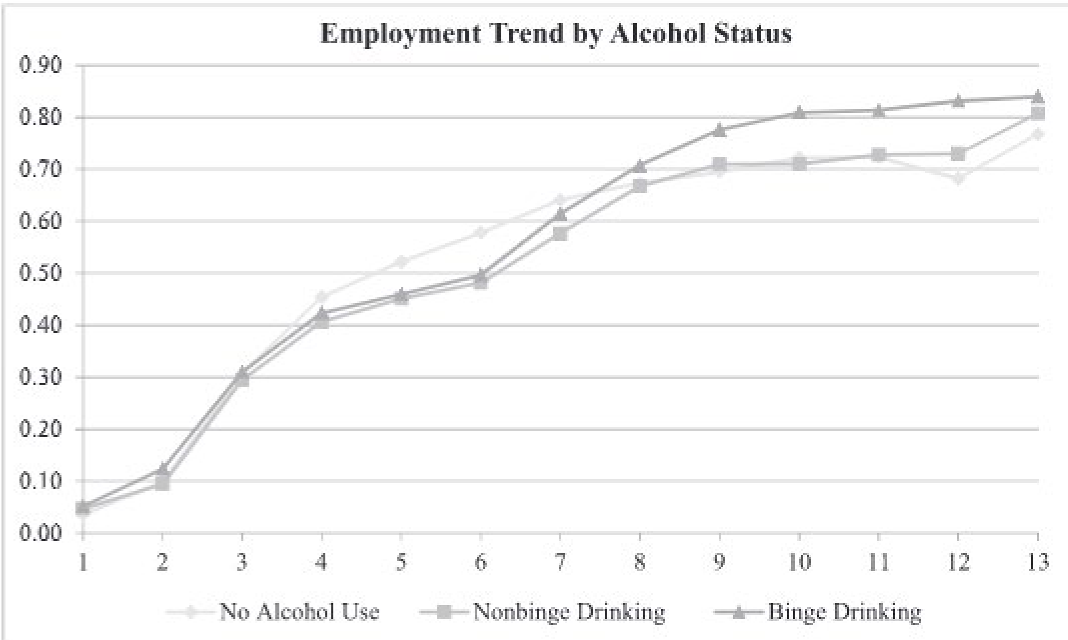

Figure 2. Employment trend by alcohol status

Figure 3 provides a trend line for the average hourly wage across study periods. Wages in our analysis sample exhibit the expected upward trend over time. Wages increase more rapidly for drinkers than for abstainers but at about the same rate for nonbinge and binge drinkers. By Period 8, the difference between drinkers and abstainers is striking, with drinkers earning about 16% higher wages than abstainers in Period 8, a premium that grows to almost 40% by Period 13. This pattern is broadly consistent with the drinker's bonus but is also consistent with the greater educational attainment implied by the trend lines in Figure 1.

Table 2 presents estimation results for the empirical model presented in Section 3 from OLS, IV, FE, RE-IV, and DFM. Summarizing the key results across all models, any alcohol use appears to lower the efficiency of both work experience and education in forming human capital, but binge drinking more than offsets this result. Although almost always statistically insignificant, estimates for the alcohol stock variables exhibit the opposite sign pattern: Prior use of alcohol increases wages whereas binge use decreases wages. In general, we see the expected positive and significant wage returns to work experience and education. The estimated returns to education are on the lower end or below the range of estimates from the broader labor economics literature (Card, 1999), except for the IV and RE-IV estimates, which are implausibly large. When significant, the coefficients on the demographic control variables are generally of the expected sign and magnitude.

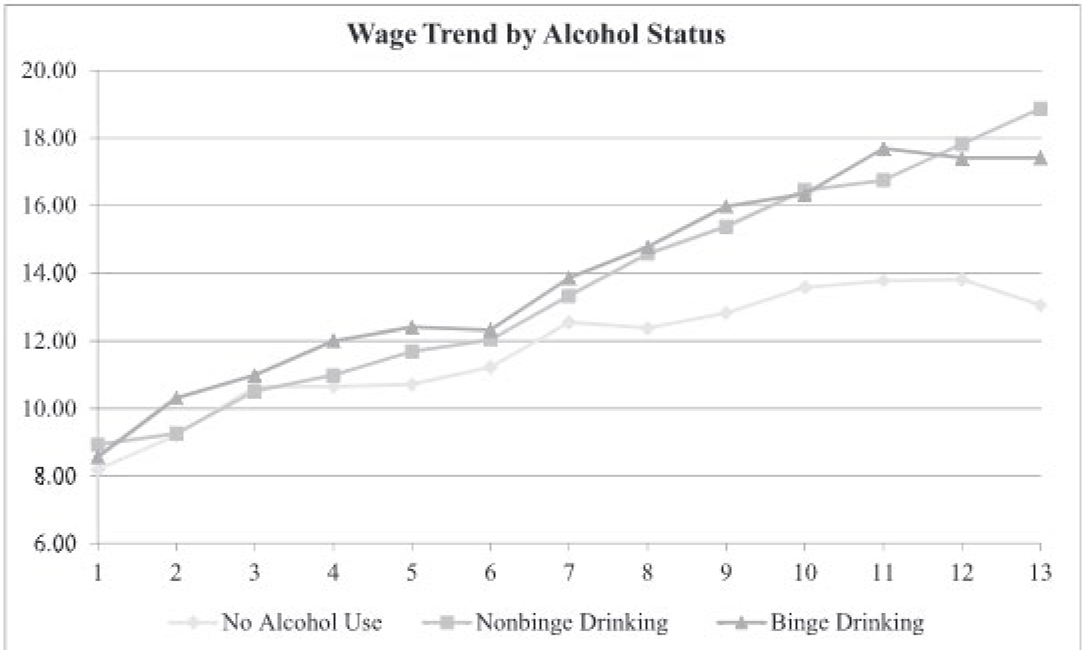

Figure 3. Wage trend by alcohol status

Table 2. Results

	Cross-sectional models		Panel models		
	OLS[a]	IV	FE	RE-IV	DFM
Endogenous stock variables					
Work experience	0.035** (0.0121)	1.440 (0.818)	0.049** (0.011)	1.230 (2.900)	0.054** (0.006)
Work experience, squared	0.0132 (0.0071)	−0.063 (0.615)	−0.008 (0.006)	−0.315 (1.563)	−0.011** (0.003)
Work experience and any alcohol use	−0.0034 (0.0122)	−2.513 (1.745)	−0.012 (0.010)	−1.547 (2.541)	−0.004 (0.005)
Work experience and binge drinking	0.0117 (0.0130)	1.718 (0.926)	0.002 (0.010)	1.986 (2.038)	0.016** (0.005)
School enrollment	0.0272** (0.0070)	0.218 (1.022)	0.050** (0.015)	1.102 (1.494)	0.033** (0.004)
School enrollment and any alcohol use	−0.0061 (0.0109)	−0.225 (1.538)	−0.039 (0.020)	−2.073 (3.488)	−0.019** (0.005)
School enrollment and binge drinking	0.0192 (0.0127)	0.355 (1.091)	0.043* (0.019)	2.523 (3.760)	0.031** (0.005)
Any alcohol use	0.0128 (0.0097)	0.422 (1.137)	0.016 (0.009)	1.648 (2.703)	0.004 (0.004)
Heavy alcohol use	−0.0180 (0.0117)	−0.219 (0.886)	−0.002 (0.010)	−2.129 (3.095)	−0.012** (0.004)
Education					
Has less than high school	−0.0613** (0.0130)	−0.229 (0.582)	0.015 (0.019)	−0.267 (0.393)	−0.013 (0.007)
Has associate's degree	0.2092** (0.0349)	−0.065 (0.680)	0.152** (0.047)	−0.095 (1.409)	0.112** (0.017)
Has bachelor's degree	0.3293** (0.0208)	−0.165 (0.895)	0.326** (0.053)	0.033 (1.288)	0.27** (0.011)
Has master's degree	0.5725** (0.0463)	0.047 (1.137)	0.572** (0.098)	0.218 (2.605)	0.51** (0.038)
Has doctoral or professional degree	0.8716** (0.1075)	0.392 (1.170)	0.518* (0.218)	0.633 (2.064)	0.94** (0.060)
Demographics					
Male	0.1491** (0.0104)	−0.349 (0.431)	- -	0.018 (1.277)	0.072** (0.006)
Race/ethnicity					
Black	−0.0005 (0.0135)	0.160 (0.275)	- -	0.120 (1.300)	−0.032** (0.008)
Hispanic	0.0206 (0.0137)	−0.061 (0.101)	- -	0.102 (0.637)	−0.001 (0.008)
Age at entry into the NLSY97					
12 years old	−0.0472* (0.0236)	−0.034 (0.209)	- -	−0.235 (0.428)	−0.026* (0.013)
13 years old	−0.0158 (0.0216)	−0.056 (0.143)	- -	−0.140 (0.389)	0.003 (0.011)

	Cross-sectional models		Panel models		
	OLS[a]	IV	FE	RE-IV	DFM
14 years old	−0.0249 (0.0211)	−0.102 (0.216)	- -	−0.226 (0.413)	−0.033** (0.010)
15 years old	−0.0046 (0.0216)	−0.077 (0.170)	- -	−0.175 (0.384)	−0.022* (0.011)
16 years old	−0.0197 (0.0215)	−0.113 (0.135)	- -	−0.127 (0.310)	−0.007 (0.011)
Less than 18 years	0.0796** (0.0174)	−0.084 (0.205)	0.086** (0.018)	0.401 (0.264)	0.038** (0.010)
Older than 21 years	0.0873** (0.0218)	−0.139 (0.270)	0.088** (0.021)	0.486 (0.301)	0.049** (0.011)
ASVAB percentile score	0.0012** (0.0003)	0.007 (0.006)	- -	−0.003 (0.014)	0.001** (0.000)
Geographic region					
Lived in northcentral region	−0.0498** (0.0166)	−0.048 (0.074)	−0.057 (0.038)	0.007 (0.674)	−0.088** (0.008)
Lived in southern region	−0.0952** (0.0148)	−0.146* (0.064)	−0.060* (0.030)	0.003 (0.251)	−0.088** (0.008)
Lived in western region	0.0446** (0.0164)	−0.022 (0.057)	−0.017 (0.037)	0.139 (0.331)	−0.011 (0.010)
Number of years in sample	−0.0111 (0.0065)	−0.324 (0.617)	−0.004 (0.006)	−0.819 (1.425)	−0.01** (0.003)
Local unemployment rate	−0.0060** (0.0016)	0.008 (0.011)	−0.005** (0.001)	−0.014 (0.023)	−0.002** (0.001)
Constant	2.1297** (0.0298)	2.606** (0.584)	2.140** (0.037)	2.772 (2.063)	2.423** (0.017)

Note. N = 24,129 observations on 5,330 individuals.
[a] Cluster robust standard errors.
* $p < .10$; ** $p < .05$; *** $p < .01$.

Looking specifically at cross-sectional models that ignore the panel structure of the data, the OLS results suggest a statistically significant return to work experience of approximately 3.5% in years without alcohol use, with a positive but insignificant quadratic term. Given the construction of the quadratic term, it captures nonlinearities in the production of human capital, not in the relationship between human capital and wages. Thus, it suggests an increasing marginal human capital product of work experience. The alcohol use interactions with work experience are both insignificant but suggest that any alcohol use slightly lowers the marginal human capital product of work experience, whereas binge drinking more than offsets this reduction. OLS results for education suggest a statistically significant wage return to education of approximately 2.7%, which is much smaller than other estimates in the literature. Consistent with the results found for work experience, both alcohol interactions are insignificant but suggest that any use lowers the marginal human capital product of education, which is more than offset by binge drinking. The IV results in Table 2 are imprecise but confirm the sign patterns of the OLS results except for the quadratic term on work experience; IV results suggest a diminishing marginal human capital product of work experience.

Turning to the panel data models, the FE results suggest statistically significant positive wage returns to experience of approximately 4.9% and an insignificant negative quadratic term. Consistent with the OLS and IV results, the FE results on the alcohol interactions are insignificant and suggest that any alcohol use lowers the marginal human capital product of work experience but that binge drinking more than offsets this reduction. We again see the same pattern for education: a positive and significant return to education without drinking, a negative interaction with any alcohol use, and a positive interaction with binge drinking. The RE-IV results are the least precise of all results but suggest the same sign pattern as the IV and FE results.

The preferred DFM model is fit using the upward testing approach suggest by Mroz (1999), resulting in a final specification with 11 mass points for the time-invariant random effect and 25 mass points for the time-varying random effect. The DFM results show a positive and significant return to work experience of approximately 5.4%, with a significant and negative quadratic term of approximately −1.1%. The incremental impact of any alcohol use is negative, but insignificant and essentially zero. The DFM estimates suggest that binge drinking increases the wage return to experience by about 1.6 percentage points. The return to education in years without alcohol use is significant and approximately 3.3%. Both alcohol interactions with education are significant, suggesting that any drinking reduces the return to education by about 1.9 percentage points whereas binge drinking increases the return by an additional 3.1 percentage points. The coefficient on the binge drinking stock variable is significant and negative, implying that each year of binge drinking reduces wages by about 1.2%.

To better understand better the implications of our estimates, we use the DFM results to predict wages for an individual who entered college with no prior alcohol use but drank all 4 years of college, holding all other covariates constant at the mean for continuous variables or the mode for categorical variables. Drinking, but not bingeing, for all 4 years of college reduces the predicted wage by 5.6% compared to not drinking, or about $0.86 per hour using the sample average wage of $15.43. In contrast, binge drinking all 4 years of college increases the predicted wage by 2.1% or about $0.32 per hour compared to not drinking. Comparing a college graduate working 35 hr/week (0.875 FTE) who does not drink during his first 4 years of work experience to one who drinks all 4 years, nonbinge drinking increases wages by about 0.3% or $0.04 per hour. Binge drinking increases the predicted wage by about 1% or $0.16 per hour.

5.1 Instrument validity and strength

We assess the validity of the identifying instruments using standard tests. First, we test the joint significance of the identifying instruments in the first-stage regressions on the endogenous stocks using Wald tests with cluster robust standard errors. The identifying instruments are jointly significant at the .01 level or better in all first-stage regressions, indicating that they significantly predict the endogenous variables. Both the Sargan and Basman tests fail to reject the null of overidentification at the .05 level, suggesting our instruments are valid; however, both tests have *p* values of .053. We also estimate Shea's partial R^2 and adjusted partial R^2 for all endogenous stocks to assess the amount of variation in the endogenous variables explained by the instruments. In all cases, the partial R^2s are less than 0.01 and adjusted partial R^2s are negative for all but the linear work experience variable. Although potentially problematic, Shea's partial R^2s are intended to reflect the ability of the identifying instruments to explain the variation in multiple, independent endogenous variables. Because our nine endogenous stock variables are highly correlated accumulations of only three underlying behaviors (alcohol use, labor supply, and school enrollment), Shea's R^2s may not be appropriate in this context. We could not conduct a minimum eigenvalue test for the strength of the instruments because test critical values are not available for nine endogenous regressors.

Finally, we reestimate our preferred DFM specification with the family background and contemporaneous price and tuition variables included in the wage equation and conduct a Wald test of their joint significance. We fail to reject the null hypothesis that the 14 identifying instruments from the IV models are validly excluded from the wage equation ($\chi^2 = 2.31$; $p = .99$). Because the family background variables may be related to wages through mechanisms other than the alcohol, schooling, and labor market behaviors that we model, we also test the joint significance of just the family background variables in the wage equation in the same DFM specification. Results suggest that the family background variables are validly excluded from the wage equation when the accumulated alcohol, schooling, and labor market behaviors of youth are jointly modeled ($\chi^2 = 1.09$, $p = .30$).

Although the IV testing results suggest our identifying instruments are weak, the totality of the testing evidence suggests that they are valid. Nonetheless, we urge caution when interpreting the IV estimates and consider the DFM estimates to be preferred because DFM does not rely solely on the instruments for identification and has been shown to be robust to weak instruments (Mroz, 1999).

5.2 Labor market attachment

In addition to wages, other labor market behaviors such as employment or labor supply are also potentially of interest to economists and policy makers. The employment and FTE (i.e., hours worked) equations in our DFM specification do not have precise theoretical interpretations because they do not use a theoretically motivated set of covariates. We therefore do not present formal statistical results for labor market attachment and instead discuss them qualitatively. If we assume that the endogenous stock variables only enter the FTE equation as determinants of the omitted wage variable, then the signs of their coefficients should be consistent with the role of wage in determining hours of work; that is, increasing wages should increase hours of work. Thus, we expect the signs of the endogenous stock variables in the FTE equation to match those of the wage equation. We find a matching sign pattern of coefficients in the FTE equation except for the accumulated squared work experience.

For the decision to be employed, the role of human capital is not as straightforward. On the one hand, increasing human capital should expand the set of job offers that a job-seeking individual faces and therefore should be positively related to employment. On the other hand, increasing human capital may also increase the reservation wage, which would decrease the set of viable job offers and therefore make human capital negatively related to employment. Empirical studies in this area focus on years of education as a measure of human capital and find a positive relationship between years of education and the measures of employment (Mincer, 1991; Groot & De Brink, 2000; Riddell & Song, 2011). We find that work experience and years of schooling are positively associated with employment. The coefficients on the accumulated interactions with the alcohol use variables, however, are of opposite sign from their corresponding coefficients in the wage equation. This sign pattern suggests that measures of the quantity of human capital investment expand the set of job offers that job-seeking individuals face. Measures of the quality of human capital appear to affect the probability of employment by raising the reservation wage and therefore reducing the set of viable job offers.

5.3 Robustness checks

A key omitted variable in our empirical model is the health stock. As a robustness check, we use self-reported health status as a proxy for the level of the health stock. We include an indicator for self-reported good-to-excellent health in the FE and DFM specifications and find that our results are qualitatively unchanged. This is consistent with health effects being captured in the modeled unobserved heterogeneity in both the FE model and DFM, as suggested by the theoretical derivation in Equation 4.

Another potential issue is our use of binge drinking rather than heavy drinking (defined by SAMHSA as binge drinking on five or more occasions in the past month). To explore the sensitivity of our results to our sole use of binge drinking, we add the accumulated years of work experience and education while heavily drinking to our specification and estimate the model using FE. Although not significant, we find that heavy drinking reduces the efficiency of both schooling and work experience in forming human capital. The coefficients on the accumulated interactions of any drinking with schooling and work experience in these models are unchanged from the FE results in Table 2. The binge drinking interactions are still positive, but of larger magnitudes than the FE results in Table 2.

6 DISCUSSION

Economists and policy makers have long been concerned about the possible effects of alcohol use on measures of labor force productivity, including wages and human capital investments such as schooling. Despite the widespread expectation that alcohol use should have a detrimental effect on these outcomes, the most consistent finding in the empirical literature is that individuals who drink moderately have higher wages than those who do not. In this paper, we use the NLSY97 and the theoretical model proposed by Bray (2005) to make two contributions. First, we more completely specify individuals' drinking histories using a measure of binge drinking that is consistent with the definition used by many U.S. policy makers. Second, we isolate the education- and work experience-related components of the drinker's bonus from other possible mechanisms.

In descriptive trend analyses, we find evidence of the drinker's bonus in that both moderate and binge drinkers had wages greater than those of abstainers, particularly in later time periods. We also find evidence of differential investments in schooling by drinkers that suggested a possible role for human capital formation in explaining this difference. Our preferred modeling results support this conclusion. We find evidence consistent with slight negative effects of any alcohol use on the wage returns of schooling and with positive and larger effects of binge drinking on the wage returns of schooling and work experience. Although estimates of the impact of any alcohol use on the wage returns to human capital investments are different for education than for work experience, the net incremental estimate for binge drinking is remarkably similar between the two investments at about 1.2 percentage points. This result is consistent with previous studies finding a drinker's bonus at consumption levels above those captured in our binge drinking measure (e.g., French & Zarkin, 1995; Barrett, 2002). Interpreting our results within the context of our theoretical model, estimates across all models suggest any alcohol use reduces the marginal human capital product of schooling, but binge drinking increases the marginal human capital product of schooling and work experience.

Our finding that binge drinking improves the marginal human capital product of schooling and work experience is counterintuitive, particularly because we used binge drinking as a measure of problematic alcohol use. Although our theoretical model isolates the human capital effects of alcohol use on wages, it does not specify the mechanism by which alcohol use affects human capital formation. One obvious mechanism is the direct, acute effect that intoxication has on the ability to learn and retain new skills and knowledge. We find it unlikely that this effect has any bearing on our empirical results because we do not assess alcohol use immediately before or during school or work. Rather, our measures of alcohol use are used as proxies for longer term drinking behaviors that likely have effects different from those of intoxication.

We posit that our results suggest both allocative and productive efficiency effects of alcohol use. Allocative efficiency operates through the choice of the quantity of inputs to the human capital production function, whereas productive efficiency operates through the productive capacity of the quantities chosen. Our measure of school enrollment reflects only the decision to enroll in school; it does not control for the study time invested by individuals. It therefore combines allocative and productive efficiency. Our measures of work experience and alcohol use, however, control for both the decision to work or drink and the amount of time to work or alcohol to consume. Because they control for allocation decisions, our measures of work experience and alcohol use isolate productive efficiency. Thus, the accumulated interaction between enrollment and any alcohol use reflects both allocative and productive efficiency. All other accumulated interactions in the model isolate productive efficiency.

Our finding that any alcohol use decreases the wage returns to school enrollment is consistent with negative allocative efficiency effects, whereas our finding that binge drinking increases the wage returns to both school enrollment and work experience is consistent with positive productive efficiency effects. This interpretation is supported by other evidence on the allocative and productive efficiency effects of alcohol. Studies on the effects of alcohol use on academic performance find that drinking reduces academic performance through effects on study time and other study habits (Williams et al., 2003; Powell et al., 2004), suggesting negative allocative efficiency effects. In contrast, studies on the nonintoxication effects of alcohol use on cognition find that alcohol use improves cognitive performance (Elias et al., 1999; Stampfer, Kang, Chen, Cherry, & Grodstein, 2005; Peters, Peters, Warner, Beckett, & Bulpitt, 2008; Reas, Laughlin, Kritz-Silverstein, Barrett-Connor, & McEvoy, 2016), including among youth and young adults (Peele & Brodsky, 2000). One study even finds positive cognitive effects of alcohol consumption at levels as high as two to four drinks per day for women and four to eight drinks per day for men (Elias et al., 1999). These results suggest positive productive efficiency effects of alcohol use on human capital formation.

Beyond our use of past month drinking measures to proxy for full-year drinking patterns, other limitations to our work exist. First, although they passed standard tests, our identifying instruments were weak predictors of alcohol use, schooling, and labor supply. The preferred DFM is generally robust to weak instruments (Mroz, 1999), but caution is still warranted when interpreting our results. Second, we use self-reported information on wages and hours of work. Self-reported data on wages and hours may include systematic measurement error that could bias our estimates. Although any measurement error would be orthogonal to the unobserved heterogeneity modeled by DFM, we cannot rule out that systematic reporting error exists. Third, we do not model health and instead relegate it to unobserved heterogeneity. Although our robustness check suggests that omitting health does not bias our results, the weakness of our instruments raises questions about the model's ability to fully relegate health effects to unobserved heterogeneity. Fourth, although the NLSY97 is nationally representative, our sample restrictions may limit the external validity of our results.

Despite these limitations, our paper makes an important contribution to the literature. We have found evidence suggesting that the drinker's bonus found in many previous studies may be caused, at least in part, by potentially beneficial effects of alcohol use on human capital formation. We do not claim to have identified the optimal level of drinking, nor do we think that our results suggest any benefits from intoxication. Nonetheless, we add to a long running debate in the public health literature on the harms and benefits of alcohol use. This debate includes the potential health benefits of moderate alcohol use (Stockwell et al., 2016) and the proper definition of binge drinking (White, Kraus, & Swartzwelder, 2006; Wechsler & Nelson, 2006). Our paper, combined with the previous literature on the schooling and labor market effects of alcohol use, suggests that possible cognitive benefits from alcohol use should be added to this ongoing debate.

Similarly, although we do not think that any single study should drive major changes in national policy, we nonetheless join a growing number of studies arguing for more targeted alcohol control policies. For example, Carpenter, Dobkin, and Warman (2016) suggest that policies intended to curb acute alcohol-related harms should include a focus on extreme binge drinking, which they link to alcohol-related motor vehicle accident deaths in Canada. Our work adds to this literature by further suggesting that the effects of alcohol use are complex and include both harms and benefits. Thus, although more evidence is clearly needed, alcohol control policies should be tailored to address alcohol consumption in the contexts within which it causes harm rather than attempting to deter all drinking behaviors.

ACKNOWLEDGEMENTS

Partial funding provided by the RTI Fellows Program. We thank Matthew Farrelly for the C2ER price data, Bill Wheaton and staff in RTI's Geospatial Science and Technology Program for assistance with inverse distance weighting, Stephanie Norris for research assistance, and Gary Zarkin and attendees at the 2016 American Society of Health Economists conference for helpful comments.

CONFLICT OF INTEREST

The authors have no conflicts of interest to declare.

REFERENCES

Ashenfelter, O., & Zimmerman, D. J. (1997). Estimates of the returns to schooling from sibling data: Fathers, sons, and brothers. *Review of Economics and Statistics*, **79**(1), 1– 9.

Auld, M. C. (2005). Smoking, drinking, and income. *Journal of Human Resources*, **40**(2), 505– 518.

Barrett, G. F. (2002). The effect of alcohol consumption on earnings. *Economic Record*, **78**(240), 79– 96.

Berger, M. C., & Leigh, J. P. (1988). The effect of alcohol use on wages. *Applied Economics*, **20**(October), 1343– 1351.

Böckerman, P., Hyytinen, A., & Maczulskij, T. (2017). Alcohol consumption and long-term labor market outcomes. *Health Economics*, **26**(3), 275– 291. https://doi.org/10.1002/hec.3290

Bray, J. W. (2000). Identifying inputs to the human capital production function: The effects of alcohol use on human capital formation. In *Unpublished doctoral dissertation*. Chapel Hill, NC: University of North Carolina at Chapel Hill.

Bray, J. W. (2005). Alcohol use, human capital, and wages. *Journal of Labor Economics*, **23**(2), 279– 311.

Bryant, R. R., Sumaranayake, V. A., & Wilhite, A. (1992). Alcohol use and wages of you men: Whites vs. nonwhites. *International Review of Applied Economics*, **6**(2), 184– 202.

Card, D. (1999). The causal effect of education on earnings. In O. Ashenfelter, & D. Card (Eds.), *Handbook of labor economics* (Vol. 3, pp. 1801– 1863). Amsterdam: Elsevier.

Carpenter, C. S., Dobkin, C., & Warman, C. (2016). The mechanisms of alcohol control. *Journal of Human Resources*, **51**(2), 328– 356.

Center for Behavioral Health Statistics and Quality. (2016). *Key substance use and mental health indicators in the United States: Results from the 2015 national survey on drug use and health*. MD: Retrieved from Rockville. http://www.samhsa.gov/data/

Centers for Disease Control and Prevention (2016). *Alcohol use and your health*. Retrieved from Atlanta, GA. https://www.cdc.gov/alcohol/fact-sheets/alcohol-use.htm

Cook, P. J., & Moore, M. J. (1993). Drinking and schooling. *Journal of Health Economics*, **12**(4), 411– 429.

Cook, P. J., & Moore, M. J. (2000). Alcohol. In A. Culyer, & J. Newhouse (Eds.), *Handbook of health economics*. Amsterdam: Elsevier.

Dee, T. S., & Evans, W. N. (2003). Teen drinking and educational attainment: Evidence from two-sample instrumental variables estimates. *Journal of Labor Economics*, **21**(1), 178– 209.

Elias, P. K., Elias, M. F., D'agostino, R. B., Silbershatz, H., & Wolf, P. A. (1999). Alcohol consumption and cognitive performance in the Framingham Heart Study. *American Journal of Epidemiology*, **150**(6), 580– 589.

French, M. T., & Zarkin, G. A. (1995). Is moderate alcohol use related to wages? Evidence from four worksites. *Journal of Health Economics*, **14**(3), 319– 344.

Groot, W., & De Brink, H. M. V. (2000). Education, training and employability. *Applied Economics*, **32**(5), 573– 581.

Heckman, J., & Singer, B. (1984). A method for minimizing the impact of distributional assumptions in econometric models for duration data. *Econometrica*, **52**(2), 271– 320.

Heien, D. M. (1996). Do drinkers earn less? *Southern Economic Journal*, **63**(1), 60– 68.

Kenkel, D. S., & Ribar, D. C. (1994). Alcohol consumption and young adults' socioeconomic status. *Brookings Papers on Economic Activity: Microeconomics*, **1**, 19– 61.

Koch, S. F., & Ribar, D. C. (2001). A siblings analysis of the effects of alcohol consumption onset on educational attainment. *Contemporary Economic Policy*, **19**(2), 162– 174.

Mincer, J. (1991). Education and unemployment (no. w3838). National bureau of economic research.

Mroz, T. (1999). Discrete factor approximations in simultaneous equation models: Estimating the impact of a dummy endogenous variable on a continuous outcome. *Journal of Econometrics*, **92**, 233– 274.

Mullahy, J., & Sindelar, J. L. (1993). Alcoholism, work, and income. *Journal of Labor Economics*, **11**(3), 494– 520.

National Institute on Alcohol Abuse and Alcoholism (2015). *Drinking levels defined*. Retrieved from Rockville, MD. https://www.niaaa.nih.gov/alcohol-health/overview-alcohol-consumption/moderate-binge-drinking

Peele, S., & Brodsky, A. (2000). Exploring psychological benefits associated with moderate alcohol use: A necessary corrective to assessments of drinking outcomes? *Drug and Alcohol Dependence*, **60**(3), 221– 247.

Peters, R., Peters, J., Warner, J., Beckett, N., & Bulpitt, C. (2008). Alcohol, dementia and cognitive decline in the elderly: A systematic review. *Age and Ageing*, **37**(5), 505– 512.

Powell, L. M., Williams, J., & Wechsler, H. (2004). Study habits and the level of alcohol use among college students. *Education Economics*, **12**(2), 135– 149.

Reas, E. T., Laughlin, G. A., Kritz-Silverstein, D., Barrett-Connor, E., & McEvoy, L. K. (2016). Moderate, regular alcohol consumption is associated with higher cognitive function in older, community-dwelling adults. *The Journal of Prevention of Alzheimer's Disease*, **3**(2).

Renna, F. (2007). The economic cost of teen drinking: Late graduation and lowered earnings. *Health Economics*, **16**(4), 407– 419.

Renna, F. (2008). Teens' alcohol consumption and schooling. *Economics of Education Review*, **27**(1), 69– 78.

Riddell, W. C., & Song, X. (2011). The impact of education on unemployment incidence and re-employment success: Evidence from the US labour market. *Labour Economics*, **18**(4), 453– 463.

Stampfer, M. J., Kang, J. H., Chen, J., Cherry, R., & Grodstein, F. (2005). Effects of moderate alcohol consumption on cognitive function in women. *New England Journal of Medicine*, **352**(3), 245– 253.

Stockwell, T., Zhao, J., Panwar, S., Roemer, A., Naimi, T., & Chikritzhs, T. (2016). Do "moderate" drinkers have reduced mortality risk? A systematic review and meta-analysis of alcohol consumption and all-cause mortality. *Journal of Studies on Alcohol and Drugs*, **77**(2), 185– 198.

Wechsler, H., & Nelson, T. F. (2006). Relationship between level of consumption and harms in assessing drink cut-points for alcohol research: Commentary on “Many college freshmen drink at levels far beyond the binge threshold” by White et al. *Alcoholism: Clinical and Experimental Research*, **30**(6), 922– 927.

White, A. M., Kraus, C. L., & Swartzwelder, H. S. (2006). Many college freshmen drink at levels far beyond the binge threshold. *Alcoholism: Clinical and Experimental Research*, **30**(6), 1006– 1010.

Williams, J., Powell, L. M., & Wechsler, H. (2003). Does alcohol consumption reduce human capital accumulation? Evidence from the College Alcohol Study. *Applied Economics*, **35**(10), 1227– 1239.

Willis, R. J. (1992). Wage determinants: A survey and reinterpretation of human capital earnings functions. Second edition. In O. Ashenfelter, & R. Layard (Eds.), *Handbook of labor economics* (Vol. 1. New York, NY: North Holland.

Yamada, T., Kendix, M., & Yamada, T. (1996). The impact of alcohol consumption and marijuana use on high school graduation. *Health Economics*, **5**(1), 77– 92.

Zarkin, G. A., French, M. T., Mroz, T. A., & Bray, J. W. (1998). Alcohol use and wages: New results from the National Household Survey on Drug Abuse. *Journal of Health Economics*, **17**(1), 53– 68.

Appendix 1

Taking a second order Taylor series expansion of k around a fixed point, $k(\bar{s},\bar{l},\bar{a})$ yields:

$$K_t = K_{t-1} + k(\bar{s},\bar{l},\bar{a}) + (s_{t-1}-\bar{s})k_s(\bar{s},\bar{l},\bar{a}) + (l_{t-1}-\bar{l})k_l(\bar{s},\bar{l},\bar{a}) + (a_{t-1}-\bar{a})k_a(\bar{s},\bar{l},\bar{a})$$
$$+ 0.5(s_{t-1}-\bar{s})^2 k_{ss}(\bar{s},\bar{l},\bar{a}) + 0.5(l_{t-1}-\bar{l})^2 k_{ll}(\bar{s},\bar{l},\bar{a}) + 0.5(a_{t-1}-\bar{a})^2 k_{aa}(\bar{s},\bar{l},\bar{a})$$
$$+ (s_{t-1}-\bar{s})(l_{t-1}-\bar{l})k_{sl}(\bar{s},\bar{l},\bar{a}) + (s_{t-1}-\bar{s})(a_{t-1}-\bar{a})k_{sa}(\bar{s},\bar{l},\bar{a})$$
$$+ (a_{t-1}-\bar{a})(l_{t-1}-\bar{l})k_{la}(\bar{s},\bar{l},\bar{a}) + \text{higher order terms}$$

where k_z refers to the derivative of k with respect to function argument z. Recursing human capital back to the initial period and rearranging terms gives

$$K_t = K_0 + A(t-1) + B\sum_{j=1}^{t-1} s_j + C\sum_{j=1}^{t-1} l_j + D\sum_{j=1}^{t-1} a_j + k_{sa}(\bar{s},\bar{l},\bar{a})\sum_{j=1}^{t-1} s_j a_j + k_{la}(\bar{s},\bar{l},\bar{a})\sum_{j=1}^{t-1} l_j a_j +$$
$$k_{sl}(\bar{s},\bar{l},\bar{a})\sum_{j=1}^{t-1} s_j l_j + 0.5k_{ss}(\bar{s},\bar{l},\bar{a})\sum_{j=1}^{t-1} s_j^2 + 0.5k_{ll}(\bar{s},\bar{l},\bar{a})\sum_{j=1}^{t-1} l_j^2 + 0.5k_{aa}(\bar{s},\bar{l},\bar{a})\sum_{j=1}^{t-1} a_j^2 + \nu_{t-1} \quad (A1)$$

where

$$A = k(\bar{s},\bar{l},\bar{a}) - \bar{s}k_s(\bar{s},\bar{l},\bar{a}) - \bar{l}k_l(\bar{s},\bar{l},\bar{a}) - \bar{a}k_a(\bar{s},\bar{l},\bar{a}) + 0.5\bar{s}^2 k_{ss}(\bar{s},\bar{l},\bar{a}) + 0.5\bar{l}^2 k_{ll}(\bar{s},\bar{l},\bar{a})$$
$$+ 0.5\bar{a}^2 k_{aa}(\bar{s},\bar{l},\bar{a}) + \bar{s}\bar{l}k_{sl}(\bar{s},\bar{l},\bar{a}) + \bar{s}\bar{a}k_{sa}(\bar{s},\bar{l},\bar{a}) + \bar{a}\bar{l}k_{la}(\bar{s},\bar{l},\bar{a})$$

$$B = k_s(\bar{s},\bar{l},\bar{a}) - \bar{l}k_{sl}(\bar{s},\bar{l},\bar{a}) - \bar{a}k_{sa}(\bar{s},\bar{l},\bar{a}) + \bar{s}k_{ss}(\bar{s},\bar{l},\bar{a})$$

$$C = k_l(\bar{s},\bar{l},\bar{a}) - \bar{l}k_{ll}(\bar{s},\bar{l},\bar{a}) - \bar{a}k_{la}(\bar{s},\bar{l},\bar{a}) + \bar{s}k_{sl}(\bar{s},\bar{l},\bar{a})$$

$$D = k_a(\bar{s},\bar{l},\bar{a}) - \bar{l}k_{la}(\bar{s},\bar{l},\bar{a}) - \bar{a}k_{aa}(\bar{s},\bar{l},\bar{a}) + \bar{s}k_{sa}(\bar{s},\bar{l},\bar{a})$$

The first term in equation A1 is the initial human capital stock (K_0). The second term, A(t – 1), captures the accumulated human capital product at the input vector $(\bar{s},\bar{l},\bar{a})$ net of the total differential effect of each input at $(\bar{s},\bar{l},\bar{a})$. The third, fourth, and fifth terms of equation A1 capture the cumulative inputs multiplied by their marginal human capital product. The remaining terms capture the accumulated effects of varying one input on the marginal human capital product of another input plus the accumulated higher order terms. Substituting equation A1 for K_t and a first order Taylor series expansion of the health stock for H_t into equation (1), we obtain the following wage equation:

$$\ln(w) = \beta_0 + \beta_1 X_1 + \beta_2 K_0 + \beta_3 H_0 + \pi_1(t-1) + \pi_2 \sum_{j=1}^{t-1} s_j + \pi_3 \sum_{j=1}^{t-1} l_j + \pi_4 \sum_{j=1}^{t-1} m_j + \pi_5 \sum_{j=1}^{t-1} a$$
$$+\pi_6 \sum_{j=1}^{t-1} s_j\, a_j + \pi_7 \sum_{j=1}^{t-1} l_j\, a_j + \pi_8 \sum_{j=1}^{t-1} s_j\, l_j + \pi_9 \sum_{j=1}^{t-1} s_j^2 + \pi_{10} \sum_{j=1}^{t-1} l_j^2 + \pi_{11} \sum_{j=1}^{t-1} a_j^2 + \epsilon_t \quad (4)$$

where

$\pi_1 = \beta_2 A + \beta_3[h(\bar{m},\bar{a}) - \bar{m}h_m(\bar{m},\bar{a}) - \bar{a}h_a(\bar{m},\bar{a})$

$\pi_2 = \beta_2 B$

$\pi_3 = \beta_2 C$

$\pi_4 = \beta_3 h(\bar{m},\bar{a})$

$\pi_5 = \beta_2 D + \beta_3 h_a(\bar{m},\bar{a})$

$\pi_6 = \beta_2 k_{sa}(\bar{s},\bar{l},\bar{a})$

$\pi_7 = \beta_2 k_{la}(\bar{s},\bar{l},\bar{a})$

$\pi_8 = \beta_2 k_{sl}(\bar{s},\bar{l},\bar{a})$

$\pi_9 = 0.5\beta_2 k_{ss}(\bar{s},\bar{l},\bar{a})$

$\pi_{10} = 0.5\beta_2 k_{ll}(\bar{s},\bar{l},\bar{a})$

$\pi_{11} = 0.5\beta_2 k_{aa}(\bar{s},\bar{l},\bar{a})$

and ε_t captures ξ_t and accumulated higher order terms from the Taylor series expansions.

Appendix 2

The empirical wage model is:

$$\begin{aligned}\ln(w_{it}) = \gamma_0 + \gamma_1 EXP_{it} + \gamma_2 EXP2_{it} + \gamma_3 EXPDRK_{it} + \gamma_4 EXPBNG_{it} + \gamma_5 ED_{it} + \gamma_6 EDDRK_{it} \\ + \gamma_7 EDBNG_{it} + \gamma_8 DRK_{it} + \gamma_9 BNG_{it} + \gamma_{10} \boldsymbol{deg}_{it} + \gamma_{11} \boldsymbol{x}_{it} + u_{it}\end{aligned} \quad (5)$$

Let $SCHOOL_{it}$ be an indicator for person i being enrolled in period t, EMP_{it} be an indicator for person i being employed in period t, $HOURS_{it}$ be the hours worked by person i in period t, $ALC1_{it}$ be an indicator for any drinking by person i in period t, and $ALC2_{it}$ be an indicator for binge drinking by person i in period t. Then

$$SCHOOL_{it} = \begin{cases} 1 & \text{if } \beta_{\mathrm{S},0} + + \beta_{\mathrm{S},1}\boldsymbol{x}_{it} + \beta_{\mathrm{S},2}\boldsymbol{f}_{it} + \beta_{\mathrm{S},3}\boldsymbol{p}_{it} + \beta_{\mathrm{S},4}\boldsymbol{STOCKS}_{it} + \mu_{S,it} > 0 \\ 0 & otherwise \end{cases}$$

$$EMP_{it}|_{SCHOOL_{it}=0} = \begin{cases} 1 & \text{if } \beta_{\mathrm{E},0} + + \beta_{\mathrm{E},1}\boldsymbol{x}_{it} + \beta_{\mathrm{E},2}\boldsymbol{f}_{it} + \beta_{\mathrm{E},3}\boldsymbol{p}_{it} + \beta_{\mathrm{E},4}\boldsymbol{STOCKS}_{it} + \mu_{E,it} > 0 \\ 0 & otherwise \end{cases}$$

$$HOURS_{it}|_{EMP_{it}=1} = \beta_{\mathrm{H},0} + + \beta_{\mathrm{H},1}\boldsymbol{x}_{it} + \beta_{\mathrm{H},2}\boldsymbol{f}_{it} + \beta_{\mathrm{H},3}\boldsymbol{p}_{it} + \beta_{\mathrm{H},4}\boldsymbol{STOCKS}_{it} + \mu_{H,it}$$

$$ALC1_{it} = \begin{cases} 1 & \text{if } \beta_{A1,0} + + \beta_{\mathrm{A1},1}\boldsymbol{x}_{it} + \beta_{\mathrm{A1},2}\boldsymbol{f}_{it} + \beta_{A1,3}\boldsymbol{p}_{it} + \beta_{A1,4}\boldsymbol{STOCKS}_{it} + \mu_{A1,it} > 0 \\ 0 & otherwise \end{cases}$$

$$ALC2_{it}|_{ALC1_{it}=1} = \begin{cases} 1 & \text{if } \beta_{\mathrm{A2},0} + \beta_{\mathrm{A2},1}\boldsymbol{x}_{it} + \beta_{\mathrm{A2},2}\boldsymbol{f}_{it} + \beta_{\mathrm{A2},3}\boldsymbol{p}_{it} + \beta_{\mathrm{A2},4}\boldsymbol{STOCKS}_{it} + \mu_{A2,it} > 0 \\ 0 & otherwise \end{cases}$$

where the *it* subscript refers to the *i*th individual in period *t*, $\boldsymbol{f}_{it}$ is a vector of family-background variables, $\boldsymbol{p}_{it}$ is a vector of price variables, and $\boldsymbol{STOCKS}_{it}$ comprises accumulation of education, work experience, alcohol use, and their interactions, as defined in equation 5.

Redefining notation to be more compact yields

$$ln(w_{it}) = \boldsymbol{X}_{Wit}\boldsymbol{\beta}_W + \epsilon_{Wit}$$

$$SCHOOL_{it} = \begin{cases} 1 & \text{if } \boldsymbol{X}_{Sit}\boldsymbol{\beta}_S + \mu_{Sit} > 0 \\ 0 & otherwise \end{cases}$$

$$EMP_{it}|_{SCHOOL_{it}=0} = \begin{cases} 1 & \text{if } \boldsymbol{X}_{Eit}\boldsymbol{\beta}_E + \mu_{Eit} > 0 \\ 0 & otherwise \end{cases}$$

$$HOURS_{it}|_{EMP_{it}=1} = \boldsymbol{X}_{Hit}\boldsymbol{\beta}_W + \mu_{Hit}$$

$$ALC1_{it} = \begin{cases} 1 \text{ if } \boldsymbol{X}_{A1it}\boldsymbol{\beta}_{A1} + \mu_{A1it} > 0 \\ 0 \ \ otherwise \end{cases}$$

$$ALC2_{it}|_{Alc1_{it}=1} = \begin{cases} 1 \text{ if } \boldsymbol{X}_{A2it}\boldsymbol{\beta}_{A2} + \mu_{A2it} > 0 \\ 0 \ \ otherwise \end{cases}$$

$$\mu_{jit} = \eta_{ji} + \gamma_{jit} + \varepsilon_{jit}; \ \ j = W, S, E, H, A1, A2$$

The resulting likelihood function is

$$\mathrm{L} = \prod_{i=1}^{N}\sum_{j=1}^{J}\pi_{\eta j}\prod_{t=1}^{T}\sum_{k=1}^{K}\pi_{\upsilon k}\left[\frac{1}{\sigma_w}\phi\left(\frac{\ln(w_{it}) - \boldsymbol{X}_{Wit}\boldsymbol{\beta}_W - \eta_{wj} - \upsilon_{wk}}{\sigma_w}\right)\right]^{EMP_{it}(1-SCHOOL_{it})}$$

$$\mathrm{x}\left[\frac{1}{\sigma_\mathrm{H}}\phi\left(\frac{HOURS_{it} - \boldsymbol{X}_{Hit}\boldsymbol{\beta}_H - \eta_{Hj} - \upsilon_{Hk}}{\sigma_\mathrm{H}}\right)\right]^{EMP_{it}(1-SCHOOL_{it})}$$

$$\mathrm{x}\left[\Lambda\left(\boldsymbol{X}_{A1it}\boldsymbol{\beta}_{A1} + \eta_{A1j} + \upsilon_{A1k}\right)\right]^{ALC1_{it}}\left[1 - \Lambda\left(\boldsymbol{X}_{A1it}\boldsymbol{\beta}_{A1} + \eta_{A1j} + \upsilon_{A1k}\right)\right]^{(1-ALC1_{it})}$$

$$\mathrm{x}\left[\Lambda\left(\boldsymbol{X}_{A2it}\boldsymbol{\beta}_{A2} + \eta_{A2j} + \upsilon_{A2k}\right)\right]^{ALC2_{\ it}(1-ALC1_{it})}$$

$$\mathrm{x}\left[1 - \Lambda\left(\boldsymbol{X}_{A2it}\boldsymbol{\beta}_{A2} + \eta_{A2j} + \upsilon_{A2k}\right)\right]^{(1-ALC2_{it})(1-ALC1_{it})}$$

$$\mathrm{x}\left[\Lambda\left(\boldsymbol{X}_{Sit}\boldsymbol{\beta}_S + \eta_{Sj} + \upsilon_{Sk}\right)\right]^{SCHOOL_{it}}\left[1 - \Lambda\left(\boldsymbol{X}_{Sit}\boldsymbol{\beta}_S + \eta_{Sj} + \upsilon_{Sk}\right)\right]^{(1-SCHOOL_{it})}$$

$$\mathrm{x}\left[\Lambda\left(\boldsymbol{X}_{Eit}\boldsymbol{\beta}_E + \eta_{Ej} + \upsilon_{Ek}\right)\right]^{EMP_{it}(1-SCHOOL_{it})}$$

$$\mathrm{x}\left[1 - \Lambda\left(\boldsymbol{X}_{Eit}\boldsymbol{\beta}_E + \eta_{Ej} + \upsilon_{Ek}\right)\right]^{(1-EMP_{it})(1-SCHOOL_{it})}$$

where φ(·) is the normal probability distribution function and Λ(·) is the normal cumulative distribution function. $\pi_{\eta j}$ and $\pi_{\upsilon k}$ are subsequently modified with a logit transformation with α as a parameter to be estimated.

$$\pi_{\eta j} = \frac{e^{\alpha_j}}{1 + \sum_{j=1}^{J-1} e^{\alpha_j}}, \pi_{\upsilon k} = \frac{e^{\alpha_k}}{1 + \sum_{k=1}^{K-1} e^{\alpha_k}}$$