THE UNIVERSITY OF NEW SOUTH WALES

School of Economics

ECON3208/5408, Applied Econometric Methods, 2020 T3

# **Project Report Cover Sheet**

|  |  |  |
| --- | --- | --- |
| Full Name | Student Number | Class Number |
| George Bai | Z5254903 | 7369 |

## 1 Introduction

As of the current day, excessive alcohol consumption still remains a significant public health problem within modern society, adding excess risk to consumers and whom they interact with. The US Department of Health and Human Services has estimated that by 1998, alcohol related mortality accounted for over 100000 deaths that year. Moreover, higher levels of alcohol consumption have been linked to social consequences such as greater domestic and criminal violence as it is believed to impair cognitive function and increase aggressiveness (US Department of Health and Human Services, 1993).

Here, we assess physician advice as a potential cost-effective mechanism in curbing excessive alcohol consumption and in doing so these health and social consequences. The literature indicates that within controlled environments, physician advice has demonstrated strong signs of effectiveness, and so should be studied further within real world environments.

As an example, a multinational trial was conducted by the World Health Organization (WHO) Brief Intervention Study Group (1996) and found that male patients subject to the physician advice drank approximately one standard drink less per day, representing approximately 17% less average daily alcohol consumption compared to those in the control group whom received no special treatment. Moreover, Fleming et al. (1997) found that twelve months after a physician advice intervention, problem drinkers that received physician advice had on average 4 less drinks within the last seven days compared to those with control-group interventions.

What we find is that the analysis requires dealing with major econometric issues. Here, we note that the dependent variable “drinks” must be non-negative and that there is a relatively large probability mass for no drinks within the last two weeks (21%). We also note that there may be potential endogeneity surrounding physician advice as the data does not capture health conscientiousness, [peer pressure] and other such factors affecting the probability of receiving advice and levels of alcohol consumption.

As such, we model alcohol consumption using a tobit model so as to accommodate the non-negative nature of the dependent variable as well as the large probability mass at zero drinks. The model produces non-negative fitted values which prove easier to interpret compared to an OLS model. It also helps provide a non-linear interpretation around the zero value. We also deal with endogeneity through a control function method for the tobit models, as well as instrumental variables in the 2SLS model.

We find that under the tobit model where the regression coefficients are estimated with the control function method, the presence of physician advice will have an average partial effect of increasing drinks by 12.59 compared to people who have not received such an intervention, holding all other factors fixed. However, this is a rather counterintuitive result, suggesting that the model has trouble correcting for endogeneity.

## 2 Data

Here, the data reflects 2467 responses from a 1990 National Health Interview survey core questionnaire and special supplements (Kenkel DS, Terza JV. 2001) . This is a sub-sample that consists solely of males whom drink currently and have reported being assessed as having high-blood pressure

Variable Definitions, Means, Standard Deviations and Key Features

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Variable | Definition | Mean | Std. Dev. | Key Features |
| Drinks | Number of alcoholic beverages drunk in the last two weeks | 14.70 | 22.75 | This is the dependent variable. It is non-negative, with 21.3% of the observations being 0. |
| Advice | Told by physician to drink less | 0.279 | 0.448 | This is a binary variable. |
| Editinc | Monthly income | 2.575 | 5.008 | Measured in American dollars ($1000) |
| Unemploy | If a person is unemployed | 0.666 | 0.168 | This is a binary variable. |
| Educ | Years of formal education | 12.93 | 3.087 | This is non-negative. |
| Hhrtcond | If they have a heart condition | 0.146 | 0.352 | This is a binary variable. |
| Hlthins | If they have health insurance | 0.815 | 0.389 | This is a binary variable. |
| Dr1 | They see the same doctor | 0.721 | 0.449 | This is a binary variable |

## 3 Conceptual Model

In our model of the consumer, they are a utility maximizing agent constrained by their budget, health and peers. Alcohol provides immediate utility from general enjoyment, the satisfaction of relieving stress, acquiescence to social pressure, and other such factors. Consumers balance this with the tradeoff from the disutility of the short and long health consequences as well as the potential social risks of drinking such as reduced inhibitions, amongst other costs.

As such, physician advice may be instrumental in allowing the consumer to reassess the health risks of alcohol consumption as much higher than before (Kenkel DS, Terza JV. 2001). Moreover, it helps provide a rationale that consumers can use to resist social pressures to drink, helping to reduce the disutility from failing to participate in social activities. It may also reduce the general enjoyment of drinking and change alcohol consumption from a stress-relieving mechanism to a stress-inducing mechanism as the physician advice acts as a persistent reminder of their moral and social failings. In totality, the consumer is expected to reduce their alcohol consumption as they adjust their drinking levels to reflect updated utility production parameters.

In determining the effects of physician advice on alcohol consumption, it is necessary to control for socio-economic factors such as income, education and employment amongst other factors. We control for income as lower levels of income may decrease a person’s willingness to seek advice due to the need to spend the income on other goods/services. It may also increase a person’s willingness to drink so to ease the pressure of day to day life. Moreover, unemployment should be controlled for as the unemployed may tend to be demotivated and so less willing to seek and hence receive advice. They may also drink more due to an excess of time and demotivation to care about long-term health effects due to their present-grounding crisis of unemployment. Education is controlled for as people that are more educated may be more willing to seek advice due to a better understanding of the value of that advice, and they may drink less as they understand the health risks better.

Finally, health factors must be controlled for as people with health conditions may be more likely to receive physician advice due to a potential heightened risk of excess alcohol consumption. Moreover, they may also be less likely to drink, due to this heightened risk.

However, there are some elements in the demand function for alcohol that cannot be controlled for due to an absence of data/proxy variables. Hence, a potential source of problems lies in how physician advice may be correlated with the random error term, generating endogeneity issues within methods of estimation such as OLS.

As one example, it is plausible that physician advice may be correlated with a higher willingness to seek that advice from people who are more health conscientious, who simultaneously are less willing to drink for that same reason (Kenkel DS, Terza JV. 2001). We also note that peer pressure may cause people to drink more than they ordinarily would, and it may also be that same peer pressure that normalizes their levels of alcohol consumption such that they resist seeking and receiving advice.

## 4 Econometric Models

Initially the simplest possible linear model: is estimated using OLS in order to assess the correlation between Drinks and Advice, and so as to help in interpreting the effects of controlling for further variables

The main econometric model we will inspect is the linear model:

In this model and the model that follows, the socio-economic factors and health factor are controlled for due to the reasons posed in the theory. Some factors such as age or race are not included in the model as we do not believe that these factors affect levels of alcohol consumption, except possibly for the very old or insofar as race is correlated with some culture that prioritizes drinking.

The first way this model is estimated is through OLS. Here, the key assumption made is that the error term is uncorrelated with the explanatory variables, such that the zero conditional mean assumption holds true (Woolridge JM. 2013).

This is likely false as the data does not contain observations regarding the influence of peer pressure in increasing the amount of alcohol consumed, as well as potentially reducing the rate at which advice is sought due to a normalization of the consumer’s alcoholic consumption levels. Moreover, it doesn’t take into account health conscientiousness, where people who are more concerned about health tend to seek advice more while simultaneously tending to drink less. As such, there is the possibility that these effects are falsely attributed to physician advice.

There is also the possibility of measurement error within the model. Although there may be little measurement errors for those that haven’t drunk within the last two weeks, it is possible that people who receive advice tend to report drinking less than their true level due to a fear of being looked down upon, social stigma, shame, or other such factors. Moreover, there is the issue of simultaneity. It is extremely plausible that for example, as people who drink more may be more likely to receive advice to pause drinking due to the higher-health concerns.

As a result, it does not appear that the basic OLS model will be effective due to the presence of endogeneity. Still, we utilize it here as it serves as a linear approximation of the true model and is thus helpful in enabling comparisons to be made across models. Moreover, if it does so happen that Advice is exogenous, then OLS will be consistent. We assume it is likely that the socio-economic factors and health factor chosen in the model are minimally endogenous as they may be unrelated to peer pressure or health conscientiousness. It may be strange to consider there to be little correlation with health conscientiousness, but seeking advice is a passive activity compared to say exercise, which would actually correlate with Hhrtcond for example.

In an attempt to control for endogeneity, we again use the linear model, using Hlthins and Dr1 as the instruments. Here, the model is estimated through 2SLS. The main assumption is that Hlthins is a valid instrumental variable. It should be uncorrelated with the error term as there should be no relation to the levels of peer-pressure an individual experiences. It should also not be correlated with health conscientiousness as health insurance is a form of risk-management where the insurance company charges different premiums based on a person’s level of health risk, and so the price discrimination balances out the correlation between health conscientiousness and better health. It should also be uncorrelated with any measurement error. Health insurance should also be correlated with advice negatively, as health insurance prices are generally discriminatory based on health status (lower for people in better health). As such, people who do not have health insurance are more likely to be in poor health, which would likely be correlated with requiring higher levels of advice. However, there is indeed the possibility that health insurance makes it easier to seek and receive advice, and so correlated positively. Still, we believe the negative correlation is stronger as the price discrimination affects all people that have health insurance.

We also assume that Dr1 (whether a person sees the same doctor) is correlated with advice positively insofar as that when familiarity with a doctor is achieved, the doctor is more likely to give advice as they understand the patient’s situation better. It should also be uncorrelated with the error term as the choice of doctor is often forced on a patient due to either geographic or time constraints (the doctor may be seeing other patients) by the medical facility, so peer pressure, health conscientiousness, etc are unrelated factors, and by similar reasoning it should not be correlated with drinking.

Hence, the three basic requirements for IV estimation are likely satisfied – that of instrument exogeneity, instrument relevance and instrument exclusion.

One problem with constructing alcohol demand as a linear model is that we assume that the dependent variable can be any real value, and so it gives no guarantees that the fitted values for “drinks” are non-negative, which may cause interpretation issues (Woolridge JM. 2013). Moreover, at the corner solution, there likely exists a large probability mass of zero drinks within the sample. In order to address this, a tobit model is used where we find the observed response Drinks to be: . This constrains the dependent variable to be non-negative and helps to enable a non-linear interpretation around the corner solution.

However, for this to be consistent, we must assume that there is no endogeneity, which as explained in the OLS section is unlikely. Hence, in order to correct for it, the control function method is used. We use the probit model , where to estimate the residual , which may estimate Advice better by constraining it within the interval [0, 1]. Then, estimating the tobit model , allows the coefficients to be consistently estimated if the model is correct

However, that the standard errors from the Tobit model estimation are incorrect as they are based on the probit model estimates. We assume that they are good approximations for the correct standard error, but this is nevertheless an issue.

## 5 Empirical Results

Here, we present the estimation results for our empirical models, with coefficients estimated using OLS, 2SLS, a basic tobit model and a tobit model that utilizes the control function method.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Independent Variables | Simple Model (OLS) | Linear (OLS) | 2SLS with instrumental variable for advice (hlthins) | Tobit | Tobit with control function method |
| Advice | 7.672  (1.010) | 7.621  (1.168) | 21.32  (16.82) | 9.515  (1.370) | 19.20  (19.29) |
| Editinc |  | 0.094  (0.062) | 0.108  (0.067) | 0.133  (0.088) | 0.142  (0.089) |
| Unemploy |  | 9.480  (4.539) | 8.465  (4.872) | 10.85  (4.969) | 10.13  (5.28) |
| Educ |  | -0.227  (0.147) | -0.014  (0.295) | 0.0006  (0.183) | 0.152  (0.345) |
| Hhrtcond |  | -4.779  (1.090) | -5.638  (1.667) | -7.014  (1.497) | -7.621  (2.04) |
| Constant | 12.56  (0.533) | 15.69  (2.109) | 9.231  (7.973) | 8.547  (2.540) | 3.976  (9.254) |
| v |  |  |  |  | -9.737  (19.21) |
| R-squared | 0.0229 | 0.0344 |  | 0.0337 | 0.034 |
|  | 22.50 | 22.38 | 23.169 | 26.51 | 26.50 |

When using OLS on the model , we find that the coefficients for Advice and Editinc are positive when we previously theorized them to be negative.

We note that Advice is both counter-intuitively estimated to be positive with a large magnitude in a statistically significant manner, where in using a t-test to test the null hypothesis against the alternate hypothesis , we reject the null hypothesis at the 5% significance level. This is highly suggestive of endogeneity, as some omitted variable, e.g. addictive personality, may be may be correlated with requiring more advice and simultaneously a desire to drink more. As such, OLS may be an inappropriate model due to the ZCM assumption being untenable.

At the very least however, we surprisingly find that when OLS estimation is used on , that we find that using a F-test, we fail to reject the null hypothesis at the 35% significance level, indicating that even if Editinc and Unemploy are endogenous, they would not bias advice due to a lack of correlation.

In our assumption for the 2SLS model, we assume that the instruments are exogenous. However, as there are more instruments than endogenous regressors, we can test whether at least one may be endogenous. We find that when regressing the 2SLS residuals on all exogenous variables and then multiplying by the sample size, that and there is a p value of 0.0006 when comparing with the Chi-squared distribution, rejecting the null hypothesis that all instruments are exogenous.

As such, we drop dr1 from the equation as we believe that the economic theory is weaker for it. It is possible that someone who shops around for doctors is more health-conscientious for example. Hence, it is removed in order to prevent bias and inconsistency.

Now, we also check for instrument relevance for hlthins. We use OLS estimation on , Using a t-test to test the null hypothesis against the alternate hypothesis , we find that the t value is -3.58. As such, we reject the null hypothesis that Hlthins is exogenous at the 1% significance level, and so it is likely correlated with Advice.

Saving the residuals from this regression as , we further test that Advice is endogenous through another OLS regression on:

Using the residual based-test for endogeneity, we test the null hypothesis against the alternate hypothesis . Here, we find that the t-value is -0.85 and so we fail to reject the null hypothesis at the 10% significance level. However, as the 2SLS estimates for Advice are vastly different from the estimates using OLS, we treat Advice as endogenous.

As such, the 2SLS assumptions of an endogenous variable and instrument relevance is satiated, though we are unable to determine if there is instrument exogeneity. There, we must rely on economic theory.

However, when we control for endogeneity, we find that the magnitude of Advice becomes even greater than before, counter to expectations. However, the standard error also increases, meaning that we fail to reject the null hypothesis that the coefficient on advice is 0 at the 20% significance level.

When we further control for the probability mass at zero as well as the non-negative nature of drinks alongside endogeneity through a tobit model, we find that the average partial effect of advice on drinks is that receiving advice correlates with 12.59 extra drinks consumed within the last two weeks. This is less than the 2SLS effects but more than the OLS estimates, implying that accounting for endogeneity actually further supports the counterintuitive result.

Comparing across all the results, we find that the signs and magnitudes of the estimates for advice are robust to the variations in assumptions surrounding dealing with endogeneity and data features , with a strong positive correlation between receiving advice and drinking more. These results are likely too positive to be credible, going against previous literature.

## 6 Conclusion

Here, we estimate the effects of physician advice on alcohol consumption to be positive and large in magnitude, where our tobit model indicates that receiving advice has an average partial effect of increasing alcohol consumption by 12.59 drinks compared to people who have not received such an intervention. This is robust to changes in our assumptions and models, however we suspect that there may be a problem with accounting for endogeneity as these results go against other literature on the topic.

Although these results goes against previous literature, it may be possible that the real-world effects of advice are different than in controlled experiments. In the unlikely event that further evidence supports the idea that physician advice actually increases alcohol consumption, this may have the policy implication that physicians should rely less on granting advice and possibly more on intrusive methods such as counseling. In order to control alcohol consumption, it may also be more cost-effective to rely on methods to change the country’s culture such that people view alcohol poorly, akin to smoking. This may be done with advertisements associating alcohol with a lower-class lifestyle for example.

There are issues that are not resolved within this paper. As the instrumental variable is suspect, we would like to collect data on the effects of peer-pressure on alcohol consumption using proxy variables such as self-assessed ordinal data on the frequency with which other people influence their levels of alcohol consumption. Moreover, we desire data on the effects of health-conscientiousness on alcohol consumption. It may be possible to collect data on a person’s average level of exercise per week in order to act as another a proxy variable.

## 7 References

Fleming MF, Lawton Barry K, Baier Manwell L, Johnson K, London R. 1997. Brief physician advice for problem alcohol drinkers: a randomized controlled trial in community-based primary care practices. *JAMA* **277**: 1039–1045.

Kenkel DS, Terza JV. 2001. The effect of physician advice on alcohol consumption: count regression with an endogenous treatment effect. *Journal of Applied Econometrics* **16**: 165-184

US Department of Health and Human Services. 1993. Eighth Special Report to the U.S. Congress on Alcohol and Health.

World Health Organization Brief Intervention Study Group. 1996. A cross-national trial of brief interventions with heavy drinkers. *American Journal of Public Health* **86**(7): 948–955.

Woolridge JM. 2013. *Introductory Econometrics A Modern Approach.* 5th Edition. Cengage Learning: Mason