Factors that Influence the Frequency and Quantity of Tobacco Use Among U.S. Youth

Xuedong Wu
Ph.D. Student
Department of Agricultural and Applied Economics
University of Georgia
bestwu@uga.edu

Gregory Colson
Assistant Professor
Department of Agricultural and Applied Economics
University of Georgia
gcolson@uga.edu

J. Scott Shonkwiler
Professor
Department of Agricultural and Applied Economics
University of Georgia
jss1@uga.edu

Selected Paper prepared for presentation at the Agricultural & Applied Economics Association's 2013 AAEA & CAES Joint Annual Meeting, Washington, DC, August 4-6, 2013.

Abstract

Despite declines in US smoking rates tobacco remains among the leading causes of preventable deaths in the US and is directly linked to chronic and mortal diseases including cancers and heart disease. With a focus on youth smoking behaviors, this paper considers contributory influences for two related, but distinct, measures of youth tobacco addiction - the frequency and quantity of cigarettes consumed per month. Using the 2011 National Youth Tobacco Survey a bivariate ordered probit model is estimated along with associated conditional probabilities controlling for both youth-specific factors and peer group influences. Results indicate strict controls on tobacco access, parental involvement, and school initiatives can significantly reduce the frequency and quantity of smoking by adolescents and decrease the probability of youth becoming heavier smokers.

Key Words: National Youth Tobacco Survey, Smoking frequency, Smoking quantity, Bivariate ordered probit model, Conditional probability.

Factors that Influence the Frequency and Quantity of Tobacco Use Among U.S. Youth

Introduction

Tobacco is among the leading causes of preventable deaths in the US, accounting for approximately 440,000 deaths each year (Marshall et al. 2006), and is directly linked to chronic and mortal diseases including cancers and heart disease (Mokdad et al. 2004, Danaei et al. 2009). While smoking rates have declined in recent decades following widespread information campaigns, private and government outreach efforts, new laws prohibiting smoking in many locations, and substantial cigarette taxes, it is estimated that in 2010 approximately 19% of adults are still regular smokers (CDC 2011). Furthermore, despite significant efforts to reduce smoking rates among U.S. youth, it is estimated that 20% of high school students and 5% of middle school students have tried smoking (CDC 2010, 2011). Efforts to curb smoking have particularly focused on reducing smoking rates among children for a variety of well founded reasons. Research has shown that a significant proportion of adult tobacco users began smoking in their adolescence, rather than as an adult, and this habit tends to continue over a lifetime (Mathers et al. 2006). Furthermore, in addition to the health consequences directly linked to smoking at an early age, research has shown that tobacco is often the earliest stage of drug involvement in adolescence and can be a gateway drug for illicit substances including cocaine and cannabis (Kandel and Yamaguchi 1993, Mathers et al. 2006).

In order to develop programs and policies to reduce childhood smoking it is critical to assess the behavioral and situational factors that are associated with adoption, frequency, and magnitude of smoking among adolescents. In this study we build upon previous research assessing factors influencing childhood smoking with a focus on two key smoking behaviors

among children (ages older than 9) - the total quantity of cigarettes smoked per day and the number of days per month cigarettes are smoked. While these two smoking behaviors - quantity and frequency - are expected to be positively related, they capture two distinct features of childhood smoking and addiction important for policy analysis. We explore using the 2011 National Youth Tobacco Survey (NYTS) conducted by the Centers for Disease Control (CDC) whether youth-specific characteristics, school and peer influences, and situational factors influence not only overall tobacco consumption, but also the frequency (e.g., consistent vs. infrequent binge behavior).

Our empirical methodology offers a number of new insights and improvements to previous research focusing on youth tobacco consumption. In contrast to works that use single independent equation models (e.g., Mowery et al. 2004)to analyze either the quantity or frequency of tobacco use, we employ a bivariate ordered probit model to jointly model youths' tobacco decisions. This approach has two distinct advantages. First, failure to jointly model can lead to biased estimates of the impact of factors influencing youth tobacco use. Second, jointly modeling the quantity and frequency of tobacco usage enables detecting the potential relationship between these two behaviors. Furthermore, by computing the conditional probability, this bivariate model reveals different smoking habits of adolescent which cannot be seen in a univariate model. Additionally, in contrast to previous work using the NYTS, our analysis explicitly controls for school and peer effects. Due to confidentially concerns, the NYTS (like many other youth data sets) does not disclose any identifiable information (e.g., location, racial background, income level) regarding the schools in which survey participants are enrolled. This is problematic for analysis because failure to control for these factors can lead to biased estimates. Using the limited identifiable information available in the NYTS we develop and

incorporate in our analysis a number of school and peer controls (e.g., ratio of smokers in the school, minority ratio of the school, etc.) that have not previously been included in youth tobacco analysis.

Overall, our results indicate a significant positive relationship between smoking frequency and quantity, and identify several important factors that affect those two behaviors. As well, the magnitude of the effects of several key factors related to the conditional probabilities are also analyzed. Results from this study suggest several policy avenues to efficiently reduce adolescent tobacco use.

Data

To assess the impact of individual, peer, and school influences on the quantity and frequency of tobacco consumption among US youth we employ the 2011 National Youth Tobacco Survey (NYTS) conducted by Center of Disease Control and Prevention (CDC). First conducted in 1999, this survey is designed to provide national data on long-term, intermediate, and short-term indicators key to the design, implementation, and evaluation of comprehensive tobacco prevention and control programs (CDC 2011). The survey employs a stratified three-stage cluster sample design, and participation in the survey is voluntary at both the school and student level (CDC 2011). With an 83.2% school participation rate and 88% student participation rate, the data includes 18,866 student respondents of ages older than 9. Given our focus on understanding the factors that influence the gravity of smoking behaviors among US youth, we exclude all individuals who have never tried smoking. This results, after further exclusions of incomplete surveys, in a final sample of 1,293 students (719 males and 574 females).

The key survey responses that we focus upon are “During the past 30 days, on how many days did you smoke” and “During the past 30 days, on the days you smoked, how many cigarettes did you smoke per day”, which describes the frequency and quantity behavior of young smokers, respectively. For the frequency question, answered are coded into four groups listed from “rarely smokers” to “regular smokers”: 1 or 2 days, 3-9 days, 10-29 days and all 30 days per month. Similarly the answers to the quantity question are grouped into 4 groups distinguishing the level of smoking: 1 cigarette per day, 2-5 per day, 6- 20 per day and more than 20 cigarettes per day. Each response variable is ordinally denoted from 1 to 4 with 1 denoting the smallest frequency or number and 4 denoting the highest frequency or number. Table 1 provides a summary of the percentage of the sample falling in each category. As can be seen from the table, for the smoking frequency, there are basically the same number of respondents in each group; while for the quantity, about 75% of young smokers smoke less than 5 cigarettes a day.

From the NYTS a number of survey responses related to factors that are hypothesized to influence smoking behaviors are included in the analysis including demographic information, student awareness of smoking risks, and social environmental factors. These variables are defined and summarized in Table 2. Demographic variables include the age, gender, race, and weekly discretionary income of each student. As well, we construct a variable, GradeBelow, indicating whether the student is behind the normal grade level for their age level. In the sample, about 13% percent of students were below their normal grade level.

To control for smoking history, exposure to smoking, and access and attitudes towards cigarettes, a number of variables are considered. These include AgeFirstSmoke denoting the age when the respondent first tried smoking capturing the important initiation to the smoking habit (Kiernan 2002). In the sample, the average age respondents first tried smoking was about 12

years old. Three measures of how students acquire cigarettes are included indicating whether they purchase their own tobacco or whether other individuals provide access. In the sample, the most common way that students acquired cigarettes was purchasing themselves. To capture self-awareness of the dangers of smoking we consider a measure of how dangerous each student perceives tobacco, whether parents have given advice on the dangers of tobacco, and whether the student feels that tobacco companies promote tobacco to adolescents. In addition, we include a variable to capture whether the student has considered quitting.

Finally, to control for peer and school effects, we construct several variables capturing social environmental factors that may affect smoking behaviors. This includes the percentage of survey respondents in the same school as the survey taker who (a) are willing to wear clothing bearing tobacco company logos, (b) the average number of school days that are skipped, (c) the ratio of minorities, and (d) smoke. The latter variable measuring the percentage of students that smoke in a school, SmokeRatio, is particularly important to consider because it captures the peer influences of attending a school where smoking is more common. As we will further discuss in later sections, previous analyses that do not consider these school-specific effects can lead to biased estimates of the impact of factors including smoking behavior.

Methods

To analyze factors that influence smoking frequency and magnitude we employ a bivariate ordered probit model which can be derived from a latent variable model (Sajaia 2008). In this section we briefly discuss the model. Assume there are N total observations with each individual observation denoted by *i*, and the two interested behavior variables, frequency and quantity of

cigarette smoking, are decided by latent variables y_{i1}^* and y_{i2}^* with the following system equations (Greene and Hensher 2010):

$$\begin{aligned} y_{i1}^* &= \mathbf{x}_{\mathbf{i1}}^{'}\boldsymbol{\beta}_{\mathbf{1}} + \varepsilon_{i1}, y_{i1} = j \text{ if } \mu_{j-1} < y_{i1}^* < u_j, j = 0,..., J_1 \\ y_{i2}^* &= \mathbf{x}_{\mathbf{i2}}^{'}\boldsymbol{\beta}_2 + \varepsilon_{i2}, y_{i2} = j \text{ if } \delta_{j-1} < y_{i2}^* < \delta_j, j = 0,..., J_2 \end{aligned} \text{ and } \begin{pmatrix} \varepsilon_{i1} \\ \varepsilon_{i2} \end{pmatrix} \sim N\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}\right]. \quad (1)$$

In the system of equations above, latent variables y_{i1}^* and y_{i2}^* measure the underlying utilities one obtains from certain smoking behaviors, which are revealed and observed in the forms of categories. x_{i1} and x_{i2} are vector of exogenous variables which may share some common components. and ε_{i1} and ε_{i2} are random disturbances which are assumed to be bivariate normal distributed with mean zero and variance 1. The correlation coefficient of these two random terms is ρ which captures the possible relationship between the two latent variables. The unknown cutoffs that divide the latent utility into several groups satisfy $\mu_0 < \mu_1 < \cdots < \mu_{J_1}$ and $\delta_0 < \delta_1 < \cdots < \delta_{J_2}$. Each of these parameters must be estimated.

To estimate the model maximum likelihood methods are used. Based on the assumptions above, the probability that $y_{i1} = p$ and $y_{i2} = q$ is:

$$\begin{aligned} \Pr(y_{i1} = p, y_{i2} = q \mid \mathbf{x_1}, \mathbf{x_2}) &= \Pr(\mu_{p-1} < y_{i1}^* < u_p,\ \delta_{q-1} < y_{i2}^* < \delta_q \mid \mathbf{x_1}, \mathbf{x_2}) \\ &= \Pr(y_{i1}^* < u_p, y_{i2}^* < \delta_q \mid \mathbf{x_1}, \mathbf{x_2}) \\ &- \Pr(y_{i1}^* < u_{p-1}, y_{i2}^* < \delta_q \mid \mathbf{x_1}, \mathbf{x_2}) \\ &- \Pr(y_{i1}^* < u_p, y_{i2}^* < \delta_{q-1} \mid \mathbf{x_1}, \mathbf{x_2}) \\ &+ \Pr(y_{i1}^* < u_{p-1}, y_{i2}^* < \delta_{q-1} \mid \mathbf{x_1}, \mathbf{x_2}) \end{aligned} \quad (2)$$

Each probability in (2) can be calculated using the bivariate normal distribution function. The log-likelihood of observation *i* is:

$$\ln \ell_i = \sum_{p=1}^{J_1} \sum_{q=1}^{J_2} 1(y_{i1} = p, y_{i2} = q) \ln \Pr(y_{i1} = p, y_{i2} = q \mid \mathbf{x_1}, \mathbf{x_2}) \quad (3)$$

Aggregating over the individual likelihood functions for the entire sample, we derive coefficient estimates using ML. For most cases of discrete choice models the marginal effect of each exogenous variable should be computed to reveal the exact effects on the dependent variable. However in this case, the choice of appropriate margin needs to be considered, since the common bivariate probability $\Pr(y_{i1} = p, y_{i2} = q \mid \mathbf{x_1}, \mathbf{x_2})$ and the derived marginal effects based on this probability may not have much implication on the goal of the analysis. Instead, the following conditional probability and related marginal effects are more interesting:

$$\Pr(y_{i1} = p \mid y_{i2} = q, \mathbf{x_1}, \mathbf{x_2}) = \frac{\Pr(y_{i1} = p, y_{i2} = q \mid \mathbf{x_1}, \mathbf{x_2})}{\Pr(y_{i2} = q \mid \mathbf{x_2})}. \quad (4)$$

The probability in the denominator can be calculated directly fitting univariate ordered Probit model on y_2 since the marginal distribution of bivariate normal distribution is the common univariate normal distribution. The standard error of this conditional probability is also of interest and can be calculated analytically using the delta method. However, since the sample size is large enough in this study, the consistent estimator of standard errors can be achieved by bootstrap methods which is also relatively easy to implement.

Model Estimates

The model estimation results are shown in Table 3. Notice that the explanatory variables for smoking frequency and quantity are not the same and the given values in the table are levels of

estimated coefficients. Although for the ordered probit model the magnitude of coefficient does not have an immediate interpretation, the sign of the coefficient reveals the direction of the effect of certain explanatory variables of interest. However this effect is only determined for the first and last categories (Greene and Hensher 2010), i.e., for ($y_1 = 0$, $y_2 = 0$) (both in the first category), the marginal effect of certain variable is opposite with the sign of its coefficient; while for (($y_1 = J_1, y_2 = J_2$) (both in the last category), the partial effect of explanatory variable is consistent with the sign of its estimated coefficient. For any categories in the middle groups the direction of partial effects cannot be directly determined from the sign of the estimated coefficient. However, it is enough here to reach some brief conclusions based on the analysis of two extreme categories.

For the convenience of discussion, we call those who belong to the fourth group in smoking frequency (y_1) "regular smokers" and those in the same position in daily smoking quantity (y_2) "heavy smokers". The first important implication is the significant negative coefficient of age when smoking was first tried. This indicates, as would be expected, the later in life that children first try smoking the lower the probability they will become both regular and heavy smokers. This could be attributed to the more developed nervous system of older youth which may resist the effect of tobacco. As well, this result tends to agree with previous evidence that has found for adults that as individuals become older, gain more education, and have more life experiences, they will be more likely to smoke less. This indicates that efforts that prevent or delay children from trying or learning about any form of tobacco product can be beneficial. If they can avoid trying tobacco use before 18 years of age, it is unlikely they will ever start smoking (Kiernan 2002).

Besides the age of first trying smoking, there are two more noticeable factors – whether respondent think tobacco use is dangerous and whether they ever think about quitting smoking. The coefficient estimates show the importance of establishing self-confidence and awareness of the harmfulness of tobacco. Regular cigarette use is related to Nicotine addiction which may develop to further serious disease, so it is necessary to strengthen perceptions of the dangers of tobacco, as well as the feasibility of quitting. This will lower the probability of young people becoming heavy and regular smokers.

Although the legal age for purchasing cigarettes is 18, it is clear from the data summary presented in the previous section that a significant percentage of youth smokers are able to purchase cigarettes or easily find others to purchase for them. Looking at the impact of cigarette access on the frequency and quantity of cigarettes smoked by adolescents, it is clear that those individuals who purchase cigarettes are more likely to be heavier smokers whereas those who are given cigarettes are lighter smokers. This result helps reinforce the push for strict controls of minor access to tobacco products and substantial penalties for stores and individuals who sell or give tobacco to underage smokers.

The third types of explanatory variables considered in the analysis are “environmental factors” which include factors related to family and school environment. Not surprisingly, children with people smoking at home have a higher probability of becoming regular and heavy smokers compared to those who do not. Parent advice on tobacco use significantly reduces the probability of children become regular and heavy smokers. These two results indicate the importance of family involvement and habits on controlling youth tobacco use. Parents can help reduce the likelihood of students becoming heavy frequent smokers by setting good habits in

their home and giving appropriate advice and instruction to their children to let them fully understand the risk of tobacco use.

Schools also play a significant role in supervising the behavior of youth. Foremost, the model results indicate that there are significant school peer smoking effects. The likelihood of a student being a heavy or frequent smoker is positively related to the percentage of smokers in the same school. This indicates that schools that are unable to minimize the number of active smokers will increase the probability of other students becoming regular and heavy smokers. So schools should pay attention to their management and regulations on smoking. Not only making non-smoking environment for their students, schools should also strengthen their education to counteract the misconceptions of teens about the harmful effect of tobacco, which should not be restricted to class. For example, health care providers like physicians, nurses, or dentists often give information about preventive health care and anticipatory guidance, but adolescents report they rarely receive information about smoking or tobacco use from these health care providers (Kiernan 2002). The other significant variable related to school environment is the average class skipping rate in each school. The positive coefficient also indicates the importance of school management – strict enforcement of school attendance - is critical for reducing smoking behaviors.

Finally, looking at the impact of gender and racial backgrounds, the model estimates suggest that there is no significant difference between boys and girls in terms of the frequency or quantity of smoking behaviors, but there are some differences across racial lines. The coefficient of Hispanic on tobacco use frequency is significant, as well as that of minority ratio of school. Although the others race variables are not significantly different from zero, the negative sign of each coefficient states that the minority race student have lower probability to be heavy and

regular smokers than their white counterparts. This conclusion agrees with the results from previous studies (Faulkner, Farrelly, and Hersey 2000). However, previous studies have not been able to explain this difference well. Suggestions have made that the gap may related to socioeconomic status. Although this kind of information is not included in this study, it is unlikely that this race gap can be explained by socioeconomic variables since researchers have found that the cigarette smoking difference in race persists even after adjustment for socioeconomic status (Faulkner and Merritt 1998, Flint, Yamada, and Novotny 1998).

Conditional Probabilities

As previously discussed, the probability of two response variables belonging to a certain category conditional on all explanatory variables, like the probability usually computed in an univariate ordered probit, is informative but does not fully reveal the relationship between smoking behaviors. Instead, as we explore in this section, the probability of one response variable conditional on the other response variable and all explanatory variables yields additional insights. All such conditional probabilities are listed in Table 2.

Looking at the conditional probabilities in Table 2 it is clear that the frequency and quantity of smoking are positively correlated since almost all the large probabilities occur when both response variables are in same or adjacent categories. This is consistent with the estimation of $\rho = 0.66$ in the bivariate ordered probit model. For the detailed information, we can focus the following cases: low frequency and low quantity smokers and high frequency and high quantity smokers.

The conditional probability $\Pr(y_2 = 1 | y_1 = 1, x_1, x_2) = 0.75$, corresponds to a student who only smokes 1 cigarette per day during the past month, given $y_1 = 1$ (smoking 1 or 2 times per

month). For childern who fall in this frequency category they are highly likely to only smoke a single cigarette on that day. This smoking situation corresponds to (relatively) a lesser smoking problem reflecting the first step of smoking initiation for young people. Thus it is important to discover and stop the problem in its early stage. Table 4 gives the marginal effects of explanatory variables on this probability. Smoker ratio in a school contributes the largest positive effect to this probability, which strengthen again the importance of school management. Also notice that the main access method to cigarettes of these "early smokers" is by someone else. This indicates how adolescents start smoking – access and initiation through other smokers. Again, this result reinforces the importance of educating children and preventing this initiation to smoking.

Considering the opposite case, $\Pr(y_1 = 4 \mid y_2 = 4, x_1, x_2) = 0.9149$ corresponds to the conditional probability of a student smoking every day conditional on the student smoking more than 20 cigarettes a day. Adolescents in this group are heavily addicted to cigarettes and may need urgent intervention therapy. The marginal effects of explanatory variables are given in Table 5. The variable that positively contributes the most to this probability is the average days of skipped classes in a student's school. Again, this indicates the necessity of school involvement for controlling tobacco use. This can also be revealed by the relatively large positive effect of smoking ratio in schools. On the other hand, the main access method to cigarettes of these young people is via purchases by themselves or others on their behalf This result, as mentioned before, indicates that substantial penalties for stores and individuals who sell or give tobacco to underage smokers may be needed.

Conclusion

Despite substantial progress, tobacco use at all age levels, but particularly among adolescents, remains a significant health issue. Using data from the 2011 National Youth Tobacco Survey, this study has explored the impact of student demographics, awareness of smoking risks, and social environmental factors to explore jointly two measures of tobacco addiction - the frequency and quantity of cigerettes consumed per month. As expected, both measures of smoking habits are positively correlated, but to a lesser degree than might have been hypothesized. This indicates that there is a spectrum of smoking habits among US youth ranging from infrequent low quantities to frequent high quantities. Analyzing the factors that contribute to youth frequency and magnitude of consumption controlling for both individual-specific and school-specific factors reveals a number of areas for policy improvement. Our results indicate that strict controls on access to youth tobacco access and parental involvement can significantly reduce smoking habits. At the school level, strict policies to reduce absenteeism and prohibition of tobacco affiliated clothing both have the potential to decrease the quantity and frequency of youth tobacco use. Finally, our results indicate that there is a double dividend for schools when they reduce the tobacco consumption of their students. Not only for a given student tobacco use reduction is there benefit for that student, but there is a spillover effect on the rest of the school. As a whole, these results indicate that schools are critical in the fight against youth tobacco use.

References

CDC, U.S. 2011. "2011 National Youth Tobacco Survey Methodology Report."

Danaei, Goodarz, Eric L Ding, Dariush Mozaffarian, Ben Taylor, Jürgen Rehm, Christopher JL Murray, and Majid Ezzati. 2009. "The preventable causes of death in the United States: comparative risk assessment of dietary, lifestyle, and metabolic risk factors." *PLoS medicine* no. 6 (4):e1000058.

Faulkner, Dorothy L, Matthew C Farrelly, and James C Hersey. 2000. "Race, grade level, and cigarette smoking: the 1999 National Youth Tobacco Survey." *Journal of the National Cancer Institute* no. 92 (16):1360-1360.

Faulkner, Dorothy L, and Robert K Merritt. 1998. "Race and cigarette smoking among United States adolescents: the role of lifestyle behaviors and demographic factors." *Pediatrics* no. 101 (2):e4-e4.

Flint, Alan J, Eileen G Yamada, and Thomas E Novotny. 1998. "Black-white differences in cigarette smoking uptake: progression from adolescent experimentation to regular use." *Preventive medicine* no. 27 (3):358.

Greene, William H, and David A Hensher. 2010. *Modeling ordered choices: A primer*: Cambridge University Press.

Kandel, Denise, and Kazuo Yamaguchi. 1993. "From beer to crack: developmental patterns of drug involvement." *American Journal of Public Health* no. 83 (6):851-855.

Kiernan, Michael P. 2002. "Adolescent tobacco use: Prevention and cessation." *Journal Information* no. 4 (2).

Marshall, LaTisha, Michael Schooley, Heather Ryan, Patrick Cox, Alyssa Easton, Cheryl Healton, Kat Jackson, Kevin C Davis, and Ghada Homsi. 2006. "Youth tobacco surveillance—United States, 2001–2002." *MMWR Surveill Summ* no. 55 (3):1-56.

Mathers, M, JW Toumbourou, RF Catalano, J Williams, and GC Patton. 2006. "Consequences of youth tobacco use: a review of prospective behavioural studies." *Addiction* no. 101 (7):948-958.

Mokdad, Ali H, James S Marks, Donna F Stroup, and Julie L Gerberding. 2004. "Actual causes of death in the United States, 2000." *JAMA: the journal of the American Medical Association* no. 291 (10):1238-1245.

Mowery, Paul D, Matthew C Farrelly, M Lyndon Haviland, Julia M Gable, and Henry E Wells. 2004. "Progression to established smoking among US youths." *Journal Information* no. 94 (2).

Sajaia, Zurab. 2008. "Maximum likelihood estimation of a bivariate ordered probit model: implementation and Monte Carlo simulations." *The Stata Journal* no. 4 (2):1-18.

Table 1. Summary of Frequency and Quantity of Tobacco Use

Frequency of Tobacco Use (Days per month)y_1	**Quantity of Tobacco Use (Cigarettes per month)y_2**				
	1 Per Day	**2-5 Per Day**	**6-20 Per Day**	**20+ Per Day**	**Total**
1-2 Days	14.3%	6.1%	0.3%	0%	20.7%
3-9 Days	8.7%	14.6%	2.2%	0.2%	25.7%
10-29 Days	3.5%	17.9%	4.5%	0.3%	26.2%
30 Days	0.5%	10.0%	11.7%	5.2%	27.4%
Total	27.0%	48.6%	18.7%	5.7%	100%

Table 2. Summary of Survey Respondent Characteristics

Variable	Definition	Mean	Stdev.
Age	Age in years	18.726	1.0890
Male	Gender is male	0.5565	0.4970
GradeBelow	1 if Age behind normal grade	0.1339	0.3407
Hispanic	1 if Race is Hispanic	0.1989	0.3993
Black	1 if Race is Black	0.0952	0.2936
AgeFirstSmoke	Age when first tried smoking	12.323	2.6855
BuyOwnCig	1 if Buy cigarettes themselves	0.3181	0.4659
OtherBuyCig	1 if Other person buy for them	0.3104	0.4628
SomeoneGiveCig	1 if Other person give them	0.3173	0.4656
Promotion	1 if Believe Tobacco companies promote to adolescents	1.4559	0.4982
SmokeInHome	1 if smoker lives in home	0.5797	0.4938
Dangerous	1 if believe tobacco use is dangerous	0.7175	0.4504
ParentAdvice	1 if parent gave advice on tobacco	0.5325	0.4991
SmokeRatio	Ratio of smokers in school of respondent	0.1246	0.0804
ThinkQuit	1 if ever think about quitting smoking	0.1409	0.3480
FashionSchool	# of respondent in school willing to use product with tobacco logo	9.1563	6.0149
SchoolAveSkip	Average # of days skipping classes at respondents' school in past 30 days	5.4247	2.4965
MinorityRatio	Ratio of minorities in respondents' school	0.0540	0.0365
WeeklyIncome	Amount of money spent discretionally in last month	37.1331	24.7276

Table 3. ML Estimation of Bivariate Ordered Probit Model

Variable	Level	Level
	Frequency	**Quantity**
Age	0.0850 ***	0.0615 ***
	(0.0216)	(0.0211)
Male	0.0750	
	(0.0528)	
GradeBelow	-0.0035	
	(0.0781)	
Hispanic	-0.1853 ***	-0.1213
	(0.0803)	(0.0852)
Black	-0.0050	0.1649
	(0.1070)	(0.1081)
AgeFirstSmoke	-0.1383 ***	-0.1332 ***
	(0.0132)	(0.0131)
BuyOwnCig	0.6917 ***	0.4787 ***
	(0.0823)	(0.0807)
OtherBuyCig	0.3959 ***	0.2767 ***
	(0.0723)	(0.0721)
SomeoneGiveCig	-0.2482 ***	-0.2123 ***
	(0.0710)	(0.0715)
Promotion	0.0361	
	(0.0517)	
SmokInHome	0.2486 ***	
	(0.0524)	
Dangerous	-0.1299 **	-0.2480 ***
	(0.0708)	(0.0707)
ParentAdvice	-0.2346***	-0.1520 ***
	(0.0622)	(0.0625)
SmokeRatio	2.1710 ***	1.8508 ***
	(0.5311)	(0.4411)
ThinkQuit	-0.2323 ***	
	(0.0736)	
FashionSchool	0.0028	
	(0.0061)	
SchoolAveSkip	0.0362 ***	0.0395 ***
	(0.0126)	(0.0127)
MinorityRatio		-1.3623 **
		(0.8548)
WeeklyIncome		0.0018 **
		(0.0011)
LR test for independent equations:	Chi2=464.19, p=0.0000	
Log Likelihood:	-2747.8697	

Table 3. Continued

Estimation	**Value**	**90% CI**
Shared. Con	0.7938	0.7321, 0.8555
	(0.0375)	
μ_1	-0.6936	-1.0565, -0.3307
	(0.2206)	
μ_2	0.1667	-0.1963 , 0.5295
	(0.2206)	
μ_3	1.0104	0.6451, 1.3757
	(0.2221)	
δ_1	-0.8487	-1.1843, -0.5131
	(0.2040)	
δ_2	0.6198	0.2828, 0.9568
	(0.2049)	
δ_3	1.6398	1.2969, 1.9826
	(0.2085)	
ρ	0.6606	0.6244, 0.6939
	(0.0211)	

Table 4. Conditional Probability

Probability	**Value**	**Probability**	**Value**
$\Pr(y_1=1 \mid y_2=1, x_1, x_2)$	0.4768	$\Pr(y_2=1 \mid y_1=1, x_1, x_2)$	0.7449
	(0.0075)		(0.0101)
$\Pr(y_1=1 \mid y_2=2, x_1, x_2)$	0.1306	$\Pr(y_2=1 \mid y_1=2, x_1, x_2)$	0.3401
	(0.0027)		(0.0033)
$\Pr(y_1=1 \mid y_2=3, x_1, x_2)$	0.0207	$\Pr(y_2=1 \mid y_1=3, x_1, x_2)$	0.1367
	(0.0006)		(0.0016)
$\Pr(y_1=1 \mid y_2=4, x_1, x_2)$	0.0026	$\Pr(y_2=1 \mid y_1=4, x_1, x_2)$	0.0320
	(0.0001)		(0.0006)
$\Pr(y_1=2 \mid y_2=1, x_1, x_2)$	0.3789	$\Pr(y_2=2 \mid y_1=1, x_1, x_2)$	0.3964
	(0.0069)		(0.0073)
$\Pr(y_1=2 \mid y_2=2, x_1, x_2)$	0.3084	$\Pr(y_2=2 \mid y_1=2, x_1, x_2)$	0.6138
	(0.0026)		(0.0044)
$\Pr(y_1=2 \mid y_2=3, x_1, x_2)$	0.1349	$\Pr(y_2=2 \mid y_1=3, x_1, x_2)$	0.6610
	(0.0028)		(0.0033)
$\Pr(y_1=2 \mid y_2=4, x_1, x_2)$	0.0374	$\Pr(y_2=2 \mid y_1=4, x_1, x_2)$	0.4674
	(0.0012)		(0.0058)
$\Pr(y_1=3 \mid y_2=1, x_1, x_2)$	0.1898	$\Pr(y_2=3 \mid y_1=1, x_1, x_2)$	0.0173
	(0.0050)		(0.0005)
$\Pr(y_1=3 \mid y_2=2, x_1, x_2)$	0.3494	$\Pr(y_2=3 \mid y_1=2, x_1, x_2)$	0.0787
	(0.0022)		(0.0012)
$\Pr(y_1=3 \mid y_2=3, x_1, x_2)$	0.3460	$\Pr(y_2=3 \mid y_1=3, x_1, x_2)$	0.2163
	(0.0034)		(0.0022)
$\Pr(y_1=3 \mid y_2=4, x_1, x_2)$	0.1948	$\Pr(y_2=3 \mid y_1=4, x_1, x_2)$	0.4429
	(0.0041)		(0.0048)
$\Pr(y_1=4 \mid y_2=1, x_1, x_2)$	0.0444	$\Pr(y_2=4 \mid y_1=1, x_1, x_2)$	0.0003
	(0.0017)		(0.00001)
$\Pr(y_1=4 \mid y_2=2, x_1, x_2)$	0.2184	$\Pr(y_2=4 \mid y_1=2, x_1, x_2)$	0.0036
	(0.0040)		(0.00009)
$\Pr(y_1=4 \mid y_2=3, x_1, x_2)$	0.5533	$\Pr(y_2=4 \mid y_1=3, x_1, x_2)$	0.0222
	(0.0058)		(0.0004)
$\Pr(y_1=4 \mid y_2=4, x_1, x_2)$	0.9149	$\Pr(y_2=4 \mid y_1=4, x_1, x_2)$	0.1574
	(0.0109)		(0.0025)

1. Standard errors are computed by bootstrap method with replication 399;
2. All the probabilities are significantly different from zero under 5% level;

Table 5. Marginal Effects of $\Pr(y_2 = 1 | y_1 = 1, x_1, x_2)$

Variable	Marginal Effect		
	Frequency	**Quantity**	**Total**
Age	-0.0692	0.0127	-0.0565
Male	-0.0610		-0.0610
GradeBelow	0.0028		0.0028
Hispanic	0.1508	-0.0251	0.1257
Black	0.0040	0.0340	0.0038
AgeFirstSmoke	0.1125	-0.0275	0.0850
BuyOwnCig	-0.5628	0.0989	-0.4639
OtherBuyCig	-0.3221	0.0571	-0.2650
SomeoneGiveCig	0.2020	-0.0438	0.1582
Promotion	-0.0294		-0.0294
SmokInHome	-0.2023		-0.2023
Dangerous	0.1057	-0.0512	0.0545
ParentAdvice	0.1909	-0.0314	0.1595
SmokeRatio	-1.7663	0.3822	1.3841
ThinkQuit	0.1890		0.1890
FashionSchool	-0.0023		-0.0023
SchoolAveSkip	-0.0294	0.0082	-0.0212
MinorityRatio		-0.2814	-0.2814
WeeklyIncome		0.0004	0.0004

Table 6. Marginal Effects of $\Pr(y_1 = 4 \mid y_2 = 4, x_1, x_2)$

Variable	**Marginal Effect**		
	Frequency	**Quantity**	**Total**
Age	0.0193	-0.0136	-0.0057
Male	0.0171		0.0171
GradeBelow	-0.0008		-0.0008
Hispanic	-0.0421	0.0268	-0.0153
Black	-0.0011	-0.0365	-0.0376
AgeFirstSmoke	-0.0314	0.0295	-0.0019
BuyOwnCig	0.1573	-0.1059	0.0514
OtherBuyCig	0.0900	-0.0612	0.0288
SomeoneGiveCig	-0.0564	0.0470	-0.0142
Promotion	0.0082		0.0082
SmokInHome	0.0565		0.0565
Dangerous	-0.0295	0.0549	0.0254
ParentAdvice	-0.0534	0.0336	-0.0198
SmokeRatio	0.4937	-0.4094	0.0843
ThinkQuit	-0.0528		-0.0528
FashionSchool	0.0006		0.0006
SchoolAveSkip	0.0082	0.3014	0.3096
MinorityRatio		-0.0004	-0.0004
WeeklyIncome		-0.0087	-0.0087