Catchy Title?

Quantifying the Effects of Alcohol Consumption on Labor Force Participation

Abstract goes here

Introduction

With the United States’ economic recovery from the Great Recession in full swing, it is often of interest to researchers to look at labor force participation and unemployment rates as signs of economic recovery. However, there exist factors external to the economy which affect the rate at which people are employed. Factors such as alcoholism, family size, or geographic area can greatly affect an individual’s decision to participate in the labor force. Moreover, there may be inferences that can be drawn about a person’s employment status based on some of their demographic features. It is therefore of interest to construct a model which allows for the examination of employment rate individuals based on various demographic effects.

Specifically, this paper aims to examine the relationship between excessive alcohol consumption and employment rates of individuals using data from a longitudinal study conducted by the National Longitudinal Survey of Youth (NLSY). In the context of this paper, employment rates are defined as the rate of people who are employed or seeking employment as a proportion of the total number of observations. To accomplish this, a model will be constructed which models the effects of various frequencies of excessive alcohol consumption and the effect each level has on the probability of an individual being employed. This model will be contrasted with simpler models to better capture the effects of excessive alcohol consumption. Following the construction of the model, the model’s fit will be evaluated through a variety of basic methods to assess the accuracy of this model moving forward. Finally, future considerations on improvements to the model will be discussed.

**Data**

The data used in this paper comes from a longitudinal study conducted by the NLSY in both 1989 and 1994. Specifically, this paper focuses on data from 1994 in order to present a more modern model of the relationship between alcohol consumption and employment.

The primary labor variable of interest is labor force participation. Originally coded as a four-category variable (employed, unemployed, out of labor force, in active armed forces), the variable was condensed into a single, binary output variable. There were no observations in the “In Active Armed Forces” group, so the “Employed” (*emp*) dummy variable was coded as a 1 if the original employment status was employed or unemployed and a 0 if the original status was out of the labor force. This codification allows us to take into account natural unemployment and capture the probability that an individual is in the labor force (i.e., either employed or unemployed and seeking work) given their alcohol consumption tendencies. It should be noted that in the original documentation, there was no distinction made about whether “Unemployed” status meant unemployed, but still seeking labor, so the assumption was made that “Unemployed” encompassed people who were indeed seeking work and not out of the labor force.

The primary alcohol consumption variable was a six-category variable entitled “drnk6m” which indicated the number of times in the past month that an individual had 6 or more drinks (noted as excessive alcohol consumption) in one sitting. This variable was split into 7 dummy variables, each named drnk\_freqx where x took the values of 0 to 6. A value of 0 meant that an individual had not had 6 or more drinks in one sitting in the past month, a value of 1 indicated once, a value of 2 indicated 2 or 3 times, a value of 3 indicated 4 or 5 times, a value of 4 indicated 6 or 7 times, a value of 5 indicated 8 or 9 times, and a value of 6 indicated 10 or more times.

Individual characteristics captured in the regressions included the following characteristics: age (coded in years), gender (coded as a dummy variable where a male individual is recorded as 1), race (coded as three dummy variables where a value of 1 for *race1* indicates a Hispanic individual, a value of 1 for *race2* indicates a Black individual, and a value of 1 for *race3* indicates all other races), health (coded as a dummy variable where a value of 1 indicates that the individual has a health issue limiting the type or amount of work they can do), the number of years of education completed (coded as an integer from 0 to 20, inclusive), family size (coded as an integer), intelligence test percentile (coded as a value between 1 and 99, inclusive) and marital status (coded as three dummy variables where a value of 1 for *marst1* indicates an individual has never been married, a value of 1 for *marst2* indicates an individual is married with a spouse present, and a value of 1 for *marst3* encompasses all other marital statuses). Descriptive statistics of these variables can be found in **Table 3** in the appendix.

Geographic characteristics captured in the regressions included the following characteristics: Region (coded as a dummy variable where a value of 1 in *reg1* indicates an individual living in the Northeast, a value of 1 in *reg2* indicates an individual living in the North Central, a value of 1 in *reg3* indicates an individual living in the South, and a value of 1 in *reg4* indicates an individual living in the West), and local unemployment rate (coded as a midpoint of a range to avoid individual identification). Descriptive statistics of these variables can be found in **Table 4** in the appendix.

Family characteristics captured in the regressions included the following characteristics: Father’s level of education (coded as an interaction term which takes the value of 0 if the individual did not live with their father at age 14 or the number of years of education an individual’s father has if the individual lived with their father), mother’s level of education (coded as an interaction term which takes the value of 0 if the individual did not live with their mother at age 14 or the number of years of education an individual’s mother has if the individual lived with their mother), father’s working status (an interaction term taking the value of 0 if the individual did not live with their father at age 14 or 1 if the individual lived with their father and their father worked at at age 14), and mother’s working status (an interaction term taking the value of 0 if the individual did not live with their mother at age 14 or 1 if the individual lived with their mother and their mother worked at at age 14). Descriptive statistics of these variables can be found in **Table 5** in the appendix.

**Appendix**

**Table 1**. Descriptive statistics for labor force participation.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Variable** | **Mean** | **SD** | **Min.** | **Max.** |
| In Labor Force | 28.80908 | .362836 | 0 | 1 |

**Table 2**. Descriptive statistics for frequency of excessive alcohol consumption. Variables recorded as frequency of having six or more drinks in one sitting in the past month.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Variable** | **Mean** | **SD** | **Min.** | **Max.** |
| 0 times | .6670156 | .4713163 | 0 | 1 |
| 1 time | .0936005 | .2912939 | 0 | 1 |
| 2 or 3 times | .117823 | .3224225 | 0 | 1 |
| 4 or 5 times | .0506878 | .2193758 | 0 | 1 |
| 6 or 7 times | .0212321 | .1441678 | 0 | 1 |
| 8 or 9 times | .0098684 | .0988559 | 0 | 1 |
| 10 or more times | .0397727 | .1954394 | 0 | 1 |

**Table 3**. Descriptive statistics for individual characteristic variables used in the labor-alcohol model.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Variable** | **Mean** | **SD** | **Min.** | **Max.** |
| Age | 32.89653 | 2.220882 | 29 | 37 |
| Male | .4982057 | .5000342 | 0 | 1 |
| Black (*race1*) | .1848086 | .3881712 | 0 | 1 |
| Hispanic (*race2*) | .2873804 | .4525743 | 0 | 1 |
| Other Race (*race3*) | .527811 | .4992633 | 0 | 1 |
| Health | .0783493 | .2687405 | 0 | 1 |
| Family Size | 3.190789 | 1.604193 | 1 | 13 |
| Never Married | .2654007 | .4415794 | 0 | 1 |
| Married (with Spouse) | .5337919 | .4988941 | 0 | 1 |
| Intelligence Test Percentile | 40.63382 | 28.80908 | 1 | 99 |

**Table 4**. Descriptive statistics for geographic characteristic variables used in the labor-alcohol model.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Variable** | **Mean** | **SD** | **Min.** | **Max.** |
| Northeast | .1719498 | .377365 | 0 | 1 |
| North Central | .2444677 | .4298033 | 0 | 1 |
| South | .3779904 | .4849215 | 0 | 1 |
| West | .2055921 | .4041638 | 0 | 1 |
| Local Unemployment | 7.327303 | 2.763158 | 1.5 | 16.5 |

**Table 5**. Descriptive statistics for family characteristic variables used in the labor-alcohol model.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Variable** | **Mean** | **SD** | **Min.** | **Max.** |
| Dad’s Education | 7.529157 | 6.200652 | 0 | 20 |
| Mom’s Education | 9.76256 | 4.548345 | 0 | 20 |
| Dad working at 14 | .6695574 | .4704077 | 0 | 1 |
| Mom working at 14 | .5134569 | .4998562 | 0 | 1 |