Nils Berzins

ECON 339

Dr. Kuehn

December 2nd, 2022

Final Paper

Statistical Analysis of Alcohol Consumption’s Impacts on Labor Market Outcomes

**Introduction:**

If one were told to predict the productivity of working individuals given their drinking habits, they would likely construct a heuristic stating that the regular consumption of alcohol should indicate a heightened reluctance to work. This essay will analyze the impacts of regular alcohol consumption on labor market participation while taking into consideration demographic differences between surveyed individuals (dataset provided by the National Longitudinal Survey of Youth). More specifically, this essay will attempt to answer the question: Does regular alcohol consumption have an impact on an individual’s actual number of hours worked in a calendar year? The results indicate that regular monthly alcohol consumption is a significant predictor of how many hours are worked yearly, however its affects don’t point towards the anticipated direction. After accounting for race, sex, family income, marital status, whether or not the individual has children, and individual health, for every one day increase in number of days alcohol was consumed in the past month, the number of yearly hours worked increases by 2.055 hours.

**Data Description:**

Dependent Variable - Hours Worked (Yearly): Total number of hours worked in the past calendar year

Key Independent Variable - Days: Number of days in the last month that the individual has had at least 1 drink

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Variable** | **Mean** | **Median** | **Min** | **Max** | **2.5%** | **97.5%** |
| Hrswrk | 1689 | 2070 | 0 | 6240 | 1674.395 | 1704.39 |
| Days | 4.862 | 2 | 0 | 31 | .0003 | 4.109 |

Chart, scatter chart, box and whisker chart

Description automatically generated

**Empirical Approach:**

After accounting for relevant demographic variables, this regression model is trying to decide whether the number of days in the past month someone has consumed alcohol is a statistically significant predictor of the number of hours they work yearly. The demographic variables included were race, family income (transformed for easier comparison to the rest of the data, marital status, whether the individual was a mother/father (further explanation of this variable below), and whether the individual had a health condition that impeded their ability to work regularly. The number of hours worked yearly would be impacted on whether an individual had dependents in their life (spouse/kids) and/or if the individual had a health condition limiting their ability to work, making these obvious additions to the model. Race and family income, however, have more nuanced justifications for their inclusion in the model. Individuals who come from higher income families likely have an easier time finding work opportunities than those from lower income families. And the same logic applies behind individuals of different races: Black or Hispanic individuals likely have a harder time finding work opportunities than individuals with other racial backgrounds.

Demographic Variable Descriptions:

Race: Dummy Var, 1 = Hispanic, 2 = Black, 3 = Other

Log(Faminc + 1): Individual’s family income, log transformation applied to adjust for extremely high income families (typically making over $1,000,00 annually)

Marital Status: Whether individual is married, 1 = never married, 2 = currently married, 3 = other

MoFo: Combined variable, 0 = Man No Kid, 1 = Woman No Kid, 2 = Father, 3 = Mother. Assumption made that most individuals had only 1 child (median value of numkids was 1), so this was easily combined with binary sex variable.

Health: =1 if the individual has a health problem that limits amount of work that can be done.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Summary Statistics of Demographic Variables | | | | |
| **Variable** | **Mean** | **Median** | **Min** | **Max** |
| Race | 2.385 | 3 | 1 | 3 |
| Family Income | 45698 | 29000 | 0 | 1057448 |
| Log(faminc + 1) | 10.043 | 10.275 | 0 | 13.871 |
| Marital Status | 1 | 2 | 1 | 3 |
| MoFo | 1.388 | 1 | 0 | 3 |
| Health | .05765 | 0 | 0 | 1 |

Chart

Description automatically generated

**Estimation Results:**

|  |  |  |
| --- | --- | --- |
| 95% CIs on Demographic Variables | | |
| **Variable** | **2.5%** | **97.5%** |
| Intercept | -905.811 | -695.176 |
| Days | .0003 | 4.109 |
| Race | 22.377 | 58.740 |
| Log(faminc+1) | 190.325 | 208.793 |
| Marital Status | 6.093 | 47.782 |
| MoFo | 291.653 | 315.850 |
| Health | -905.811 | -695.176 |

Table

Description automatically generated

As seen in the confidence interval and dependent variable table, all variables included in the model are significant predictors of the number of hours worked annually by at least the .05 alpha level (including the key dependent variable days). However, the limited number of included predictors produced an adjusted R2 of .281, implying that only 28.1% of the data is explained by the proposed model. So, while I am confident that this model’s usage of monthly drinking frequency is statistically significant when predicting hours worked annually, I am weary to conclude that this is the most optimized model given the data. Furthermore, the statistical significance of the days variable most likely suffers from omitted variable bias.

The variable race sees that individuals who are Hispanic work 40.599 hours more than individuals who are black, while those who identified as “other” worked 40.599 hours more than those who are Hispanic. For every 1% increase in family income, the predicted number of hours worked increased by 199.559 hours. Individuals who had never married worked 26.938 hours less than those who are married and 53.876 hours less than those who identified their marital status as “other”. Men with no children worked 303.751 hours less than women with no children, 607.502 hours less than fathers, and 911.253 hours less than mothers. Those with medical conditions that inhibited their ability to work were predicted to work 434.696 hours less than those without medical conditions.

**Conclusion:**

For every additional day in a month where at least one alcoholic drink was consumed, the predicted number of hours worked annually increased by 2.055 hours. This may seem contradictory to our common perception of heavier drinkers being lazier people, but another explanation could prove useful here: overworked individuals could be using alcohol as a stress reliever or relaxant at the ends of their workdays/workweeks. Unfortunately, as mentioned in the estimation results section, the days variable most likely suffers from omitted variables bias and therefore may not actually produce a positive coefficient in a more powerful model.