CHAPTER 5

Reinvestigating the Effect of Alcohol Consumption on Hypertension Disease

This chapter is developed from the original paper "Reinvestigating the Effect of Alcohol Consumption on Hypertension Disease" by Suknark, Sirisrisakulchai and Sriboonchitta (2015) presented at the 8th Conference of the Thailand Econometric Society, published in "Causal Inference in Econometrics", Springer International Publishing Switzerland. This paper can be found in the appendix C.

5.1. Introduction

Hypertension is a major risk factor for cardiovascular disease. Tobacco consumption and alcohol consumption are the most important avoidable causes of cardiovascular diseases worldwide (Pechacek, et al, 2003). In Thailand, cardiovascular disease and hypertension are major health problems. The cardiovascular morbidity rate and Hypertension have been reported as an important cause of morbidity for past several years (Antonakis & Lalive, 2011) The cardiovascular disease was ranked fourth largest cause of death in 2012 (32.9 people per 100,000) (Fuchs et al, 2001) Hypertension was the third largest cause of death (37.4 people per 100,000). Both causes are increasing in frequency (The permanent Secretary Ministry of Thailand, 2012). Table 5.1 presents the causes of death in Thailand.

Table 5.1 Causes of death in Thailand (per 100,000 inhabitants)

Cause groups	2008	2009	2010	2011	2012
1. Malignant neoplasm, all forms	87.6	86.3	91.2	95.2	98.5
2. Accident and poisonings	55.1	55.6	51.6	52.8	51.6
3. Hypertension and cerebrovascular disease	24.7	24.7	31.4-		37.4
4. Disease of the heart	29.8	29.0	28.9	31.4	32.9
5. Pneumonia and other diseases of lung	23.0	22.9	25.7	26.3	26.1

Source: Bureau of Policy and Strategy, Office of the Permanent Secretary, Ministry of Public Health

Epidemiological and experimental investigations have established a close association between alcohol consumption and hypertension (Klatsky, 1996). A number of population studies have almost unanimously shown an empirical link between high levels of alcohol consumption and hypertension. In many former studies, risk factors were assessed using multiple regression models; i.e. age , gender, family history , weight and height for hypertension (Puddey, 1985), (Fuchs et al ,2001), (Xin, et al, 2001), (Mlunde, 2007), (Yadav et.l, 2008), (Venkataraman, 2013) and (Oyunbileg, 2015).

In 1985, Puddey et al studied the direct effects of alcohol consumption (Puddey, 1985). This chapter received reponse from 46 male drinkers (22-55 years) to a questionaire. Similar research has been conducted by Taraman, Yadav and others (Venkataraman, 2013) and (Yadav, 2008). (Venkataraman, 2013) They studied both tobacco and alcohol consumption among people in Southern India. The results suggest that hypertension becomes more frequent among women of advancing age who are alcohol users, especially if they smoke (Yadav, 2008). Yadav studied the prevalence of hypertension in northern India among subjects of approximately 30 years of age. The results indicate that increasing age, body mass index, obesity and impaired glucose tolerance were significantly associated with hypertension (Yadav, 2008). They all used regression analysis in their research. Oyunbileg used regressive methods to assess the correlation between alcohol consumption and arterial hypertension among the population of the Gobi region of China (Oyunbileg, 2015).

In Thailand, multiple regression modeling research shows an empirical link between high levels of alcohol consumption and hypertension. Leelarassme et al studied a sample in Bangkok (Leelarassme, et al, 1978). Pati and Siviroj studied a sample in Northern Thailand (Pati, 2004) and (Siviroj et al, 2013). Howteerakul studied people living in the rural areas in Thailand (Howteerakul et al, 2006). In these studies the researchers used data collected in a national survey (Pati, 2004) and (Siviroj et al, 2013).

In this chapter, the researchers reinvestigate the effect of alcohol consumption on the levels of hypertension. Moreover, the researchers investigate whether alcohol consumption may be a cause of hypertension by introducing the switching regression model for observational data analysis. This investigation of causality has been done using the Neyman-Rubin counterfactual framework (Guo & Fraser, 2010). In this framework, groups were compared to extract the differences that are due to changes in the level of alcohol consumption and the affect on levels of hypertension.

Ideally, we just want to compare two groups that only differ with regard to whether they consumed alcohol or not. Two groups should be identical in all relevant characteristics, so that we can deduce changes due to the effect of alcohol consumption on hypertension level. This idea can be done in practice by randomly assigning individuals to each group by using the same probability of being assigned to either the control or treatment groups. The characteristics of the individuals in each group may be viewed as interchangeable and are the same, at least on average, if we have a large enough sample.

The potential outcomes framework (Morgan & Winship, 2007) uses a what-if scenario as the baseline for making causal inferences. This baseline allows each individual the same chance for potential outcomes. However, we can only observe the individual in one treatment condition. Under the interchangeable assumption and given that each group only differs with regard to the treatment condition, the causal effect can then be statistically estimated as the difference in the means of an outcome of interest between the control and treatment groups (Antonakis & Lalive, 2011).

To consistently estimate the average treatment effect, we have to satisfy the ignorable treatment assignment assumption (Rosenbaum & Rubin, 1983). This condition ensures that the outcome of interest is independent of the treatment assignment mechanism (Morgan & Winship, 2007). This assumption can be satisfied when we assign each individual randomly to the treatment and control groups. In other words, we can only consistently estimate the average treatment effect if there is no selection bias in the treatment assignment.

In our study, the survey participants self-selected themselves into treatment condition (whether to drink alcohol or not), thus it is hard to believe that the independence assumption holds in this situation. In the observed data, the treatment assignment is not ignorable, thus using treatment as a dummy variable in the statistical model will lead to the bias estimation of the treatment effect. Factors affecting the self-selection might cause the dummy variable of treatment condition to be correlated with the random errors in the outcome model, which leads to the biased parameters estimation. By introducing the endogenous switching regression model for ordered outcomes, the new treatment effect estimation should give us more accurate and reliable results for causal inference.

5.2. Data

This study uses the Thai National Health Examination Survey, No.4 (NHES IV) data of 2009. The ordered dependent variables are the blood pressure levels of participants. Blood pressure levels may be classified into three groups of individuals as follows (Thailand context). First, those with an average diastolic blood pressure (DBP) and systolic blood pressure (SBP) of < 80/120 mmHg as normal blood pressure level; second, those with DBP and SBP of 80-90/120-140 mmHg as pre-hypertension level; and third those with DBP and SBP of greater than 90/140 mmHg as regarded as having hypertension. The independent variables are gender, age, income, chronic diseases (if any), marital status, level of education, and occupation. Table 5.2 presents a description of the variables and related statistics.

Table 5.2 Description of variables and statistics

Variables	Description	N	Mean	SD	Min	Max
BP	Level of blood presure	20450	0.446	0.697	0	2
Gender	1 if individual is male; 0 otherwise	20450	0.524	0.499	0	1
Age	In Year	20450	52.917	18.236	14	98
Income	In 1,000 Baht	20450	3.310	5.698	0	32,480
Bachelor	1 if individual graduated from Bachalor degree or higher; 0 otherwise	20450	0.061	0.24	0	1
Agr	1 if individual works in agricultural sector; 0 otherwise	20450	0.176	0.381	0	1
Whi	1 if individual is white-collar worker	20450	0.035	0.184	0	1
Police	1 if individual works as police or soldier; 0 otherwise	20450	0.012	0.108	0	1
Labor	1 if individual is in labor sector; 0 otherwise	20450	0.480	0.499	0	1
Married	Marital status where 1 indicates married; 0 otherwise	20450	0.636	0.481	0	1
Pe_bmi25	1 if individual has body mass index more than 25; 0 otherwise	20450	0.348	0.476	0	1
Pe_tc200	1 if individual has chloresterol level more than 200; 0 otherwise	20450	0.561	0.496	0	1
Qlhealth	Self health quality assessment, where 5 is the highest level	20450	3.708	0.867	0	5
NCD	Number of chronic diseases	20450	0.632	0.959	0	10

5.3 Empirical Results

A total of three models were estimated using the Independence copula, the Normal copula, and the Frank copula. We selected the best fitted model based on Akaike Information Criteria (AIC), which is the Frank copula model. Table 5.4 shows the log-likelihood values for the Independence copula and the Frank copula models. A likelihood ratio test rejects the Independence copula model in the Frank copula model. Therefore, the results provided here are only from the Independence copula and the Frank copula models. In the following subsection, we will discuss the results from the Frank copula model for the policy implications.

5.3.1 Binary Choice Equation for Alcohol Consumption

Table 5.3 gives the results of the selection equation. The results of the binary outcome equation of self-selected alcohol consumption provide the effects of the variable on the propensity toward alcohol consumption relative to non-alcohol consumption. All the parameter estimates were statistically significant at the standard level. The coefficient interpretations are: 1) young individuals, individuals income, individuals who are married are more likely to consume alcohol; 2) males are more likely to use alcohol than females; 3) individuals who work in the agricultural sector and work in high risk occupations such as the police and military are more likely to use alcohol than white-collar workers and those in the labor sector.

Table 5.3 Estimation results of selection equation for alcohol consumption

Variables	Independent		Frank	
	Coeff.	Std.err	Coeff.	Std.err
Selection equation				
Intercept	-0.347	0.050	-0.343	0.050
Age	-0.014	0.001	-0.014	0.001
Income	1.34E-05	1.88E-06	1.36E-05	1.88E-06
Sex	0.946	0.02	0.946	0.020
Agr	0.345	0.032	0.347	0.032
Whi	0.200	0.061	0.203	0.061
Police	0.315	0.088	0.317	0.088
Labor	0.181	0.029	0.184	0.029
Bachelor	-0.044	0.044	-0.042	0.045
Married	0.048	0.022	0.048	0.022

5.3.2 Factors Affecting Hypertension level for Non-alcohol Users

Table 5.4 presents the model estimation results of hypertension levels for non-alcohol users. The significant explanatory variables included in the model are age, income, high cholesterol, gender, non-communicable diseases, occupation, education, and Body Mass Index. The coefficient interpretations are: 1) older individuals, individuals with lower income, individuals who have higher cholesterol 200 mg/dl, or the number of chronic diseases, or obese (BMI < 25) individuals, or individuals with education lower than bachelor degree level are more likely to develop hypertension; 2) females are more likely to have hypertension than males; 3) individuals who work in the agricultural sector are more likely to have hypertension.

5.3.3 Factors Affecting Hypertension level for Alcohol Users

Table 5.4 also presents the model estimation results of hypertension levels for alcohol users. The explanatory variables included in the model that are significant are age, income, high cholesterol, gender, non-communicable diseases, occupation, education, and Body Mass Index (BMI). The coefficient interpretations are: 1) older individuals, individuals who have a higher health quality assessment, individuals who have higher cholesterol 200 mg/dl, or higher numbers of chronic diseases, or individuals who are obese (BMI < 25), individuals whose educational levels are lower than bachelor degree, and individuals who are married are more likely to have hypertension; 2) males are more likely to hypertension than females; 3) individuals who work in the agricultural sector are more likely to have hypertension.

Table 5.4 Estimation results of blood pressure level equations

Regime	Independent				Frank			
	Alcohol user		Non-alcohol user		Alcohol user		Non-alcohol user	
Variables	Coeff.	Std.err	Coeff.	Std.err	Coeff.	Std.err	Coeff.	Std.err
Intercept	-1.696	0.075	-1.718	0.100	-1.838	0.073	-1.701	0.126
Age	0.026	0.001	0.022	0.001	0.027	0.001	0.022	0.001
Income	-1.68E-06	2.38E-06	4.26E-06	2.44E-06	-4.91E-06	2.28E-06	4.12E-06	2.51E-06
Sex	0.198	0.023	0.538	0.032	-0.138	0.046	0.528	0.051
Agr	-0.010	0.035	-0.040	0.047	-0.083	0.036	-0.044	0.051
Whi	-0.066	0.075	0.033	0.086	-0.102	0.072	0.031	0.087
Police	-0.038	0.126	0.041	0.108	-0.112	0.115	0.038	0.109
Labor	0.080	0.029	0.124	0.044	0.039	0.029	0.122	0.045
Bachelor	-0.176	0.053	-0.195	0.061	-0.156	0.051	-0.195	0.062
Married	-0.032	0.022	-0.085	0.032	-0.042	0.022	-0.085	0.032
Qlhealth	0.006	0.012	-0.040	0.017	0.006	0.011	0.040	0.017
NCD	0.149	0.011	0.160	0.019	0.143	0.011	0.160	0.019
pe_bmi25	0.481	0.022	0.511	0.031	0.460	0.022	0.512	0.031
Pe_tc200	0.252	0.022	0.250	0.028	0.236	0.021	0.245	0.028
$k_{s,1}$	10.28	0.014	1.045	0.018	0.976	0.022	1.045	0.018
θ					-2.583	0.588	-0.097	0.412
LL	-30334.15				-30324.16			

5.3.4 Effect of Alcohol Consumption on Blood Pressure Level

From Table 5.4, the dependence parameter θ tells us about the direction of self-selection biases. The t-test is used for hypothesis testing. The null hypothesis that $\theta = 0$ implies that there is no self-selection bias. If the null is rejected, the quantification of the selection effects can be computed by comparing the outcome distribution of $Pr(Y_0 = j|S=1)$ with the counterfactual predicted distribution $Pr(Y_0=j|S=0)$ of an individual who chooses to consume alcohol but is hypothetically allocated to non-alcohol user regime (Luechinger, 2010).

The parameter θ for alcohol user regime is negative and significant. This indicates that the two random errors (ε_0 and ν) tend to move in the opposite direction. The negative correlation means that the alcohol user counterfactual blood level of those who actually chose not to drink are below than that of an average. For the non-alcohol user regime, the dependence parameter is not significant, indicating that the blood pressure level distribution of those who are non-alcohol users do not differ from the distribution of an arbitrary individual with the same characteristics.

To quantify the effect of alcohol consumption on the probability of each level of blood pressure, we compute the average of predicted probabilities of each level of blood pressure for alcohol user and non-alcohol user regimes. Table 5.5 shows the results of the mean of predicted probabilities of outcome within the sample. The results indicate that alcohol consumption is more likely to lead to higher level of blood pressure and hypertension disease.

Table 5.5 Predicted probabilities of blood pressure level within sample for the Frank copula model

Outcome	Mean of predicted probabilities						
	Sober			Drink			Difference
	Min	Mean	Max	Min	Mean	Max	of Mean
Normal	0.035	0.52	0.942	0.009	0.404	0.907	0.116
Moderate	0.053	0.289	0.375	0.084	0.34	0.399	-0.051
High	0.005	0.191	0.797	0.009	0.256	0.904	-0.065

5.4 Concluding Remarks

This chapter applied a copula-based endogenous switching regression for ordinal outcomes to examine the effect of alcohol consumption on levels of hypertension, using the data from National Health Examination Survey in 2009. We present the Frank copula and Independence copula models in this paper. We found statistical evidence for positive self-selection on alcohol users.

From the empirical results previously discussed, the following are the recommended policy designs to reduce the levels of hypertension in alcohol users. In non alcohol users, other policies may help reduce levels of hypertension:

a) The protection and prevention program for hypertension for non-alcohol users, should focus more on the needs of women, individuals from lower income groups and those with lower educational levels.

b) For alcohol users, the protection and prevention program for hypertension should be focus more on male alcohol users, undergraduates, and manual workers.

c) The high risk group includes individuals who are older, have higher numbers of chronic diseases, high cholesterol, and who may be obese. These people should receive regular follow-up medical examinations and take appropriate measures, including life-style changes to prevent hypertension.