How Alcohol Consumption Affects Labor Supply?
Junming Hu1, a, *, †, Dian Sheng2, b, *, †, Yao Zhang3, c, *, † and Ning Ding4, d, *, †

1Department of Economics, University of Birmingham
2Department of Business, East China University of Science and Technology
3Department of Business, Penn State University
4Department of Science, Xi’an Jiaotong-liverpool University

*Corresponding author. Email: ‘jjxh106@student.bham.ac.uk; †19002519@mail.ecust.edu.cn; †ykz5379@psu.edu; †Ning.ding19@student.xjtlu.edu.cn
†These authors contributed equally.

ABSTRACT
The purpose of this research is to understand the impact of drinking on the labor market. The data set comes from the National Longitudinal Survey of Youth 1989 and 1994 (NLSY89, NLSY94), which includes information about 12,686 individuals restricted to young adults with the age between 24 and 32 in 1989. The data was collected by a sample survey, and it was mainly investigated by dependent variables such as the amount of alcohol consumed. It is found through our research that moderate drinking is beneficial to labor supply by using Ordinary Least Squares, Two-Stage Least Squares, and Instrumental Variable-Probit models. However, excessive drinking still has a huge impact on labor. The main importance of this research is to explore the impact of alcohol on the labor market.

Keywords: drinking, labor market.

1. INTRODUCTION

Liquor permeates many civilized histories, from literary and artistic creation, cultural entertainment to food and cooking, health care, and other aspects, which occupies an important position in the history. Until now, alcohol is still a drink that most people like. Whether in business conference, class reunion, or gathering of friends, alcohol is almost inseparable. Some people even drink alcohol every day. Alcohol has almost become a part of most people’s lives. However, drinking has a big impact on people’s work, which may cause work errors and decrease the efficiency of people around them.

Due to the popularity of alcohol, many people have ever drunk alcohol. As a result, the data of the 1979 Youth Special (NLSY79) is used to investigate the impact of drinking on the labour market.

Although many scholars have discussed the issues of labour supply and drinking, our research achieves more progress. We use R² as a basis to assign weights to the six drinking-related variables, and finally integrate a new multi-dimensional variable grade that incorporates drinking conditions. Because of the advantages of defining variables in this way, it is found that moderate drinking is actually beneficial to labour supply, which has been ignored by previous studies. This is also the core part of our entire research. In addition to the influence of gender and health on drinking, the difference in drinking status between people living in cities and rural areas is also explored, which is not investigated by scholars in previous studies.

2. DATA & VARIABLES

The data set used in this article comes from the National Longitudinal Survey of Youth (NLSY89 and NLSY94) provided by a sample of employed persons, including information about 12,686 individuals restricted to young adults who are between the ages of 24 and 32 in 1989 (and hence 29-37 in 1994). Each individual has a unique identifier (variable named id) and the year is indicated by the variable named year.

We selected the variable hrswrk as the dependent variable to observe the labor supply. Its meaning is the number of working weeks of the individual in the past year. Obviously, the key influencing factors of our entire study are related to drinking, so the six variables related to drinking are all within our consideration. The six variables reflect respectively whether the individual has drunk alcohol, whether he has drunk alcohol in the last month, the number of times he has drunk in the last
month, the number of days he drank in the last month, the cycle of drinking once, and the number of times he drank more than expected. Picking whichever one of the six variables for research seems too one-sided. Any one of them cannot balance time, frequency, and drinking psychology at the same time. Consequently, we generate a new variable named grade combining the six features. Initially, we need to find a method to define how heavy a drinker is. From the NLSY89 and NLSY94, we regressed hrswrk which means the total number of hours worked in the past calendar year, with respect to 6 variables who indicate alcohol consumption respectively, and all models contained only hrswrk, alcohol consumption, and a constant, without any control variables, to get the R²’s. As we know, the R² could be regarded as a goodness of fitness, namely, the power of explanation to the dependent variable, so it could be reasonable to be the weight. But it could not be used directly for the large difference, and we need to amend it. So we regard

\[ \text{weight}_i = \log(1000 \times R_i^2) \]  

However, the weight does not take dimension into consideration, and we need to make variables dimensionless as well. We use grade to be the only variable that indicates alcohol consumption, and

\[ \text{grade}_i = \sum_{j=1}^{6} \text{weight} \times \left( \frac{X_{ij}}{\max\{X_i\}} \right) \]

where i means individual i and j means the variables that indicates alcohol consumption including drnkev, drnkmo, days, drnk6m, perday and gtint. The detail of generating grade is shown in Table 1.

![Fig. 1. Distribution Density of Grade](image)

### Table 1 How grade generated

| variable   | mean                     | How we define in our model                        |
|------------|--------------------------|--------------------------------------------------|
| **drnkev** | a dummy variable = 1 if the individual has ever had a drink, = 0 otherwise | No change                                       |
| R²=0.008   | weight=2.014             |                                                  |
| **drnkmo** | a dummy variable = 1 if the individual has had a drink in the last month, 0 otherwise | No change                                       |
| R²=0.025   | weight=3.238             |                                                  |
| **drnk6m** | a categorical variable indicating the number of times in the past month the individual has had 6 or more drinks in one sitting. It is defined as follows:

0 = Never  
1 = Once  
2 = 2 Or 3 Times  
3 = 4 Or 5 Times  
4 = 6 Or 7 Times  
5 = 8 Or 9 Times  
6 = 10 Or More Times | six dummy variables for drnk6m equals 1~6. When we calculate grade, we need to take the value of six dummy variables for an individual divided by the maximum

\[ \text{drnk6m}(=6) \in \left( \frac{X_{ij}}{\max\{X_i\}} \right) \] |
| R²=0.006   | weight=1.757             |                                                  |
| **days**   | the number of days in the last month the individual has had at least 1 drink | No change                                       |
| R²=0.013   | weight=2.533             |                                                  |
| **perday** | the average number of drinks per day on a day when the individual drinks (this is 0 if the individual doesn’t drink) | No change                                       |
| R²=0.001   | weight=0.000             |                                                  |
A categorical variable that answers the question of whether the individual has ever drunk more than intended. It is defined as follows:

- $g_{t} = 0$ if $gtint = 0$
- $g_{t} = 1$ if $gtint = 1$
- $g_{t} = 2$ if $gtint = 2$
- $g_{t} = 3$ if $gtint = 3$
- $g_{t} = 4$ if $gtint = 4$
- $g_{t} = 5$ if $gtint = 5$

Generate 5 dummy variables for $g_{t} = 0 \sim 4$.

When we calculate grade, we need to take the value of $gt$ of an individual divided by the maximum $gt$ (=5) in

$$grade = \frac{gtint}{5}$$

Of course, in addition to the explained variables and the main explanatory variables, there are still some control variables and other influencing factors that we must consider. Since many primitive variables such as region, race, gender are represented by serial numbers, we use dummy variables to represent categorical variables in the regression. Variables named starting with the letters ‘drc’ reflect race. Variables named starting with the letters ‘drln’ reflect the religion the respondents believe in when the investigation. Similarly, the variable at the beginning of ‘drg’ indicates the region. At the same time, to simplify the model, we integrated the height and weight into one variable according to the BMI calculation formula. Family income has also been reduced to one-thousandths in advance. To verify our conjecture, the labor supply of people who drink moderately is better (will be described in detail later).

3. MODEL

3.1. Captions/numbering

To detect the association between alcohol consumption and labor market supply, we would regress with OLS originally. And to avoid autocorrelation, although data from two years, we would regress year by year instead of using a mixed effect model.

We begin with fitting the total number of hours worked in the past calendar year by OLS.

$$hrswrk = \alpha_{1} + X_{2} \beta_{1} + grade_{1} \gamma + grade_{2} \delta + \varepsilon_{i}$$

(3)

where $X_{1}$ is a vector of observable characteristics for individual $i$ which are assumed to be predetermined to current employment and alcohol consumption. These include the usual characteristics: age, marital status, ethnicity, year of education, family size, having health problems that limits the amount or kind of work that can be done, degree of urbanization, parents’ working status and education, IQ level, and poverty status. In the work equation, we also control for body-mass index (BMI), which we enter as a quadratic term. Some former researchers regarded these individual and demographic characteristics as the exogeneous variable.

3.2. 2SLS

However, former researchers reckon that we should use the instrumental variable to allow for the possibility of unobserved heterogeneity[1]. The choice of IVs is a difficult task in the literature that has investigated the effect of alcohol. We need instruments which affects the working time only through the mechanism of increasing or decreasing the alcohol consumption. However, different instruments may cause very different results of the effect of alcohol consumption on working time, and the choice of IV might be different for different genders. But here we investigate the overall effect of alcohol consumption for both males and females. We regard the grade and grade$^2$ as the instrumented variable with some of the religions, mom’s years of schooling, and whether living in the urban area in 2SLS. The equation of first-stage regression is

$$grade_1 = \alpha_{2} + X_{1} \beta_{2} + I_{phi} + \varepsilon_{i1}$$

(4)

$$grade_2 = \alpha_{3} + X_{1} \beta_{3} + I_{phi} + \varepsilon_{i2}$$

(5)

3.3. Probit and IV-Probit

When we consider labor supply, the basic theorem to measure the amount of labor supply is Solow Model, which measures labor supply in working time. However, from the aspect of individual, a longer working time does not imply a higher income. For example, a high-level engineer may earn more but work less than a construction worker, but how much they drink is hard to guess. So we have to take employment status into consideration. The NLSY classified employment status into 4 categories with variable empst, which defined as $1 = Employed$ $2 = Unemployed$ $3 = Out Of Labor Force$ $4 = In Active Armed Forces$. So we define $demp = 1$ if $empst = 1$ and $= 0$ if $empst = 2$ or $3$. 

We could use a probit model to test how drinking affects the probability of an individual is employed, but the choice of control variables from that in OLS. The equation is

\[ demp_i = a_4 + X_i' \beta_4 + grade_i \gamma + grade_i^2 \delta + \varepsilon_{4i} \]  

(6)

And the instruments we choose are different as well.

\[ grade_i = a_5 + X_i' \beta_5 + I_i' \phi + \varepsilon_{5i} \]  

(7)

\[ grade_i^2 = a_6 + X_i' \beta_6 + I_i' \psi + \varepsilon_{6i} \]  

(8)

4. EMPIRICAL RESULTS

4.1. OLS

Table 1 reports the OLS results of the estimations for working weeks and alcohol consumption in 1989 and 1994. From the coefficient, it is palpable to find that moderate drinkers are always with a longer working time in a year than binge drinkers' and abstainers' working time, for the coefficient of grade is positive and that of the quadratic term is negative, namely, the curve would be downward bending, and the relationship is significant most of the time. Some of the demographic variables, family factors, religions, regions, and macroeconomic factors would affect working time. We would omit the control variable in the following table.

| Dependent variable | Total number of hours worked in the past calendar year |
|--------------------|------------------------------------------------------|
| grade              | 46.752**                                             |
|                    | (13.359)                                             |
| grade2             | -2.902**                                            |
|                    | (1.79)                                               |
| Control Variables Omitted |
| N                  | 5356                                                 |
| \( R^2 \)          | 0.290                                                |

Standard errors in parentheses
\* p < 0.1, \** p < 0.05, \*** p < 0.01

When we separate that element affecting the alcohol consumption, the alcohol consumption itself does not have such a strongly effect on the working time.

| Dependent variable | total number of hours worked in the past calendar year |
|--------------------|------------------------------------------------------|
| grade              | -172.964                                             |
|                    | (291.410)                                            |
| grade2             | 18.138                                               |
|                    | (28.359)                                             |
| Control variables omitted |
| N                  | 5356                                                 |
| \( R^2 \)          | 0.161                                                |

Standard errors in parentheses
\* p < 0.1, \** p < 0.05, \*** p < 0.01
4.3. Probit Model

The result showed in Table 4 implies that moderate drinkers are more likely to be employed. The employment status is more likely to be affected by parents’ working status when the respondents were 14, other religions which are not included in the survey, and other marital status like divorce.

Table 5 Probit Model

| Dependent Variable | 1989 | 1994 | 1989 | 1994 |
|-------------------|------|------|------|------|
| grade             | 0.091*** | 0.091*** | (0.028) | (0.031) |
| grade2            | -0.006** | -0.007*** | (0.003) | (0.003) |

Control Variables Omitted

N 5256 3808

Standard errors in parentheses

* p < 0.1, ** p < 0.05, *** p < 0.01

Also, we try to use IV-Probit model to solve endogeneity in the ordinary Probit model. But unfortunately, this result is not that reliable because we cannot solve endogeneity even such amount of instruments are used. Former literature used religion [2], job repetitiveness, obesity [3] and parental smoking, partner smoking and non-chronic health condition [4], beer tax [5], long-term, non-acute illnesses [6]. Here for lack of data, we cannot say the last result is concise enough, but it is worthy to show the IVs we choose, especially those significant ones.

Table 6 First Stage Regression

| year | grade 1989 | grade 1994 | grade2 1989 | grade2 1994 |
|------|-----------|-----------|-------------|-------------|
| south14 | -0.291*** | -0.220** | -3.162*** | -2.388*** |
| (0.098) | (0.112) | (1.135) | (1.277) |
| dadhc | 0.061*** | 0.047*** | 0.699*** | 0.569*** |
| (0.010) | (0.012) | (0.118) | (0.135) |
| drln2 | -0.454*** | -0.195 | -5.455*** | -2.526 |
| (0.124) | (0.150) | (1.419) | (1.696) |
| drln4 | 0.790*** | 0.634*** | 8.564*** | 6.830*** |
| (0.185) | (0.207) | (2.116) | (2.378) |
| drln7 | 0.393*** | 0.257** | 3.838*** | 2.474*** |
| (0.109) | (0.129) | (1.256) | (1.464) |
| bmi | 0.073 | -0.050 | 1.090 | -0.617 |
| (0.060) | (0.063) | (0.692) | (0.712) |
| bmi2 | -0.002 | 0.000 | -0.024** | 0.002 |
| (0.001) | (0.001) | (0.012) | (0.012) |
| urbrur | 0.189 | 0.200 | 1.745 | 1.572 |
| (0.102) | (0.119) | (1.174) | (1.386) |
| finc | 0.005* | 0.048*** | 0.046 | 0.439*** |
| (0.003) | (0.014) | (0.033) | (0.165) |
| drg1 | 0.240*** | 0.023 | 2.156 | 0.631 |
| (0.115) | (0.132) | (1.321) | (1.502) |
| dma2 | -0.795*** | -0.494*** | -9.728*** | -7.057*** |
| (0.106) | (0.149) | (1.214) | (1.837) |
| cons | 6.862*** | 8.537*** | 59.421*** | 84.504*** |
| (0.914) | (1.007) | (10.482) | (11.476) |

Control Variables and Some of IVs Omitted

N 4969 3546 4969 3546

Standard errors in parentheses

* p < 0.1, ** p < 0.05, *** p < 0.01
sensitivity to the depressing effects of alcohol is a feasible scheme [3]. In particular, with the global spread of COVID-19, there are several papers on the implications of alcohol for the market economy in the context of the epidemic. Some researchers [8] and [9] delve into the issues involved. Of course, the relationship between gender, age, and the volume of alcohol abuse cannot be ignored. Relevant issues have been discussed in depth as well[11].

In our research, there are some deficiencies caused by supervisory and objective reasons, which is remained to be corrected. For example, our data is too old. At present, we have not found any newer data related to labor and drinking status. The current world is completely different from that of 1984 and 1994. The entertainment habits (including drinking) of people are undergoing more than subtle changes. Let us make a bold conjecture: perhaps in a much more modern society where business is more prosperous, businessmen’s demand for drinking and entertainment has increased, and consequently drinking is conducive to the supply of labor. We have no way of knowing it. And although our indicators for judging drinking are more comprehensive than previous studies, considering multiple factors such as time, alcohol consumption, frequency, the weights we assign to the six drinking-related variables may not be very scientifically reliable.

REFERENCES

[1] Renna, F., Alcohol Abuse, Alcoholism, and Labor Market Outcomes: Looking for the Missing Link. Industrial and Labor Relations Review, 2008. 62(1): p. 92-103.
[2] Heien, D.M., Do Drinkers Earn Less? Southern Economic Journal, 1996. 63(1): p. 60-68.
[3] Berger, M.C. and J.P. Leigh, The effect of alcohol use on wages. Applied Economics, 1988. 20(10): p. 1343-1351.
[4] MacDonald, Z. and M.A. Shields, Does problem drinking affect employment? Evidence from England. Health Economics, 2004. 13(2): p. 139-155.
[5] Mullahy, J. and J. Sindelar, Employment, unemployment, and problem drinking. Journal of Health Economics, 1996. 15(4): p. 409-434.

[6] MacDonald, Z. and M.A. Shields, The Impact of Alcohol Consumption on Occupational Attainment in England. Economica, 2001. 68(271): p. 427-453.

[7] Allen, H.C., et al., Acute rewarding and disinhibiting effects of alcohol as indicators of drinking habits. Psychopharmacology, 2021. 238(1): p. 181-191.

[8] Stockwell, T., et al., The burden of alcohol on health care during COVID-19. Drug and Alcohol Review, 2021. 40(1): p. 3-7.

[9] Callinan, S., et al., Purchasing, consumption, demographic and socioeconomic variables associated with shifts in alcohol consumption during the COVID-19 pandemic. Drug and Alcohol Review, 2021. 40(2): p. 183-191.

[10] Zhang, X., et al., Alcohol consumption and risk of cardiovascular disease, cancer and mortality: a prospective cohort study. Nutrition Journal, 2021. 20(1): p. 13.

[11] Bellia, F., et al., Transient serotonin depletion at adolescence, but not at early infancy, reduced subsequent anxiety-like behavior and alcohol intake in female mice. Psychopharmacology, 2021. 238(1): p. 215-225.