DOES THE US CIGARETTE TAX IMPOSED ON TOBACCO EXPENDITURE WORK

Katawoot Jeambua [1]
Lecturer at Faculty of Business Administration,
Rajamangala University of Technology Thanyaburi, Thailand
katawoot.j@rmutt.ac.th

Abstract

The cigarette and smoking are the one of the important health and socioeconomic issues in both national and global aspects. The centralized policy such as government policy and education system can influence the people to be healthy and active in self-prevention. The cigarette excise tax is one of the state-level tools to capture the externalities of being smoker. Also, the revision of the effect of cigarette tax on tobacco expenditure should be concerned. Therefore, we deal with the data from the Consumer Expenditure from 2009 to 2011. We find that most of variables are significant with exception of quarterly and yearly dummy variables which do not seem to influence tobacco consumption. The results yield that the tobacco spending of CUs is affected by the cigarette tax significantly. Especially, in 2011, some state had cigarette tax increased, the results indicated that the household decreased the tobacco expenditure significantly. Furthermore, the differences in gender, ethnic group, marital status and region had the unique effect on the tobacco expenditure. Age, education and family size inversed U-shape pattern to tobacco expenditure. Nevertheless, income level progressively increased the tobacco expenditure. Finally, we found that most of explanatory variables and especially cigarette tax can significantly affect the probability of being smoker obviously. This paper considers the policy implications to reduce the cigarette spending by influencing the members of the family to avoid consuming tobacco, increase the cigarette tax to reduce the smoke consumption, communicate the hazardous of smoking and being smoker to the earners of the family to cut the tobacco expenditures, as well as take care of the middle age, middle of education, and the smokers who came from the middle size family whenever starting the government campaign.

Keywords: Cigarette Tax, Tobacco Expenditure

[1] *Corresponding author*: Rajamangala University of Technology Thanyaburi,
Rangsit-Nakhonnayok Rd. (Klong6), Thanyaburi, Pathum Thani, Thailand 12110

Introduction

Cigarette smoking has been identified as the leading cause of preventable morbidity and premature mortality in the United States. Smoking is responsible for approximately one in five deaths in the United States. From 2000 to 2004, smoking had killed an average of approximately 443,000 people each year in the United States alone. This included an estimated 269,655 male and 173,940 female deaths annually. Among adults, most smoking-attributable deaths were due to lung cancer (125,522), coronary heart disease (80,005) and chronic obstructive pulmonary disease and other airway obstruction (78,988) (American Lung Association, 2011). One study from Bureau of Economic Analysis estimated that a greater decline in the smoking rate would offer significant reductions in the costs of smoking. Decreasing the smoking rate to 15 percent by 2023, instead of the 19 percent predicted by current trends, would offer $31.4 billion in savings on pulmonary conditions due to smoking and an increase in productivity of $79 billion (Centers for Disease Control and Prevention, 2009). Moreover, consumers spending on tobacco rose from $80 billion in 2008 to $98 billion in 2011 in inflation-adjusted dollars — even though the amount of tobacco purchased fell 11%. Higher taxes accounted for about half that spending increase. The rest went to tobacco companies and retailers.

However, the effect of cigarette excise taxes on the behavior of smokers have been long debated. The government policy and education system to influence the people to be healthy and active in self-prevention. Nevertheless, the individuals may not well behave in consumptions to select the optimal choice. The risk of undesirable manner in tobacco consumption results in the health diseases problem, for instance heart-attack; cancer and lung diseases. Also, these issues play the important roles in both national level and worldwide aspects. Since the assumption of the consumers has changed their preferences over time, the revision of policy should be in concern. In this paper, therefore, we deal with the data of the Consumer Expenditure from 2009 to 2011 and consider some detail of household characteristic issues. The regional fixed-effect is in consideration in our analysis. Finally, we will apply the model to estimate the probability of being smoker.

Literature Reviews

There is an extensive amount of literature that has investigated the factors affecting cigarette consumption. Traditionally, the addiction models which analyze based on the current consumption decisions depend on past choices and gradual response to the price that implying long-run price elasticity will exceed the short-run elasticity. There are two basic approaches which are Myopic demand (Mullahy, 1986) that dependence of current consumption is on past behavior, but ignores the future consequences of addiction model and Rational demand models (Becker & Murphy, 1988) which allows the farsighted behavior, implying the increases in future "costs" of addiction. The study of Van Walbeek (2003) indicated that real cigarette excise tax rates had fallen because tax increases had not kept pace with inflation. Increases in the price of cigarettes decrease

smoking, particularly by the younger. One result was that taxes had no significant effect on the percentage of adult smokers in a state population. However, for underage smokers, cigarette taxes do curb behavior by one-year lagged tax. However, two studies, by Wasserman et al. (1991) about the youth and Chaloupka (1991) about the young adults, found no significant effect of price on youth smoking.

The researches on race/ethnicity and socioeconomic status of Evan, et al. (1978) found significant differences in price elasticity across different racial/ethnic populations. Moreover, racial/ethnic differences may reflect differences due to socioeconomic status which are supported by U.K-based data of Townsend, et al. (1994) and U.S.-based data of Farrell, et al. (1998) which revealed the similar evidence that at or below median income person has sensitivity to price more than the median income. Moreover, Chaloupka (1991) finds that less educated person is fairly sensitive to price, while more educated is virtually insensitive to price. Additionally, mixed evidence on differences in price sensitivity by gender which are mostly studied in U.S. concluded that men are more sensitive to price than women, but the result is controversy in U.K. papers. However, relatively small literature concludes that increases in taxes and prices on other tobacco products would lead to similar reductions in prevalence and consumption (Chaloupka & Wechsler, 1995) - prevalence and frequency of smokeless tobacco used among young males in U.S. The study is similar to the Ohsfeldt & Boyle (1994; 1998) found that the prevalence of smokeless tobacco used among adult males in U.S. In general, these studies also conclude that various tobacco products are substitutes for one another.

Methodology

1. Research design

In this paper we employed tobit model (Tobin, 1958) to find the relationship between a non-negative tobacco expenditure and independent variables. The case that an observed zero spending can occur either because the household genuinely does not purchase the good, or because, for one reason or another, a zero is incorrectly reported. Which is in fact the case is not known in advance so that the contamination has to be coped with statically by employing tobit regression as the result from Deaton and Irish (1984) paper. The data both household characteristics (Consumer Units: CUs) and tobacco expenditure based on the Consumer Expenditure Survey 2009-2011 (CEX2009-CEX2011) were used. The CEX data was quarterly gathered by Bureau of Labor Statistics in each year. We employed the state cigarette excise tax 2009-2011 data from the centers for disease control and prevention. Firstly, we estimated tobacco expenditure of US household, the basic model we applied from the model in Deaton and Irish (1984). Also, we extended to study on both different methods which are OLS, tobit with lower limit and upper limit and dealt with the model specification which square terms of explanatory variables and the difference in the specification of the model in lower and upper limit (left and right-censored) censored. Second, we

predicted the probability of being smoker by both different methods which are OLS, logistic regression and probit regression and differences in specification which have additional square terms to obtain the conditional marginal effect.

First, we found that the quarterly and yearly dummy variables are not significant in our estimation; we implied that the quarterly and yearly dummy variables had not influenced the tobacco expenditure. Alternatively, tobit regression with square terms and with lower and upper limit seemed to fit to our data more than OLS. After we extended the upper limit from tobacco expenditure from percentile of 90 to 99.99, we found that the model seemed to fit the data well, but the larger in magnitude of the coefficient and the higher standard error will be the issues. Age, education level and family size had the same pattern that inversed u-shape and positive effect on tobacco consumption. As we focused on marital status, being married and widowed tended to consume less tobacco than single, but divorced and separated CUs tended to spend on tobacco more than single. We found the negative impact on tobacco demand from being the Black and the Asian comparing to the White, but being Multiracial had the positive impact on tobacco spending. Especially, the tobacco demand is progressive with respect to income. Cigarette tax had a negative impact on tobacco expenditure. Moreover, the dummy variable which indicated CUs in 2011 who lived in the state that increased cigarette tax had less tobacco expenditure than those CUs in the state which did not increase cigarette tax. CU who lived in urban area tended to spend less than CUs who lived in rural area. The last but not least, CUs who lived in Northeast, Midwest and South tended to have tobacco spending more than CUs who live in West.

Second, we developed the model to predict the probability of being smoker. We began with the estimation by using OLS, logistic regression and probit regression. With assumption of non-linearity of probabilities of being smoker, the probit regression yields the maximum likelihood. Therefore, we developed probit regression with square terms and calculate the conditional marginal effect of the explanatory variables on probabilities of being a smoker. We found that being male had more probability of being a smoker than female approximately 0.0256. Married and widowed CUs had probability of being a smoker more than single CUs by 0.052 and 0.028 respectively. The Black and the Asian had less probability of being a smoker than the White about 0.0686 and 0.0515, respectively, comparing to the White, but Multiracial CUs had more probability of being a smoker than White CUs by 0.0816. Urban life had possibility to be a smoker less than rural life about 0.1029. The CU's household who lived in Northeast, Midwest and South had probability of being smoker more than CUs who lived in West by 0.0508, 0.0728 and 0.0512, respectively.

We concluded that the explanatory variables in the models influenced the tobacco expenditure as we expected. The quarters and year did not significantly affect tobacco expenditure. The cigarette tax is the one of the important factors in consideration of the tobacco expenditure. The different race/ethnic and marital status had different tobacco consumption pattern. There is the exist of regional fixed-effect on CUs tobacco expenditure. We discussed a number of influential factors for

this result in terms of household characteristics. Ultimately, we explored the variables and concluded that cigarette taxes had significantly an effect on tobacco expenditure.

2. Data and Model Specification

2.1 Household tobacco expenditure

$$Tobacco_i = \alpha + \beta X_i + \gamma C_i + \Omega I_i + \delta V_i + \lambda Q_r + \theta Y_t + \mathcal{E}_i$$

Where, $Tobacco_i$ is the household tobacco expenditure in *r*th quarter and year *t*th of *i*th CUs and, X_i the vector of CUs characteristics such as education level, number of members in family, C_i the cigarette excise tax in each state at the particular year. I_i is annual financial income after tax of CUs household, V_i a dummy variable indicates the geographical factor-region, Q_r a dummy variable indicate quarter that CUs data collected. Y_t a dummy variable indicates the year that CUs data collected, $\alpha, \beta, \gamma, \Omega, \delta, \lambda, \theta$ the parameters to be determined, and $\mathcal{E}_i$ the unobserved random errors.

2.2 Probability of being smoker model

$$P(Smoker)_i = \pi + \tau X_i + \chi C_i + \rho I_i + \sigma V_i + \kappa Q_r + \psi Y_t + \eta_i$$

Where, X_i the vector of CUs characteristics such as education level, number of members in family, C_i the cigarette excise tax in each state at the particular year. I_i is annual financial income after tax of CUs household, V_i a dummy variable indicates the geographical factor-region, Q_r a dummy variable indicate the quarter that CUs data collected. Y_t a dummy variable indicates the year that CUs data collected, $\pi, \tau, \chi, \rho, \sigma, \kappa, \psi$ the parameters to be determined, and η_i the unobserved random errors.

2.3 Data

The Consumer Expenditure Survey (CEX) 2009-2011 is the main data source in our analysis which provides a continuous and comprehensive flow of data on the buying habits of American consumers. It is gathered annually by the U.S. Census Bureau for the Bureau of Labor and Statistics (BLS). Also, it has varieties of new releases, reports, articles in the Monthly Labor Review. Their public used microdata file presenting detailed expenditure and income data for the diary components of the CE and the interview components of the CE for 2009-2011. The former includes weekly expenditure (EXPN), annual income (DTB) and imputed income (DTID) files. These files are categorized by a Universal Classification Code (UCC). The latter yields data on up to 95 percent of total household expenditure which are FMLY, MEMB, MTBI, FPAR, MCHI, ITBI and ITLL files.

We used EXPN and DTBD to perform and conducted the income analysis and its relevance. The FMLY and MEMB file are composed of demographic of consumer units (CUs) and CU members. These two files contain summary level of expenditure and income on the FMLY files granted relatively to consumer spending, by general expenditure category. We utilized its data and selected the substantial household characteristics and demographics to implement in our analysis.

The FPAR and MCHI datasets were grouped as 2-year datasets (2010 and 2011), plus the first quarter of the 2012 and contained paradata about the interview survey. We also realized it as the source of the interview survey process.

However, some collected data are not reported. The CE implemented multiple imputations process of income data since 2004. Many income variables and other income related variables are included in FMLY, MEMB, ITBI and ITII files. The topcoding refers to the replacement of data in cases where the value of the original data exceeds prescribed critical value or the cases that CE will not able to identify CUs who participated. Moreover, the CE concerns about geographical issues, some state has less people comparing to its area. In terms of characteristic and imputed data, we employed only completed data to be analyzed in our paper. Also, we assumed that the analysis of the state in the same region can imply the same pattern by the inferential statistics process. This means that the behavior, matter and evidence of other Western states (California, Nevada, Oregon, Utah, and Washington) can be implied to the pattern in Nevada. Additionally, the state cigarette excise tax 2009-2011 data comes from the centers for disease control and prevention.

All of the independent variables are provided by the CEX2009, 2010 and 2011. In addition to the vector of characteristics of consumer unites such as variables on age, race, gender, family size, number of earners, marital status and urban or rural residence, the following variables need some explanation.

The data on the CUs' level of education is compiled into the following discrete groups: never attended school; 1st through 8th grade; 9th through 12th grade, but no high school diploma; high school graduate; some college, less than college graduate; Associate's degree (occupation, vocational, or academic); Bachelor's degree, Master's degree; or Professional or Doctorate degree. We ranked the levels in order from lowest level of education to the highest one (from 0 to 17).

3. Description of Variables:

Dependent Variables:

1) Basic model

   The CEX 2009-2011 reports the tobacco expenditure of CUs in quarter in unit of dollars

   $$Tobacco\ expenditure_i\ , unit\ of\ dollars\ in\ quarter$$

2) Probability of being smoker

   We derive probability of being smoker by identify as a dummy variable, DummyT.

   $DummyT_i$ =1 if the *i*th CU have tobacco expenditure more than 0 dollars, 0 if CU have reported no tobacco expenditure.

Independent Variables:

$Male_i$ = 1 if the *i*th CU is male, 0 otherwise

Age_i = the age of the *i*th CU in years

$Education_i$	= the level of education of the *i*th CU
$Married_i$	= 1 if the *i*th CU married, 0 otherwise
$Widowed_i$	= 1 if the *i*th CU widowed, 0 otherwise
$Divorced_i$	= 1 if the *i*th CU divorced, 0 otherwise
$Separated_i$	= 1 if the *i*th CU separated, 0 otherwise
$Family\ size_i$	= the number of members in the *i*th CU's household
$Earners_i$	= the number of earners in the *i*th CU's household
$Black_i$	= 1 the *i*th CU is African American, or Black, 0 otherwise
$Native_i$	= 1 the *i*th CU is American Indian, or Alaskan Native, 0 otherwise
$Asian_i$	= 1 if the *i*th CU is Asian, 0 otherwise
$Hawaiian_i$	= 1 if the *i*th CU is Hawaiian, 0 otherwise
$Multirace_i$	= 1 if the *i*th CU is Multi-race, 0 otherwise
$Income_i$	= the annual financial income after tax of the *i*th CU's household
$Ln(income)_i$	= the logarithmic term of the annual financial income after tax of the *i*th CU's household
$Cigtax_i$	= state cigarette tax for *i*th CU's household in the particular year, unit of dollars per pack.
$Increasecigtax2011_i$	= 1 if the particular state had increase cigarette tax in 2011, 0 otherwise
$Urban_i$	= 1 if the *i*th CU resides in urban area, 0 otherwise
$Northeast_i$	= 1 if the *i*th CU resides in Northeast region, 0 otherwise
$Midwest_i$	= 1 if the *i*th CU resides in Midwest region, 0 otherwise
$South_i$	= 1 if the *i*th CU resides in South region, 0 otherwise
$Quarter_r$	= 1 if *i*th CU's household information collected in quarter *r*th, 0 otherwise, where r= 1, 2, 3, 4
$Year_t$	= 1 if *i*th CU's household information collected in year *t*th, 0 otherwise, where t = 2009, 2010, 2011

Research Finding

1. Tobit Regression

First, we had tested the independent variables which were significant with the exception of the *Native, Hawaiian, Quarter1, Quarter2, Quarter3, Year 2010,* and *Year 2011* variables. Therefore, we employed the tobit regression with differences in upper-limit (right-censored) which will give us more uncensored observation than we did in previous session. In other words, we have more variation in our model and we would like to see the changes in magnitude of the effect of

explanatory variable. Therefore, we considered the model at 95th, 99th and 99.99th percentile as the following

Table 1 Results of tobit regression restricted to lower limit but different specification in upper limit.

Model	**Percentile 95**			**Percentile 99**			**Percentile 99.99**		
Observation	41,961			41,961			41,961		
Uncensored Observations	6,001			8,148			8,523		
Left-Censored Obs/Tobacco ≤ Right-Censored Obs/Tobacco ≥	33,434/0 2,526/260			33,434/0 379/606.67			33,434/0 4/3,289.167		
Sigma	336.1954			348.1729			397.619		
Pseudo R^2	0.0232			0.0186			0.0188		
	Coeff.	Robust S.E.	P-Value	Coeff.	Robust S.E.	P-Value	Coeff.	Robust S.E.	P-Value
Constant	-851.65	(93.05)	0.000***	-851.65	(93.05)	0.000***	-961.79	(110.99)	0.000***
Male	24.99	(4.91)	0.000***	24.99	(4.91)	0.000***	29.29	(5.67)	0.000***
Age	16.33	(0.92)	0.000***	16.33	(0.92)	0.000***	18.26	(1.08)	0.000***
Age^2	-0.18	(0.01)	0.000***	-0.18	(0.01)	0.000***	-0.20	(0.01)	0.000***
Education	91.36	(12.91)	0.000***	91.36	(12.91)	0.000***	100.91	(15.49)	0.000***
Education^2	-5.30	(0.49)	0.000***	-5.30	(0.49)	0.000***	-5.88	(0.59)	0.000***
Married	-69.00	(7.83)	0.000***	-69.00	(7.83)	0.000***	-76.39	(8.98)	0.000***
Widowed	-44.12	(11.93)	0.000***	-44.12	(11.93)	0.000***	-51.95	(13.50)	0.000***
Divorced	46.35	(8.20)	0.000***	46.35	(8.20)	0.000***	52.19	(9.39)	0.000***
Separated	-1.71	(13.83)	0.902	-1.71	(13.83)	0.902	-3.54	(15.52)	0.820
Family size	38.88	(6.23)	0.000***	38.88	(6.23)	0.000***	41.17	(7.04)	0.000***
Family size^2	-4.03	(0.83)	0.000***	-4.03	(0.83)	0.000***	-4.02	(0.96)	0.000***
Black	-107.25	(7.64)	0.000***	-107.25	(7.64)	0.000***	-118.49	(8.84)	0.000***
Native	27.31	(37.13)	0.462	27.31	(37.13)	0.462	35.38	(42.72)	0.407
Asian	-81.61	(12.81)	0.000***	-81.61	(12.81)	0.000***	-90.22	(15.03)	0.000***
Hawaiian	18.00	(37.52)	0.631	18.00	(37.52)	0.631	16.01	(42.02)	0.703
Multirace	88.04	(17.85)	0.000***	88.04	(17.85)	0.000***	105.69	(21.84)	0.000***
Ln(income)	7.04	(2.07)	0.001***	7.04	(2.07)	0.001***	7.93	(2.37)	0.001***
Cigtax	-13.50	(4.01)	0.001***	-13.50	(4.01)	0.001***	-13.85	(4.57)	0.002***
Increasecigtax2011	-30.72	(13.23)	0.020**	-30.72	(13.23)	0.020**	-36.93	(15.21)	0.015**
Urban	-161.17	(27.48)	0.000***	-161.17	(27.48)	0.000***	-178.07	(31.04)	0.000***
Northeast	66.70	(9.11)	0.000***	66.70	(9.11)	0.000***	73.02	(10.41)	0.000***
Midwest	81.06	(7.28)	0.000***	81.06	(7.28)	0.000***	90.61	(8.53)	0.000***
South	63.24	(6.60)	0.000***	63.24	(6.60)	0.000***	70.63	(7.71)	0.000***

Note: *** Significant level at $\alpha = 0.01$

Table 1 shows that all the independent variables are significant with exception of the *Separated*, *Native* and *Hawaiian* variables. However, *Increasecigtax2011* variable is significant at 90% significant level in the right-censored at tobacco ≥ 151.67 dollars (90th percentile) but it is 95% significant level in another models. The direction of the effect of explanatory variables on tobacco expenditure are the same as we did. The magnitude of coefficient of *Ln(income)* are pretty higher

which are 5.96, 7.04 , 7.04 and 7.93 respectively. But, the magnitude of coefficient of *Cigtax* tends to be lower as we increase the i^{th} percentile which are 17.27, 13.50, 13.50 and 13.85 respectively. Moreover, dummy variable *Increscigtax2011* tends to be higher which are 22.12, 30.72, 30.72 and 36.93, respectively. The ancillary statistic /sigma is analogous to the square root of the residual variance in OLS regression. The values are 327.10, 336.19, 348.17 and 397.61, respectively which indicate that when we try to increase uncensored observations in the tobit model, we have substantially increased variation to the models.

Table 2 The test statistics for parameters in tobit regressions

Group of variables	**F-statistic**			**P-value**		
	95^{th} percentile	99^{th} percentile	99.99^{th} percentile	95^{th} percentile	99^{th} percentile	99.99^{th} percentile
Marital status	79.29	78.61	74.45	0.0000	0.0000	0.0000
Races	29.43	29.48	26.40	0.0000	0.0000	0.0000
Regions	3.96	3.90	3.81	0.0191	0.0203	0.0223

Note: Marital status = H_0 : Married=Widowed=Divorced=Separated;

Races = H_0: Black=Native=Asian=Hawaiian=Multirace;

Regions = H_0:Northeast=Midwest=South

To test the null hypotheses of the coefficients of the diffence in marital status, races and regional fixed-effect, we obtain F-statistics and p-values as shown in Table 2. We have enough evidence to reject the null hypothesese. Therefore, we can conclude that being Married, Widowed, Divorced and Seperated have different effects on tobacco expenditure at 99% significant level. Being Black, Asian, Native, Multiracial, Hawaiian have differrent effects on tobacco expenditure at 99% significant level. However, we have the difference of regional fixed-effect impact on tobacco consumption at 95% significant level in all models with exception for the model which 99^{th} percentile upper limit that the regional fixed-effect have different impacts on tobacco consumption at 99% significant level.

2. Probability of being smoker model.

We introduced the model which predicts the probability of being smoker. Therefore, we generated a dummy variable as dependent variable. *DummyT* defined as 1 if CUs have expenditure more than zero and 0, otherwise.

Table 3 Results of the model of probability of being smoker.

Model	**Logistic Regression**			**Probit Regression**		
Observation	41,961			41,961		
Loglikelihood	-20202.22			-20191.9		
F-test / Chi-square	1784.08			1864.63		
Prob>F / Prob.> X^2	0.000			0.000		
R^2 / Pseudo R^2	0.0683			0.0699		
	Coefficient	Robust S.E.	P-Value	Coefficient	Robust S.E.	P-Value
Constant	1.629	(0.195)	0.000***	0.959	(0.116)	0.000***
Male	0.149	(0.026)	0.000***	0.085	(0.015)	0.000***
Age	-0.015	(0.001)	0.000***	-0.009	(0.001)	0.000***
Education	-0.181	(0.008)	0.000***	-0.107	(0.004)	0.000***
Earners	0.036	(0.006)	0.000***	0.020	(0.003)	0.000***
Married	-0.168	(0.040)	0.000***	-0.096	(0.023)	0.000***
Widowed	-0.318	(0.066)	0.000***	-0.180	(0.036)	0.000***
Divorced	0.497	(0.042)	0.000***	0.288	(0.025)	0.000***
Separated	0.250	(0.073)	0.001***	0.147	(0.043)	0.001***
Family size	0.040	(0.010)	0.000***	0.024	(0.006)	0.000***
Black	-0.416	(0.041)	0.000***	-0.237	(0.023)	0.000***
Native	0.241	(0.183)	0.187	0.137	(0.107)	0.199
Asian	-0.480	(0.071)	0.000***	-0.254	(0.038)	0.000***
Hawaiian	0.064	(0.196)	0.743	0.025	(0.112)	0.821
Multirace	0.501	(0.093)	0.000***	0.287	(0.056)	0.000***
Ln(income)	0.049	(0.011)	0.000***	0.028	(0.006)	0.000***
Cigtax	-0.137	(0.020)	0.000***	-0.072	(0.011)	0.000***
Urban	-0.720	(0.137)	0.000***	-0.436	(0.085)	0.000***
Northeast	0.389	(0.048)	0.000***	0.202	(0.027)	0.000***
Midwest	0.474	(0.038)	0.000***	0.263	(0.022)	0.000***
South	0.353	(0.035)	0.000***	0.194	(0.020)	0.000***

Note: *** Significant level at $\alpha = 0.01$

The Table 3 shows that all of the independent variables are significant with the exception of the *Native and Hawaiian* variables. Being male tends to be smoker higher than female. The higher age and additional year of education tend to decrease the probability of being smoker. The increases in number of earner and family size of CUs tend to increase the probability of being smoker. Being Black and Asian tends to decrease the probability of being smoker compare to White. Conversely, being Multiracial tends to increase the probability of being smoker compare to White. The percentage increase of income tends to increase the probabilities of being smoker. The increase in cigarette excise tax tends to decrease probability of being smoker. CUs who live in urban area tends to be smoker less than CUs who live in rural area. The CU's household who live in Northeast, Midwest and South tends to be the smokers more than CUs who live in West.

Finally, we establish probit regression with the square term of age, education and family size variables and conditional marginal effects by employing Delta-method. The estimation equation and the result are the following.

Table 4 Results of the probit regression with square term and its marginal effect.

	Probit			**Conditional Marginal Effect**			
Observation	41,961			41,961			
Loglikelihood	-19783.355			-19783.355			
Wald X^2 (23)	2359.58			2359.58			
Prob.> X^2	0.000			0.000			
Pseudo R^2	0.0661			0.0661			
AIC	0.944			N/A			
LR (23)	2799.112			N/A			
	Coefficient	Robust S.E.	P-Value	dF/dX	Robust S.E.	P-Value	Mean of Variables
Constant	-2.4994	(0.2624)	0.000***	-	-	-	-
Male	0.0958	(0.0149)	0.000***	0.0256	(0.0040)	0.000***	0.4750
Age	0.0495	(0.0028)	0.000***	0.0132	(0.0008)	0.000***	49.9676
Age^2	-0.000585	0.000029	0.000***	-0.00016	0.00000764	0.000***	2797.5090
Education	0.2549	(0.0351)	0.000***	0.0680	(0.0093)	0.000***	13.2836
$Education^2$	-0.0150	(0.0013)	0.000***	-0.0040	(0.0004)	0.000***	179.9807
Earners	0.0330	(0.0034)	0.000***	0.0088	(0.0009)	0.000***	3.6297
Married	-0.1957	(0.0237)	0.000***	-0.0524	(0.0064)	0.000***	0.5166
Widowed	-0.1122	(0.0363)	0.002***	-0.0287	(0.0089)	0.002***	0.0924
Divorced	0.1470	(0.0253)	0.000***	0.0410	(0.0074)	0.000***	0.1509
Saparted	0.0104	(0.0433)	0.811	0.0028	(0.0117)	0.811	0.0302
Family size	0.0838	(0.0180)	0.000***	0.0224	(0.0048)	0.000***	2.5369
Family $size^2$	-0.010181	(0.0023)	0.000***	-0.00272	(0.0006)	0.000***	8.7080
Black	-0.2846	(0.0235)	0.000***	-0.0686	(0.0050)	0.000***	0.1220
Native	0.0905	(0.1065)	0.395	0.0251	(0.0307)	0.395	0.0041
Asian	-0.2111	(0.0383)	0.000***	-0.0515	(0.0085)	0.000***	0.0522
Hawaiian	0.0300	(0.1129)	0.791	0.0081	(0.0309)	0.791	0.0039
Multirace	0.2742	(0.0559)	0.000***	0.0816	(0.0183)	0.000***	0.0139
Ln(income)	0.01953	(0.0064)	0.002***	0.005212	(0.0017)	0.002***	10.4253
Cigtax	-0.0694	(0.0111)	0.000***	-0.0185	(0.0030)	0.000***	1.4425
Urban	-0.3374	(0.0867)	0.000***	-0.1029	(0.0295)	0.000***	0.9945
Northeast	0.1820	(0.0268)	0.000***	0.0508	(0.0078)	0.000***	0.2144
Midwest	0.2557	(0.0219)	0.000***	0.0728	(0.0066)	0.000***	0.1977
South	0.1865	(0.0200)	0.000***	0.0512	(0.0056)	0.000***	0.3236

Note: *** Significant level at α =0.01

The ***italic*** variables in the Table 4 indicate that marginal effect refers to the discrete change of dummy variable from 0 to 1 and P>|z| correspond to the test of the underlying coefficient being 0.All of the independent variables are significant with the exception of the *Separated, Native and Hawaiian v*ariables. Being male have probability of being smoker higher than female approximately 0.0256. Married and Widowed CUs have probability of being smoker more than single CUs by 0.052 and 0.028, respectively. Black and Asian have probability of being smoker less than White about 0.0686 and 0.0515 respectively, comparing to White but Multiracial CUs have probability of being smoker higher than White CUs by 0.0816. The CUs who live in urban area have possibility to be smoker less than CUs who live in rural area by 0.1029. The CU's household who live in

Northeast, Midwest and South have probability of being smoker more than CUs who live in West by 0.0508, 0.0728 and 0.0512 respectively.

Test Statistic for estimation of probability of being smoker.

Since, the logistic regression measures the relationship between a categorical dependent variable, being smoker, and independent variables, by using probability scores as the predicted values of the dependent variable. The regression coefficients are usually estimated using maximum likelihood estimations which have shown in the Table 5. The likelihood-ratio test which assess model fit is also the recommended procedure to assess the contribution of individual "predictors" to a given model which are all reject the null hypothesis that the parameters are significant at 99% significance level. The Akaike information criterion (AIC) is a measure of the relative quality of a statistical model, for a given set of data. As such, AIC provides a mean for model selection.

Table 5 Test statistics for logistic regression and probit regression model

Model	**Test Statistics**					
	Log-likelihood	Log-likelihood Ratio	Prob > LR	AIC	AIC*n	BIC
Logistic Regression	-20202.220	1961.361	0.0000	0.964	40446.441	-406025.717
Probit Regression	-20191.898	1982.026	0.0000	0.963	40425.796	-406046.362
Probit Regression with square term	-19783.355	2799.112	0.0000	0.944	39614.710	-405831.515

From the Table 5, we found that the lower AIC, the better of quality of the model. Therefore, probit regression with square has the lowest AIC, 0.944. Bayesian information criterion (BIC) simply reduces to maximum likelihood selection because the number of parameters is equal for the models of interest. The larger BIC, the better of model fit to the data. The probit regression with square-term has the highest value of BIC,-405831.515. Therefore, we can conclude that probit regression with square term is preferable.

To test the null hypotheses of the coefficients of the diffence in marital status, races and regional fixed-effect, we obtain Chi-square and p-values as in the Table 6. We have enough evidence to reject the null hypotheses. Hence, we can conclude that being Married, Widowed, Divorced and Seperated have diffent effects on being smoker at 99% significant level. Being Black, Asian, Native, Multiracial, Hawaiian have differrent effects on being smoker at 99% significant level. Also, we have the difference of regional fixed-effect impact on probabilities of being smoker at 99% significant level.

Table 6 Test statistic for parameters for logistic regression and probit regression

Group of variables	Chi-square			P-value		
	Logistic Regression	Probit Regression	Probit Regression with square term	Logistic Regression	Probit Regression	Probit Regression with square term
Marital status (H_0 : Married=Widowed =Divorced=Separated)	392.21	387.57	227.98	0.0000	0.0000	0.0000
Races (H_0: Black=Native=Asian =Hawaiian=Multirace)	102.92	93.74	98.96	0.0000	0.0000	0.0000
Regions (H_0:Northeast=Midwest =South)	12.52	13.01	14.36	0.0019	0.0015	0.0008

Summary and Conclusions

This paper investigated the impact of cigarette excise tax on tobacco expenditure of US household from 2009 to 2011. We identified and up-to-date tobacco expenditure by employing the dummy variables and studies in 1980s and 1990s. We found that most of variables were significant with exception of quarterly and yearly dummy variables which did not seem to influence tobacco consumption. The results yielded that the tobacco spending of CUs were affected by the cigarette tax significantly. The difference in gender, ethnic group, marital status and region had the unique effect on the tobacco expenditure. Age, education and family size has inversed U-shape pattern to tobacco expenditure. However, income level progressive increased the tobacco expenditure.

Possible Extension and Limitation of the Study

There are exogenous variables effect on tobacco expenditure such as price of another kind of tobacco, personal preferences, advertising and promotion or even anti-smoking propaganda, both in local and national aspect. Cross sectional data have the limitation of time to forecast in the long-run tobacco consumption so that we cannot track the household along the time. Moreover, the use of total expenditure as an exogenous variable is theoretically inconsistent with our formulations of 'reporting' bias which imply that total expenditure is a random variable determined by the sum of all the reporting effects over all goods. This issue can only really be dealt with in the context of a system of demand equations and is central in Kay, Keen (1986) paper. Also, AIDS (Deaton & Muellabauer, 1980) and Quadratic Engel Curves and Consumer Demand (Banks, Blundell & Lewbel 1997) are possible.

References

American Lung Association. (2011). *Trends in tobacco use*. Washington, DC: American Lung Association.

Banks, J., Blundell, R., & Lewbel, A. (1997). Quadratic Engel curves and consumer demand. *Review of Economics and statistics, 79*(4), 527-539.

Becker, G. S., & Murphy, K. M. (1988). A theory of rational addiction. *Journal of political Economy, 96*(4), 675-700.

Centers for Disease Control and Prevention. (2009). *State Cigarette Excise Taxes, United states, 2009-2011*. Accessed November 5, 2013 from http://www.cdc.gov/mmwr/preview/mmwrhtml/mm5819a2.htm

Chaloupka, F. (1991). Rational addictive behavior and cigarette smoking. *Journal of political Economy, 99*(4), 722-742.

Chaloupka, F. J., & Wechsler, H. (1995). Price, tobacco control policies and smoking among young adults (No. w5012). *National Bureau of Economic Research*.

Deaton, A., & Irish, M. (1984). Statistical models for zero expenditures in household budgets. *Journal of Public Economics, 23*(1-2), 59-80.

Deaton, A., & Muellbauer, J. (1980). An almost ideal demand system. *The American economic review, 70*(3), 312-326.

Evans, R. I., Rozelle, R. M., Mittelmark, M. B., Hansen, W. B., Bane, A. L., & Havis, J. (1978). Deterring the Onset of Smoking in Children: Knowledge of Immediate Physiological Effects and Coping with Peer Pressure, Media Pressure, and Parent Modeling 1. *Journal of applied social psychology, 8*(2), 126-135.

Farrell M, et el., 1998, Substance misuse and psychiatric comorbidity: An overview of the opcs national psychiatric morbidity. Additive Behaviors, Vol. 23, pp. 909-918

Keen, M. (1986). Zero expenditures and the estimation of Engel curves. *Journal of Applied Econometrics, 1*(3), 277-286.

Mullahy, J. (1986). Specification and testing of some modified count data models. *Journal of econometrics, 33*(3), 341-365.

Ohsfeldt, R. L., & Boyle, R. G. (1994). Tobacco excise taxes and rates of smokeless tobacco use in the US: an exploratory ecological analysis. *Tobacco Control, 3*(4), 316.

Ohsfeldt, R. L., Boyle, R. G., & Capilouto, E. L. (1998). Tobacco taxes, smoking restrictions, and tobacco use (No. w6486). National bureau of economic research.

Tobin, J. (1958). Estimation of relationships for limited dependent variables. *Econometrica: journal of the Econometric Society*, 24-36.

Townsend, J., Roderick, P., & Cooper, J. (1994). Cigarette smoking by socioeconomic group, sex, and age: effects of price, income, and health publicity. *Bmj*, 309(6959), 923-927.

Van Walbeek, C. (2003). *Tobacco excise taxation in South Africa*. Geneva: World Health Organization.

Wasserman, J., Manning, W. G., Newhouse, J. P., & Winkler, J. D. (1991). The effects of excise taxes and regulations on cigarette smoking. *Journal of health economics, 10*(1), 43-64.