

# ALCOHOL CONSUMPTION IN THAILAND: A STUDY OF THE ASSOCIATIONS BETWEEN ALCOHOL, TOBACCO, GAMBLING, AND DEMOGRAPHIC FACTORS

Pannapa Changpetch<sup>1</sup>, Dominique Haughton<sup>2</sup>, Mai T. X. Le<sup>3</sup>,  
Sel Ly<sup>4</sup>, Phong Nguyen<sup>5</sup>, Tien T. Thach<sup>6</sup>

<sup>1</sup> Assistant Professor, Department of Mathematical Sciences, Bentley University, MA, <sup>2</sup> Professor, Department of Mathematical Sciences, Bentley University, MA, USA; <sup>2</sup> Affiliated Researcher, Paris 1 University (SAMM), Paris, <sup>2</sup> Affiliated Researcher, Toulouse 1 University (TSER), Toulouse, FRANCE; <sup>3</sup> Assistant Professor, Faculty of Mathematics-Informatics, The University of Sciences, Ho Chi Minh City, <sup>4</sup> Lecturer, Faculty of Mathematics and Statistics, Ton Duc Thang University, Ho Chi Minh City, <sup>5</sup> Director, General Statistics Office, Hanoi, <sup>6</sup> Lecturer, Faculty of Mathematics and Statistics, Ton Duc Thang University, Ho Chi Minh City, VIETNAM.

## ABSTRACT

*This paper provides a thorough study of alcohol consumption in Thailand in terms of the relationships between this activity and tobacco consumption, gambling consumption, and demographic factors. Logistic regression and Treenet are used to analyze datasets drawn from a socio-economic survey of 43,830 Thai households conducted in 2009. Using logistic regression to analyze the likelihood of alcohol consumption, we found that household tobacco expenditure, the region in which the household is located, household gambling expenditure, household income, and the educational level, religion, sex, age, work status, and marital status of the head of household are all associated with a household's alcohol consumption. The strongest predictors of household alcohol consumption are tobacco expenditure, the religion and sex of the head of household. Using Treenet to analyze household expenditure on alcohol, we found that both the proportion of household expenditure spent on tobacco and the proportion of household expenditure spent on gambling are important factors. The results of our study show that alcohol, tobacco, and gambling co-occur. Given these relationships, programs that successfully target households with alcohol-dependent members will not only reduce alcohol consumption in Thailand but will reduce tobacco consumption and gambling consumption likewise.*

**Keywords:** Alcohol consumption; Tobacco consumption; Gambling consumption; Logistic regression; Treenet

## INTRODUCTION

### Alcohol Consumption in Thailand

Alcohol has long been a major problem in Thai society and has become more problematic overtime. In particular, excessive alcohol consumption is related to the leading causes of death in Thailand, including malignant neoplasm, heart disease, and hypertension with cerebrovascular disease [1]. According to the latest alcohol-consumption data collected by the World Health Organization (WHO), Thailand is ranked first among ASEAN countries for alcohol consumption followed closely by Laos and the Philippines.

In the literature on alcohol consumption, several articles focused on the relationship between alcohol and tobacco show that alcohol use and smoking frequently co-occur [2,3,4,5,6]. Aekplakorn et al. (2008) [7] investigated the association between the prevalence of cigarette smoking, the prevalence of alcohol consumption, and socioeconomic factors in Thailand through a logistic regression analysis using a nationally representative cross-sectional survey

in 2004 of 39,290 individuals aged 15 and over. The results show that the strongest predictor of alcohol misuse is smoking and vice versa.

A number of studies investigate the associations between alcohol and gambling [8, 9] and found that gambling and alcohol consumption are related. In a review of studies on the associations between gambling and the use of alcohol, tobacco, and illicit drugs among US youth between 2000-2014 [10], most of the studies included found gambling to be associated with the use of these substances. The methods used in most of the studies are bivariate analysis, multiple linear regression, and logistic regression analysis.

This paper provides a thorough study of alcohol consumption in Thailand by exploring relationships between alcohol consumption, tobacco consumption, gambling consumption, and demographic factors. We apply two techniques to analyze datasets drawn from a socio-economic survey of 43,830 Thai households conducted in 2009. In addition to logistic regression analysis, the method most commonly used in the literature on alcohol, we also implement a new method that has never been used before in this context, i.e., Treenet analysis. The objective of Treenet is to obtain a more precise understanding of any non-linear relationship between response and predictors.

Of the two techniques used here—logistic regression analysis and Treenet—we began by using logistic regression analysis to investigate the association between the likelihood of alcohol consumption in relation to demographic factors, household tobacco expenditure, and household gambling expenditure. We studied at-home alcohol consumption and away-from-home alcohol consumption separately; as we considered it likely that the relationships between the factors implicated in alcohol consumption would differ between these two types of consumption.

For the second analysis (Treenet), we investigated the association between the proportion of household expenditure spent on alcohol and demographic factors, the proportion spent on tobacco, and the proportion spent on gambling. For this analysis, we built two Treenet models to separately capture the non-linear dependence of the proportion of alcohol consumed at home and the proportion consumed away from home on the predictors. Note that in Treenet, all expenditure proportions are computed relative to total household expenditure. The main advantages of this method are that it reveals non-linear relationships between response and predictors and that the non-parametric approach adopted here makes it possible to handle a response variable with a large number of zero values.

Overall, we expect our results to be useful to both researchers and government practitioners in their efforts to tailor programs that target households with alcohol-dependent members in order to reduce alcohol consumption in Thailand.

This paper is organized as follows. Section 2 reviews the dataset used in this study. Section 3 presents the logistic regression analysis. Section 4 presents the Treenet analysis. Section 5 offers a discussion and concluding remarks.

## **Dataset**

We used a dataset collected via a socio-economic survey of Thai households conducted in 2009. Of the 43,830 households included in the survey, 6,857 consumed alcohol at home and 4,255 consumed alcohol away from home, representing 15.64% and 9.7% of the full sample, respectively. Note that the original data comprise 43,844 household records. However, there are 14 records with missing values so we decided to omit these records from our analysis. The factors included in our analyses are shown in Table 1. A stratified two-stage sampling was implemented for the survey. The primary sampling units were blocks for municipal areas

and were villages for non-municipal areas. The secondary sampling units were private households.

**Table 1. Factors of Interest**

| <i>Predictor</i>            | <i>Details for each categorical variable</i>  |
|-----------------------------|---|
| Region                      | Note: Region of household<br>1. Bangkok Metropolis (6.2%),2. Central (excluding Bangkok) (29.4%),3. North (24.4%),4. Northeast (25.7%),5. South (14.4%)   |
| Area                        | Note: Area of household<br>1. Municipal area (61.7%),2. Non-municipal area (38.3%)  |
| Number of household members | Note: Number of members in household<br>min = 1, median = 3, max = 17, mean = 3.18, standard deviation = 1.63   |
| Income                      | Note: Average monthly total income per household (Thai Baht)<br>min = -103,988, median = 14,424, max= 2,821,572, mean = 22,392, standard deviation = 38,062   |
| Sex                         | Note: Sex of head of household<br>1. Male (64.8%),2. Female (35.2%)   |
| Age                         | Note: Age of head of household (years)<br>min = 15, median = 51, max= 99, mean = 51.70, standard deviation = 14.76  |
| Marital status              | Note: Marital status of head of household<br>1. Single (8.9%),2. Married (68.5%),3. Widowed (16.7%),4. Other (6.0%)   |
| Religion                    | Note: Religion of head of household<br>1. Buddhist (94.9%),2. Islamic(4.3%),3. Christian and other (0.8%)   |
| Disability                  | Note: Whether head of household is disabled<br>0. No (97.5%), 1. Yes (2.5%)   |
| Welfare                     | Note: Whether head of household receives welfare or medical services<br>0. No (2.0%), 1. Yes (98%)  |
| Gambling expenditure        | Note: Average monthly expenditure on lottery tickets and other kinds of gambling per household (Thai Baht)<br>min = 0, median = 0, max= 23,833, mean = 160.5, standard deviation = 508.5  |
| Tobacco expenditure         | Note: Average monthly expenditure on tobacco products per household (Thai Baht)<br>min = 0, median = 0, max= 14,964, mean = 112.2, standard deviation = 316.5   |
| Amount debt                 | Note: Total debt at end of previous month<br>min = 0, median = 10,000, max= 57,000,000, mean = 155,045, standard deviation = 616,968  |
| Government fund             | Note: Whether head of household borrowed money from a government fund<br>0. No (84.1%), 1. Yes (15.9%)  |
| Education                   | Note: Educational level of head of household<br>1. Primary (58.2%),2. Lower secondary (10.0%),3. Upper secondary (10.7%),4. Post-secondary (3.7%), 5. Bachelor's degree (10%),6. Master's degree (1.5%),7. Doctoral degree (0.05%),8. Other (0.12%), 9. Missing values (5.8%) |

|                                    |   |
|------------------------------------|---|
| Work status                        | Note: Work status of head of household<br>1. Employer (6.3%),2. Own-account worker (36.9%),3. Contributing family worker (2.3%),4. Government employee (10.7%), 5. State enterprise employee (1.0%),6. Private company employee (21.5%),7. Member of producers' cooperative (0.03%),8.Housewife (4.3%),9. Student (0.7%),10. Child or elderly person (12.2%),11. Ill or disabled person (1.4%),12. Looking for a job (0.1%), 13. Unemployed (0.4%),14. Other (2.2%) |
| Proportion of tobacco expenditure  | Note: Proportion of monthly expenditure on tobacco products per household by total monthly expenditure<br>min = 0, median = 0, max= 0.2907, mean = 0.0077, standard deviation = 0.0194  |
| Proportion of gambling expenditure | Note: Proportion of monthly expenditure on lottery tickets and other kinds of gambling by total monthly expenditure<br>min = 0, median = 0, max= 0.6195, mean = 0.0096, standard deviation = 0.0218   |

Note that for each household, the household head is selected by members of the household. Further details on the distribution of alcohol expenditure via histograms and box plots are given in Changpetch et al. [11]. Note that in 2009 the exchange rate ranged from 30.35 to 35.22 Bahts to the US dollar.

### Logistic Regression Analysis

For the first analysis, we used logistic regression analysis to investigate the associations between the likelihood of alcohol consumption and demographic factors, household tobacco expenditure, and household gambling expenditure. We studied the likelihood of at-home and away-from-home alcohol consumption in 2009 separately, as we thought it likely that these two types of consumption would differ in terms of the relationships between the factors. The factors for the two logistic regression models are all given in Table 1, with the exception of the last two factors, i.e., the proportion of household tobacco expenditure and the proportion of household gambling expenditure. Note that this dataset does not separate tobacco consumption and gambling consumption into at-home and away-from-home consumption as it does for alcohol consumption. The results are shown in Table 2.

In Model 1a, the household's demographic background is regressed on whether alcohol is consumed at home, followed by adding tobacco expenditure and gambling expenditure in Model 1b. Similarly, in Model 2a, the household's demographic background is regressed on whether alcohol is consumed away from home, followed by adding tobacco expenditure and gambling expenditure in Model 2b. Our purpose is to investigate if the factors remain significant when more predictors are included in the model.

Model 1a shows that households with a Buddhist head are about 13.6 times more likely to consume alcohol at home than are households with a Muslim head, whereas households with a male head are about twice more likely to consume alcohol at home than are households with a female head. When tobacco expenditure and gambling expenditure are added, Model 1b shows that the coefficients and odds ratios do not change drastically in comparison to Model 1a. Households with a Buddhist head are about 14.26 times more likely to consume alcohol at home than are households with a Muslim head, whereas households with a male head are about 1.92 times more likely to consume alcohol at home than are households with a female head. Also note that, as is the case for Model 1a, all the demographic factors in Model 1b are significant, with a significance level of 0.05, with the exceptions of the amount of

debt, disability status, welfare status, and government fund. Tobacco expenditure and gambling expenditure, which are the added factors, are also significant for Model 1b.

Based on the Chi-square values in Model 1b, the strongest predictor of at-home alcohol consumption is tobacco expenditure, followed by the religion and then the sex of the head of household. The probability of consuming alcohol at home is higher on the basis of the following conditions, given in order of the strength of the predictors: higher tobacco expenditure, a Buddhist head of household, a male head of household, a location in the central region (Bangkok excluded), a younger head of household, a larger household size, a head of household with active work status, a head of household with a lower educational level, higher gambling expenditure, a location in a non-municipal area, a married head of household, and higher income.

**Table 2. Logistic Regression Models for Alcohol Consumption at Home and Alcohol Consumption away from Home**

|                                 | Consumption at home         |                         |                             |                         | Consumption away from home  |                         |                             |                         |
|---------------------------------|-----------------------------|-------------------------|-----------------------------|-------------------------|-----------------------------|-------------------------|-----------------------------|-------------------------|
|                                 | Model 1a                    |                         | Model 1b                    |                         | Model 2a                    |                         | Model 2b                    |                         |
|                                 | coefficient<br>(odds ratio) | Chi-square<br>(p-value) | coefficient<br>(odds ratio) | Chi-square<br>(p-value) | coefficient<br>(odds ratio) | Chi-square<br>(p-value) | coefficient<br>(odds ratio) | Chi-square<br>(p-value) |
| Constant                        | -1.048                      |                         | -1.289                      |                         | -1.674                      |                         | -1.861                      |                         |
| Number household                | 0.12386                     | 189.790<br>(0.000)      | 0.10547                     | 134.430<br>(0.000)      | 0.0932                      | 71.650<br>(0.000)       | 0.0775                      | 49.090<br>(0.000)       |
| Income                          | 0.000001                    | 9.620<br>(0.002)        | 0.000001                    | 4.230<br>(0.040)        | 0.000001                    | 8.460<br>(0.004)        | 0.000001                    | 4.520<br>(0.033)        |
| Age                             | -0.02091                    | 233.540<br>(0.000)      | -0.02012                    | 209.830<br>(0.000)      | -0.01368                    | 66.800<br>(0.000)       | -0.01322                    | 61.310<br>(0.000)       |
| Amount debt                     | 0                           | 0.600<br>(0.439)        | 0                           | 0.000<br>(0.983)        | 0                           | 1.290<br>(0.255)        | 0                           | 0.400<br>(0.528)        |
| <b>Region</b><br>[ref: Bangkok] |                             | 268.750<br>(0.000)      |                             | 238.670<br>(0.000)      |                             | 139.040<br>(0.000)      |                             | 172.600<br>(0.000)      |
| Central                         | 0.2548<br>(1.2902)          |                         | 0.2962<br>(1.3477)          |                         | 0.1449<br>(1.1559)          |                         | 0.1772<br>(1.1938)          |                         |
| North                           | -0.0111<br>(0.989)          |                         | 0.1<br>(1.1052)             |                         | 0.6216<br>(1.8618)          |                         | 0.7129<br>(2.0399)          |                         |
| Northeast                       | -0.3449<br>(0.7083)         |                         | -0.2459<br>(0.782)          |                         | 0.283<br>(1.327)            |                         | 0.3681<br>(1.445)           |                         |
| South                           | -0.1611<br>(0.8512)         |                         | -0.1666<br>(0.8465)         |                         | 0.2899<br>(1.3363)          |                         | 0.293<br>(1.3404)           |                         |
| <b>Area</b><br>[ref: Municipal] |                             | 16.510<br>(0.000)       |                             | 27.440<br>(0.000)       |                             | 0.150<br>(0.702)        |                             | 1.420<br>(0.233)        |
| Non-municipal                   | 0.1212<br>(1.1288)          |                         | 0.1582<br>(1.1714)          |                         | 0.014<br>(1.0141)           |                         | 0.0439<br>(1.0449)          |                         |
| <b>Sex</b><br>[ref: Male]       |                             | 361.330<br>(0.000)      |                             | 311.720<br>(0.000)      |                             | 229.210<br>(0.000)      |                             | 265.830<br>(0.000)      |
| Female                          | -0.6961<br>(0.4985)         |                         | -0.6534<br>(0.5203)         |                         | -0.7753<br>(0.4606)         |                         | -0.7342<br>(0.4799)         |                         |

|  | Consumption at home |                    |                     |                    | Consumption away from home |                    |                     |                    |
|--|---------------------|--------------------|---------------------|--------------------|----------------------------|--------------------|---------------------|--------------------|
|  | Model 1a            |                    | Model 1b            |                    |                            |                    | Model 1a            |                    |
|  | coefficient         | Chi-square         |                     | coefficient        | Chi-square                 |                    | coefficient         | Chi-square         |
|  | (odds ratio)        | (p-value)          | (odds ratio)        | (p-value)          | (odds ratio)               | (p-value)          | (odds ratio)        | (p-value)          |
| <b>Marital status</b><br>[ref: Single] |                     | 26.300<br>(0.000)  |                     | 27.300<br>(0.000)  |                            | 58.110<br>(0.000)  |                     | 57.220<br>(0.000)  |
| Married                                | 0.2768<br>(1.3189)  |                    | 0.2887<br>(1.3347)  |                    | -0.3396<br>(0.7121)        |                    | -0.3403<br>(0.7116) |                    |
| Widowed                                | 0.2671<br>(1.3061)  |                    | 0.2674<br>(1.3065)  |                    | 0.0464<br>(1.0475)         |                    | 0.0411<br>(1.0419)  |                    |
| Other                                  | 0.1141<br>(1.1209)  |                    | 0.1262<br>(1.1345)  |                    | -0.0802<br>(0.923)         |                    | -0.0817<br>(0.9215) |                    |
| <b>Education</b><br>[ref: Primary]     |                     | 111.120<br>(0.000) |                     | 88.750<br>(0.000)  |                            | 22.010<br>(0.005)  |                     | 23.840<br>(0.002)  |
| Lower-secondary                        | -0.0836<br>(0.9198) |                    | -0.1078<br>(0.8978) |                    | -0.1385<br>(0.8706)        |                    | -0.1615<br>(0.8508) |                    |
| Upper secondary                        | -0.1629<br>(0.8497) |                    | -0.1887<br>(0.8281) |                    | -0.1859<br>(0.8304)        |                    | -0.2109<br>(0.8098) |                    |
| Post-secondary                         | -0.3527<br>(0.7028) |                    | -0.3494<br>(0.7051) |                    | -0.0582<br>(0.9434)        |                    | -0.0589<br>(0.9428) |                    |
| Bachelor                               | -0.5163<br>(0.5967) |                    | -0.4603<br>(0.6311) |                    | -0.1837<br>(0.8322)        |                    | -0.1436<br>(0.8662) |                    |
| Master                                 | -0.825<br>(0.438)   |                    | -0.721<br>(0.4865)  |                    | -0.278<br>(0.7571)         |                    | -0.204<br>(0.8152)  |                    |
| Doctoral                               | -1.67<br>(0.1891)   |                    | -1.43<br>(0.2382)   |                    | 0.063<br>(1.0654)          |                    | 0.223<br>(1.2495)   |                    |
| Other                                  | 0.13<br>(1.1394)    |                    | 0.136<br>(1.146)    |                    | -1.253<br>(0.2856)         |                    | -1.236<br>(0.2906)  |                    |
| Missing Values                         | 0.0881<br>(1.0921)  |                    | 0.1075<br>(1.1134)  |                    | 0.0865<br>(1.0903)         |                    | 0.1052<br>(1.1109)  |                    |
| <b>Religion</b><br>[ref: Buddhist]     |                     | 468.750<br>(0.000) |                     | 470.750<br>(0.000) |                            | 255.410<br>(0.000) |                     | 248.570<br>(0.000) |
| Islamic                                | -2.610<br>(0.0735)  |                    | -2.658<br>(0.0701)  |                    | -2.504<br>(0.0817)         |                    | -2.492<br>(0.0827)  |                    |
| Other                                  | -0.506<br>(0.603)   |                    | -0.481<br>(0.6179)  |                    | -0.277<br>(0.7578)         |                    | -0.247<br>(0.7809)  |                    |
| <b>Disability</b><br>[ref: No]         |                     | 0.670<br>(0.414)   |                     | 0.670<br>(0.415)   |                            | 4.140<br>(0.042)   |                     | 4.050<br>(0.044)   |
| Yes                                    | -0.086<br>(0.9172)  |                    | -0.087<br>(0.9167)  |                    | -0.266<br>(0.7666)         |                    | -0.264<br>(0.768)   |                    |
| <b>Welfare</b><br>[ref: No]            |                     | 0.040<br>(0.832)   |                     | 0.070<br>(0.786)   |                            | 0.910<br>(0.339)   |                     | 1.050<br>(0.307)   |
| Yes                                    | 0.0204<br>(1.0206)  |                    | 0.0265<br>(1.0268)  |                    | 0.12<br>(1.1272)           |                    | 0.129<br>(1.1374)   |                    |

|                                     | Consumption at home |                  |                     |                    | Consumption away from home |                  |                     |                    |
|-------------------------------------|---------------------|------------------|---------------------|--------------------|----------------------------|------------------|---------------------|--------------------|
|                                     | Model 1a            |                  | Model 1b            |                    |                            |                  | Model 1a            |                    |
|                                     | coefficient         | Chi-square       |                     | coefficient        | Chi-square                 |                  | coefficient         | Chi-square         |
|                                     | (odds ratio)        | (p-value)        | (odds ratio)        | (p-value)          | (odds ratio)               | (p-value)        | (odds ratio)        | (p-value)          |
| <b>Government fund</b><br>[ref: No] |                     | 0.090<br>(0.759) |                     | 0.020<br>(0.886)   |                            | 1.100<br>(0.295) |                     | 0.850<br>(0.358)   |
| Yes                                 | 0.0117<br>(1.0118)  |                  | 0.0055<br>(1.0055)  |                    | 0.0486<br>(1.0498)         |                  | 0.0429<br>(1.0439)  |                    |
| Producers cooperative               | -0.62<br>(0.5386)   |                  | -0.77<br>(0.4653)   |                    | -0.33<br>(0.7186)          |                  | -0.4<br>(0.6694)    |                    |
| Housewife                           | -0.0301<br>(0.9703) |                  | -0.0478<br>(0.9533) |                    | 0.073<br>(1.0759)          |                  | 0.066<br>(1.0682)   |                    |
| Student                             | -0.745<br>(0.4748)  |                  | -0.646<br>(0.5241)  |                    | -0.579<br>(0.5603)         |                  | -0.502<br>(0.6053)  |                    |
| Child elderly                       | -0.1706<br>(0.8432) |                  | -0.1359<br>(0.8729) |                    | -0.4419<br>(0.6428)        |                  | -0.4067<br>(0.6658) |                    |
| Ill disabled                        | -0.562<br>(0.57)    |                  | -0.545<br>(0.5796)  |                    | -0.636<br>(0.5292)         |                  | -0.613<br>(0.5419)  |                    |
| Looking jobs                        | -0.666<br>(0.514)   |                  | -0.636<br>(0.5293)  |                    | -1.69<br>(0.1852)          |                  | -1.67<br>(0.1883)   |                    |
| Unemployed                          | -0.09<br>(0.9136)   |                  | -0.243<br>(0.784)   |                    | 0.107<br>(1.1129)          |                  | -0.01<br>(0.9904)   |                    |
| Other                               | -0.09<br>(0.9141)   |                  | -0.079<br>(0.9237)  |                    | -0.46<br>(0.6314)          |                  | -0.449<br>(0.6386)  |                    |
| <b>Tobacco expenditure</b>          |                     |                  | 0.000948<br>(0.000) | 645.320<br>(0.000) |                            |                  | 0.000686<br>(0.000) | 256.040<br>(0.000) |
| <b>Gambling expenditure</b>         |                     |                  | 0.00013<br>(0.000)  | 30.950<br>(0.000)  |                            |                  | 0.000152<br>(0.000) | 36.520<br>(0.000)  |

Model 2a shows that households with a Buddhist head are about 12.24 times more likely to consume alcohol away from home than are households with a Muslim head, whereas households with a male head are 2.17 times more likely to consume alcohol away from home than are households with a female head. In Model 2b, the coefficients and odds ratios do not change drastically compared to Model 2a. The households with a Buddhist head are about 12.09 times more likely to consume alcohol away from home than are households with a Muslim head, whereas households with a male head are about 2.08 times more likely to consume alcohol away from home than are households with a female head. Also, note that, as is the case for Model 2a, all the demographic factors in Model 2b are significant, with a significance level of 0.05, with the exceptions of amount of debt, area, welfare status, and government fund. Tobacco expenditure and gambling expenditure, which are the added factors, are also significant for Model 2b.

Based on the Chi-square values in Model 2b, the strongest predictor that a household will consume alcohol away from home is the sex of the head of household, followed by household expenditure on tobacco, and then the religion of the head of household. The probability of consuming alcohol away from home is higher on the basis of the following conditions, given

in order of the strength of the predictors: a male head of household, higher tobacco expenditure, a Buddhist head of household, a head of household with active work status, a location in the North, a younger head of household, a widowed head of household, a larger household size, higher gambling expenditure, higher income, and a non-disabled head of household.

Between the logistic regression models for the likelihood of consuming alcohol at home and away from home, the differences in terms of significant factors are area (significant in Models 1a and 1b, but not significant in Models 2a and 2b) and disability (not significant in Models 1a and 1b, but significant in Models 2a and 2b).

**Treenet**

In this section, we refine our understanding of alcohol consumption by employing data-mining model, which capture non-linearities and interactions automatically.

In this second analysis, we investigate the association between the proportion of total household expenditure allocated to alcohol and demographic factors, the proportion allocated to tobacco, and the proportion allocated to gambling. We constructed Treenet models ([www.salford-systems.com/treenet.html](http://www.salford-systems.com/treenet.html) and [12]) for the proportion of household expenditure spent on alcohol consumed at home (Model 1), and the proportion of household expenditure spent on alcohol consumed away from home (Model 2). Compared to the multiple linear regression model, this method has two main advantages: it reveals non-linear relationships between response and predictors, and the non-parametric approach adopted here makes it possible to handle a response variable with a large number of zero values (about 80% of the dataset).

With the exceptions of tobacco expenditure and gambling expenditure, all the factors in Table 1 are included in each of the two Treenet models.

Figure 1 shows that five variables are important for explaining the proportion of alcohol expenditure consumed at home: the proportion of tobacco expenditure, the region of the household, the sex of the head of household, the proportion of gambling expenditure, and the age of the head of household. Note that Treenet derived the relative importance of the predictors (with the most important variable assigned an importance of 100 for reference), as shown in Figure 1.

Figure 1: Variable importance in Treenet Model 1

| Variable      | Score  |  |
|---------------|--------|--|
| PROP_TOBACCO  | 100.00 |  |
| REGION\$      | 60.22  |  |
| SEX\$         | 35.84  |  |
| PROP_GAMBLING | 34.51  |  |
| AGE           | 28.05  |  |

Figure 2 (a-e) shows the partial effect of each predictor on the estimated response (while holding other predictors constant). In Figure 2(a), we can see that the effect of the proportion of tobacco expenditure on the proportion of alcohol expenditure “kicks in” at about an



approximate value of 0.001, with no further effect beyond 0.001. These results suggest that a proportion of tobacco expenditure of 0.001 or above is associated with a jump in the proportion of total alcohol expenditure of about 0.0005 units. Figure 2(b) shows the negative effect of regions 3 (North), 4 (Northeast) and 5 (South). Note that the effect of religion is probably captured by region, as there is a high proportion of Muslims in the South of Thailand. Figure 2(c) reveals the negative effect of a female head of household on the proportion of alcohol expenditure consumed at home. In Figure 2(d), we can see that the effect of the proportion of gambling expenditure on the proportion of total alcohol expenditure “kicks in” at an approximate value of 0.001, with no further effect beyond 0.001. These results suggest that a proportion of gambling expenditure of 0.001 or above is associated with a jump in proportion of total alcohol expenditure of about 0.00007 units. In Figure 2(e), the effect of the age of the household head on the proportion of alcohol expenditure consumed at home appears only when the head of household’s age reaches approximately 47; at this age estimated home proportions drop and do not drop again at older ages.

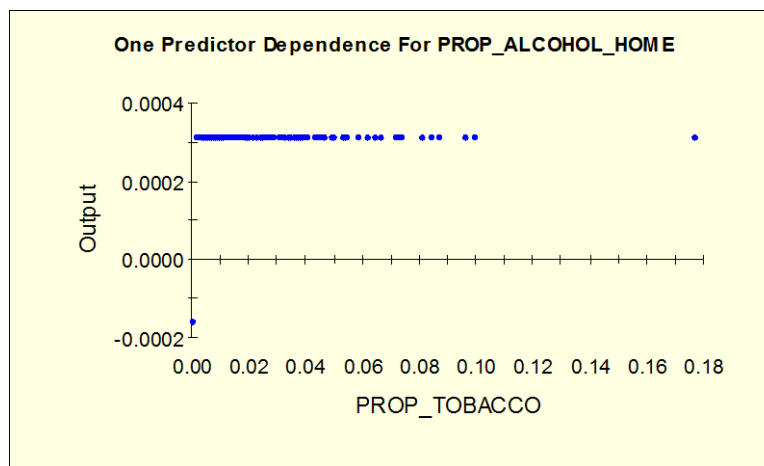


Figure 2(a): Proportion of tobacco expenditure and proportion of alcohol expenditure consumed at home

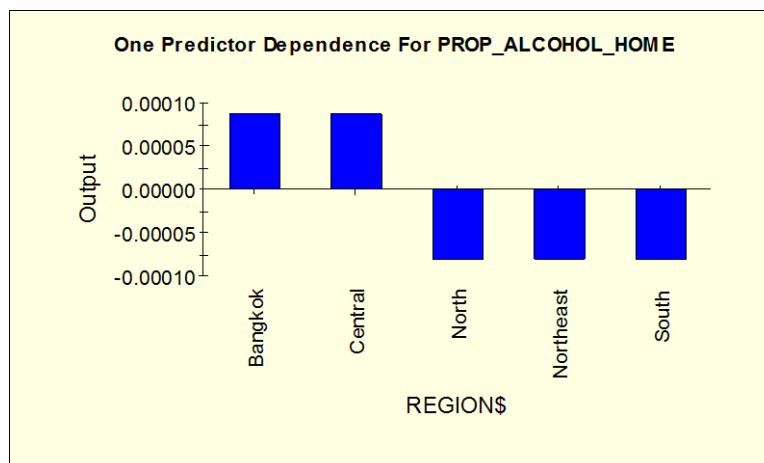


Figure 2(b): Region of household and proportion of alcohol expenditure consumed at home

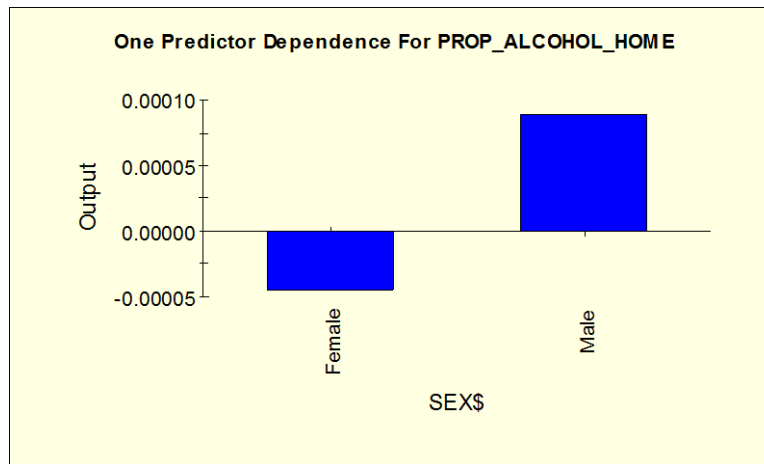


Figure 2(c): Sex of head of household and proportion of alcohol expenditure consumed at home

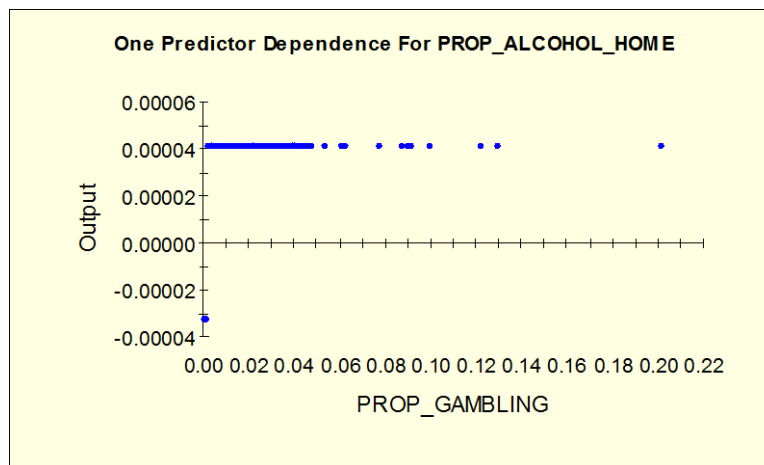


Figure 2(d): Proportion of gambling expenditure and proportion of alcohol expenditure consumed at home

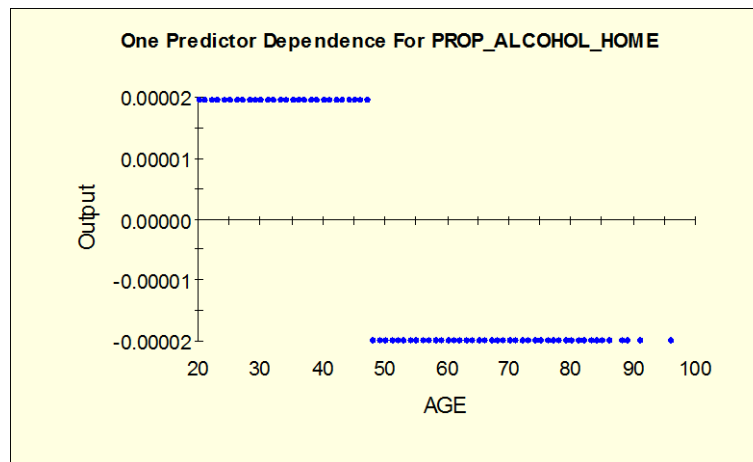


Figure 2(e): Age of head of household and proportion of alcohol expenditure consumed at home

Figure 3 shows that five variables are important for predicting the proportion of alcohol expenditure consumed away from home: the proportion of tobacco expenditure, the sex of the head of household, the proportion of gambling expenditure, the work status of the head of household, and religion of the head of household.

Figure 3: Variable importance in Treenet Model 2

| Variable      | Score  |  |
|---------------|--------|--|
| PROP_TOBACCO  | 100.00 |  |
| SEX\$         | 53.73  |  |
| PROP_GAMBLING | 48.41  |  |
| WORK_STATUS\$ | 45.17  |  |
| RELIGION\$    | 38.36  |  |

Figure 4(a) shows a similar pattern to that in Figure 2(a). However, in Figure 4(a), the trigger number is 0.002 instead of 0.001. Figure 4(b) reveals the negative effect of having a female head of household on the proportion of alcohol expenditure consumed at home. Figure 4(c) shows a similar pattern to that in Figure 2(d). These results suggest that a proportion of gambling expenditure of 0.001 or above is associated with a jump in the proportion of total alcohol expenditure of about 0.000004 units. Figure 4(d) shows the positive effect of work status 1 (Employers), 4 (Government employees), and 5 (State enterprise employees), whereas Figure 4(e) reveals the negative effect of the Muslim religion on the proportion of alcohol expenditure consumed at home.

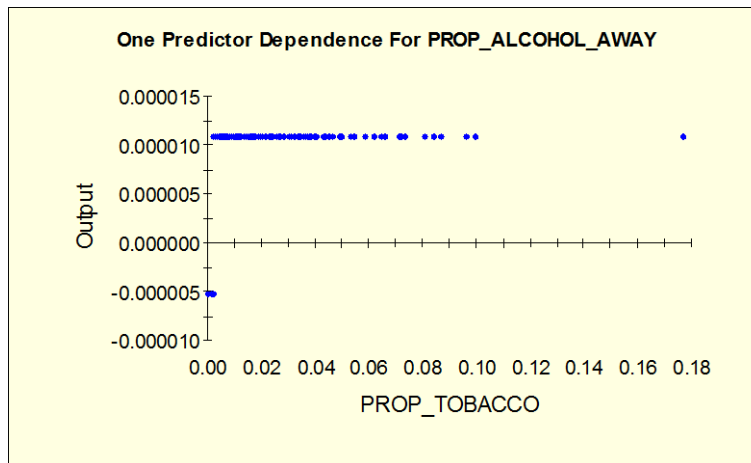


Figure 4(a): Proportion of tobacco expenditure and proportion of alcohol expenditure consumed away from home

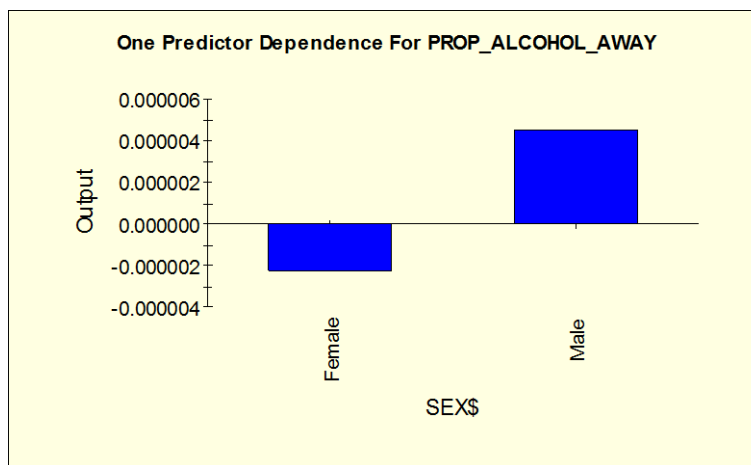


Figure 4(b): Sex of head of household and proportion of alcohol expenditure consumed away from home

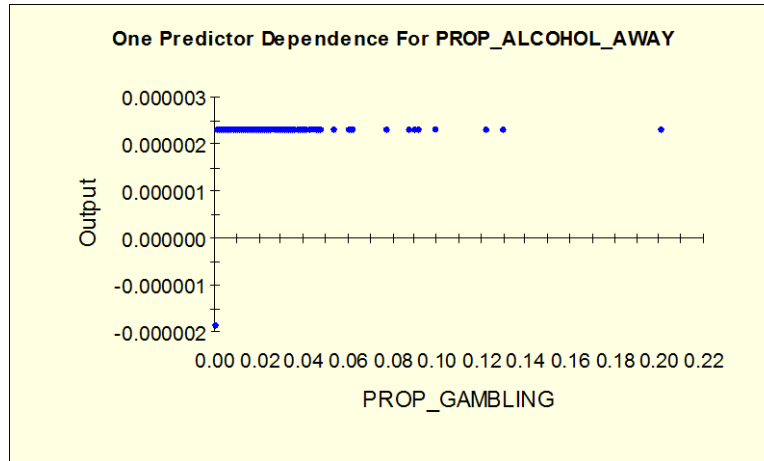


Figure 4(c): Proportion of gambling expenditure and proportion of alcohol expenditure consumed away from home

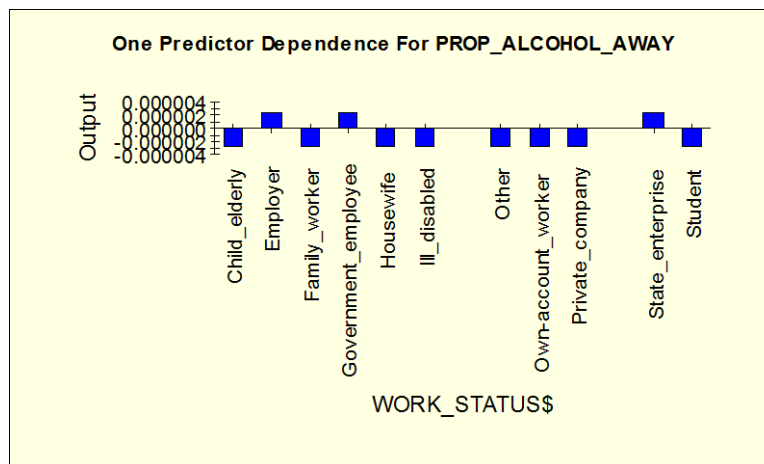


Figure 4(d): Work status of head of household and proportion of alcohol expenditure consumed away from home

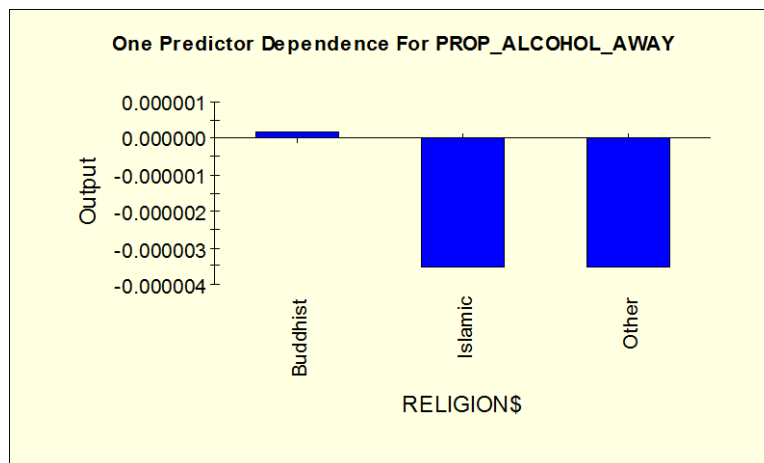


Figure 4(e): Religion of head of household and proportion of alcohol expenditure consumed away from home

From both models, the proportion of tobacco expenditure is the most important variable. The sex of the head of household and the proportion of gambling expenditure are also important variables for both models, whereas the age of the head of household is an important variable for Model 1. The religion of the head of household is an important variable for Model 2,

whereas region (which acts as a proxy for religion) is an important variable for Model 1. The work status of the head of household is important variables for Model 2.

The Treenet results suggest that the proportion of expenditure on alcohol consumed at home is higher for households that consumed tobacco at a proportion of higher than 0.001, households in Bangkok and the central area, households with a male head of household, households that spend money on gambling at a proportion of higher than 0.001, and households whose head of household is younger than 47 years old. The proportion of expenditure on alcohol consumed away from home is higher for households that consumed tobacco at a proportion of higher than 0.002, households with a male head of the household, households that spend money on gambling at a proportion of higher than 0.001, households whose head of household is an employer, works for the government, or is a state enterprise employee, and households whose head of the household is non-Muslim.

## **CONCLUSION**

This paper provides a thorough study of alcohol consumption in Thailand in regard to the relationships between alcohol consumption, tobacco consumption, gambling consumption, and demographic factors. We applied logistic regression and Treenet to the analysis of data drawn from a socio-economic survey of 43,830 Thai households conducted in 2009. Beyond the bivariate, multiple linear regression, and logistic regression analyses typically used in literature related to the study of alcohol, we implemented a method that had never been used in this context before, i.e., Treenet, a method that uncovers non-linear associations between responses and predictors. Note that we used logistic regression to analyze the likelihood of consuming alcohol (at home and away from home) and we used Treenet to study the proportion of total household expenditure on alcohol (at home and away from home).

The results of our study show that alcohol, tobacco, and gambling co-occur. Through logistic regression, the strongest predictors of household alcohol consumption are tobacco expenditure, followed by the religion, and then the sex of the head of household. Through Treenet, we found that the proportion of household expenditure spent on tobacco is the most important predictor of the proportion of household expenditure spent on alcohol.

In conclusion, a female head of household is associated with a lower likelihood of a household consuming alcohol and a lower proportion of household expenditure spent on alcohol. The religion of the head of household is associated with both likelihood of alcohol consumption and spending on alcohol. Compared with those with a non-Muslim head, households with a Muslim head are less likely to consume alcohol and spend a smaller proportion of household expenditure on alcohol. The age of the head of household is also related to both the likelihood of alcohol consumption and the proportion of household expenditure spent on alcohol. Households with a younger head have a higher likelihood of consuming alcohol and spend a higher proportion of their expenditure on alcohol than do households with an older head. The educational level of the head of household is also associated with alcohol consumption: Households with a head with a lower educational level are more likely to consume alcohol than are households with a head with a higher educational level (from logistic regression). In regard to work status, both the logistic regression and Treenet models show that households headed by a government or state enterprise employee are more likely to consume alcohol and to spend a higher proportion of their expenditure on alcohol than are households with a head in other kinds of employment. This result is surprising, as government employees generally have lower income than those with other kinds of work status, e.g., employers and private company employees.

Similar to the study by Aekplakorn et al. (2008) [7], our study shows that smoking is strongly related to the likelihood of alcohol consumption. However, in our research, we used household data whereas Aekplakorn et al. (2008) used individual data. The use of household data is considered a limitation in our research, as the data do not allow us to offer an interpretation and conclusion based on individuals, which is more conventional as a unit of analysis. On the other hand, in our study, we investigated more demographic factors and also included gambling consumption. We not only used logistic regression to study the likelihood of alcohol consumption, but we also studied the proportion of household expenditure spent on alcohol using Treenet, a method that has never been used before in any of the literatures in this field.

With all the associations from our study, we expect our results to be useful to both researchers and government practitioners in their efforts to tailor programs to households that include alcohol-dependent members in order to reduce problems related to alcohol consumption in Thailand.

### **ACKNOWLEDGEMENT**

We would like to thank Dr. Robin Room for his guidance in interpreting and communicating the results to readers.

## REFERENCES

- [1] Kamsa-ard, S., et al. (2014). Alcohol consumption and mortality: The Khon Kaen Cohort Study. *Journal of Epidemiology*, 24 (2), 154–60.
- [2] Kahler, C. W. et al. (2008). Cigarette smoking and the lifetime alcohol involvement continuum. *Drug Alcohol Depend*, 93, 111–20.
- [3] Burger, M., Mensink, G., Bronstrup, A., Thierfelder, W., & Pietrzik, K. (2004). Alcohol consumption and its relation to cardiovascular risk factors in Germany. *Eur J Clin Nutr*, 58, 605–614.
- [4] Chioloro, A., Wietlisbach, V., Ruffieux, C., Paccaud, F., & Cornuz, J. (2006). Clustering of risk behaviors with cigarette consumption: A population-based survey. *Prev Med*, 42, 348–53.
- [5] Clausen, T., Charlton, K. E., & Holmboe-Ottesen, G. (2006). Nutritional status, tobacco use and alcohol consumption of older persons in Botswana. *Journal of Nutrition, Health & Aging*, 10, 104–110.
- [6] Bonevski, B., Regan, T., Paul, C., Baker, A. L., & Bisquera, A. (2014). Associations between alcohol, smoking, socioeconomic status and comorbidities: Evidence from the 45 and Up Study. *Drug and Alcohol Review*, 33 (2), 169-76.
- [7] Aekplakorn, W., Hogan, M. C., Tiptaradol, S., Wibulpolprasert, S., Punyaratabundhu, P., & Lim, S. S. (2008). Tobacco and hazardous or harmful alcohol use in Thailand: joint prevalence and associations with socioeconomic factors. *Addictive Behavior*, 33, 503-14.
- [8] Duhig, A. M. et al. (2007). Characteristics of adolescent past-year gamblers and non-gamblers in relation to alcohol drinking. *Addict Behavior*, 32, 80-89.
- [9] Barnes, G. M. et al. (2009). Gambling alcohol, and other substance use among youth in the United States. *J Stud Alcohol Drugs*, 70, 134-42.
- [10] Peters, E. N. et al. (2015). Relationship of gambling with tobacco, alcohol, and illicit drug use among adolescents in the USA: Review of the literature 2000-2014. *American Journal on Addiction*, 24, 206-16.
- [11] Changpetch, P., Haughton, D., Le, M. T. X., Ly, S., Nguyen, P., & Thach, T. T. (2016). *Alcohol consumption in Thailand: A study of the associations between alcohol, tobacco, gambling, and demographic factors*. Retrieved from hal-01271254.
- [12] Friedman, J. (2001). Greedy function approximation: a gradient boosting machine. *Annals of Statistics*, 29 (5), 1189-232.