The Effect of Binge Drinking Frequency on Unemployment

Larry Liao, Edward Harvey, Shuo Zheng, Vivek Manjunath

This paper seeks to determine the effect of the frequency of binge drinking on unemployment in young adults. Binge drinking is defined as 6 or more drinks in one sitting and frequency is split into three categories: non-binge drinkers, infrequent binge drinkers, and frequent binge drinkers. Young adults who binge drink fewer than four times a month are deemed infrequent binge drinkers, while those who drink more are frequent binge drinkers and those who abstain from binge drinking are non-drinkers. Our three models use the infrequent and frequent binge-drinking dummies as independent variables and weeks unemployed annually as the dependent variable. Our first two models estimate by OLS with various demographic controls while the final regression employs an individual fixed-effects regression. A significant positive relationship was found between frequency of binge drinking and weeks unemployed in the first two models but not with the fixed effects regression.

Introduction

As both college students and people seeking future employment, we are interested in how an individual's drinking habits might affect their ability to maintain employment. We identify two different schools of thought regarding this relationship. The first is the more commonly held and seemingly rational notion that engaging in binge drinking is associated with poor workplace performance and other behaviors related to higher probability of unemployment. The contrasting notion is that more frequent binge drinking may be associated with greater social capabilities or a work-hard, play-hard mentality, both factors that may be related to lower probability of unemployment. We seek to identify if either of these perceptions can possibly be supported by analyzing the direction, magnitude, and significance of the relationship between binge drinking frequency and weeks of unemployment.

To explore this relationship, we looked at data from the National Longitudinal Survey of Youth from the years 1989 and 1994 for individuals between the ages of 24 and 32 in 1989, with a total of 16,000 individuals surveyed. The dependent variable in our analysis is the total number of weeks spent unemployed in the past calendar year. We used frequencies of binge drinking as the independent variables. The data set includes a variable for the number of times an individual has engaged in binge drinking over the last month (defined as having 6 or more drinks). We decided to break this variable into dummy variables: non-binge drinkers (0 times per month), infrequent binge drinkers (greater than 0 and less than 4 times per month), and frequent binge drinkers (greater than 3 times per month).

We began with a main regression of weeks of unemployment on the dummy independent variables for binge drinking frequency. Then, we added various controls sequentially. Finally, we performed a fixed effects regression involving all controls.

After running our regressions, our most plausible interpretation was that a positive relationship between frequency of binge drinking and unemployment most likely exists. Seven different OLS regressions with various demographic controls returned significant results. In fact, in these OLS regressions, almost all of the coefficients related to the frequency of binge drinking are significant are the 1% level – even as we add controls.

The regression controlling for individual-level fixed effects did not produce significant results, but this result does not necessarily invalidate the results from the OLS regressions. Rather, the 95% confidence for the dummies related to frequent and infrequent binge drinking encompass the results found in the seven different OLS regressions. The large interval reflects the imprecision regarding the fixed effects regression. Therefore, our most plausible interpretation is that there is a significant, positive relationship between frequency of binge drinking and weeks unemployed.

Methodology

To determine the relationship between frequency of binge drinking and employment, we used a series of OLS regressions with various controls and also a fixed effect regression. The OLS with no controls except for year takes the form:

$$(1)\quad Y_i = \beta_0 + \beta_1 bingefreq_i + \beta_2 bingeinfreq_i + \beta_3 D_i^{1994} + u_i$$

where Y_i represents weeks unemployed in a year, $bingefreq_i$ is a dummy for frequent binge drinkers, $bingeinfreq_i$ is a dummy for infrequent binge drinkers, and D^{1994} controls for year effects. Thereby, the coefficients β_1 and β_2 may respectively be interpreted as the difference in unemployment between frequent binge drinkers and non-drinkers and the difference in unemployment between infrequent binge drinkers and non-drinkers. The constant term, β_0, is the weeks of unemployment for the average non-drinker. This model likely suffers from omitted variable bias, however, as there are variables that affect both frequency of drinking and unemployment that are not accounted for. For example, with something like age, older persons may drink less and be more employable for reasons due to their advanced age. This positive bias is not accounted for in the above regression. An OLS regression that considers this form of bias takes the form:

$$(2)\quad Y_i = \beta_0 + \beta_1 bingefreq_i + \beta_2 bingeinfreq_i + \beta_3 D_i^{1994} + \beta_4 D_i^{sex} + \beta_5 D_i^{race} + \beta_{6-233} D_i^{industry} + \beta_{234} age_i + \beta_{235-242} D_i^{religion} + \beta_{243} yrsed_i + \beta_{244} D_i^{badhealth} + \beta_{245} famsz_i + \beta_{246} urate_i + u_i$$

where controls for sex, race, industry, age, religion, years of education, health, family size, and local unemployment rate are all added to (1). This approach aims to reduce the bias in the estimates of β_1 and β_2. However, sources of bias may still exist from unobservable characteristics across individuals. These characteristics' unobservable nature means surveys like the National Longitudinal Survey of Youth cannot collect data on them. Things like propensity to addiction are not observable or at least not perfectly measurable and, in this particular case, may bias the estimates upward. Nonetheless, there is a way to eliminate the bias from these unobserved effects if they are fixed across time. This regression, a fixed effects regression, takes the form:

$$(3)\quad Y_i = \beta_0 + \beta_1 bingefreq_{it} + \beta_2 bingeinfreq_{it} + \beta_3 D_t^{1994} + \beta_{4-229} D_{it}^{industry} + \beta_{230} age_{it} + \beta_{231} yrsed_{it} + \beta_{232} D_{it}^{badhealth} + \beta_{233} famsz_{it} + \beta_{234} urate_{it} + v_{it}$$

where the dummy controls for sex, race, and religion are dropped because of collinearity. Since these controls are fixed across individuals, conducting fixed effects on

Table 1. Mean Characteristics in 1989 and 1994 Cross-Section, by Frequency of Binge Drinking

Characteristics	Frequent Binge Drinkers	Infrequent Binge Drinkers	Non-Drinkers	All
Weeks Unemployed	3.223	2.331	1.779	2.036
	(7.930)	*(6.746)*	*(6.171)*	*(6.486)*
Year 1994 Survey Wave	0.476	0.466	0.502	0.490
	(0.500)	*(0.499)*	*(0.500)*	*(0.500)*
% Male	0.834	0.683	0.419	0.523
	(0.373)	*(0.465)*	*(0.493)*	*(0.500)*
% Hispanic	0.177	0.185	0.176	0.179
	(0.382)	*(0.389)*	*(0.381)*	*(0.383)*
% Other Race	0.550	0.591	0.538	0.554
	(0.498)	*(0.492)*	*(0.499)*	*(0.497)*
Age	30.020	30.080	30.460	30.320
	(3.361)	*(3.381)*	*(3.366)*	*(3.375)*
Years of Education	12.140	12.820	13.440	13.170
	(1.931)	*(2.249)*	*(2.364)*	*(2.337)*
% with Health Problem	0.026	0.035	0.049	0.043
	(0.160)	*(0.185)*	*(0.215)*	*(0.204)*
Family Size	2.740	2.861	3.099	3.007
	(1.749)	*(1.545)*	*(1.563)*	*(1.577)*
Local Unemployment Rate	6.485	6.395	6.375	6.389
	(2.589)	*(2.605)*	*(2.653)*	*(2.635)*
Frequent Binge Drinkers				0.070
				(0.254)
Infrequent Binge Drinkers				0.284
				(0.451)
Observations	727	2,968	6,764	10,459

(standard deviations in parentheses)

Notes: Weeks unemployed is shown as the total weeks unemployed in the past calendar year. % with health problem indicates those who have a health problem that limits the amount or kind of work that can be done. % other race indicates those who are not Hispanic and not Black. Frequent Binge Drinkers and Infrequent Binge Drinkers are defined as those who drink more than six drinks in one sitting four times or more in the last month, and greater than zero but less than four times in the last month, respectively.

individuals in this regression eliminates the need for these controls. Only two of the two hundred twenty-eight industry dummies in (2) are dropped, suggesting that there are shifts in industry employment from 1989 to 1994.

Unobservable effects that are not fixed across time may remain in the composite error term, v_{it}, but this idiosyncratic error is assumed to average to zero, meaning the estimates of β_1 and β_2 should be unbiased.

Data Description

The data used in our analysis are from the National Longitudinal Survey of Youth. This data set includes information on labor market outcomes, alcohol consumption, and various demographics. The data were taken in 1989 and 1994 for individuals aged 24 to 32 in 1989. In each regression, we dropped observations that had missing values for any included variables or did not have data for both 1989 and 1994. Thus, we ran each regression on a sample size of 10,459 observations.

To study unemployment, our dependent variable, we chose a variable in the data set that recorded the total number of weeks spent unemployed in the past calendar year (*wksue*). In looking at binge drinking frequency – our right-hand side variable – we chose a categorical variable that indicated the number of times in the past month that an individual had six or more drinks in one sitting; we split this variable into three dummy variables. The first was a dummy for individuals who did not engage in binge drinking (*bingenever*), which we call non-drinkers. The second was a dummy for individuals who engaged in binge drinking infrequently, or greater than zero but less than four times in the last month (*bingeinfreq*), which we call infrequent binge drinkers. The third was a dummy for individuals who engaged in binge drinking four or more times in the last month (*bingefreq*), which we call frequent binge drinkers. The regressions incorporate controls for year (*yr94*, a year dummy), gender (*male)*, years of education (*yrsed)*, health status (*badhealth,* a dummy for whether or not individuals have a health problem that limits the amount or kind of work that can be done), family size (*famsz)*, religion, industry, local unemployment rate (*urate)*, age, and race (where *other race* denotes those who are not Hispanic and not Black). Means and standard deviations of the variables used in our analysis, including controls, are shown in Table 1.

Results

The balance test in *Table 2* provides a justification for each of the controls by determining the significance of the controls' relationship with the frequency-of binge-drinking-explanatory variables. *Table 2* shows a significant, negative relationship between both frequency-of-binge-drinking explanatory

Table 2. Balance Test Regressions

Explanatory Variables	male	yrsed	badhealth	famsz	urate	age	hispanic	other race
bingefreq	0.415 ***	-1.295 ***	-0.023 ***	-0.359 ***	0.110	-0.446 ***	0.001	0.012
	(0.017)	*(0.088)*	*(0.007)*	*(0.074)*	*(0.109)*	*(0.128)*	*(0.017)*	*(0.023)*
bingeinfreq	0.264 ***	-0.616 ***	-0.013 ***	-0.239 ***	0.020	-0.385 ***	0.009	0.053 ***
	(0.012)	*(0.059)*	*(0.004)*	*(0.036)*	*(0.060)*	*(0.074)*	*(0.010)*	*(0.013)*
Observations	10,459	10,459	10,459	10,459	10,459	10,459	10,459	10,459
R-squared	0.084	0.028	0.001	0.007	0.000	0.003	0.000	0.002
F-stat	413.40	125.40	8.20	28.37	0.51	17.19	0.45	9.08
p-value	0.00	0.00	0.00	0.00	0.60	0.00	0.64	0.00

Notes: Table shows balance test regressions of each control on bingefreq and bingeinfreq dummy independent variables, defined as drinking more than six drinks in one sitting four times or more in the last month, and greater than zero but less than four times in the last month, respectively. Controls include gender, years of education, health, family size, local unemployment rate, age, and race. Religion and industry dummies, used later as controls, have been omitted for space. Badhealth is a dummy that indicates those who have a health problem that limits the amount or kind of work that can be done. Other race is a dummy that indicates those who are not Hispanic and not Black. Standard errors (in parentheses) were computed to be robust to arbitrary error correlation within households as well as heteroskedasticity.

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

variables and the *yrsed*, *badhealth*, *famsz*, and *age* controls. Increasing levels of years of education, poor health, family size, and age, therefore, are associated with a decreased probability of binge drinking—a notion that seems entirely plausible. It is plausible that these factors also affect unemployment, so it is prudent to control for these variables in the context of reducing omitted variables bias. For example, *yrsed* and *age* are likely negatively related to weeks unemployed, so controlling for them reduces upward biases on β_1 and β_2. Controlling for poor health and family size – factors that seem to be positively associated with weeks unemployed – on the other hand, accounts for negative bias. The dummy for gender shows that men engage in binge drinking more frequently than women, and this variable is likely to be positively related to unemployment due to a greater number of women dropping out of the labor force while unemployed, perhaps due to childcare. As a result, controlling for gender would likely reduce an upward bias on our estimates. In cases where *Table 2* fails to display any significant relationship between controls and the explanatory variables, like with *urate* and the dummy for Hispanic race, we elected to include the variables as controls in an effort to reduce the standard error. Since these variables likely affect unemployment but do not have a significant relationship with binge drinking frequency, controlling for them will help to produce more precise estimates.

Beginning with our OLS regression in equation (1) of Table 3, we regress *wksue* on *bingefreq* and *bingeinfreq*, accounting for year fixed effects with a 1994 year dummy and adjusting our standard errors to be robust to arbitrary error correlations within households as well as heteroskedasticity. Our estimates of the explanatory variables, then, show that frequent binge drinkers, on average, are associated with 1.444 more weeks of unemployment than non-drinkers. Infrequent binge drinkers are also associated with more weeks of unemployment than non-drinkers; in this case, the estimate is smaller in magnitude, at .553 more weeks of unemployment than non-drinkers. Both estimates are significant at the 1% level, implying at first glance that these could be causal estimates.

However, we must control for omitted variables bias. From equations (2) through (7), we gradually add in controls, starting with a male dummy in equation (2). Controlling for gender, our estimates show that frequent binge drinkers and infrequent binge drinkers are now associated with 1.291 and .455 more weeks of unemployment than non-drinkers, respectively. In this case, we note that males upwardly biased our original OLS estimates, resulting in the reduction of the magnitude of our estimates when gender is controlled. These estimates are still significant at the 1% level as well.

From equations (3) to (6), we follow the same pattern, adding in one new control in each regression. In equation (3), we add in controls for race in our regression, before adding controls for industry, age, and religion in equations (4), (5), and (6), respectively. Since race is not significantly associated with our independent variables – as seen in our balance test in Table 2 – we note that controlling for race helps to make our estimates more precise, as shown from the slight reduction in the standard errors of equation (3) estimates: .316 to .313 for *bingefreq* and .154 to .153 for *bingeinfreq*.

In equations (4) and (5), the magnitude of our estimates is revised further downward after we control for industry (4) and age (5). Although our estimates now show that frequent binge drinkers are associated with .866 more weeks of unemployment than non-drinkers, and that infrequent binge drinkers are associated with .435 more weeks of unemployment than non-drinkers, these estimates are significant at the 1% level.

Table 3. OLS Regressions of Binge Drinking Frequency on Weeks Unemployed 1984 and 1994

Explanatory Variable	(1) *wksue*	(2) *wksue*	(3) *wksue*	(4) *wksue*	(5) *wksue*	(6) *wksue*	(7) *wksue*
bingefreq	1.444 ***	1.291 ***	1.335 ***	0.889 ***	0.866 ***	0.872 ***	0.756 **
	(0.309)	*(0.316)*	*(0.313)*	*(0.315)*	*(0.315)*	*(0.315)*	*(0.315)*
bingeinfreq	0.553 ***	0.455 ***	0.592 ***	0.450 ***	0.435 ***	0.424 ***	0.377 **
	(0.148)	*(0.154)*	*(0.153)*	*(0.156)*	*(0.156)*	*(0.156)*	*(0.158)*
yr94	0.026	0.026	0.040	0.070	0.462 **	0.469 **	0.264
	(0.118)	*(0.118)*	*(0.117)*	*(0.120)*	*(0.188)*	*(0.188)*	*(0.196)*
male		0.369 ***	0.331 **	0.024	0.021	0.011	0.089
		(0.141)	*(0.139)*	*(0.154)*	*(0.154)*	*(0.153)*	*(0.153)*
age					-0.078 ***	-0.079 ***	-0.083 ***
					(0.030)	*(0.030)*	*(0.030)*
yrsed							-0.157 ***
							(0.032)
badhealth							1.093 ***
							(0.378)
famsz							0.028
							(0.048)
urate							0.128 ***
							(0.028)
Controls for race	No	No	Yes	Yes	Yes	Yes	Yes
Controls for industry	No	No	No	Yes	Yes	Yes	Yes
Controls for religion	No	No	No	No	No	Yes	Yes
Constant	1.766 ***	1.612 ***	3.027 ***	0.000	3.157 ***	11.480 ***	5.766 ***
	(0.098)	*(0.107)*	*(0.193)*	*(0.000)*	*(0.874)*	*(1.015)*	*(1.077)*
Observations	10,459	10,459	10,459	10,459	10,459	10,459	10,459
R-squared	0.004	0.005	0.024	0.065	0.066	0.066	0.072

Notes: Table shows OLS estimates of weeks unemployed (in the last calendar year) on bingefreq and bingeinfreq dummies. All variables listed above are the same as those defined in the Table 2 footnotes. Standard errors (in parentheses) were computed to be robust to arbitrary error correlation within households as well as heteroskedasticity.

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

In equation (6) we control with religion dummies, increasing the magnitude of our frequent binge drinker estimate to .872 more weeks of unemployment than non-drinkers while decreasing the magnitude of our infrequent binge drinker estimate to .424 more weeks of unemployment than non-drinkers. These estimates are still significant at the 1% level.

In equation (7), we add our last set of controls, either to control for omitted variables bias (with *yrsed*, *badhealth*, *famsz*) or to help reduce our standard errors (with *urate*). In this case, estimates show that frequent binge drinkers are associated with .756 more weeks of unemployment than non-drinkers, while infrequent binge drinkers are associated with .377 more weeks of unemployment than non-drinkers. While these estimates are no longer

Table 4. Fixed Effects Regression of Weeks Unemployed on Binge Drinking Frequency

Explanatory Variables	wksue	Confidence Interval 95%	
bingefreq	0.135	-1.225	1.495
	(0.694)		
bingeinfreq	-0.042	-0.714	0.631
	(0.343)		
yr94	-0.431		
	(3.413)		
age	0.059		
	(0.537)		
yrsed	0.276		
	(0.321)		
badhealth	0.132		
	(0.787)		
famsz	0.125		
	(0.144) *		
urate	0.115		
	(0.064)		
Constant	-2.996		
	(15.720)		
Industry Controls	Yes		
Observations	10,459		
R-squared	0.689		

Notes: Table shows Fixed Effects estimates of weeks unemployed (in the last calendar year) on bingefreq and bingeinfreq dummies. Variables are the same as the ones described in the Table 2 footnotes. The controls listed above were ones that were not absorbed by individual fixed effects. Standard errors (in parentheses) were computed to be robust to arbitrary error correlation within households as well as heteroskedasticity.

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

significant at the 1% level, they are still significant at the 5% level.

To control for individual-level fixed effects, we employ a fixed-effects regression in Table 4, employing the same controls as in equation (7), noting that the controls listed in Table 4 are the ones that were not absorbed by individual fixed effects. In these estimates, frequent binge drinkers are associated with .135 more weeks of unemployment than non-drinkers, while infrequent binge drinkers are associated with .042 fewer weeks of unemployment than non-drinkers. It is important to note that neither of these estimates is significant at the 10% level. A possible explanation for this sudden reversal in significance can be found through the fixed effects standard errors. The large standard errors for both estimates, .694 for frequent binge drinkers and .343 for infrequent binge drinkers – compared to our OLS standard errors in Table 3 – may imply that our fixed effects regression is too noisy to precisely estimate the effects of binge drinking frequency on weeks unemployed. This view is reinforced by the fact that the 95% confidence interval for *bingefreq* and *bingeinfreq* in our fixed effects regression encompasses all of our OLS regression estimates in Table 3.

Discussion

The National Longitudinal Survey of Youth gathered self-reported data, which is vulnerable to measurement error. In the case of misreporting, we expect classical measurement error, since we see no reason for systematic misreporting in weeks unemployed, frequency of binge drinking, or any of the controls for this particular survey. Measurement error in weeks unemployed, the dependent variable, will not affect the values of any slope estimates, but if misreporting occurs in any of the right-hand side variables (the binge drinking frequency variables), then our slope estimates will be attenuated. Therefore, any positive effect of binge drinking on weeks unemployed produced from the models above may in fact be larger. In addition, because we are working with panel data, we compute our standard errors to be robust to arbitrary error correlation within households by clustering on households. This also helps account for any potential heteroskedasticity.

While the fixed effects regression shows no significant relationship between binge drinking and unemployment, the estimate is too imprecise to completely discredit the significant relationship found in our previous OLS estimates. Because of this, we argue that it is worth considering the implications of this significant result in the lives of college students and young professionals.

Although we cannot conclude that this relationship is casual, it is possible that more frequent binge drinking could be detrimental to seeking or maintaining employment. Our estimates indicate that frequent binge drinking is associated with only .756 weeks more unemployment, which amounts to just over 5 days. However, the implications of these 5 days appear more serious after considering that the mean number of weeks unemployed for all individuals surveyed is 2.036, just over 14 days.

It is also possible that people who are unemployed are more likely to binge drink, potentially introducing the phenomenon of reverse causality into the regression. Factors such as greater amounts of free time, fewer responsibilities, or feelings of sadness or depression could contribute to more frequent binge drinking amongst unemployed people.

There are areas that our analysis did not explore that may be important for further research. One of these is that our data and analysis categorized binge drinking as consuming six or more drinks in one sitting, but this may definition may be too broad. An analysis in which binge drinking is defined separately for men and women and by different weight categories could produce more accurate estimates of the effects on unemployment.

As college students make the transition from school into their professional lives, they might be well served to leave their binge drinking habits behind. Securing a job after graduation is already a challenging task, and graduates may act in their own self-interest to reduce binge drinking in the hopes of more successfully maintaining employment.

Conclusion

Based on our regressions, the most plausible interpretation is that there is a significant, positive relationship between binge drinking and unemployment. Seven different OLS regressions with various different demographic controls and one fixed effects regression were run to determine this relationship. Each of the OLS regressions returned significant results for binge drinking variables at the 5% significance level.

Although the fixed effects regression was not significant, it is likely that this resulted from an imprecise estimation. The 95% confidence interval is quite large, which reflects this imprecision. It also encapsulates the estimates found in the OLS regressions, reaffirming the significant results found with the seven OLS regressions.

The strong association between binge drinking and unemployment suggests that students may have it in their best interests to leave college drinking habits behind as they transition into their professional lives.