



AgEcon SEARCH
RESEARCH IN AGRICULTURAL & APPLIED ECONOMICS

The World's Largest Open Access Agricultural & Applied Economics Digital Library

This document is discoverable and free to researchers across the globe due to the work of AgEcon Search.

Help ensure our sustainability.

Give to AgEcon Search

AgEcon Search
<http://ageconsearch.umn.edu>
aesearch@umn.edu

*Papers downloaded from **AgEcon Search** may be used for non-commercial purposes and personal study only. No other use, including posting to another Internet site, is permitted without permission from the copyright owner (not AgEcon Search), or as allowed under the provisions of Fair Use, U.S. Copyright Act, Title 17 U.S.C.*

Analysis of Alcohol Consumption in the United States: Probability and Level of Intake

Rodolfo M. Nayga, Jr., Oral Capps, Jr.

This article examines the impact of socio-demographic factors on individual consumption of alcohol in the United States using the Heckman procedure and logit analysis. Factors considered are urbanization, race, ethnicity, region, weight, height, sex, food stamp participation, employment status, diet status, day of consumption, household size, age, and income. All the variables, with the exception of ethnicity and food stamp participation, significantly influence the decision to consume alcohol. Results also indicate that urbanization, region, sex, food stamp participation, and household size significantly affect the level of alcohol consumption.

Introduction

The United States has experienced increasing movement away from alcohol consumption. In fact, recent trends in alcohol consumption reflect patterns toward moderation and safety (Boyd, 1989). Concerns over drinking and driving and alcohol and health appear to have been partly responsible for the declining trend in alcohol consumption. Liquor sales, for instance, are continuing their steady decline but sales of non-alcoholic beverages and "healthier" alcoholic drinks are increasing (DeLuca, 1991; Maxwell, 1987).

Even with the declining trend in alcohol consumption, however, certain population groups are still facing risks associated with alcohol use. For instance, previous research has expressed concern over the potential for alcohol abuse among Hispanic youth (Patterson, 1992). Older people in the United States also tend to drink more hard liquor (Exter, 1991). Moreover, previous studies reveal that 25 percent more blacks than whites die from cancer, citing alcohol consumption as a contributing factor for increasing death rates among blacks (Djata, 1987). Young bachelors and single parents also have been found to exhibit heavy usage of alcoholic beverages in restaurants (Danko, 1990).

Knowledge of the effect of socio-demographic factors on alcohol consumption is needed in the design of alcohol education programs. In addition, a study related to alcohol consumption would be useful in developing production and marketing programs that

would reflect consumption patterns and preferences of individual consumers. Little attention has been paid, however, to the analysis of the effect of socio-demographic factors on alcohol consumption in the United States (Godfrey, 1989; Akutsu et al., 1989). The purpose of this paper is to assess the impacts of various socio-demographic characteristics on the probability and level of individual alcohol intake. By examining the participation and consumption decisions separately, policy makers could use the information derived from this study in determining the factors affecting the decision to consume alcohol and the level or volume of alcohol consumption.

Model Specification

The exogenous variables used in the analysis include urbanization, region, race, sex, employment, household size, age, height, weight, and income. Dummy variables pertaining to whether the individual receives food stamps or not; whether the individual is on special diet or not; and whether the three-day intake of the individual occurred mostly during a weekend or a weekday are also included in the analysis. Two models are developed in this study to analyze both the probability and level of alcohol intake. Mathematically, the empirical models are specified as:

$$\text{Alc}_{ki} = b_0 + b_1\text{urban1} + b_2\text{urban2} + b_3\text{region1} + b_4\text{region2} + b_5\text{region4} + b_6\text{race2} + b_7\text{race3} + b_8\text{race4} + b_9\text{hisp1} + b_{10}\text{sex1} + b_{11}\text{employ1} + b_{12}\text{fstamp1} + b_{13}\text{diet1} + b_{14}\text{hsize} + b_{15}\text{weight} + b_{16}\text{height} + b_{17}\text{age} + b_{18}\text{agesq} + b_{19}\text{weekend} + b_{20}\text{income} + b_{21}\text{incomesq}$$

for $k=1, 2$. Alc_{ki} (dependent variable of Model 1, the probability equation) is equal to 1 if individual i consumed alcohol during the three-day survey period and 0 otherwise. Alc_{2i} (dependent variable of Model 2,

Assistant Professor, Department of Agricultural Economics and Marketing, Rutgers University, New Brunswick, NJ; Professor, Department of Agricultural Economics, Texas A&M University, College Station, TX.

The authors wish to thank anonymous journal referees for their helpful comments and suggestions.

the consumption equation) is the average daily intake of alcohol by individual *i* in grams per 1,000 kilocalories. The names and descriptions of the independent variables are shown in Table 1.

One classification is eliminated from each group of variables for estimation purposes so as to avoid the problem of perfect multicollinearity. The base group includes individuals who satisfy the following description: reside in a nonmetro area (urban3); in the South (region3); white (race1); nonhispanic (hisp2); female (sex2); not employed (employ2); not participating in the food stamp program (fstamp2); not on a special diet (diet2); and the three-day intake occurred mostly during a weekday (weekday). Household income is used instead of individual income because the data set used only provides income information for the household and not for an individual.

The anthropomorphic measurements of the individual—age, sex, height, and weight—are included as exogenous variables to account for physical differences between individuals. Squared terms are included for income and age in order to investigate possible nonlinearities in the relationships with alcohol consumption. Previous alcohol research has often reported significant differences in alcohol consumption due to age, sex, and weight (Akutsu et al., 1989).

Table 1
Names and Description of the Independent Variables

Name	Description
urban1	1 if individual resides in a central city; 0 otherwise
urban2	1 if individual resides in a suburban area; 0 otherwise
region1	1 if individual is in the Northeast; 0 otherwise
region2	1 if individual is in the Midwest; 0 otherwise
region4	1 if individual is in the West; 0 otherwise
race2	1 if individual is black; 0 otherwise
race3	1 if individual is Asian or Pacific Islander; 0 otherwise
race4	1 if individual is of some other race; 0 otherwise
hisp1	1 if individual is hispanic; 0 otherwise
sex1	1 if individual is male; 0 otherwise
employ1	1 if individual is employed; 0 otherwise
fstamp1	1 if individual is receiving food stamps; 0 otherwise
diet1	1 if individual is on a special diet; 0 otherwise
hsize	household size
weight	weight of the individual in pounds

height	height of the individual in inches
age	age of the individual in years
agesq	square of the age of the individual
weekend	1 if the three-day intake of the individual occurred mostly during a weekend; 0 otherwise
income	household income
incomesq	square of household income

To investigate the effect of various socio-demographic variables on the probability of consuming alcohol, logit analysis is employed in the estimation of Model 1. The Heckman procedure, on the other hand, is used to estimate Model 2 because some individuals report no consumption of alcohol over the survey period. If only nonzero consumption observations are used in parameter estimation, ordinary least squares procedures would yield inconsistent estimates from selectivity bias (Maddala, 1983). Heckman (1976, 1979) described sample selection bias as a type of a specification error or omitted-variable problem. Subsequently, Heckman proposed a technique that amounts to estimating the inverse of Mill's ratio for each observation using probit analysis at the first stage of the estimation process. The inverse of Mill's ratio is defined as the ratio of the value of the standard normal density function to the value of the standard normal distribution function (for additional details, see Capps and Cheng, 1986). The probit analysis employs all available observations; the dependent variable equals to one if the individual reports an intake and zero otherwise. The second stage involves the use of the inverse of Mill's ratio as a regressor in the original model specification and the nonzero intake observations.

Data Source and Description

The data set used in this study is the Individual Intake phase of the United States Department of Agriculture's (USDA) 1987-88 Nationwide Food Consumption Survey (NFCS). This data set is the most recent of the national household food consumption surveys conducted by USDA and was just made available to researchers in 1991.

The individual intake phase of the 1987-88 NFCS data set provides data on three days of alcohol intake by individuals of all ages surveyed in the 48 contiguous states. These individuals were asked to provide three consecutive days of data. The first day's data were collected using 24-hour dietary recall. The time period for this one-day recall was from midnight to 11:59 p.m. on the day preceding the interview. This collection was done using an in-home personal interview. Data for the second and third days were col-

lected using a self-administered two-day dietary record.

The original data set contains 11,045 individuals. However, individuals with missing socio-demographic information are deleted from the sample. After the elimination of individuals with missing socio-demographic information, the data set contained 6,219 observations. About 18.7 percent of the individuals in the sample consumed alcohol during the three-day survey period. Average daily alcohol consumption for the whole sample considered is 0.68 grams per 1,000 kilocalories. Average daily alcohol consumption in the sample, however, ranged from zero to 22.8 grams per 1,000 kilocalories.

The means of the independent variables used in the analysis are exhibited in Table 2. About 21 percent of the sample reside in central city areas (Urban1); 49 percent in suburban areas (Urban2); and 30 percent in nonmetro areas (Urban3). Most of the individuals (35 percent) included in the sample come from the South (Region3). Eighty-six percent are white (Race1); 96 percent are nonhispanic (Hisp2); 45 percent are male (Sex1); 58 percent are employed (Employ1); 95 percent are non-recipients of the food stamp program (Fstamp2); 14 percent are on a special diet (Diet1); and about 16 percent consumed alcohol mostly on a weekend during the three-day survey period (Weekend). Moreover, the average age of the individuals is about 43 years and average household size is approximately three. Average weight is about 159 pounds and average height is about 67 inches. Average household income is approximately \$29,500. Other descriptive statistics of the variables are available from the authors upon request.

Table 2
Means of the Independent Variables Used in the Analysis

Variable	Mean
Urbanization	
Urban1	0.21
Urban2	0.49
Urban3 ^a	0.30
Region	
Region1	0.20
Region2	0.27
Region3 ^a	0.35
Region4	0.18
Race	
Race1 ^a	0.86
Race2	0.10

Race3	0.01
Race4	0.03
Hispanic Origin	
Hisp1	0.04
Hisp2 ^a	0.96
Sex	
Sex1	0.45
Sex2 ^a	0.55
Employment Status	
Employ1	0.58
Employ2 ^a	0.42
Food Stamp Participation	
Fstamp1	0.05
Fstamp2 ^a	0.95
Special Diet	
Diet1	0.14
Diet2 ^a	0.86
Week	
Weekend	0.16
Weekday	0.84
Age	43.33
Hsize	3.03
Weight	159.48
Height	66.72
Income	29,486.80

^aExcluded category in the regression model.

Empirical Findings

Model 1

The parameter estimates of Model 1 (probability equation) are exhibited in Table 3. This equation correctly classifies individuals as having positive or zero alcohol intake about 81.5 percent of the time using the (0.50, 0.50) classification scheme. For this criterion, a correct classification means that the predicted probability of intake is equal to or greater than 0.50 for individuals who actually consume alcohol and a predicted probability less than or equal to 0.50 for those who do not consume alcohol. As indicated in Table 3, all of the variables, except ethnicity and food stamp participation, are statistically significant at the 0.05 level. The urbanization, regional, and race variables significantly contribute to the explanatory power of the model as indicated by the results of the likelihood ratio tests in Table 4. Specifically, the empirical results of Model 1 indicate that individuals from urban and suburban areas are more likely to consume alcohol than individuals from nonmetro areas. This result is consistent with Blaylock and Blisard's (1993) finding

which indicates that among males, urban residents are most likely to drink alcohol.

In terms of region, individuals from the South are more likely to consume alcohol than individuals from other regions of the United States. The reason for this result is not clear. Moreover, this result is not in agreement with Blaylock and Blisard's (1993) result which indicates that the most likely drinkers in men are those not from the South.

Table 3
Estimated Coefficients of Model 1 and Model 2

Variable	Model 1 Parameters	Change in Probability ^b	Model 2
Intercept	6.043* (0.901)	0.815	2.619 (2.404)
Weight	0.003* (0.001)	0.0004	0.001 (0.003)
Height	-0.038* (0.013)	-0.005	-0.011 (0.036)
Urban1	0.639* (0.105)	0.086	0.804* (0.276)
Urban2	0.318* (0.091)	0.043	0.699* (0.250)
Region1	-0.659* (0.099)	-0.089	-0.437* (0.263)
Region2	-0.273* (0.097)	-0.037	-0.266 (0.266)
Region4	-0.740* (0.100)	-0.099	-0.274 (0.262)
Race2	0.513 (0.159)	0.069	0.201 (0.446)
Race3	0.523 (0.458)	0.071	0.108 (1.261)
Race4	0.742* (0.320)	0.100	-0.849 (0.881)
Hisp1	-0.138 (0.223)	-0.019	-0.526 (0.565)
Sex1	0.589* (0.101)	0.079	0.893* (0.272)
Employ1	-0.194* (0.088)	-0.026	-0.175 (0.240)
Fstamp1	0.246 (0.240)	0.033	1.499* (0.726)
diet1	0.286* (0.107)	0.039	-0.178 (0.283)
Hsize	0.303* (0.030)	0.041	-0.148* (0.079)
Age	-0.088* (0.001)	-0.012	0.012 (0.035)
Agesq	0.0009* (0.0001)	0.0001	-0.000179 (0.000377)
Weekend	-0.191* (0.0001)	-0.026	0.093

	(0.090)		(0.228)
Income	-0.00002*	-0.000003	0.000003
	(2.89E-6)		(0.000008)
Incomesq	7.11E-11*	9.59E-12	-2.37E-11
	(1.68E-11)		(1.000E-8)
Imratio			0.773*
			(0.319)

% Right Pred.* 81.5

Adjusted R-SQ

0.300

*Statistically significant at the 0.05 level

*Percentage of right predictions based on 50-50 classification scheme.

^bEqual to the product of the parameter estimates times the value of the logistic density function [$\beta \cdot f(z)$]. At the sample means, the value of this density function ($f(z)$) is 0.135.

Note: Standard errors are in parentheses.

Table 4
Results of the Likelihood Ratio Tests

Variables	Model 1	Model 2
Urbanization	37.36*	5.011*
Region	72.62*	0.943
Race	17.77*	0.394

*Statistically significant at the 0.05 level.

Results indicate that blacks are more likely to consume alcohol than whites. Djata (1987) revealed that blacks are more likely to die from cancer due to alcohol consumption than are whites. Alcohol companies have targeted black consumers by sponsoring events in black communities and by employing black-owned advertising agencies to develop campaigns (Djata, 1987). Hispanics, on the other hand, are less likely to consume alcohol than are whites. Previous research has expressed concern over the potential for alcohol abuse among Hispanics (Patterson, 1992). The reasons for these results concerning ethnic/racial differences in alcohol consumption are not clear but Akutsu et al. (1989) indicated that they may be due to physiological reactivity and cultural values.

Males are more likely to drink alcohol than are women. This result is consistent with the finding disclosed by the Interagency Board for Nutrition Monitoring and Related Research (1993). This finding, which was based on data from the Alcohol Supplement

of the 1988 National Health Interview Survey, indicates that males are more likely to be heavy drinkers and less likely to be abstainers than are females.

Interestingly, employed individuals are less likely to consume alcohol than are unemployed individuals. It may be possible that unemployed individuals have more time to drink alcohol either at home or at bars than do employed individuals. Results also indicate that individuals who are on a special diet are more likely to consume alcohol than are those who are not on a special diet. This result is surprising considering the general belief that individuals who are on a special diet are more health conscious than are individuals who are not on a special diet.

Household size and the probability of consuming alcohol are positively related. This result may be related to the importance of social cues and culture in reinforcing certain drinking practices. Generally, drinking occurs in social or group-oriented situations and it may be possible that this phenomenon is reflected in some larger households.

The likelihood of consuming alcohol decreases initially with age and then increases with successive increments of age as indicated by the negative and positive signs of the age and agesq variables. This result is consistent with Exter's (1991) report indicating that older people in the United States tend to drink more liquor than do younger people. Interestingly, individuals who were interviewed during the weekends are less likely to consume alcohol than are those who were interviewed during the weekdays. The likelihood of consuming alcohol, on the other hand, decreases initially with income and then increases with successive increments of income.

Model 2

The ordinary least squares estimates of the coefficients obtained in the second stage of the Heckman procedure are found in Table 3¹. As indicated by the results of the likelihood ratio tests in Table 4, the urbanization variables as a group contribute significantly to the explanatory power of the model. Based on the statistically significant estimates (0.05 level), results indicate that individuals residing in central cities consume more alcohol than those residing in non-metro areas. Likewise, individuals residing in suburban areas consume more alcohol than those residing in non-metro areas. These results are consistent with those from Model 1 and with the finding of Blaylock and Blisard (1993) wherein men residing in urban areas were more likely to consume wine than those residing in non-urban areas. Lutz et al.'s (1992, p. 41) descriptive analysis also revealed the same result.

In terms of region, individuals from the Northeast consume significantly less alcohol than do individuals from the South. This result is consistent with the finding from Model 1. However, Blaylock and Blisard (1993) reported that men from the South consume less wine than do those not from the South.

None of the coefficients of the race variables is statistically significant. However, males consume more alcohol than do females. This result is consistent with the finding from Model 1 and with the Inter-agency Board for Nutrition Monitoring and Related Research report that men are more likely to drink alcohol than are women. Patterson (1992), on the other hand, reported that no significant differences exist between males and females in regard to their perceptions of alcohol warnings and the risks associated with alcohol consumption. Results of this study also indicate that employed individuals do not significantly consume less alcohol than do unemployed individuals. Empirical results indicate, however, that food stamp recipients consume more alcohol than do non-food stamp recipients. This result might have some important implications in the design of alcohol education programs geared toward food stamp recipients who are basically low-income individuals.

In contrast to Model 1's estimates, age and the weekend variables do not significantly affect the consumption of alcohol. Income is also not statistically significant in Model 2. As expected, the income elasticity, when evaluated at the means of the variables, is relatively inelastic at 0.13.

Interestingly, household size is negatively related to the consumption of alcohol. This result is consistent, however, with the descriptive results disclosed by Lutz et al. (1992, p. 21) concerning the average annual household use of alcoholic beverages. The coefficient associated with the inverse of Mill's ratio (IMRATIO) is statistically significant indicating that sample selection bias could have been introduced in the estimates if the observations with zero alcohol consumption were deleted and not used in the analysis. This result further justifies the use of the Heckman procedure in the estimation of Model 2.

Concluding Remarks

Although considerable literature exists on nutrient intake models, little attention has been paid to the analysis of alcohol intake in the United States. In this light, this study examined the impact of socio-demographic factors on the probability and level of alcohol intake. The information obtained from this study could be used by educators and policy makers in the development of educational programs and tools that could help Americans decide on the appropriate levels

of alcohol to consume. The empirical results of this study could also be used as a guide in directing alcohol education programs toward certain population groups. For example, results of this study indicate that males and those who reside in central cities and suburban areas consume more alcohol and are more likely to consume alcohol. Therefore, to be more effective, outreach and education programs dealing with the intake of alcohol may need to be directed toward these groups of individuals.

The information derived from this study could also be used by the alcoholic beverage industry as a guide in the design of alcohol misuse campaigns. In a social climate that is increasingly hostile toward drunk driving, beer and liquor companies have initiated public service campaigns that encourage moderate and responsible alcohol consumption. In fact, Miller Brewing Co., Anheuser-Busch, The Coors Brewing Co., Austin Nichols, and Joseph E. Seagrams and Sons have allocated marketing dollars for alcohol misuse education programs (Penzer, 1990).

Results of this study could also be useful to processors, wholesalers, and retailers of alcoholic products in the design of marketing programs and in the marketing of alcoholic products targeted at specific population groups. For example, based from the empirical results, marketers may be able to sell beverages that claim low or zero amounts of alcohol to females and individuals who reside in non-metro areas.

The analysis conducted in this study, however, does not differentiate between the consumption of various alcohol products and does not include an examination of interdependencies among various alcohol products. Future research could be focused on estimating separate models for various alcohol products (e.g., beer, spirits, and wine). Estimates of income elasticities in the past have varied between different beverages and across studies (Walsh, 1982).

Clearly, further analysis of alcoholic products merits attention. More definitive and disaggregate data (e.g., scanner data), however, are needed in the estimation of demand models for disaggregate alcohol products. Further improvements and/or refinements in data collection and model formulation are necessary to provide more definitive results to policy makers and to the alcoholic beverage industry. Moreover, as Blisard and Blaylock (1993) suggested, consumption surveys should include questions on both market participation and frequency of purchase so that an individual with observed zero consumption may be identified through the data as either a market non-participant or an infrequent consumer or purchaser.

The results of this study are subject to the limitations of the data used. The General Accounting Office (GAO) has expressed reservations about the representativeness of the data. However, the GAO

concluded that it is not possible to determine if nonrespondents differed systematically from respondents. Lutz et al. (1992, p. 13) also indicated that, in most instances, the consumption data were fairly consistent with the data from Continuing Consumer Expenditure Surveys.

Endnote

¹No heteroskedasticity test was performed because the error structure of the equation used in the second stage is explicitly heteroskedastic. However, the technique developed to correct for heteroskedasticity broke down, making the implementation of generalized least squares procedure not possible. Consequently, ordinary least squares is used in the second stage of the procedure to produce consistent estimates.

References

- Akutsu, P. D., Sue, S., Zane, N. W. S., and Nakamura, C. Y. (1989). "Ethnic Differences in Alcohol Consumption among Asians and Caucasians in the United States: An Investigation of Cultural and Physiological Factors." *Journal of Studies on Alcohol*, 50, 261-267.
- Blaylock, J. R. and Blisard, W. N. (1993). "Wine Consumption by US Men." *Applied Economics*, 25, 645-651.
- Blisard, N. and Blaylock, J. (1993). "Distinguishing Between Market Participation and Infrequency of Purchase Models of Butter Demand." *American Journal of Agricultural Economics*, 75, 314-320.
- Boyd, G. D. (1989). "Beverage Alcohol, 1989 and Beyond." *Restaurant Hospitality*, 73, 85-116.
- Capps Jr., O. and Cheng, H. (1986). "The Missing Income Problem in Analyses of Engel Functions." *Western Journal of Agricultural Economics*, 11, 31-39.
- Danko, W. D. (1990). "An Empirical Evaluation of the Gilly-Enis Updated Household Life Cycle Model." *Journal of Business Research*, 21, 39-57.
- DeLuca, M. (1991). "Coffee, Tea for Me." *Restaurant Hospitality*, 75, 148-150.
- Djata, B. (1987). "The Marketing of Vices for Black Consumers." *Business and Society Review*, 62, 47-49.

- Exter, T. (1991). "Boozing Boomers." *American Demographics*, 13, 6.
- Godfrey, C. (1989). "Factors Influencing the Consumption of Alcohol and Tobacco: The Use and Abuse of Economic Models." *British Journal of Addiction*, 84, 1123-1138.
- Heckman, J. J. (1976). "The Common Structure of Statistical Models of Truncation, Sample Selection and Limited Dependent Variable and a Simple Estimator for Such Models." *Annals of Economic and Social Measurement*, 5, 475-492.
- Heckman, J. J. (1979). "Sample Selection Bias as a Specification Error." *Econometrica*, 47, 153-161.
- Interagency Board for Nutrition Monitoring and Related Research (1993). *Nutrition Monitoring in the United States: Selected Findings from the National Nutrition Monitoring and Related Research Program*. Ervin, B. and Reed, D., eds., Hyattsville, Maryland.
- Lutz, S. M., Smallwood, D. M., Blaylock, J. R., and Hama, M. Y. (1992). *Changes in Food Consumption and Expenditures in American Households during the 1980s*. U.S. Department of Agriculture Statistical Bulletin No. 849, Human Nutrition Information Service and Economic Research Service, Washington, DC.
- Maddala, G. S. (1983). *Limited Dependent and Qualitative Variables in Econometrics*. Cambridge University Press, Cambridge.
- Maxwell, J. C. (1987). "Excise Tax Hurts Liquor Sales." *Advertising Age*, 58, 52.
- Patterson, L. (1992). "Young Adults' Perception of Warnings and Risks Associated with Alcohol Consumption." *Journal of Public Policy and Marketing*, 11, 96-103.
- Penzer, E. (1990). "Ringling in a New Era." *Incentive*, 164, 48-51.
- Walsh, B. M. (1982). "The Demand for Alcohol in the UK: A Comment." *Journal of Industrial Economics*, 30, 439-446.