Understanding and Modifying Health Behaviors

Kristina N. Piorkowski

University of New Mexico

Follow this and additional works at: https://digitalrepository.unm.edu/econ_etds

Part of the Economics Commons

Recommended Citation

This Dissertation is brought to you for free and open access by the Electronic Theses and Dissertations at UNM Digital Repository. It has been accepted for inclusion in Economics ETDs by an authorized administrator of UNM Digital Repository. For more information, please contact disc@unm.edu.
Kristina N. Piorkowski
Candidate
Department of Economics

This dissertation is approved, and it is acceptable in quality and form for publication:

Approved by the Dissertation Committee:

Alok K. Bohara, Chairperson

Gabriel Sanchez

Melissa Binder

José Pagán
UNDERSTANDING AND MODIFYING HEALTH BEHAVIORS

By

KRISTINA N. PIORKOWSKI

B.A., Economics, University of New Mexico, 2013
M.A., Economics, University of New Mexico, 2015

DISSERTATION

Submitted in Partial Fulfillment of the
Requirements for the Degree of
Doctor of Philosophy

Economics

The University of New Mexico
Albuquerque, New Mexico

July, 2018
DEDICATION

To all those who helped me open doors and gave me a love of learning.

To Kody, thank you for supporting me, feeding me, and keeping me sane in this process.
ACKNOWLEDGEMENTS

I would first like to thank Dr. Alok Bohara for all of his advice and comments on my work these past seven years. You have supported my research and the many ways I have wanted to take it all while giving me the skills to conduct the best research possible in whatever setting I end up in.

I also acknowledge all my committee members Dr. Melissa Binder, Dr. Gabriel Sanchez, and Dr. José Pagán. I am very grateful for their guidance, ideas, and support in this and other aspects of my graduate career.

I would not be here today without the support of the Robert Wood Johnson Foundation Center for Health Policy at the University of New Mexico. This fellowship provided me with a community of scholars which challenged me to become a better person and more attentive economist. I am so grateful of the myriad of research opportunities I was presented and took advantage of which greatly impacted this dissertation. Thank you Dr. Sanchez for providing me the opportunity to be a RWJF fellow.

This dissertation was mostly written alone in persistent isolation, but my journey was by no means a lonely one. There were countless people who have taught me, inspired me, encouraged me, and supported me. I would not be here today without all of their consorted efforts. Thank you.
UNDERSTANDING AND MODIFYING HEALTH BEHAVIORS

By

Kristina N. Piorkowski

B.A., Economics, University of New Mexico, 2013
M.A., Economics, University of New Mexico, 2015
Ph.D., Economics, University of New Mexico, 2018

ABSTRACT

The rates of non-communicable diseases are growing and now are the largest cause of death worldwide. The most prevalent non-communicable diseases (cardiovascular disease, cancer, chronic respiratory disease, and diabetes) all have the same behavioral risk factors in common. This dissertation explores the drivers behind two risky health behaviors – sugar-sweetened beverage consumption and youth tobacco use – in order to improve policies tailored to mitigate these behaviors.

Chapter 2 uses primary survey data representative of New Mexico’s adult population to estimate the willingness to pay (WTP) for a tax on sugar sweetened beverages (SSBs) in New Mexico, a state marked by high rates of obesity and a history of failed SSB taxes. We examine the direct and indirect roles that eating habits, knowledge and awareness around food, and related policies and attitudes have on the preferences for
SSB taxes. Traditional contingent valuation regression methods were reformulated as a three-equation simultaneous model to address issues of protest responses and endogeneity. Results indicate that respondents who have healthier eating habits have statistically significantly higher preferences for SSB taxes. Knowledge that poor diet can lead to being overweight increases the preferences for SSB taxes. Respondents who have heard of other SSB taxes before or who are a conservative are statistically significantly more likely to be a protest response. The estimated individual median WTP 95% confidence interval was 0.002-0.961 pennies-per-ounce, which is lower than previously proposed taxes in other localities and could help explain why a recent 2-penny-per-ounce SSB tax failed in the city of Santa Fe, New Mexico.

Chapter 3 continues the exploration into SSB taxation by asking a more basic question: why would someone support a SSB tax? Using primary data of a statewide representative sample of New Mexico, we assess how media coverage; knowledge, attitudes, and behaviors; and political ideology impact support for expanding the Health Diné Nation Act (which is, in part, a SSB tax implemented by the Navajo Nation) to all of New Mexico. From a partial proportional odds model, we find that respondents who have heard of the Healthy Diné Nation Act know that drinking too much soda can cause obesity, and believe obesity is a major problem are more likely to strongly support expansion of the SSB tax portion of the Healthy Diné Nation Act. However, hearing about other SBB taxes is associated with a 11-percentage point increase in being strongly against the expansion.

The fourth chapter transitions from the United States to Nepal to investigate youth smoking behaviors. With a trivariate ordered probit model with independence we
estimate the likelihood youths will fall into one of three mutually exclusive smoking categories – never smoker, former smoker, and level of cigarette consumption by current smokers – and the impact proximity to other smokers; pro-tobacco exposure and social perceptions; and anti-tobacco education have on these smoking statuses. We also calculate the average marginal effect to determine the full impact of each variable on each smoking status and take into account possible differences in gender through the use of interactions. We find that having a close friend who smokes greatly increases the likelihood of smoking for youths, whereas having a parent who smokes only increases the likelihood of smoking for boys. Perversely, exposure to anti-tobacco media greatly increases the likelihood of smoking. On the other hand, high quality anti-tobacco education reduces the level of cigarette consumption but does not impact smoking initiation.

This dissertation provides insights into ways that taxation, information, advertising, and education can be used to reduce risky health behaviors. By implementing policies guided by the empirical results, we will be able to make major strides towards reducing the main risk factors for non-communicable diseases. Reducing SSB consumption can reduce rates of obesity which itself is associated with diabetes, certain cancers, and cardiovascular disease. The second and third chapter demonstrate was taxation, education, and media coverage can be leveraged to reduce consumption of SSBs. The reduction of global tobacco use has been a major goal of the WHO and with our findings in chapter four we have identified policy strategies to reduce youth smoking behaviors in Nepal. This can serve as an example for other developing nations looking to implement programs to reduce their own youth’s smoking behaviors.
Table of Contents

List of Figures ........................................................................................................... xii

List of Tables ............................................................................................................. xiii

List of Abbreviations ............................................................................................... xv

Chapter 1: Introduction .............................................................................................. 1

1.1. Global Burden of Non-Communicable Diseases ............................................. 1

1.2. Health Behaviors and Economics ................................................................. 2

1.3. Global Scourges: Obesity and Smoking ......................................................... 3

1.3.1. Obesity .................................................................................................. 4

1.3.2. Smoking ................................................................................................. 5

1.4. Contributions of this Dissertation ................................................................. 7

Chapter 2: Less Than a Penny for Your Thoughts: Estimating the Willingness to
Pay for a Sugar-Sweetened Beverage Tax ............................................................... 11

2.1. Introduction .................................................................................................. 11

2.2. Background .................................................................................................. 12

2.3. Data ............................................................................................................. 15

2.3.1. Survey .................................................................................................. 15

2.3.2. Willingness to Pay Question .................................................................. 16

2.3.3. Variables ................................................................................................. 17

2.3.4. Hypotheses ............................................................................................. 20

2.4. Empirical Model .......................................................................................... 21

2.4.1. Random Utility Framework .................................................................. 21

2.4.2. Empirical Model ..................................................................................... 21
### Chapter 3: Elements of Support for Sugar-Sweetened Beverage Taxation: Empirical Evidence from New Mexico

#### 3.1. Introduction

#### 3.2. Methods

- **3.2.1. Model**
- **3.2.2. Hypotheses**

#### 3.3. Data

- **3.3.1. Survey**
- **3.3.2. Measures**

#### 3.4. Empirical Results

- **3.4.1. Descriptive Statistics**
- **3.4.2. Partial Proportional Odds Model Results**
- **3.4.3. Average Marginal Effects**
- **3.4.4. Robustness Check**

#### 3.5. Discussion
### Chapter 4: To Light Up or Snuff Out? Estimating Adolescent Smoking Behavior in Nepal

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.1</td>
<td>Introduction</td>
<td>82</td>
</tr>
<tr>
<td>4.2</td>
<td>Background Literature</td>
<td>85</td>
</tr>
<tr>
<td>4.3</td>
<td>Trivariate Ordered Probit Model with Independence</td>
<td>90</td>
</tr>
<tr>
<td>4.4</td>
<td>Data</td>
<td>92</td>
</tr>
<tr>
<td>4.4.1</td>
<td>Global Youth and Tobacco Survey</td>
<td>92</td>
</tr>
<tr>
<td>4.4.2</td>
<td>Smoking Status</td>
<td>93</td>
</tr>
<tr>
<td>4.4.3</td>
<td>Independent Variables</td>
<td>94</td>
</tr>
<tr>
<td>4.4.4</td>
<td>Hypotheses</td>
<td>96</td>
</tr>
<tr>
<td>4.5</td>
<td>Estimation Strategy</td>
<td>98</td>
</tr>
<tr>
<td>4.5.1</td>
<td>Principal Component Analysis</td>
<td>98</td>
</tr>
<tr>
<td>4.5.2</td>
<td>Maximum Likelihood Estimation</td>
<td>98</td>
</tr>
<tr>
<td>4.5.3</td>
<td>Coefficient Interactions</td>
<td>99</td>
</tr>
<tr>
<td>4.5.4</td>
<td>Average Marginal Effects</td>
<td>100</td>
</tr>
<tr>
<td>4.5.5</td>
<td>Models</td>
<td>102</td>
</tr>
<tr>
<td>4.6</td>
<td>Results</td>
<td>102</td>
</tr>
<tr>
<td>4.6.1</td>
<td>Descriptive Statistics</td>
<td>102</td>
</tr>
<tr>
<td>4.6.2</td>
<td>Trivariate Ordered Probit Regression Results</td>
<td>103</td>
</tr>
<tr>
<td>4.6.3</td>
<td>Average Marginal Effects Results</td>
<td>105</td>
</tr>
<tr>
<td>4.6.4</td>
<td>Review of Results</td>
<td>109</td>
</tr>
<tr>
<td>4.7</td>
<td>Discussion</td>
<td>110</td>
</tr>
</tbody>
</table>
Chapter 5: Conclusion ................................................................. 135

5.1. Dissertation Summary ...................................................... 135
5.2. Avenues for Future Research ............................................ 139

Appendices ................................................................................. 141

Appendix A: Principal Component Analysis .......................... 142
Appendix B: Supplemental Tables ........................................... 152
Appendix C: New Landscapes of a Majority-Minority State ........ 158
Appendix D: Stata Code .............................................................. 190

References .................................................................................. 338
List of Figures

Figure 2.1.: Histogram of Bootstrapped Median WTP ........................................... 48
Figure 4.1.: Smoking Decision Tree ........................................................................ 121
Figure A.1.: Scree Plot of Eigenvalues after PCA ..................................................... 148
List of Tables

Table 2.1.: Summary Statistics ................................................................. 37
Table 2.2.: Three Equation System Results for Preferences of SSB Taxes .......... 40
Table 2.2.A.: Probit Results for Preferences of SSB Taxes ............................. 42
Table 2.2.B.: Probit with Selection Results for Preferences of SSB Taxes .......... 43
Table 2.2.C.: Probit with Selection with Control Function Approach Results for Preferences of SSB Taxes ................................................................. 44
Table 2.3.: Hypothesis Table of Three Equation System and Sensitivity Analyses Raw Results and Conditional Average Marginal Effects for Willingness to Pay for a SBB Tax for Model 2 ................................................................. 46
Table 2.4.: Confidence Interval for the Bootstrapped Individual Median WTP Results With and Without Selection Median Value Reported and Trimmed at Varying Levels . 47
Table 3.1.: Descriptive Statistics by Level of Support Mean and Standard Deviation ... 72
Table 3.2.: Partial Proportional Odds Model for Level of Support for Expanding Healthy Diné Nation Act ................................................................. 73
Table 3.2.A.: Ordered Logit Results for Level of Support for Expanding Healthy Diné Nation Act ................................................................. 75
Table 3.2.B.: Generalized Ordered Logit Results for Level of Support for Expanding Healthy Diné Nation Act ................................................................. 77
Table 3.3.: Hypothesis Table with Average Marginal Effects on the Level of Support For PPOM Model 4 ................................................................. 81
Table 4.1.: Descriptive Statistics of Smoking Behavior and Model Covariates by Sex of Analytical Model ................................................................. 123
Table 4.2.: Regression Results of Trivariate Ordered Probit with Independence Model
…………………………………………………………………………………………………….. 125
Table 4.3.A.: Average Marginal Effects for Never Smoking ………………….. 127
Table 4.3.B.: Average Marginal Effects for Quitting Smoking ………………… 129
Table 4.3.C.: Average Marginal Effects of the Conditional Mean of Level of Cigarettes Smoked …………………………………………………………………………………………………….. 131
Table A.1.: Correlation Matrix Between Social Perception Variables ………… 146
Table A.2.: Social Perception Principal Components/Correlation, Unrotated … 147
Table A.3.: Principal Components (Eigenvectors) Unrotated …………………….. 149
Table A.4.: Principal Component (Eigenvectors), Orthogonal Varimax Rotation … 150
Table B.4.4.: Multinomial Logit Results for Current Smokers of Both Sex Combined and Boys and Girls Separately ………………………………………………………………………. 152
Table B.4.5.: Multinomial Logit Results for Former Smokers of Both Sex Combined and Boys and Girls Separately ………………………………………………………………………. 155
**List of Abbreviations**

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AIC</td>
<td>Akaike Information Criterion</td>
</tr>
<tr>
<td>CDF</td>
<td>Cumulative Density Function</td>
</tr>
<tr>
<td>CFA</td>
<td>Control Function Approach</td>
</tr>
<tr>
<td>CV</td>
<td>Contingent Valuation</td>
</tr>
<tr>
<td>FCTC</td>
<td>Framework Convention on Tobacco Control</td>
</tr>
<tr>
<td>GYTS</td>
<td>Global Youth Tobacco Survey</td>
</tr>
<tr>
<td>HDNA</td>
<td>Healthy Diné Nation Act</td>
</tr>
<tr>
<td>HETIP</td>
<td>Health Education and Tobacco Intervention Program</td>
</tr>
<tr>
<td>IID</td>
<td>Independent and Identically Distributed</td>
</tr>
<tr>
<td>KMY</td>
<td>Kasterdis, Munkin, and Yen (2010)</td>
</tr>
<tr>
<td>NCDs</td>
<td>Non-Communicable Diseases</td>
</tr>
<tr>
<td>NLMSS</td>
<td>New Landscapes of a Minority-Majority State</td>
</tr>
<tr>
<td>NOAA</td>
<td>National Oceanic Atmospheric Administration</td>
</tr>
<tr>
<td>PCA</td>
<td>Principal Component Analysis</td>
</tr>
<tr>
<td>PDF</td>
<td>Probability Density Function</td>
</tr>
<tr>
<td>PPOM</td>
<td>Partial Proportional Odds Model</td>
</tr>
<tr>
<td>SSBs</td>
<td>Sugar-Sweetened Beverages</td>
</tr>
<tr>
<td>USDA</td>
<td>United States Department of Agriculture</td>
</tr>
<tr>
<td>USDHHS</td>
<td>United States Department of Health and Human Services</td>
</tr>
<tr>
<td>WHO</td>
<td>World Health Organization</td>
</tr>
<tr>
<td>WTP</td>
<td>Willingness to Pay</td>
</tr>
</tbody>
</table>
Chapter 1: Introduction

1.1 Global Burden of Non-Communicable Diseases

Non-communicable diseases (NCDs), which are non-infectious diseases and sometimes called chronic diseases, are the largest cause of death worldwide (Yach et al. 2004). The most common types of NCDs are cardiovascular disease, cancer, chronic respiratory disease, and diabetes. Although NCDs used to be concentrated in developed nations, developing countries are increasingly burdened by these diseases. Not only are NCDs the leading cause of death in developing nations (Beaglehole and Yach 2003), around 80% of all premature deaths\(^1\) from NCDs occur in low and middle-income countries (Abegunde et al. 2007). These nations are faced with a dual burden of communicable and non-communicable diseases placing increasing strain on health systems, poverty, and economic development.

However, as most NCDs are largely preventable, the ensuing human and economic loss is unnecessary and can be greatly reduced. The most prevalent NCDs share the same risk factors and all these risk factors are behavioral: tobacco use, harmful alcohol consumption, unhealthy diets, and lack of physical activity (Yach et al. 2004). Tobacco use alone accounts for 1 in 6 of all NCD related deaths (Beaglehole et al. 2011). If we can understand why people engage in these risky health behaviors, we can create legislation, policies, and programs to both encourage healthy behaviors and reduce unhealthy behaviors. In this dissertation, we explore this idea through an economic lens.

\(^1\) Deaths amongst people between ages 30-69.
to strengthen our understanding of the motivations behind engaging in risky health behaviors and propose ways to modify these behaviors.

1.2 Health Behaviors and Economics

The field of health economics is relatively new. Ken Arrow’s seminal 1963 article detailed how the demand and supply of medical care differs from traditional economic goods and hence deserves special attention within economics. Since then topics such as health care spending, market structures, competition, and regulation; demand for pharmaceuticals; health equity; medical workforce; demand for health insurance; and risky health behaviors\(^2\) have been studied with increasing frequency in economics.

Although many of these areas in health economics overlap with trying to understand different aspects of NCDs, we concentrate on the specific contribution that risky health behaviors have on the NCD epidemic because of their important role in the spread of NCDs.

The economics toolkit contains multiple strategies to mitigate risky health behaviors (Cawley and Ruhm 2012). Taxes can be placed on goods with negative externalities or subsides can be used to promote desired behavior like fruit and vegetable consumption. Purchasing restrictions on some goods, such as needing to be 21 years or older in the United States to purchase alcohol, are made to reduce their accessibility. Attempts have been made to provide information to general and targeted populations about the consequences of certain health behaviors such as the dangers of sexually

\(^2\) This includes alcohol use, tobacco use, obesity, sexual behavior, illicit drug use, and unhealthy diets.
transmitted diseases resulting from unprotected sex. Restrictions on advertising on certain goods are common. For instance, in the United States (US) tobacco companies are prohibited from advertising to children and across the world direct to consumer advertising of pharmaceuticals is prohibited, except in the US and New Zealand. Additionally, taking a page from behavioral economics, one approach to reducing risky health behaviors centers on changing the default of choices such that picking the healthier or better for a consumer is unnoticeably easier than the unhealthy option.

Within the field of health behaviors, two of the most studied areas are obesity and tobacco use (Cawley and Ruhm 2012). Although much has been learned about their economic consequences and ways to mitigate these behaviors, more work is needed to improve strategies aimed at reducing the spread and prevalence of NCDs.

SSBs are beverages which have no nutritional value and whose increased levels of consumption across all age groups in the US have played an important role in the obesity epidemic (Brownell and Frieden 2009). When it comes to tobacco use, massive strides have been made in reducing consumption in the US. However, the same is not true for developing nations. Furthermore, little attention has been devoted in the economics literature to youth smoking behaviors in developing countries especially regarding how there could be differences in smoking behaviors by gender. Economics would stand to benefit from having a better understanding of the drivers of SSB consumption in industrialized diets and youth tobacco use in developing nations.

1.3 Global Scourges: Obesity and Smoking
1.3.1 Obesity

The adult obesity rate in the US has grown from 13% in 1962 to close to 34% in 2008 (Ogden and Carroll 2010). There are many elements of the obesogenic environment, from structural factors – such as food deserts and relative prices of healthy versus unhealthy foods – to individual factors – like fast food consumption and reduced physical activity – that have played a role in rising obesity rates in the United States. These factors can work alone or in conjunction with other elements of the obesogenic environment creating a complex network of causes and correlates for growing obesity rates. Sugar-sweetened beverage (SSB) consumption, defined as all liquid beverages that have an added caloric sweetened (Chriqui et al., 2013) have certainly played a role in the obesity epidemic. Some even consider it to be the single largest driver of this epidemic (Brownell and Frieden 2009). Between 1977 and 1996, soft drink consumption among US adults ages 19 to 29 years old increased by 71% (Kim and Kawachi 2006). Despite recent reduction in regular soft drink consumption, the rate of SSB consumption remains high due to increased sport and energy drink consumption (Han and Powell 2013).

SSBs are considered an empty calorie food: they provide no nutritional value. Consumption of these drinks account for 39% of daily added sugar of the average diet for Americans 2 years and older (U.S. Department of Health and Human Services [USDHSS] and U.S. Department of Agriculture [USDA] 2015). The health consequences of SSBs has been well documented in the literature. Increased SSB intake has been shown to have a positive association with increased body weight – for adults and youth – adiposity, and type II diabetes (Brownell et al. 2009). One policy intervention that both economists and public health experts have suggested to reduce SSB consumption is
through the taxation of the good. However, the number of failed SSB tax attempts far outnumber successful taxes thereby limiting the viability of this policy option. In this dissertation we explore potential reasons these taxes are failing and ways to fine-tune the tax value to increase its likelihood of successful passage.

1.3.2 Smoking

The smoking epidemic is shifting from the developed to the developing world. Since 1984, cigarette smoking rates in the developing world have surpassed that of the developed world (Food and Agriculture Organization of the United Nations [FAO] 2003) and it is estimated currently 80% of smokers resided in developing nations (WHO n.d.). A unified global effort to address the spread of tobacco was seen in 2003 with the adoption of the WHO Framework Convention on Tobacco Control. The FCTC calls for a robust approach to tobacco control acknowledging the need to address supply and demand issues (WHO 2005). As of August 2017, there are 181 Parties, covering 90% of the world’s population, that are bound to the FCTC’s provisions (WHO 2018).

The FCTC explicitly acknowledges the dangers of children’s exposure to tobacco smoke and the increased consumption by adolescents and children worldwide (WHO 2005). Finding ways to mitigate youth smoking behavior is critical because most smokers begin smoking before age 18 (USDHHS 2012) and the tobacco industry has used strategies to actively cultivate young adults (aged 18 to 24) through life transitions into becoming and remaining a smoker (Ling and Glantz 2002). Economics has begun to seriously consider ways to model and study youth risky behavior because the driving forces behind why people engage in these behaviors might not be the same for youth and
adults. Developmental psychology studies find that youth have lower decision-making competence, are more impacted by the social reactions, and less able to understand future risks and consequences of actions (Gruber 2001).

Furthermore, there are gender differences in smoking behaviors. Worldwide, there are fewer female smokers than male smokers. Even though the prevalence of daily smokers has decreased for men and women since 1980, because of global population growth the absolute number of smokers has increased (Ng et al. 2014). There is especially cause for concern among young female smokers. Across 120 sites in which the Global Youth Tobacco Survey (GYTS) has been implemented, 61 showed no difference in boys and girls cigarette usage (Global Youth Tobacco Survey [GYTS] Collaborating Group 2003). A few previous studies have demonstrated that the following factors impact boys and girls differently: the desire to control weight (Tsai et al. 2008), school personnel who smoke on smoking prevalence (Nikaj and Chaloupka 2015), and taxes (Nonnemaker and Farrelly 2011). This increase in tobacco use by boys and girls in developing nations does not bode well for the future prevalence of NCDs.

Despite the major gains by Big Tobacco in developing nations, we cannot forget that substantial progress has previously been made against Big Tobacco in countries like the US. We can take these lessons learned and with careful consideration of cultural differences adapt them to developing nations specifically targeting youth. We contribute to this effort in Chapter 4 which examines youth smoking behaviors in Nepal. As of 2001, the boys to girls smoking ratio in Nepal was 6.7:1.0 (GYTS Collaborating Group 2003). There is still time to act to prevent girls from reaching this undesirable gender parity.
1.4 Contributions of this Dissertation

The research in this dissertation contributes to the ongoing conversation about ways to mitigate risky health behaviors both nationally and internationally. The first two analyses (found in Chapters 2 and 3) utilize primary statewide representative data to focus on SSB taxation in New Mexico: a state with high rates of diabetes and ethnic/racial disparities in obesity rates. The third analysis in Chapter 4 transitions from the United States to the Nepal to focus on another risky health behavior: youth smoking. We use two waves of data from the Nepal Global Youth and Tobacco Surveillance Survey for this chapter.

The first analysis, Chapter 2, is one of the first to empirically estimate the willingness to pay (WTP) for a SSB tax. We do so by using a hypothetical referendum to empirically estimate the willingness to pay for a tax on sugar sweetened beverages in New Mexico. The roles that eating habits; knowledge and awareness around food and related policies; and attitude have both directly and indirectly on the preferences for SSB taxes are assessed. Traditional contingent valuation regression methods were reformulated as a three-equation simultaneous system to address issues of protest responses and endogeneity. The first equation in the system separately estimates the endogenous variable – the number of SSBs consumed in the average week. The next equation of the system is a selection equation that serves the role of identifying protest respondents – people who would say no to the tax regardless of the value. The final equation in the system estimates the preferences for SSB taxation. After the three-equation system is simultaneously estimated, the individual median WTP is calculated following Cameron (1988) using bootstrapping. The results indicate that respondents who
have healthier eating habits or know that a poor diet can lead to being obese or overweight have statistically significantly increased preferences for SSB taxation. The estimated willingness to pay for a sugar-sweetened beverage tax was less than one-penny-per-ounce, which can help explain why a recent two-penny-per-ounce sugar-sweetened beverage tax failed in city of Santa Fe, New Mexico.

The second analysis, in Chapter 3, addresses a simple, but vital, question regarding SSB taxes: why would someone support this tax in the first place? This question becomes increasingly important to answer as more local and state governments seek to implement such a tax citing the positive impacts these taxes can have on health outcomes to revenue generation. Yet, the number of successful SSB taxes in the United States is far outpaced by the number of failed attempts. This analysis assesses the roles media coverage; knowledge, attitudes, and behaviors; and political ideology have on a respondent’s level of support for expanding the Healthy Diné Nation Act of 2014 to all of New Mexico. This law was signed by the Navajo Nation Council and it calls for a 2 percent tax to be imposed on all minimal-to-no-nutritional value foods on the Navajo Nation. With the use of primary data, we are able to focus in on the taxation of SSBs instead of the broader minimal-to-no-nutritional value foods.

A partial proportional odds model is used to assess the level of support for expansion. This modeling technique has the benefit of allowing the coefficients on the variables to vary across categories only when there is a statistically justified reason to do so. As such the effect of one variable can have a different magnitude or direction of impact for the outcome categories. The results find that increasing awareness that SSBs can lead to being obese/overweight or awareness that obesity is a major problem in the
state reduces strong opposition and increases strong support for expansion. Because our finding that respondents who have heard of other SSB taxes before strongly reduces the likelihood of supporting expansions, we suggest that newly proposed taxes do their best to differentiate themselves from these previous taxes.

The final analysis, in Chapter 4, moves from SSBs to tobacco use in the developing world. The smoking epidemic is moving from the developed to the developing world. This is especially true in the Nepal, which has the highest rate of female smokers in all of South East Asia (World Bank 2014). However, most smokers begin smoking before their eighteenth birthday (USDHHS 2012). As such, this chapter focuses on adolescent smoking behaviors in Nepal. Using two waves of nationally representative data, we assess how proximity to other smokers, pro-tobacco marketing and social perception, and anti-tobacco media and knowledge influence smoking status of youth. This is one of the few papers to estimate the smoking statuses of never smoking, level of smoking by current smokers, and former smoker simultaneously. We accomplish this by estimating a trivariate ordered probit model with independence as is detailed in Kasteridis, Munkin, and Yen (2010). Guided by preliminary analyses, we include interaction terms between the variables of interest and gender in order to understand how these variables can impact the sexes differently. The overall impact of each variable on each smoking status is determined by calculating the average marginal effect for never and former smokers as well as the average marginal effect of the conditional mean of the level of cigarettes consumed. The findings demonstrate that having a close friend who smokes or being exposed to anti-tobacco media greatly increases smoking behaviors. On the other hand, high quality formal anti-tobacco education reduces level of cigarette
consumption, but not initiation. Greater exposure to pro-tobacco media increases girls’ level of cigarette consumption. We also find that having a parent who smokes increases only boys’ smoking behaviors and further boys are much more likely to abstain from smoking behaviors if they think it will change their weight.

This dissertation takes commonly studied risky health behaviors – SSB consumption and smoking – and addresses novel research questions missing from the literature at large. The first two chapters provide policy makers insight onto the elements that increase support for SSB taxation and cautions policy makers from going with the status quo one-penny-per ounce tax value. The last chapter rigorously studies a population that is increasingly bearing the burden of the tobacco epidemic: youth in developing nations. Based upon the findings from this analysis we make recommendations to augment existing policies and programs, especially around formal and informal education, to better protect the people they are trying to serve.
Chapter 2: Less Than a Penny for Your Thoughts: Estimating the Willingness to Pay for a Sugar-Sweetened Beverage Tax

2.1 Introduction

The rate of adult obesity in the United States has more than doubled in the past five decades from 13% in 1962 to close to 34% in 2008 (Ogden and Carroll 2010). Meanwhile, between 1977 and 1996, soft drink consumption in the US increased by 71% for people 19 to 39 years old (Kim and Kawachi 2006). Today, the consumption of sugar-sweetened beverages (SSB) -- all liquid beverages which contain an added caloric sweetener (Chriqui et al. 2013) -- remains high due to the increased consumption of sport and energy drinks despite recent decreases in regular soda consumption (Han and Powell 2013). The public health literature suggests multiple ways to address the obesity epidemic (Fletcher, Frisvold, and Tefft 2011; Kim and Kawachi 2006). One increasingly popular policy recommendation is taxing SSBs at one-penny-per-ounce (Brownell et al., 2009). Currently, this is the most commonly passed SSB tax value in the US.

However, this value was initially decided arbitrarily, harkening to its potential impact on health outcomes, revenue generation, and political feasibility; this value was neither reached from any sort of optimization nor does it necessarily reflect the reality of what most consumers are willing to pay. We speculate that this penny-per-ounce tax might be too high and could have played a role in the failure of dozens of SSB tax initiatives. The previous literature focuses on the hypothetical impact of SSB taxes on consumption behavior (see Andreyeva, Chaloupka, and Brownell 2011; Smith, Lin and
Lee 2010; and Wang et al. 2012) where, by necessity, they assume that their proposed tax passes. However, the realization of these taxes relies heavily upon voters’ support. As such, the goal of this paper is to provide a robustly estimated willingness to pay (WTP) for a SSB tax and to expose what factors influence the preferences for SSB taxation.

This is accomplished using primary statewide representative survey data of adult New Mexicans which includes a referendum contingent valuation (CV) study assessing respondents’ WTP for a SSB tax in New Mexico. The CV posed the WTP question as a specific-excise tax consistent with public health recommendations (Brownell et al. 2009; Chriqui et al. 2013). The empirical modeling incorporates concerns of protest responses and endogeneity to estimate how eating behaviors; knowledge around diet and health; awareness of SSB policy; and political ideology impact the preferences for SSB taxes.

The results from a three-equation, endogenous recursive system indicate that healthy eating habits strongly and positively influence the preference for SSB taxes. Knowing that a poor diet can lead to being obese or overweight increases SSB preferences. Awareness of previous SSB policy has mixed impacts: knowing about a local SSB tax indirectly increased preferences whereas knowledge of other SSB taxes indirectly decreased preferences. Further, hearing of other SSB taxes and being a conservative statistically significantly increases the likelihood of a protest response. The 95% confidence interval for the median of the bootstrapped, individual median WTP estimates was between 0.002-0.965 pennies-per-ounce. This range is below the recently proposed SSB tax in Santa Fe, New Mexico and might be part of reason why it failed.

2.2 Background
In 2014, the first SSB tax in the US was passed in Berkeley, California at a rate of one-penny-per-ounce. Since then, an increasing number of states, counties, and cities have proposed, often unsuccessfully, their own SSB taxes. For instance, the State of New Mexico has failed to pass a half-penny-per-ounce SSB tax twice (State Senate of New Mexico 2010, 2011), and voters in Santa Fe, New Mexico, rejected a 2017 initiative for a two-penny-per ounce SSB tax. Whereas the Navajo Nation, whose geography partially overlaps with New Mexico’s, successfully passed a SSB tax. In 2014, the Healthy Diné Nation Act enacted a 2% tax on the gross receipts of the consumer for all foods and beverages of minimal-to-no nutritional value, which includes SSBs. The generated funds support community wellness projects that, for instance, improve the food environment or recreational facilities (The Healthy Diné Nation Act of 2014). The SSB tax trend in New Mexico is echoed nationwide: most SSB taxes fail.

A national study on Americans’ opinions of SSB-related policies found that a penny-per-ounce SSB tax had the lowest overall support of six potential policies (Gollust, Barry, and Niedereppe 2014). Despite the lack of public support, SSB taxes could raise substantial revenue and reduce weight (Brownell et al., 2009). As such, it is increasingly important to understand what factors influence preferences for SSB taxes; in this paper, we focus on eating behaviors; knowledge about health and diet; SSB policy awareness; and political ideology.

Preferences for SSB taxes are likely to correspond to both healthy (i.e., fruit and vegetable consumption) and unhealthy (i.e., SSB consumption) eating behaviors. The tax would impact individuals who consume more SSBs directly and would be more likely to oppose it. On the other hand, others might have preferences for eating healthy foods over
junk foods and part of that reason could be because they value their health or are more knowledgeable about health issues. A previous study found an inverse relationship between fruit and vegetable consumption and SSB consumption (Park et al. 2014b).

One of the economic arguments for implementing a SSB tax is addressing the market failure of imperfect information about the health consequences of SSB consumption. SSB consumption has been linked to higher rates of heart disease, type 2 diabetes, and increased risk of obesity (Brownell et al. 2009). It is possible that those who do not know or fully understand the health implications of SSB consumption could consume more SSBs. Park et al. (2014a) found that those who disagree that SSBs can contribute to weight gain were 1.68 times more likely to consume two or more SSBs a day compared to their informed counterparts. Additionally, there is a time inconsistency issue: sustained SSB consumption will not impact health today, but it will overtime. As such what was good today will not be good tomorrow.

SSB taxes lack a well-received public media message; a 2011 public opinion survey found that no pro-SSB tax argument had majority support (Barry, Niedereppe, and Gollust 2013). Most newspaper and television stories on SSBs present pro-SSB tax arguments (Niederdeppe et al. 2013), but consideration is needed on how the argument is framed. A pro-SSB tax argument framed to address an individual-level problem, as was found to be the case about news stories on obesity in Philadelphia (Jeong et al. 2014), could lead people to view an SSB tax as an attack on individual liberty. Learning about SSB taxes from either formal news outlets or from SSB tax opposition sources might push citizens on the margin against these taxes. Further, large funds are used to promote targeted and poignant anti-SSB arguments in local elections (Nixon et al. 2015). There is
also a question if the respondent perceives local and non-local taxes similarly. A SSB tax in another state might not impact have a direct impact on the respondent which could lead to more mixed feelings towards the tax. Whereas a local SSB tax could have a much greater impact on them and this personal connection to the tax might result in a different feeling towards a SSB tax. As such, the implication of SSB policy awareness on tax preferences is not straightforward, let alone when awareness of SSB taxes is separated into local and non-local SSB taxes.

Across the political spectrum, there is agreement that obesity is a societal problem, however, only 27% of Republicans desire government intervention to address obesity compared to 82% of Democrats (Gallup 2012). This divide reflects the differing views political parties have regarding government intervention. Since SSB taxes are a tax on an action in which people elect to participate, those who support hands-off government, often Republicans, are less likely to support them or, by extension, have preferences for them. Nearly three quarters of Republicans surveyed in a 2011 opinion poll agreed that SSB taxes were “…an unacceptable intrusion of government into people’s personal lives and individual choices” (Barry, Niederdeppe, and Gollust 2013).

2.3 Data

2.3.1 Survey

The data for this study comes from the New Landscapes of a Minority-Majority State (NLMMS) Survey (found in Appendix C) conducted by the Robert Wood Johnson Foundation Center for Health Policy and Latino Decisions between September 3 and 27,
2016. All adult residents of New Mexico were eligible to participate. The NLMMS survey was a mixed-modes survey where half of the respondents completed the survey over the telephone (603 on landlines and 150 on cellphones) via random-digit dial and the other half (752) responded on the web for a total of 1,505 responses. The telephone sample had an AAPOR response rate of 17.7 percent. Respondents completed the survey in English or in Spanish, based on their discretion and it took between 20-30 minutes to complete.

The NLMMS Survey investigates how the places where New Mexicans live, work, and play impact a wide variety of health measures, views on policies, and lived experiences. Some of those policy viewpoints which focused on SSB taxes were evaluated, in part, by a referendum choice CV. The CV portion of the survey follows National Oceanic and Atmospheric Administration (NOAA) recommendations to maximize validity, including the use of probability sampling, measuring willingness to pay and not to accept, providing a “would not vote” option, not conducting the survey through the mail, posing the question as a hypothetical referendum, and piloting the CV questionnaire (Arrow et al. 1993). We also follow more recent recommendations for conducting CV analysis outlined in Whitehead (2006).

2.3.2 Willingness to Pay Question

The WTP question poses the tax as a specific excise tax, which is consistent with the recommendation in the public health literature (Brownell et al. 2009; Chriqui et al. 2013) and with what respondents would hear on the news. Further, since the tax appears to be small, as low as one penny, we presented a hypothetical scenario that demonstrates the
pre- and post-tax price of a 12-ounce can of soda to reduce the cognitive burden of calculating the tax’s effect. The WTP questions reads:

*Suppose a referendum will be held next week in New Mexico on a sugar-sweetened beverage tax initiative that is designed to fight obesity. The obesity-targeted policies would be financed by a ___ penny-per-ounce tax on all sugar-sweetened beverages (for example, regular soda, sweetened iced teas, sport drinks, and energy drinks). This means that if a 12-ounce can of soda originally cost one dollar, after the tax the same soda would now cost {insert bid’s corresponding dollar value here}. Would you vote for or against this referendum?*

[Randomly select option: 1 ${1.12}$, 2 ${1.24}$, 5 ${1.60}$, 10 ${2.20}$, 15 ${2.80}$ and 25 ${4.00}$]

**Responses: For, Against, Would Note Vote, Don’t Know, and Refused**

Existing SSB tax values and the results of the survey pilot influenced the final bid values. After the WTP question, the survey asked each respondent about their level of certainty for their response on a scale of one (extremely uncertain) to ten (extremely certain).

**2.3.3 Variables**

Columns 1 and 2 of Table 2.1 summarize the variable definitions. There are three dependent variables. The first, YesNo, is a binary indicator equaling 1 if the respondent
accepted their bid value. *YesNo* was coded following Champ et al. (1997) and Johannesson, Liljas, and Johansson (1998) where only respondents who said yes to their bid value with a certainty of 10, the most certain, were kept as a yes. Others who said yes with less than complete certainty were recoded as no. Due to the small percentage of respondents who said they would not vote, at 7%, and to match this behavior in the modeling they were coded as missing along with those who said do not know and refused. This left us with a maximum sample size of 529 since this question was randomly asked to half of the respondents as part of a split-sample design. The next dependent variable is a binary indicator used to identify protest response; respondents who strongly disagreed with expanding the Healthy Diné Nation Act were coded as 0 in *Support_HDNA* and the rest are coded as 1. The last dependent variable is the endogenous SSB consumption variable, *Num_SSB*, which is a continuous measure for the log-transformed average number of SSBs consumed in an average week.

The system uses the following hypothesis-related, independent variables. The bid value *SSB_Tax* was log-transformed. The original less-than-one-penny-per-ounce bid values were dropped from the analysis (resulting in 46 removed observations), and 1½ pennies-per-ounce was recoded as one-penny-per-ounce. The number of fruits and vegetables, *Num_FV*, and the log-transformed number of SSBs consumed in an average week *Num_SSB*, are the two eating habits variables. Two binary variables capture knowledge about health and diet; *Poor_Diet* equals 1 if the respondent strongly agreed that poor diet can lead to being obese or overweight, and *Too_Many_SSB* takes the value of 1 if a respondent strongly agreed that drinking too many SSBs could lead to being obese or overweight. There are two dummy variables for SSB tax policy awareness; the
first one, *Heard_HDNA*, is if the respondent has heard of Healthy Diné Nation Act
before, and *Heard_Other* is if the respondent has heard of any other SSB taxes. These
variables take the value of 1 if the respondent has heard of the respective tax before. We
assessed political ideology on a scale of one to seven with one being very liberal and
seven being very conservative. We coded respondents who answered with a six or seven
as conservative and the rest were coded as non-conservative in the variable *Conservative*.

Other independent binary variables include *Obese* if the respondent has ever been
told by a health professional that they are obese\(^3\); *N_Trust_Gov* if the respondent never
trusts the government to do what is right; and *Obesity_Problem* if the respondent strongly
agrees that both childhood obesity and obesity, more generally, are major problems in
New Mexico.

Lastly, the system uses the following controls: sex (female=1), *Female*; age, *Age*,
and age squared in years, *Age2*; race/ethnicity (non-Hispanic white=1), *non-HW*; mode of
survey (phone=1, web=0), *Phone*; a binary measure for income if the respondent’s
household income in 2015 was over $50,000\(^4\), *over_50k*; three categories of education

\(^3\) Nearly 17% of the sample has been told by a medical professional that they are obese.

Although this number is much lower than the state average adult obesity rate of 28.8%
(Centers for Disease Control and Prevention 2017), our measure captures something
different. This is a measure of awareness of weight status, not of actual weight. Although
self-reported weight and height were collected, which would allow for calculation of
BMI, they were not used due to well-known issues in self-reporting weight.

\(^4\) All respondents who did not respond to the categorical income question answered if
their income was above or below $50,000. However, within the categorical income
where less than high school is the base category, high school graduate or GED, and any type of college is the highest group, *Edu_Low, Edu_Med, Edu_High*, respectively; and what metropolitan statistical area (MSA) the respondent lives including Albuquerque *Abq_MSA*, Farmington *Farm_MSA*, Santa Fe (base category) *SF_MSA*, Las Cruces *LC_MSA*, and non-MSA location *non_MSA*.

### 2.3.4 Hypotheses

The three main hypotheses of this study center on the impact that eating behaviors (hypothesis 1); knowledge and awareness of food and related policies (hypothesis 2); and attitudes (hypothesis 3) have on the preferences for SSB taxes. Formally;

\[ H1: \text{Respondents who have healthier eating habits (lower SSB consumption and higher fruits and vegetable consumption) will have increased preferences for SSB taxes.} \]

\[ H2.1: \text{Knowing that a poor diet can lead to being obese or overweight will increase the preference for SSB taxes.} \]

\[ H2.2: \text{The preferences for SSB taxes will be indirectly increased by knowledge of the Healthy Diné Nation Act before or believe that obesity is a major problem.} \]

\[ H2.3: \text{Hearing of other SSB taxes before will indirectly decrease preferences for SSB taxes.} \]

variable, $50,000 was within the category of $40,000-$60,000. This category was coded as being above $50,000. All regressions were rerun with this category being coded as below $50,000 and the findings were unchanged.
21

H3: Hearing of other SSB taxes before or being a conservative will increase the likelihood of being a protest response.

2.4 Empirical Model

2.4.1 Random Utility Framework

Let the indirect utility for individual $i$ be written as $u_{ij} = v_j(w_i, z_i) + \varepsilon_{ij}$, where $w_i$ represents the respondent’s income; $z_i$ is a vector of personal characteristics; and $\varepsilon_{ij}$ is a stochastic error term that assumed to be additively separable. Further, the subscript $j$ takes the value of 1 if the respondent accepts the bid value and 0 if they do not. A respondent will accept their bid value when it exceeds their intrinsic willingness to pay, and this occurs when their indirect utility, with the tax included, exceeds their indirect utility from the status quo as is seen in equation 1:

$$v_1(w_i - bid_i, z_i) + \varepsilon_{1i} > v_0(w_i, z_i) + \varepsilon_{0i} \quad (1)$$

We can estimate the willingness to pay after assuming a distribution and standardizing the equation accordingly.\(^5\)

2.4.2 Empirical Model

The system of equations modeled in this analysis is found in equations 2a through 2c.

\(^5\) We assume an exponential willingness to pay function where the resulting probability of acceptance is standardized by $\sigma$, the unknown error term. This results in an error term that is distributed log-normally with a mean of $e^{\frac{1}{2}\sigma^2}$ and a variance of $e^{2\sigma^2} - e^{\sigma^2}$. 
Equation 2a is a linear regression of the endogenous behavior $y_{1i}$, $Num_{SSB}$, which is log-transformed. Equation 2b is the selection equation whose dependent variable, $y_{2i}^*$, is $Support_{HDNA}$. It takes the value of 1 if $y_{2i}^* > 0$, else 0. The final equation in the system, 2c, is the outcome equation where the dependent variable, $y_{3i}^*$, is $YesNo$. This variable takes the value of 1 if $y_{3i}^* > 0$, else 0. The error term for each equation is $u_{ki}$ for $k=1,2,3$.

Equation 2c follows the random utility framework and estimates both the WTP and the preferences for SSB taxes. The independent variables are added into $X_{3i}$ over three models. Model 1 contains only the bid value, $SSB_{Tax}$. Model 2 includes $SSB_{Tax}$ plus $Num_{SSB}$, $Num_{FV}$, $Obese$, $Poor_{Diet}$, and $N_{Trust}_{Gov}$. Finally, Model 3 includes the variables from Model 2 plus the following controls: $Female$, $Age$, $Age2$, $non-HW$, education ($Edu_{Med}$ and $Edu_{High}$), Phone, MSA ($Abq_{MSA}$, $Farm_{MSA}$, $LC_{MSA}$, and $non-MSA$), and $over_{50}$. There are still two additional modeling considerations that need attention before doing this estimation: endogeneity and protest responses.

To the first point, within equation 2c, we seek to understand how eating behaviors, specifically $Num_{SSB}$, impact preferences for SSB taxes. Yet, this relationship is inherently endogenous. As such, $E[y_{1i}|u_{3i}] \neq 0$. Equation 2a models the endogenous
behavior separately in which $X_{1t}$ is a vector of independent variables: $\text{Age}$, $\text{Age2}$, $\text{non\_HW}$, $\text{over\_50k}$, and $\text{Too\_Many\_SSB}$. The inclusion of the last independent variable satisfies the exclusion restriction.

The next modeling consideration regards the sample selection bias incurred from the protest responses. The data shows that a large proportion of respondents -- 67% -- said no to their bid value (before recoding for certainty). When we examined the refusal of the bid value by the level of support expressed for expanding the Healthy Diné Nation Act, 100% of respondents who strongly opposed expansion refused their bid value. We propose that this group of respondents are giving protest responses; not only are they strongly against expanding an existing SSB tax but they also refused their bid value, regardless of its value. As such, we see these respondents as protesting the payment vehicle of taxes. Altogether this means that there are two data generating processes underlying the actualization of a no value. First there are people who reject their bid value because it exceeds what they are WTP and second, there are individuals who are fundamentally opposed to taxes on SSBs and will say no to any bid value.

If these protest responses were treated as real nos, then the results in Equation 2c would be biased. To address this, we first partition the data into protest responses and non-protest responses similar to the method implemented by Garcia et al. (2009). Then to correct the potential bias, we use a sample selection model (see Strazzera et al. 2003b and Fonta et al. 2009). Equation 2b is a function of $\text{Heard\_HDNA}$, $\text{Heard\_Other}$, $\text{Conservative}$, and $\text{Obesity\_Problem}$ encompassed in $X_{2t}$.

### 2.5 Estimation Strategy
Equation 2c is estimated using a probit regression, which is further modified to attend to the dual concerns of selection bias and endogeneity. For the selection bias issue, Equation 2b is estimated to separate the protest responses from the non-protest responses, where only the latter will be present in the outcome of Equation 2c. Since both outcomes of 2b and 2c are binary, they are jointly estimated using a probit with selection model based on van der Ven and van Pragg (1981). Jointly modeling these equations allows \( \text{cov}(u_{2i}, u_{3i}) \neq 0 \), instead it is equals \( \rho_{23} \). If \( \rho_{23} \) is statistically different from zero, then there is a selection process occurring and the estimation benefits from estimating the equations together.

To address the endogenous behavior arising in Equation 2c, we could implement a control function approach (CFA) following Rivers and Vuong (1988). The first step separately estimates the endogenous relationship using a linear regression with Equation 2a. The covariates in 2a include at least one excluded variable from 2c, such that for all but the final model, Equation 2c, is over-identified and the final model is just identified. The reduced form residuals, \( \hat{u}_{1i} \), are obtained from the OLS regression of 2a. The second step incorporates \( \hat{u}_{1i} \) and \( \gamma_{1i} \) into the probit regression of equation 2c. The success of the CFA depends upon upholding the assumption that \( (u_{3i}, u_{1i}) \) are independent of \( X_{1i} \).

However, there are several concerns with the CFA. First, the inclusion of \( \hat{u}_{1i} \) in Equation 2c results in generator regressor bias, which bootstrapping the standard errors can mitigate (Wooldridge 2010). Second, estimating the system in two steps results in ignoring the selection process and reduces the overall efficiency and robustness.

To address these limitations, we simultaneously estimate the three-equation endogenous recursive system using a full information maximum likelihood function. The
error terms \( u_{k,l} \) where \( k=1,2,3 \) have a multivariate normal distribution with the following properties:\(^6\):

\[
\begin{bmatrix}
  u_{1,l} \\
  u_{2,l} \\
  u_{3,l}
\end{bmatrix} 
\sim \mathcal{N} (0, \Sigma) \quad \text{where} \quad \Sigma = 
\begin{bmatrix}
  \sigma_1^2 & \rho_{12} & \rho_{13} \\
  \rho_{12} & 1 & \rho_{23} \\
  \rho_{13} & \rho_{23} & 1
\end{bmatrix}.
\]

After estimating the three-equation system, we move on to calculating the conditional average marginal effects and the median willingness to pay. Calculating the marginal effect changes the meaning of the covariates in Equation 2b. Previously, they indicated the propensity to be a protest response, but with the marginal effects, these coefficients now demonstrate the indirect impact they have on preferences for SSB taxes.

To calculate the WTP for the SSB tax, we use Model 2 to estimate the individual median WTP for each respondent, as is done in Cameron (1988). We focus on the median value because of the skewed nature of the WTP data. Further, we calculated the median individual WTP estimates in two ways: first, by not adjusting for protest responses and

---

\(^6\) In order to estimate the multi-equation system, the errors of the system must share a multivariate normal distribution and meet the following two conditions: (1) recursivity, where the equations can be arranged such that the matrix of coefficients of the endogenous variables are triangular and (2) the right-hand side endogenous variables are not latent variables. When all of the equations are structural, like they are here, then the estimation is full-information maximum likelihood. The log of the likelihood is computed along with the first derivatives (analytically) and the second derivatives (numerically).
then by adjusting for them which results in the unconditional and conditional individual median WTP, respectively. In both situations, we bootstrapped to obtain the individual median WTP values and stored the ensuing median value. Further, we trimmed the top 1%, 5%, and 10% of outlying values to remove extreme outliers and stored their respective median value. We repeated this process 500 times. Finally, we constructed the 95% percentile confidence interval and presented for the untrimmed and each trimmed WTP estimate from the unconditional and conditional calculations.

2.6 Results

2.6.1 Summary Statistics

We calculated the summary statistics in Table 2.1 for the observations in the final model. Note *Num_SSB* is presented in the dependent variables section only to avoid repetition.

---

7 The unconditional individual median WTP was estimated using $E[y'_{3i} | X'_{3i}, y'_{1i}] = \exp\{X'_{3i} \hat{\beta}_3 + y'_{1i} \hat{\gamma}_3\}$. The conditional individual median WTP was estimated using $E[y'_{3i} | \{y'_{2} > 0\}, X'_{3i}, y'_{1i}] = \exp\{X'_{3i} \hat{\beta}_3 + y'_{1i} \hat{\gamma}_3 + \hat{\rho}_{23} (-\hat{k}_i^{-1}) \lambda(X'_{2i} \hat{\beta}_2)\}$, where $\hat{\rho}_{23}$ is the correlation coefficient between equation (2b) and (2c), $\hat{k}_i^{-1}$ is the coefficient on the bid value in Equation (2c) which functions as the variance, and $\lambda(X'_{2i} \hat{\beta}_2) \equiv \phi(X'_{2i} \hat{\beta}_2)/\Phi(X'_{2i} \hat{\beta}_2)$ is the inverse mills ratio. If $\hat{\rho}_{23} > 0$, then the unconditional WTP estimate will be downward biased compared to the conditional WTP, whereas if $\hat{\rho}_{23} < 0$ then the unconditional WTP estimate will be upward biased.
The last three columns on the right present the mean and standard deviation for the total sample, the protest responses, and the non-protest responses, respectively.

The average bid value for the total sample, Column 3 in Table 2.1., was 8.5 pennies-per-ounce. On average, respondents reported consuming around three SSBs a week and eating fruits and vegetables nearly daily. The two measures of knowledge around food and health outcomes show that more people, at 72.73%, understand that there is a link between poor diet and weight outcomes than those who know that drinking too many SSBs can lead to being obese or overweight, at only 55.08%. More respondents were aware of other SSB taxes passed across the country, at 44.81%, than the Healthy Diné Nation Act, at 22.95%; overall SSB tax awareness overall was quite low.

When the mean of the variables that influence the WTP differ considerably between the two groups (i.e. protest versus non-protest responses) this provides a good first empirical indication that sample selection is present (Fonta and Omoke 2008; Strazzera et al. 2003a). There are several variables where there are large differences in the mean between the two groups. Non-protest respondents have an 8.6 percentage point higher rate of agreeing with Poor_Diet than protest responses, and for Too_Many_SSB, there is an 8.3 percentage point difference. There is a split on who is more aware of SSB policies, although Heard_HDNA is higher amongst non-protest responses (24.05% of non-protest vs. 16.69% of protest). Heard_Other is higher for protest responses (56.45% of protest vs. 40.98% of non-protest). Moreover, more protest responses were Conservative and N_Trust_Gov. All of these together provide the baseline warning that there indeed is a sample selection problem that needs addressing.
2.6.2 Empirical Results

Throughout the results section, we explore only the three-equation system, found in Table 2.2., in detail. Tables 2.2.A.-2.2.C. contain the full results for the probit, probit with selection model, and probit with selection with a CFA.

We estimated Equation 2b of the system to identify protest responses. Nearly all of the variables in Equation 2b were statistically significant across all models. 

*Obesity_Problem* yields a positive impact on the propensity to support SSB taxation. We find that *Conservative* is statistically significantly less likely to support expand SSB taxes, in Model 2 and 3, and hence, are more likely to be a protest response.

Previous SSB tax policy awareness had mixed impacts. On one hand, *Heard_HDNA* decreased the likelihood of being a protest response, whereas *Heard_Other* increased the likelihood of being a protest response. This echoes what was seen in the descriptive statistics: protest responses had a higher awareness of *Heard_Other* whereas non-protest respondents were more aware of *Heard_HDNA*. The data did not reject hypothesis, *H3* -stating that respondents who have heard of other SSBs and are conservative are more likely to be a protest response.

To address the endogeneity of *Num_SSB* in Equation 2c, we model *Num_SSB* separately in Equation 2a. The correlation coefficient between Equations 2c and 2a, $\hat{\rho}_{13}$, is statistically significant, indicating that the regressor is not exogenous; hence, the modeling is more efficient when estimating the equations simultaneously. The coefficient
on the $Num_{SSB}$ in 2c is negative, statistically significant, and extremely similar in
magnitude to the probit with selection with a CFA\textsuperscript{8}.

The top of Table 2.2. contains the estimates for the direct influences on SSB preferences. As expected, $SSB_{Tax}$ is negative and statistically significant in all three models. This conforms to economic theory in which the higher the bid value, the less willing a respondent would be to accept the presented bid value. There is a negative and statistically significant impact of $Num_{SSB}$ on the preferences for SSB taxes. The proposed tax would directly impact these respondents more, regardless of the size of the tax, and hence, they would be more likely to reject it. As for the other eating habit variable, we found a positive relationship between $Num_{FV}$ and $YesNo$. Together, these support the healthy eating hypothesis, $H1$. Agreeing with $Poor_{Diet}$ statistically significantly increased the preference for SSB taxes. However, in the final model this variable is no longer statistically significant. It could be that the impact of health knowledge washes out other covariates, like education and income. Together, this provides mild support for $H2.1$.

2.6.3 Marginal Effects

\textsuperscript{8} The probit with selection using a CFA also provides a test of exogeneity for the potentially endogenous regressor in Equation 2c. The t-stat on the included reduced form residual from the endogenous behavior regression is statistically different from 0 in all three models, hence we fail to accept that this regressor is exogenous (Rivers and Vuong 1988).
Not only do the covariates in Equation 2b function to identify the protest responses, by taking the conditional average marginal effects, these variables demonstrate the indirect impact on preferences for SSB taxes found in the even columns of Table 2.2.

The impact of SSB tax policy awareness maintains the same sign as the raw coefficients. Heard_HDNA indirectly increases the preferences for SSB taxes by 4.5 percentage points, whereas Heard_Other decreases these preferences by 4.4 percentage points in Model 2. As such, for the well-informed respondent, the cumulative influence of knowledge on previous SSB policies has a nearly no discernable impact on preferences for SSB taxes. Identifying as Conservative indirectly decreases the preference for these taxes by 4.7 percentage points. Lastly, respondents who agreed with Obesity_Problem indirectly increased preferences for SSB taxes by 4.4 percentage points. All together, these results support our initial hypotheses of H2.2. and H2.3.

As for Equation 2c, the magnitude of the SSB_Tax is small, but statistically significant. If the tax increased by 10%, then the preference for SSB taxes would decrease by half a percentage point. Additionally, if Num_SSB increased by a modest 10%, then preferences for these taxes would decrease by 2 percentage points, but if consumption increased by 50% the preferences would decrease by 10 percentage points. Finally, those who agreed with Poor_Diet were statistically significantly associated with an increase in preferences for SSB taxes by close to 9 percentage points.

2.6.4 Summary of Regression Results
To summarize our hypotheses and demonstrate the robustness of our results, all models were re-estimated using a simple probit regression and a probit with selection model with
survey weights and robust standard errors. For each regression method, the results were similar with the addition of covariates. As such, Table 2.3. presents the hypothesis relevant results for Model 2 of the preferred three-equation system\(^9\) along the naïve probit and probit with selection regression.

The odd numbered columns in Table 2.3. present the raw regression results and the even numbered columns contain the condition average marginal effects (for the probit regression the average marginal effects are reported). The results across the three modeling techniques produce consistent results: each respective variable has the same sign and similar level of statistical significance. Both of the sensitivity analyses results ignore endogeneity and profoundly underestimate the impact \(\text{Num}_\text{SSB}\) has on preferences for SSB taxes. The conditional average marginal effect on \(\text{Num}_\text{SSB}\) for the three-equation system is two to two-and-a-half times larger than the corresponding results from the naïve models.

2.6.5 Estimated Willingness to Pay

Figure 2.1 presents the histogram of the untrimmed conditional median of the individual median WTP. This histogram is highly skewed, so using the delta-method or other standard error estimates that rely upon normally distributed data would lead to inaccurate results. That is why we use the adistributional percentile method to calculate the 95% confidence intervals in Table 2.4.. The central tendency of the skewed data, the green line in the histogram, shows the median value of the bootstrapped values at 0.167 pennies-per-ounce.

\(^9\) This model had the lowest BIC when all three equations were estimated simultaneously.
Table 2.4 presents the 95% confidence intervals of the median value of the bootstrapped median individual WTP estimates for Model 2 of the three-equation system including and excluding the selection process. The unconditional and conditional WTP estimates are presented by the level of trimmed results (including untrimmed).

The first row of Table 2.4 presents the conditional (with selection) 95% CI for the median WTP. The untrimmed estimates range from 0.002 to 0.961 pennies-per-ounce. This range does not contain the commonly proposed SSB taxes of one-penny-per-ounce. The most conservative of estimates, trimmed of the top 10% of values, does not contain a half-penny-per-ounce. However, in the unconditional estimates in row 2, which ignore the influence of protest responses, the 95% CI becomes over inflated, which occurs because $\hat{\rho}_{23} < 0$. This range includes one-penny-per-ounce. Policy makers who disregard protest responses when determining an SSB tax value would falsely expect one-penny-per-ounce tax to pass, but this is unlikely to be the case because those values greatly exceed what New Mexicans are truly willing to pay.

2.7 Discussion and Summary

This paper is one of the first to use CV to estimate the willingness to pay for a SSB tax. We modified the basic CV framework to incorporate concerns of endogeneity and protest responses. With the use of primary data, we estimated the preferences for SSB taxes using a full-information maximum likelihood three-equation recursive endogenous system. The results, robust to sensitivity analyses, found that healthier eating habits statistically significantly increase the preferences for SSB taxes, as did knowing that a poor diet can lead to being obese or overweight. The impact of policy awareness had
mixed impacts on indirectly increasing preferences. Respondents who are conservative or aware of other SSB taxes before were more likely to be a protest response. The bootstrapped median individual WTP 95% confidence interval of 0.002-0.961 pennies-per-ounce is lower than many previously proposed SSB taxes.

On May 2, 2017, voters in Santa Fe, New Mexico, participated in a special election on a two-penny-per-ounce tax on SSBs where the revenue would go towards funding Pre-K education (Gonzales 2017). The initiative failed with 58% of voters rejecting this tax (Chacón 2017). The findings from our study (conducted several months before the SSB tax was formally proposed by the city council) provide insights on why it failed. First, our results show that the tax value was too high. The conditional 95% CI on the individual median WTP was between 0.002 to 0.961-pennies-per-ounce, which is substantially lower than the proposed two-pennies-per-ounce. However, this value was included in the unconditional 95% CI. It could have been the case that policy makers derived this value after talking to their constituents who were sympathetic towards the initiative, but overestimation of this value contributed to its downfall.

Additionally, we found that the factors associated with increased preferences for SSB taxes are low in New Mexico. Our results indicate that when a respondent knows that drinking too many SSBs can lead to being obese or overweight strongly correlated with decreased SSB consumption, which in turn increased preferences for SSB taxes. However, New Mexicans overall have a low awareness of the potential health consequences of SSB consumption; only 55% of respondents were certain about this link. This is much lower than a national study that found that 84.4% of adults agreed that SSB consumption could lead to weight gain (Park et al. 2014a). The results from the three-
equation system also demonstrate that believing that obesity is a major problem in the state indirectly increases preferences for SSB taxes. Once again, a small percentage, 40%, believe this statement to be true. Overall, with low awareness of public health concerns in the state and health knowledge compounded with an overinflated tax value, this proposal had a slim chance of passing.

Our findings regarding the impact eating behavior, knowledge and awareness around food and related policies, and political ideology have on preferences for SSB taxes are similar to a previous CV study conducted in the US. Cawley (2008) used a triple bound, dichotomous choice study to estimate the mean WTP to reduce childhood obesity by 50% in New York. The results on political ideology, opinion, and framing are in line with our results. Liberals were willing to pay more than their moderate counterparts and so were respondents who thought that obesity was amongst the most pressing issues facing youth. The study also found that respondents who thought that obesity was due to individual choices or genetics had statistically lower WTP than those who thought obesity was an environmental problem. On the other hand, findings from a study that assessed the WTP for a weight reduction in Taiwan, found that the health knowledge index (knowing that certain diseases relate to obesity) was not statistically significant (Fu, Lin and Huang 2011). The discrepancy in findings on the role of health knowledge has on preferences for obesity related polices could be due to cultural differences and the prevalence of obesity.

There are several limitations to this study. First, the data for the study is exclusively from New Mexico. As such, there are limits to the extent that the findings can be generalized, but that is not to say that our results are not generalizable since many
findings are in agreement with previous studies. Across the nation, the awareness of the Healthy Diné Nation Act is likely to be low. However, this is a strength of the study; with primary data, we are able to focus on a relevant local policy. Future studies could use this approach in other localities to determine if they also experience a discrepancy in policy support between local and non-local policies. Additionally, estimating a point value for the WTP it might be too specific to New Mexico. We attempt to mitigate this concern by presenting the 95% confidence interval instead of the point estimate.

The WTP values were sensitive to the presence of extreme outlying observations - those who consumed a larger number of fruits and vegetables and no SSBs in the average week, which is why we calculated the individual median WTP over the median WTP. The estimates were highly skewed, so we trimmed the estimates of the top 1, 5 and 10 percentile of values to obtain more conservative ranges of the WTP value.

Another limitation of the study is that the method used in recoding the dependent variable was very strict. Champ et al. (1997) found this form of recoding resulted in lower bound estimates and were more similar to actually passed entrance fees in environmental CV studies. We also do not know where and how people are hearing about SSB taxes. The literature would benefit from increased information on where people hear about SSB taxes and studies around why specific SSB taxes have passed or failed. This study is but one that partially addresses this issue.

The findings from this study have important policy implications for other SSB taxes developed across the US. Knowing that poor diet can lead to being overweight or obese greatly increased the preference for SSB taxes, as such implementing, say, a public health campaign before the proposal of a SSB tax could garner support for the policy.
Additionally, increasing the awareness around the current level of obesity in the locality and explaining the implication of this disease’s prevalence can be another means to increase support. It is critical that newly proposed SSB taxes find a way to differentiate themselves from other SSB taxes. Finally, and importantly, our results show that going with the status quo SSB tax of one-penny-per-ounce is not always a formula for success.
Table 2.1.: Summary Statistics

<table>
<thead>
<tr>
<th>Variable</th>
<th>Description</th>
<th>Total Mean (SD)</th>
<th>Protest Responses Mean (SD)</th>
<th>Non-Protest Responses Mean (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Dependent Variables</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>YesNo</td>
<td>=1 if the bid value was accepted, 0 otherwise.</td>
<td>13.76%</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Support_HDNA</td>
<td>=1 if the respondent is not strongly opposed to expanding the Healthy Diné Nation Act, 0 otherwise.</td>
<td>74.83%</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Num_SSB</td>
<td>The number of SSBs consumed in an average week.</td>
<td>3.186 (5.656)</td>
<td>3.587 (6.898)</td>
<td>3.051 (5.176)</td>
</tr>
<tr>
<td><strong>Independent Variables</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SSB_Tax</td>
<td>The bid value presented to the respondent in cents per ounce.</td>
<td>8.534 (8.313)</td>
<td>7.515 (7.529)</td>
<td>8.876 (8.544)</td>
</tr>
<tr>
<td>Num_FV</td>
<td>The number of fruits and vegetables consumed in an average week.</td>
<td>6.774 (4.501)</td>
<td>6.644 (4.352)</td>
<td>6.818 (4.556)</td>
</tr>
<tr>
<td>Poor_Diet</td>
<td>=1 if respondent strongly agrees that a poor diet can lead to being obese or overweight, 0 otherwise.</td>
<td>72.73%</td>
<td>66.28%</td>
<td>74.90%</td>
</tr>
<tr>
<td>Too_Many_SSB</td>
<td>=1 if respondent strongly agrees that drinking too many sodas can lead to being obese or overweight, 0 otherwise.</td>
<td>55.08%</td>
<td>48.85%</td>
<td>57.18%</td>
</tr>
<tr>
<td>Heard_HDNA</td>
<td>=1 if respondent has heard of the Healthy Diné Nation Act before, 0 otherwise.</td>
<td>22.95%</td>
<td>16.69%</td>
<td>24.05%</td>
</tr>
<tr>
<td>Heard_Other</td>
<td>=1 if respondent has heard of other SSB taxes before, 0 otherwise.</td>
<td>44.81%</td>
<td>56.45%</td>
<td>40.90%</td>
</tr>
<tr>
<td>Conservative</td>
<td>=1 if respondent is a conservative, 0 otherwise.</td>
<td>18.43%</td>
<td>29.49%</td>
<td>14.63%</td>
</tr>
<tr>
<td>Variable</td>
<td>Description</td>
<td>Controls</td>
<td>2019</td>
<td>2020</td>
</tr>
<tr>
<td>-------------------</td>
<td>-----------------------------------------------------------------------------</td>
<td>----------</td>
<td>------</td>
<td>------</td>
</tr>
<tr>
<td>Obese</td>
<td>=1 if told by medical profession respondent is obese, 0 otherwise.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N_Trust_Gov</td>
<td>=1 if respondent never trusts that the state government can do what is right, 0 otherwise.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Obesity_Problem</td>
<td>=1 if respondent believes both childhood and general obesity are major problems in New Mexico, 0 otherwise</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>=1 if female, 0 male.</td>
<td>51.64%</td>
<td>45.51%</td>
<td>53.70%</td>
</tr>
<tr>
<td>Age</td>
<td>Age of respondent in years.</td>
<td>49.572</td>
<td>52.403</td>
<td>48.629</td>
</tr>
<tr>
<td>(18.168)</td>
<td>(15.823)</td>
<td>(18.818)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>non_HW</td>
<td>=1 if respondent is non-Hispanic white, 0 otherwise.</td>
<td>45.50%</td>
<td>56.61%</td>
<td>41.76%</td>
</tr>
<tr>
<td>Phone</td>
<td>=1 if respondent completed survey over the phone, 0 over the web.</td>
<td>53.47%</td>
<td>53.65%</td>
<td>53.40%</td>
</tr>
<tr>
<td>Over_50k</td>
<td>=1 if respondent's household income was above $50,000 in 2015 before taxes, 0 otherwise.</td>
<td>53.57%</td>
<td>55.60%</td>
<td>52.89%</td>
</tr>
<tr>
<td>Edu_Low</td>
<td>=1 if respondent has a high school diploma, GED equivalent, or less than a high school diploma, 0 otherwise.</td>
<td>22.29%</td>
<td>19.47%</td>
<td>23.24%</td>
</tr>
<tr>
<td>Edu_Med</td>
<td>=1 if respondent has had some college education, 0 otherwise.</td>
<td>26.49%</td>
<td>29.95%</td>
<td>25.33%</td>
</tr>
<tr>
<td>Edu_High</td>
<td>=1 if respondent is a college graduate, 0 otherwise.</td>
<td>51.22%</td>
<td>50.58%</td>
<td>51.43%</td>
</tr>
</tbody>
</table>

Metropolitan Statistical Area

38
<table>
<thead>
<tr>
<th>Variable</th>
<th>Description</th>
<th>Farmington MSA</th>
<th>Santa Fe MSA</th>
<th>Las Cruses MSA</th>
<th>Albuquerque MSA</th>
<th>Non-MSA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Farm_MSA</td>
<td>=1 if respondent lives in Farmington MSA, 0 otherwise.</td>
<td>3.08%</td>
<td>2.95%</td>
<td>3.13%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SF_MSAc</td>
<td>=1 if respondent lives in Santa Fe MSA, 0 otherwise.</td>
<td>7.58%</td>
<td>10.19%</td>
<td>6.70%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LC_MSA</td>
<td>=1 if respondent lives in Las Cruses MSA, 0 otherwise.</td>
<td>10.70%</td>
<td>5.85%</td>
<td>12.33%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Abq_MSA</td>
<td>=1 if respondent lives in Albuquerque MSA, 0 otherwise.</td>
<td>46.31%</td>
<td>46.31%</td>
<td>46.31%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-MSA</td>
<td>=1 if respondent does not live in an MSA, 0 otherwise.</td>
<td>32.32%</td>
<td>34.69%</td>
<td>31.53%</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Notes: aStandard deviation only presented for continuous variables. bIndicates that the variable has been recoded such that responses of “Don't Know” and “Refuse” are given the value of 0. cBase category, in binary variables the 0 group is the base category. All summary statistics are weighted. Only observations included in the final model are presented here. Weighted sample size is 522. Source: NLMMS Survey.
Table 2.2.: Three Equation System Results for Preferences of SSB Taxes

<table>
<thead>
<tr>
<th>Equation (2c)</th>
<th>Model 1</th>
<th>Model 2</th>
<th>Model 3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Raw</td>
<td>Marginal Effects</td>
<td>Raw</td>
</tr>
<tr>
<td>SSB_Tax</td>
<td>-0.144***</td>
<td>-0.053***</td>
<td>-0.182***</td>
</tr>
<tr>
<td></td>
<td>(0.053)</td>
<td>(0.017)</td>
<td>(0.057)</td>
</tr>
<tr>
<td>Num_SSB</td>
<td>-0.732***</td>
<td>-0.204***</td>
<td>-0.846***</td>
</tr>
<tr>
<td></td>
<td>(0.184)</td>
<td>(0.043)</td>
<td>(0.204)</td>
</tr>
<tr>
<td>Num_FV</td>
<td>-0.035***</td>
<td>0.010***</td>
<td>0.037***</td>
</tr>
<tr>
<td></td>
<td>(0.013)</td>
<td>(0.004)</td>
<td>(0.014)</td>
</tr>
<tr>
<td>Obese</td>
<td>0.082</td>
<td>0.023</td>
<td>0.085</td>
</tr>
<tr>
<td></td>
<td>(0.138)</td>
<td>(0.039)</td>
<td>(0.130)</td>
</tr>
<tr>
<td>Poor_Diet</td>
<td>0.319**</td>
<td>0.087**</td>
<td>0.189</td>
</tr>
<tr>
<td></td>
<td>(0.160)</td>
<td>(0.038)</td>
<td>(0.146)</td>
</tr>
<tr>
<td>N_Trust_Gov</td>
<td>-0.065</td>
<td>-0.018</td>
<td>-0.125</td>
</tr>
<tr>
<td></td>
<td>(0.167)</td>
<td>(0.046)</td>
<td>(0.148)</td>
</tr>
<tr>
<td>Controls</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Equation (2b)

| Heard_HDNA      | 0.383*** | 0.081*** | 0.424*** | 0.045*** | 0.475*** | 0.045*** |
|                 | (0.138)    | (0.028)    | (0.157)    | (0.016)    | (0.159)    | (0.015)    |
| Heard_Other     | -0.324**  | -0.068*** | -0.399*** | -0.044*** | -0.395*** | -0.040*** |
|                 | (0.132)    | (0.023)    | (0.130)    | (0.014)    | (0.137)    | (0.013)    |
| Conservative    | -0.274    | -0.057*   | -0.401*   | -0.047**  | -0.454**  | -0.049**  |
|                 | (0.186)    | (0.033)    | (0.210)    | (0.022)    | (0.200)    | (0.02)    |
| Obesity_Problem | 0.548***  | 0.120***  | 0.389***  | 0.044***  | 0.315**   | 0.032**   |
|                 | (0.111)    | (0.024)    | (0.133)    | (0.016)    | (0.132)    | (0.013)    |

Equation (2a)

| Too_Many_SSB    | -        | -        | -0.497*** | -        | -0.462*** | -        |
|                 | (0.081)    | (0.087)    | (0.012)    | (0.012)    | (0.001)    | (0.001)    |
| Age             | -        | -        | -0.037*** | -        | -0.030**  | -        |
|                 | (0.012)    | (0.012)    | (0.0001)   | (0.0001)   | (0.088)    | (0.089)    |
| Age2            | -        | -        | 0.0002*   | -        | 0.0002    | -        |
|                 | (0.008)    | (0.008)    | (0.088)    | (0.088)    | (0.089)    | (0.089)    |
| non_HW          | -        | -        | 0.020     | -        | -0.046    | -        |
|                 | (0.089)    | (0.089)    | (0.089)    | (0.089)    | (0.089)    | (0.089)    |
| Over_50k        | -        | -        | -0.272*** | -        | -0.336*** | -        |
|                 | (0.089)    | (0.089)    | (0.089)    | (0.089)    | (0.089)    | (0.089)    |
| \(\rho_{12}\)  | -        | -        | -0.019    | -        | -0.034    | -        |
|                 | (0.075)    | (0.075)    | (0.075)    | (0.075)    | (0.075)    | (0.075)    |
| \(\rho_{13}\)  | -        | -        | 0.462**   | -        | 0.586***  | -        |
|                 | (0.153)    | (0.168)    | (0.153)    | (0.168)    | (0.153)    | (0.168)    |
| \(\rho_{23}\)  | -0.938*** | -        | -0.782*** | -        | -0.745*** | -        |
|                 | (0.053)    | (0.136)    | (0.136)    | (0.136)    | (0.136)    | (0.136)    |
| AIC             | 855.752    | 1861.56   | 1833.374  | -        | -        | -        |
|-----------|---------|----------|----------|
| Observations | 577    | 577      | 529      | 529      | 522      | 522      |

*Notes:* The dependent variable for the equation 2a, 2b, and 2c are log of number of SSBs consumed, if the respondent is not strongly opposed to expanding the Healthy Diné Nation Act, and if the respondent accepted their bid value, respectively. Controls in the outcome equation include female, age, age squared, education, non-Hispanic white, survey was conducted over the phone, MSA location, and income above $50,000. The marginal effect is probability of accepting the bid value conditional on being a non-protest response. All regressions use importance weights. Robust standard errors reported in parentheses. The asterisks indicate the level of statistical significance: ***p ≤ 0.01, **p ≤ 0.05, and *p ≤ 0.1. *Source:* NLMMS Survey.
<table>
<thead>
<tr>
<th></th>
<th>Model 1</th>
<th>Model 2</th>
<th>Model 3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
</tr>
<tr>
<td></td>
<td>Raw</td>
<td>Marginal Effects</td>
<td>Raw</td>
</tr>
<tr>
<td>SSB_Tax</td>
<td>-0.224***&lt;br&gt;(0.063)</td>
<td>-0.048***&lt;br&gt;(0.014)</td>
<td>-0.252***&lt;br&gt;(0.072)</td>
</tr>
<tr>
<td>Num_SSB</td>
<td>-</td>
<td>-</td>
<td>-0.445***&lt;br&gt;(0.105)</td>
</tr>
<tr>
<td>Num_FV</td>
<td>-</td>
<td>-</td>
<td>0.042***&lt;br&gt;(0.015)</td>
</tr>
<tr>
<td>Obese</td>
<td>-</td>
<td>-</td>
<td>0.132&lt;br&gt;(0.201)</td>
</tr>
<tr>
<td>Poor_Diet</td>
<td>-</td>
<td>-</td>
<td>0.814***&lt;br&gt;(0.187)</td>
</tr>
<tr>
<td>N_Trust_Gov</td>
<td>-</td>
<td>-</td>
<td>-0.244&lt;br&gt;(0.207)</td>
</tr>
<tr>
<td>Controls</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>AIC</td>
<td>389.448</td>
<td>-</td>
<td>312.514</td>
</tr>
<tr>
<td>BIC</td>
<td>398.163</td>
<td>-</td>
<td>342.41</td>
</tr>
<tr>
<td>Observations</td>
<td>577</td>
<td>577</td>
<td>529</td>
</tr>
</tbody>
</table>

**Notes:** The dependent variable for equation 2c is if the respondent accepted their bid value. Controls in the outcome equation include female, age, age squared, education, non-Hispanic white, survey was conducted over the phone, MSA location, and income above $50,000. All regressions use importance weights. Robust standard errors reported in parentheses. The asterisks indicate the level of statistical significance: ***p≤0.01, **p≤0.05, and *p≤0.1. **Source:** NLMMS Survey.
Table 2.2.B.: Probit Results for Preferences of SSB Taxes

<table>
<thead>
<tr>
<th></th>
<th>Model 1</th>
<th></th>
<th>Model 2</th>
<th></th>
<th>Model 3</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Raw</td>
<td>Marginal</td>
<td>Raw</td>
<td>Marginal</td>
<td>Raw</td>
<td>Marginal</td>
</tr>
<tr>
<td><strong>Equation (2c)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SSB_Tax</td>
<td>-0.143***</td>
<td>-0.052***</td>
<td>-0.183***</td>
<td>-0.049***</td>
<td>-0.200***</td>
<td>-0.052***</td>
</tr>
<tr>
<td></td>
<td>(0.055)</td>
<td>(0.017)</td>
<td>(0.066)</td>
<td>(0.015)</td>
<td>(0.0601)</td>
<td>(0.013)</td>
</tr>
<tr>
<td>Num_SSB</td>
<td>-</td>
<td>-</td>
<td>-0.373***</td>
<td>-0.101***</td>
<td>-0.245**</td>
<td>-0.063***</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.111)</td>
<td>(0.023)</td>
<td>(0.0989)</td>
<td>(0.022)</td>
</tr>
<tr>
<td>Num_FV</td>
<td>-</td>
<td>-</td>
<td>0.034**</td>
<td>0.009***</td>
<td>0.0395**</td>
<td>0.010**</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.017)</td>
<td>(0.004)</td>
<td>(0.0162)</td>
<td>(0.004)</td>
</tr>
<tr>
<td>Obese</td>
<td>-</td>
<td>-</td>
<td>0.0637</td>
<td>0.018</td>
<td>0.0734</td>
<td>0.019</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.157)</td>
<td>(0.044)</td>
<td>(0.152)</td>
<td>(0.041)</td>
</tr>
<tr>
<td>Poor_Diet</td>
<td>-</td>
<td>-</td>
<td>0.471**</td>
<td>0.112***</td>
<td>0.391**</td>
<td>0.094***</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.206)</td>
<td>(0.036)</td>
<td>(0.168)</td>
<td>(0.032)</td>
</tr>
<tr>
<td>N_Trust_Gov</td>
<td>-</td>
<td>-</td>
<td>-0.0631</td>
<td>-0.017</td>
<td>-0.112</td>
<td>-0.028</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.196)</td>
<td>(0.051)</td>
<td>(0.163)</td>
<td>(0.040)</td>
</tr>
<tr>
<td>Controls</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>X</td>
<td>X</td>
</tr>
</tbody>
</table>

| **Equation (2b)**    |         |          |         |          |         |          |
| Heared_HDNA          | 0.385*** | 0.079**  | 0.437*** | 0.055*   | 0.481*** | 0.063**  |
|                      | (0.140)  | (0.032)  | (0.161)  | (0.029)  | (0.157)  | (0.026)  |
| Heared_Other         | -0.340** | -0.070***| -0.409***| -0.051** | -0.394***| -0.051***|
|                      | (0.134)  | (0.020)  | (0.129)  | (0.020)  | (0.136)  | (0.016)  |
| Conservative         | -0.291   | -0.060** | -0.407*  | -0.051***| -0.434** | -0.057** |
|                      | (0.191)  | (0.028)  | (0.216)  | (0.018)  | (0.206)  | (0.018)  |
| Obesity_Problem      | 0.539*** | 0.115*** | 0.406*** | 0.053*   | 0.343*** | 0.047**  |
|                      | (0.113)  | (0.037)  | (0.135)  | (0.029)  | (0.128)  | (0.021)  |
| ρ                    | -0.931***| -       | -0.845***| -        | -0.909***| -        |
|                      | (0.062)  |         | (0.129)  |         | (0.076)  |          |
| AIC                  | 864.927  | -       | 760.164  | -        | 745.877  | -        |
| BIC                  | 899.790  | -       | 815.687  | -        | 852.318  | -        |
| Observations         | 577      | 577     | 529      | 529      | 522      | 522      |

Notes: The dependent variable for the equation 2b and 2c are if the respondent is not strongly opposed to expanding the Health Diné Nation Act and if the respondent accepted their bid value, respectively. Controls in the outcome equation include female, age squared, education, non-Hispanic white, survey was conducted over the phone, MSA location, and income above $50,000. The marginal effect is probability of accepting the bid value conditional on being a non-protest response. All regressions use importance weights. Robust standard errors reported in parentheses. The asterisks indicate the level of statistical significance: ***p≤0.01, **p≤0.05, and *p≤0.1. Source: NLMMS Survey.
Table 2.2.C.: Probit with Selection with Control Function Approach Results for Preferences of SSB Taxes

<table>
<thead>
<tr>
<th>Equation (2c)</th>
<th>Model 1</th>
<th>Model 2</th>
<th>Model 3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
</tr>
<tr>
<td>SSB_Tax</td>
<td>-0.143***</td>
<td>-0.052***</td>
<td>-0.193***</td>
</tr>
<tr>
<td></td>
<td>(0.055)</td>
<td>(0.017)</td>
<td>(0.065)</td>
</tr>
<tr>
<td>Num_SSB</td>
<td>-</td>
<td>-</td>
<td>-0.770***</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.235)</td>
</tr>
<tr>
<td>Num_FV</td>
<td>-</td>
<td>-</td>
<td>0.036***</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.014)</td>
</tr>
<tr>
<td>Obese</td>
<td>-</td>
<td>-</td>
<td>0.079</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.157)</td>
</tr>
<tr>
<td>Poor_Diet</td>
<td>-</td>
<td>-</td>
<td>0.403**</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.186)</td>
</tr>
<tr>
<td>N_Trust_Gov</td>
<td>-</td>
<td>-</td>
<td>-0.084</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.187)</td>
</tr>
</tbody>
</table>

Equation (2a)

| Controls      | X       | X       |

Equation (2b)

| Heard_HDNA    | 0.385*** | 0.079**  | 0.425*** | 0.054**  | 0.477*** | 0.060**  |
|               | (0.140)  | (0.032)  | (0.159)  | (0.026)  | (0.159)  | (0.025)  |
| Heard_Other   | -0.340** | -0.070*** | -0.409*** | -0.051*** | -0.402*** | -0.050*** |
|               | (0.134)  | (0.020)  | (0.129)  | (0.019)  | (0.135)  | (0.016)  |
| Conservative  | -0.291   | -0.060**  | -0.403*  | -0.051*** | -0.446**  | 0.056***  |
|               | (0.191)  | (0.028)  | (0.214)  | (0.018)  | (0.205)  | (0.017)  |
| Obesity_Problem | 0.539*** | 0.115**  | 0.384*** | 0.051**  | 0.314**  | 0.041**  |
|               | (0.113)  | (0.037)  | (0.132)  | (0.025)  | (0.131)  | (0.019)  |

Equation (2a)

| Too_Many_SSB  | -        | -        | -0.481*** | -        | -0.481*** | -        |
|               |          |          | (0.079)   |          | (0.079)   |          |
| Age           | -        | -        | -0.032*** | -        | -0.032*** | -        |
|               |          |          | (0.012)   |          | (0.012)   |          |
| Age2          | -        | -        | 0.0002    | -        | 0.0002    | -        |
|               |          |          | (0.0001)  |          | (0.0001)  |          |
| non_HW        | -        | -        | -0.029    | -        | -0.029    | -        |
|               |          |          | (0.081)   |          | (0.081)   |          |
| Over_50k      | -        | -        | -0.319*** | -        | -0.319*** | -        |
|               |          |          | (0.080)   |          | (0.080)   |          |

\( \rho \) -0.931*** -0.866*** -0.904*** -
AIC 864.927 - 754.003 - 744.747 -
BIC 899.790 - 813.797 - 855.446 -
Observations 577 577 529 529 522 522

Notes: The dependent variable for the equation 2a, 2b, and 2c are log of number of SSBs consumed, if the respondent is not strongly opposed to expanding the Health Diné Nation Act, and if the respondent accepted their bid value, respectively. Controls in the outcome equation include female, age, age squared, education, non-Hispanic white, survey was conducted over the phone, MSA location, and income above $50,000. The marginal effect is probability of accepting the bid value conditional on being a non-protest response. All regressions use importance weights. Robust standard errors reported in parentheses. The asterisks indicate the level of statistical significance: ***p≤0.01, **p≤0.05, and *p≤0.1. Source: NLMMS Survey.
Table 2.3.: Hypothesis Table of Three Equation System and Sensitivity Analyses Raw Results and Conditional Average Marginal Effects for Willingness to Pay for a SSB Tax for Model 2

<table>
<thead>
<tr>
<th>Equation (2c)</th>
<th>Three-Equation System</th>
<th>Sensitivity Analyses</th>
<th>Sensitivity Analyses</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Raw Marginal Effect</td>
<td>Probit Raw Marginal Effect</td>
<td>Probit with Selection Raw Marginal Effect</td>
</tr>
<tr>
<td>SSB_Tax</td>
<td>-0.182***  -0.049***</td>
<td>-0.252*** -0.046***</td>
<td>-0.183*** -0.049***</td>
</tr>
<tr>
<td></td>
<td>(0.057)    (0.014)</td>
<td>(0.072)   (0.013)</td>
<td>(0.066)   (0.015)</td>
</tr>
<tr>
<td>Num_SSB</td>
<td>-0.732***  -0.203***</td>
<td>-0.445*** -0.081***</td>
<td>-0.373*** -0.101***</td>
</tr>
<tr>
<td></td>
<td>(0.184)    (0.044)</td>
<td>(0.105)   (0.018)</td>
<td>(0.111)   (0.023)</td>
</tr>
<tr>
<td>Num_FV</td>
<td>0.035***   0.009**</td>
<td>0.042**   0.008***</td>
<td>0.034**   0.009***</td>
</tr>
<tr>
<td></td>
<td>(0.013)    (0.004)</td>
<td>(0.015)   (0.003)</td>
<td>(0.014)   (0.004)</td>
</tr>
<tr>
<td>Poor_Diet</td>
<td>0.319**    0.095**</td>
<td>0.814***  0.123***</td>
<td>0.471**   0.116***</td>
</tr>
<tr>
<td></td>
<td>(0.160)    (0.038)</td>
<td>(0.187)   (0.024)</td>
<td>(0.206)   (0.036)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Equation (2b)</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Equation (2a)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Too_Many_SSB</td>
<td>-0.497***</td>
</tr>
<tr>
<td></td>
<td>(0.081)</td>
</tr>
</tbody>
</table>

Notes: Results from Model 2 presented in part here. The dependent variable for equation 2a is log of number of SSBs consumed and other explanatory variables include age, age squared, non-Hispanic white, and income. The dependent variable for equation 2b is if the respondent is not strongly opposed to expanding the Health Diné Nation Act. Finally, the dependent variable for equation 2c is if the respondent accepted their bid value and other explanatory variables include if the respondent is obese and never trust the government to do what is right. The average marginal effects are conditional on being a non-protest response. All regressions use importance weights. Robust standard errors reported in parentheses. The asterisks indicate the level of statistical significance: ***p≤0.01, **p≤0.05, and *p≤0.1. Source: NLMMS Survey.
Table 2.4.: Confidence Interval for the Bootstrapped Individual Median WTP Results With and Without Selection Median Value Reported and Trimmed at Varying Levels

<table>
<thead>
<tr>
<th></th>
<th>Untrimmed</th>
<th>1% Trimmed</th>
<th>5% Trimmed</th>
<th>10% Trimmed</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Conditional (With Selection)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Median</td>
<td>(0.002, 0.961)</td>
<td>(0.002, 0.798)</td>
<td>(0.002, 0.557)</td>
<td>(0.0005, 0.427)</td>
</tr>
<tr>
<td><strong>Unconditional (Without Selection)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Median</td>
<td>(0.034, 4.737)</td>
<td>(0.023, 3.752)</td>
<td>(0.015, 2.237)</td>
<td>(0.007, 1.870)</td>
</tr>
</tbody>
</table>

*Notes:* The individual median WTP values were bootstrapped 500 times. Percentile confidence interval presented. *Source:* NLMMS Survey.
Figure 2.1.: Histogram of Bootstrapped Median WTP

Notes: WTP was bootstrapped 500 times. Median of the bootstrapped WTP reported.
Chapter 3. Elements of Support for Sugar-Sweetened Beverage Taxation: Empirical Evidence from New Mexico

3.1. Introduction

Over the past half a century in the United States, obesity rates soared; between 1962 and 2008, the adult obesity rate more than doubled from around 13% to 34% (Ogden and Carroll 2010). This increase in obesity parallels the rise in consumption of sugar-sweetened beverages (SSBs) –which are defined as all liquid beverages that contain an added caloric sweetener (Chriqui et al. 2013). Between 1977 and 1996, soft drink consumption in the US increased by 83% and 71% for people between 2-18 years old and 19-39 years old, respectively (Kim and Kawachi 2006). A systematic review of the literature found strong evidence of SSB consumption leading to weight gain for children and adolescents along with a positive association between SSB consumption and weight gain for adults (Malik, Schulze, and Hu 2006). Due to this connection, coupled with SSBs having no nutritional value, SSBs are a target of taxation. The potential taxation of these products is highly controversial.

Critics of SSB taxes question the tax’s ability to reduce population weight. In this scenario, weight loss strongly depends on both reduced SSB consumption and the caloric value of substituted beverages. If a substituted beverage contains more calories than a SSB there could be an increase in weight, all else held equal. Critics instead recommend a multi-faceted approach to addressing obesity instead of taxes alone (Fletcher, Frisvold, and Tefft 2011). SSB tax proponents contend that the tax indeed has the ability to reduce
population weight and can also generate substantial revenue (Chaloupka, Powell, and Chriqui 2011). These two factors provided the basis for leading an especially strong push for a penny-per-ounce tax on SSBs (see Brownell et al. 2009).

SSB taxes have passed with increased frequency in the United States; Berkeley, California passed the first US SSB tax in November 2014 at one-cent-per-ounce, and three years later, eight SSB taxes passed. Nevertheless, unsuccessful attempts far outnumber successful passages of SSB taxes. Between January 1, 2010 and April 4, 2017, at least 20 states proposed a statewide SSB tax (Rudd Center for Food Policy & Obesity 2017).\textsuperscript{10} None of these taxes passed. These failed attempts ranged from the widely publicized SSB tax in New York to lesser-known efforts in New Mexico.

New Mexico introduced a statewide SSB tax twice, once in 2010 and again in 2011 (State Senate of New Mexico 2010, 2011). Both bills proposed a half-penny-per-ounce SSB tax. The net revenue generated from the tax would be split between a county-supported Medicaid fund (receiving 95%) and childhood obesity prevention programs (receiving 5%). The bill failed to pass the first time in the State Legislature and died in committee in 2011 (Rudd Center for Food Policy & Obesity 2017). Although the State of New Mexico has not been successful at passing a SSB tax so far, other portions of New Mexico have been.

The Healthy Diné Nation Act (HDNA) of 2014 is, in part, a SSB tax in the Navajo Nation enacted in response to growing rates of obesity and diabetes. This law imposes a 2% tax on the gross receipts of the consumer for all foods and beverages of minimal-to-no-nutritional value defined as “…sweetened beverages and prepackaged and

\textsuperscript{10} This number does not include changes made to sales taxes or vending machine taxes.
non-prepackaged snacks stripped of essential nutrients and high in salt, saturated fat, and sugar including sweetened beverages, sweets, chips, and crisps” (*The Healthy Diné Nation Act of 2014*). The funding generated from this law goes towards supporting community wellness projects that improve the food environment (such as creating vegetable gardens and sponsoring farmers markets), recreational facilities (supporting playgrounds and swimming pools), and other community projects that improve the social and physical environment.

Seven years have passed since the last statewide SSB tax attempt in New Mexico failed and there is now renewed interest within New Mexico to pass a SSB tax. This rise in interest is by no means isolated to New Mexico but rather reflects a national trend; since the 2016 presidential election, at least 11 states have proposed SSB taxation legislation ([Rudd Center for Food Policy & Obesity 2017](https://rutter.fph.org/sites/default/files/2017-08/RuddCenter_SSB_Taxtrends.pdf)).

Although the existing literature deftly describes the potential impact of SSB taxes on consumption and revenue generation (see Andreyeva, Chaloupka, and Brownell, 2011; and Wang et al. 2012), there has been little work done to answer a more basic question: what factors influence support and opposition for SSB taxes? A previous study by Rivard et al. (2012) examined the impact respondent’s demographics had on supporting a SSB tax. However, they did not include other important factors such as policy awareness, geography, or political ideology. Additionally, they did not specify the size and type of tax in their paper.

---

We go beyond this previous study, examining the impact media coverage; knowledge, attitude, and behavior; and political ideology have on the level of support for expanding the SSB tax portion of the HDNA of 2014 to all of New Mexico with primary data. The results from a partial proportional odds model show that media coverage about the tax in question is a very important determinant of both increasing support and decreasing opposition for expansion. Respondents who are conservative or drink more SSBs are strongly against expanding the law. Whereas, respondents who know there is a link between drinking SSBs and being obese/overweight or believing that obesity is a major problem in New Mexico are more likely to support expansion. The findings of this study contribute to a larger body of literature around the determinants of support for SSB taxes that can be utilized to understand why some SSB taxes pass, others do not, and provides insight into ways to promote support or mitigate opposition for these taxes.

3.2. Methods

3.2.1. Model

Let $Y_i$ be the level of support expressed for expanding the Healthy Diné Nation Act (HDNA) to all of New Mexico for the $i^{th}$ individual. Further, let $Y$ have $j$ categories, where $j = 1, 2, 3, 4, \text{ and } 5$, ordered to express increasing levels of support for expansion. Generally, an ordered logit or probit are estimated to model this relationship. The critical assumption for these ordered models is that the parallel lines assumption -- the impact of each independent variable will be the same across all categories of the dependent variable -- will hold. The Brant test checks the validity of this assumption. If this assumption is
violated, then the model is overly restrictive and can mask underlying variations in the data.

There are several potential solutions in this situation. First, we can ignore the violation and use an ordered logit/probit anyway. However, doing so results in a loss of rich information contained within the data and can lead to a less efficient model. The next possible solution is a multinomial logit model. While this model does overcome the parallel lines violation, it does so at the cost of forfeiting the ordinal nature of the dependent variable. A third solution is a generalized ordered logit model, shown in Equation 1, where all of the independent variables vary across the categories of the dependent variable.

\[
P(y_{i} \leq j | \mathbf{X}_i) = P(y_{i}^{*} \leq \alpha_j | \mathbf{X}_i) = \Lambda(\alpha_j - \mathbf{X}_i \mathbf{\beta}_j), \quad j = 1,2,3,4
\]

Where \( \Lambda(\cdot) \) is the logit function, \( \mathbf{X}_i \) is a k by 1 vector of independent variables to explain support, \( \mathbf{\beta}_j \) is a k by 1 vector of parameters to be estimated, \( j \) denotes the dependent variable category, and \( \alpha \) represents cut points associated with the \( j^{th} \) cumulative logit regression. Nonetheless, this method has a considerably large number of parameters to estimate, and it is likely that not all variables violate the parallel lines assumption and, hence, need to vary.

A compromise between the extremes of ignoring the violation and allowing all variables to vary is a partial proportional odds model (PPOM). The PPOM allows just the variables that violate the parallel lines assumption to be non-constrained (as such, they vary across the \( j \) categories) while constraining the remaining variables. Following the
work of Peterson and Harrell (1990), we represent the results using a gamma parameterization, shown in Equation 2:

\[
P(y_i \leq j | x_i) = P(y_i^* \leq \alpha_j | x_i, z_i) = \Lambda(\alpha_j - [x_i'\beta + z_i'y_j]), \quad j = 1,2,3,4
\]  

Here \(x_i\) is a \(k\) by 1 vector containing the full set of explanatory variables; \(\beta\) is the corresponding \(k\) by 1 vector of regression coefficients for the variables in \(x_i\); \(z_i\) is a \(q\) by 1 vector, where \(q \leq k\), containing the subset of non-constrained explanatory variables; and \(y_j\) is a \(q\) by 1 vector of regression coefficients for \(z_i\) that represent deviations from proportionality where \(\gamma_1 = 0\) for a total of \(j\)-2 \(\gamma\) coefficients. In other words, one \(\beta\) coefficient represents the constrained variables and both \(\beta\) and \(\gamma\) coefficients represent the non-constrained variables. To retrieve the original \(\beta\) coefficients for equation \(j\) from the gamma parameterization, the respective \(\gamma_j\) coefficient can be added to its \(\beta\) counterpart.

We used the Stata user-written program “gologit2” (Williams 2006) to estimate the PPOM. The very nature of any logit-type regression makes the ensuing coefficients not directly interpretable. Since the PPOM is a type of ordered logit regression, the sign of the \(\beta\) coefficients might not be correct for middle categories (Wooldridge 2010). Therefore, the coefficients need transforming, and we compute the average marginal effects to do so. To test the efficiency of the PPOM, we estimated all models using an ordered logit, generalized ordered logit, and PPOM. We calculated the corrected AIC for each regression and model combination to assess which regression method had the best fit, i.e., the lowest corrected AIC.
3.2.2. Hypotheses

We identify five key characteristics of the respondent related to SSB tax support that are a mix of easy and difficult attributes for policy makers to observe and/or influence. It is critical to understand where support and opposition for a policy lies in order to determine how and to what extent it can be influenced. The first three hypotheses relate to factors which can be difficult to observe but might be modifiable. First, we postulate that hearing about previous SSB taxes will increase the support for expansion. The majority of news coverage on television and in newspapers present more pro-SSB taxes arguments than anti-SSB tax viewpoints (Niederdeppe et al. 2013). Next, respondents who know that drinking too many SSBs can lead to being overweight or obese will increase the level of support. Thirdly, respondents who believe that obesity is a major problem in New Mexico will be more likely to support expansion. A previous study found that respondents who perceived obesity to be a major national problem were more likely to support snack taxes than those who did not (Oliver and Lee 2005). With the use of primary data, we can observe these factors and policy makers can influence these factors to a certain extent with actions like public health campaigns, for example.

The next hypothesis, while still hard to observe, is also difficult to influence, and influencing this behavior is the reason for this policy. We speculate that the more SSBs a person drinks, the less likely they will be willing to support the expansion, because the tax would impact them more directly than someone who rarely or never consumes SSBs. The final hypothesis is easier to observe than to influence; respondents whose political ideology is conservative, rather than liberal, will be less likely to support expansion.
3.3. Data

3.3.1. Survey

The data for this project comes from The New Landscapes of a Majority-Minority State (NLMMS) survey, a statewide representative survey of adults living in New Mexico sponsored by the Robert Wood Johnson Foundation. The NLMMS survey is a mixed-modes survey where half of the respondents were interviewed over the phone selected using random digit dial (603 on landlines and 150 on cellphones) and the remaining completed the survey over the web (752), for a total of 1,505 respondents. Respondents completed the survey in either English or Spanish. The telephone sample had an AAPOR response rate of 17.7% and the survey has an overall margin of error of +/- 2.5%. The average survey completion time was between 20 to 30 minutes. The NLMMS survey broadly investigates how the places where New Mexicans live, work, and play impact their health, wellbeing, policy viewpoints, and lived experiences. One of policy viewpoints included was SSB taxation.

3.3.2. Measures

The survey inquired into the respondent’s knowledge around the passing of the Healthy Diné Nation Act (HDNA) of 2014, asking first: In November 2014, to combat the growing problem of diabetes and obesity, the Navajo Nation passed a new law where, in part, sugar-sweetened beverages would have an additional tax of 2 percent. Have you heard of this Navajo Nation law? The survey asked all respondents this follow-up
question: *How much do you support enacting a similar law for all of New Mexico?*

Respondents answered on a five-point Likert scale, where one was strongly against and five was strongly support. We recoded thirty-nine respondents who responded with “do not know” or refused as missing. The latter question is the dependent variable for this analysis.

The explanatory variables fall broadly into four categories in line with our hypotheses. The first category reflects media coverage on SSB taxation. We dichotomized respondents who had heard of the HDNA before with one indicating previous knowledge. We also used a similar coding method if the respondent had heard of any other SSB taxes that passed in other localities.

Next, we include variables regarding the respondent’s knowledge, attitude, and behavior around SSBs. Two variables centered on knowledge. On a five-point Likert scale the respondent answered how much they agreed with the following two questions “*A poor diet can lead to being overweight or obese*” and “*Drinking too much soda or other sugar-sweetened beverages can cause a person to be obese or overweight.*”

Respondents who strongly agreed with the respective statement were coded as one, the rest as zero. The opinion measure on obesity in New Mexico implements the same coding scheme, captured by the statement: “*Obesity is a major problem in New Mexico.*” Finally, the self-reported number of SSBs consumed during an average week represents behavior.

The third independent variable category concerns political ideology. We divided respondents into three political categories based on where they identified on a political spectrum on a scale of one to seven, where one was extremely liberal and seven is extremely conservative. We coded those who responded with 1 or 2 as a liberal (base
category); we considered those who responded with 6 and 7 to be conservatives, and the remainder are moderates.

The last category includes a set of demographic and survey controls. We included binary indicators for sex (female=1), marital status (married=1), mode of survey (phone=1, web=0), and race/ethnicity (non-Hispanic white=1). The respondent’s age in years was used as a continuous variable. We divided education into three categories where the lowest category, the base group, was for respondents who had a high school education or less; the middle group had some college education but no degree; and the highest education group consisted of those who had any sort of college degree. We classified the geographic location of the respondent into one of New Mexico’s four metropolitan statistical areas (MSAs) (Albuquerque, Farmington, Las Cruces, and Santa Fe) or outside a MSA. Santa Fe is the base category. Finally, household income was divided into three groups: the lowest group, the base group, had a 2015 pre-tax household income less than $40,000; the middle category of $40,000-$80,000; and the high-income group of a combined household income over $80,000.12

3.4. Empirical Results

12 There were a total of 189 (unweighted) or 152 (weighted) respondents who did not know their household income or refused to report it. They were omitted from the analysis. However, these respondents did answer if their household income was over or less than $50,000. When we rerun the regressions dichotomizing income to above and below this threshold, in separate regressions, we find our results to be robust across all three ways of coding income.
3.4.1. Descriptive Statistics

The descriptive statistics, found in Table 3.1., decompose respondent characteristics based upon the level of support expressed for expanding the Healthy Diné Nation Act (HDNA), and the rightmost column shows the overall average of the given variable. The first row shows the level of support expressed for expanding the HDNA. Level of support is nearly evenly split across the five categories with a low of 14.58% of respondents against expansion and a high of 23.14% of respondents neutral to the expansion. There is slightly more overall support for expansion, at 40.52%, than overall opposition, at 36.95%.

The highest levels of previous awareness of the HDNA were concentrated in the positive levels of support: 37.74%, and 24.03% of respondents who strongly support and support expansion, respectively, had heard of the HDNA before. Respondents who expressed extreme viewpoints on the expansion, both positive and negative, were the most aware of SSB taxes in other states and cities: 55.45% of those strongly against and 47.71% of those strongly support were aware of these other SSB taxes.

Those who expressed strong support for expansion had the highest levels of knowledge about the health impacts of both poor diet and too much SSB consumption on obesity (81.45% and 78.66%, respectively, compared to the total of 68.08% and 54.57%, respectively). Believing that obesity was a major problem in New Mexico was highest in the strong support category at 72.75%; this is twice as large as those who are strongly against expansion. The number of SSBs consumed during an average week for most
categories of support wavers around three, with a low of 1.84 within the strongly support category.

Across the levels of support for each category, 54.82% to 69.73% of respondents are considered moderates. The highest concentration of conservatives, at 33.54%, is among those who are strongly against expansion. Liberals have the largest concentration, at 21.98%, within the strongly support category.

As for the controls, the data shows that respondents across the support categories are relatively similar with the largest difference seen within the neutral category for expansion. This group is more likely to be younger, male, married, low-income, not high-education, and belong to the other race/ethnicity category. Those who strongly support the expansion have a higher concentration living within Albuquerque and Santa Fe.

3.4.2. Partial Proportional Odds Model Results

We estimated all models using an ordered logit, a generalized ordered logit, and a PPOM regression. Further, all regressions incorporated survey weights and contained robust standard errors. We calculated the corrected AIC for each model regression pairing.

Comparing each model across the three regression methods, we consistently find that the ordered logit has the highest corrected AIC and the PPOM has the lowest corrected AIC. This supports our initial supposition that the PPOM would be more efficient than the other regression methods.
Table 3.2 presents the results from the PPOM. As Equation 2 demonstrates, a $\beta$ coefficient represents all of the independent variables. A $\gamma$ coefficient also represents only the unconstrained variables. The unconstrained variables are heard of other SSB taxes, knowing that drinking too many SSBs can lead to being obese (except in the final model), and believing obesity is a major problem.

We consistently find that media coverage has a large impact on supporting expansion of the HDNA. Hearing about the HDNA is a constrained variable, and we interpret it like an ordered logit coefficient. Knowing about the HDNA is associated with increased likelihood of support across all models. However, hearing about other SSB taxes had an opposite impact. We find that those who have heard of other SSB taxes before are less supportive of expanding the HDNA to all of New Mexico than those who have not heard of these taxes before. Further, those who have heard of other SSB taxes before were especially, and statistically significantly, unlikely to hold positive attitudes towards expansion.

Knowledge about the impact that a poor diet can lead to being obese or overweight did not have a statistically significant impact on support for expansion. However, respondents who knew more specifically that drinking too many SSBs could lead to being overweight or obese were more likely to support expansion. In Models 2 and 3, where this variable is unconstrained, the results show that not only does knowing about this link increase support, this group of respondents is statistically significantly more likely to strongly support expansion.

---

13 The ordered logit and generalized ordered logit results are found in Tables 3.2.A. and 3.2.B.
The impact of believing that obesity is a major problem in New Mexico was not statistically significant for lower levels of support, but overall, respondents who held this opinion were more likely to support expansion over other levels of support. As for the behavior measure, respondents who consume more SSBs were less likely to express support for expansion across all models.

Lastly, we tested the impact that political ideology had on the level of support for expansion. Being a conservative, as opposed to a liberal, strongly decreased the likelihood for support; it has the largest negative impact of any variable in any of the models. Being a moderate had a similar, albeit less strong, impact on decreased support.

In the final model, Model 4, we included socio-economic and survey controls. With the introduction of these controls, we find knowing that drinking too many SSBs can lead to obesity is again statistically significant and being a moderate is no longer statistically significant. Nonetheless, the remainder of our results were robust to the addition of controls.

3.4.3. Average Marginal Effects
The raw coefficient regression results in Table 3.2. are not directly interpretable. Further, with any ordered dependent variable, we cannot observe the unique impact of the independent variables on outcomes of the dependent variable. For instance, the coefficients associated with the outcome “disagree” for \( j=2 \) use both strongly disagree and disagree as a base group against neutral, agree, and strongly agree. This implies that, as is, the researcher cannot determine the impact of the independent variables on disagree alone, but rather for the combined strongly disagree and disagree categories. By taking
the average marginal effects of the PPOM, we can determine the full impact of the independent variable, taking the sign, cut points, and $\beta$ plus $\gamma$ values into account for each dependent variable category. Table 3.3 shows the average marginal effects of hypothesis variables from Model 4, which has the lowest AIC and corrected AIC. The top row shows the five levels of support for expanding the HDNA.

The average marginal effects show that hearing about the HDNA, as opposed to not hearing about it, decreases the likelihood of being strongly against expansion by 13.6 percentage points and increases the likelihood of strongly agreeing with expansion by 11.5 percentage points, ceteris paribus. This variable has the strongest impact on decreasing respondent’s negative attitudes. Having been exposed to other SSB taxes before increases the likelihood of being strongly against expansion by 11.0 percentage points. This goes against part of our first hypothesis that there would be a positive relationship between hearing about SSB taxes and supporting expansion.

We find evidence to support our second hypothesis; not only are respondents who know the link between SSB consumption and weight more likely to strongly support expansion at 9.43 percentage points but this knowledge is also associated with a 11.1 percentage points decrease in strongly disagreeing with expansion.

Additionally, respondents who believe that obesity is a major problem are statistically significantly more likely to strongly agree with expansion by 16.2 percentage points (the largest positive impact of all the independent variables on strongly supporting), and this belief is correlated with a 4.92 percentage points decrease in the likelihood of being strongly against the expansion. This result is in line with our third hypothesis.
Respondents who consume more SSBs are statistically significantly more likely to strongly disagree with expansion and less likely to support expansion, supporting our fourth hypothesis. If SSB consumption were to increase by one standard deviation (4.843 SSBs per week), respondents would be 3.52 percentage points more likely to be strongly against expansion and 2.98 percentage points less likely to be strongly for expansion. Finally, in accordance with the final hypothesis, being a conservative, as opposed to a liberal, greatly increased the chances of being strongly against expansion by 17.8 percentage points and decreased the likelihood of supporting expansion 15.1 percentage points.

3.4.4. Falsification Test

We ran a falsification test where we replaced the dependent variable with the following tax policy question: *Tax dollars should pay for substance use prevention or intervention services.* We assessed this question using the same five-point scale as the dependent variable in the main analysis. We find that the majority of the hypothesis variables are statistically insignificant with the exception of conservative, believing that obesity is a major problem in the state, and number of SSBs consumed in an average week (only for the full model with controls). The first result is expected, if obesity is seen as an individual problem due to personal choices, so too would be drug addiction. The next result at first blush seems questionable, but we see this variable acting here as a proxy for overall awareness of health issues in the state. The last result was statistically insignificant in all models, and only marginally so in the full model. Hence, this result is not robust in predicting support for substance use and intervention programs.
3.5. Discussion

The Healthy Diné Nation Act (HDNA) of 2014 passed in the Navajo Nation as a method to address the increasing rates of obesity and diabetes. Although the HDNA taxes all minimum-to-no-nutritional quality foods, we focus in on the SSB portion of the tax using a primary survey instrument. Even though this tax is not a specific excise tax, which is recommended in the literature (Brownell et al. 2009), its successful passage provides a useful starting point for determining elements of support and opposition for the HDNA and SSB taxes more broadly.

In this study, we estimated a PPOM to assess the impact that media coverage; knowledge, attitude, and behavior; and political ideology have on expanding the HDNA to all of New Mexico. We found media coverage statistically significantly influenced the level of support, knowing the link between SSB consumption and weight outcomes, believing obesity is a major problem in New Mexico, SSB consumption, and being a conservative. These factors did not always impact the categories of support with the same magnitude nor in the same direction. This implies that some factors are more modifiable than others for gaining support around SSB taxation. This article seeks to increase the salience of these elements for policy makers looking to establish a SSB tax.

Media coverage had a mixed impact on support for SSB taxes. Hearing about the HDNA increased support, whereas hearing about other SSBs not only decreased support but further reduced the likelihood that those who were aware of these taxes would hold favorable attitudes towards support. This later finding appears to be difficult to reconcile with previous research showing that most SSB tax arguments presented in newspapers
and on television are pro-tax arguments (Niederdeppe et al. 2013). This potential inconsistency between the negative coefficient sign and the mostly positive stories on SSB taxes could be due to a matter of framing in the news; framing related to the role SSBs play in weight gain or, more broadly, how obesity itself is framed as an individual level issue, systemic issue, or a mix of both. Jeong et al. (2014) examined the framing of obesity in local news stories within Philadelphia before and after a media campaign sponsored by the CDC aimed at reducing SSB consumption. The authors found that across the study time frame most news stories framed obesity as an individual level problem, showing that obesity was caused by overeating, lack of exercise, and SSB consumption.

This framing is influential in the perception of pro- and anti-tax arguments around SSB taxes. One of the main pro-SSB-tax arguments centers on how the tax can address the health consequences of SSB consumption (Niederdeppe et al. 2013). Yet, if this argument is framed to present SSB consumption as an individual level problem, those who oppose taxes on what could be considered an individual liberty, would not see this a pro-tax argument but rather as an anti-tax argument. This, coupled with one of the main anti-tax stances being that the government should not tax SSBs (Niederdeppe et al. 2013), can explain the negative impact of hearing about other SSB taxes on level of support.

The positive impact of hearing about the HDNA on increased level of support for its expansion was first illustrated in the descriptive statistics. Even among the 77.5% of respondents who had not heard of this law before, 35.8% of them expressed support for expansion. One reason for this finding could be that the survey description of the HDNA provided motivation on why this law was passed: to reduce the rates of obesity and
diabetes. The positive support for expansion could be due to an approval of using SSB taxes to address these health issues. The association between hearing about the HDNA and support for expansion was strongly supported by the regression results.

As was identified in the hypotheses, most of the measured aspects of knowledge, attitude, and behavior are subject to influence. The results show that respondents who knew that drinking too many SSBs can lead to being overweight or obese were not only statistically significantly more likely to support expansion (by 9.43 percentage points) but knowledge also decreased the likelihood of being strongly against expansion (by 11.10 percentage points). As such, investing in ways to boost knowledge about the impacts of consuming SSBs through actions, such as public health campaigns, can be a win-win for policy makers. In Howard County, Maryland, a three-year, community-based public health campaign aimed at reducing consumption of sugary drinks, which included education about the dangers of these drinks, resulted in a 20% reduction in sugary drink consumption, compared to a control county (Schwartz et al. 2017). Beyond the success of this public health campaign, the authors suggest that other policy measures to reduce sugary beverage consumption, including SSB taxes, should still be explored.

Another method to increase support for SSB taxes is improving the public’s awareness on how obesity is a problem in their locality. We found that those who believe that obesity is a major problem in New Mexico are more likely to strongly support expansion above and beyond the other categories of support. Knowing that obesity is not just an individual level problem increases support. This result echoes a previous study which found that when people are made aware of the role the food industry plays in the
obesity epidemic, they are more likely to support food and beverage policies (Ortiz, Zimmerman, and Adler 2016).

Obesity is an increasingly pressing problem that individuals, policy makers, and taxpayers face. Although the rates of obese and overweight white, black, and Hispanic New Mexicans are below their respective national averages, New Mexico has the third highest rate of overweight and obese American Indians/Alaska Natives at 79.5% - this is 10.9 percentage points higher than the national average for this group (Kaiser Family Foundation 2016). As obesity rates and the related co-morbidities increase on a state and national scale, so do the associated medical costs. In 1998, $78.5 billion in annual medical spending within the US was attributed to obesity and overweight, further Medicare and Medicaid covered half of this cost (Finkelstein, Fiebolkorn, and Wang 2011). Awareness of the role SSBs play in the obesity epidemic and the ensuing impact are potential policy levers that can be used to influence support for SSB taxes.

We found that conservatives were very strongly against expansion of this tax. Since Oliver and Lee’s 2005 article on public opinion and obesity, political elites have started to form positions on obesity related policies, especially concerning SSB taxes. Although there is not universal agreement on this policy within either party, prominent Democrats have come out in support of SSB taxes; Michael Bloomberg spearheaded the campaigns in New York City, Berkeley, CA and Philadelphia, PA. Additionally, Republican leaders have come out against these taxes for reasons such as the unnecessary expansion of the “nanny state.”

A benefit of subscribing to a political ideology is that it reduces the information search cost on issues (Jacoby 1991). However, this mass-level ideology does not
perfectly explain conservative and liberal views on this policy; within our data, 30% of conservatives overall agree with expansion, and 28% of liberals are overall against expansion. Further support is higher among both liberals and conservatives who don’t consume SSBs during the average week than those who drink at least one per week. Our findings around political ideology are likely to transform as respondents weigh the impact this policy will have on them personally - as has happened in the past (Oliver and Lee 2005) - against the potentially contrasting views political elites might converge to in the future.

The consumption of SSBs is both difficult for policy makers to observe and influence, yet influencing this behavior is the intent of the policy. We find that respondents who consume more SSBs were less likely to support expansion. The tax would impact them more directly, which explains why they have a negative view on expansion as they are acting in their own self-interest. Although a large body of literature would say that self-interest should pay a small role in political attitudes because political elites’ framing of the issue would dominate self-interest (see Sears and Funk 1990), Green and Gerken (1989) found that indeed self-interest played a large role in supporting anti-smoking legislation. Non-smokers strongly favored expanding smoking restrictions and raising taxes over smokers. Finding ways to help consumers who drink a large amount of SSBs lower their consumption, despite their self-interested resistance, can have important implications for population health.

Even modest reductions in SSB consumption could lead to positive health and weight impacts. Using nationally representative data along with actual price data, Smith, Lin, and Lee (2010) estimated the caloric and weight impact resulting from a 20% tax on
caloric sweetened beverages. Using own and cross-price elasticities, they found that such a tax would reduce the average overall caloric intake from beverages for adults by 37 calories a day. If this calorie reduction remained constant over a year, all else held equal, and assuming that 1-pound of body fat contains 3,500 calories, the average adult could lose 3.8 pounds over the year. However, there is still a debate in the literature if substitution behaviors will ultimately undermine the entire SSB tax effort (Fletcher, Frisvold, and Tefft 2011).

There were several limitations to the project. First, as previously noted, the HDNA is a tax on all minimum-to-no-nutritional quality foods, not solely a SSB tax. It could be the case that respondents previously aware of this tax agreed to support expansion because the tax included goods beyond SSBs. However, by using primary data, we emphasized the SSB aspect of the tax. Respondents might have issues of recall bias. The HDNA was passed in 2014, two years before this survey, and respondents might not remember this law being passed. We attempt to overcome this concern by including background information about the HDNA in the question immediately preceding the dependent variable. Additionally, we do not know where the respondents heard about either the HDNA or any other SSB tax. It could be that they read about it in the newspaper, saw it on television, heard from another person, a non-traditional media source (i.e. blog or social media), or any other source. The data used in this analysis is cross-sectional and as such we are not able to determine causality of the independent variables on level of support.

3.6. Summary
With the use of primary data and a partial proportional odds model, we examined how media coverage; knowledge, attitude, and behavior; and political ideology influence the level of support for expanding the Healthy Diné Nation Act of 2014 to all of New Mexico. We found both strong and mixed evidence on how media coverage impacts the level of support for expansion; knowing about the HDNA increases the likelihood of support, whereas hearing about other SSB taxes reduces overall support, and this group is especially likely to hold negative attitudes. When respondents know about the link between SSB consumption and obesity, they are more likely to support expansion. Believing that obesity is a major problem in New Mexico statistically significant increases the likelihood of supporting expansion, and this group is particularly more likely to strongly support expansion. Finally, respondents who are conservative or have large SSB consumption are more likely to be strongly against expansion. These findings are not only applicable to the case study of New Mexico but can be applied to any locality interested in establishing a SSB tax. Our results indicate that it is important for future SSB taxes to distinguish and distance themselves from already existing SSB taxes. Increasing the knowledge about the link between SSB consumption and obesity along with awareness about the obesity problem for that area are win-win strategies for policy makers because they both increase support and decrease opposition for SSB taxes.
<table>
<thead>
<tr>
<th>Table 3.1.</th>
<th>Descriptive statistics by level of support mean and standard deviationa</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Level of Support (%)</strong></td>
<td>Strongly Against</td>
</tr>
<tr>
<td>22.37</td>
<td>14.58</td>
</tr>
<tr>
<td><strong>Media Coverage (%)</strong></td>
<td>Heard of NN Taxc</td>
</tr>
<tr>
<td>55.45</td>
<td>44.48</td>
</tr>
<tr>
<td><strong>Knowledge, Attitude, and Behavior</strong></td>
<td>Poor Diet (%)</td>
</tr>
<tr>
<td>45.59</td>
<td>40.35</td>
</tr>
<tr>
<td>Obesity Major Problem (%)</td>
<td>35.05</td>
</tr>
<tr>
<td><strong>Number of SSBs Consumed in Avg. Week (count)</strong></td>
<td>3.84</td>
</tr>
<tr>
<td><strong>Political View (%)</strong></td>
<td>Liberalb</td>
</tr>
<tr>
<td>54.82</td>
<td>63.27</td>
</tr>
<tr>
<td>Conservative</td>
<td>33.54</td>
</tr>
<tr>
<td><strong>Controls</strong></td>
<td>Female (%)</td>
</tr>
<tr>
<td>50.98</td>
<td>51.38</td>
</tr>
<tr>
<td>Age (years)</td>
<td>(15.42)</td>
</tr>
<tr>
<td>Married (%)</td>
<td>24.77</td>
</tr>
<tr>
<td>53.00</td>
<td>35.78</td>
</tr>
<tr>
<td>Phone (%)</td>
<td>51.06</td>
</tr>
<tr>
<td>Non-Hispanic White</td>
<td>51.06</td>
</tr>
<tr>
<td><strong>Education (%)</strong></td>
<td>Lowb</td>
</tr>
<tr>
<td>26.88</td>
<td>25.91</td>
</tr>
<tr>
<td>High</td>
<td>50.20</td>
</tr>
<tr>
<td><strong>Income (%)</strong></td>
<td>Lowb</td>
</tr>
<tr>
<td>34.12</td>
<td>34.19</td>
</tr>
<tr>
<td>High</td>
<td>24.59</td>
</tr>
<tr>
<td><strong>Metropolitan Location (%)</strong></td>
<td>Albuquerque</td>
</tr>
<tr>
<td>4.41</td>
<td>1.53</td>
</tr>
<tr>
<td>Farmington</td>
<td>9.32</td>
</tr>
<tr>
<td>Las Cruces</td>
<td>7.21</td>
</tr>
<tr>
<td>Santa Feb</td>
<td>35.02</td>
</tr>
</tbody>
</table>

Notes: Weighted responses reported. aStandard deviation only reported for continuous measures. bReference category. Source: NLMMS Survey. cNN stands for Navajo Nation Healthy Diné Nation Act of 2014.
Table 3.2.: Partial Proportional Odds Model for Level of Support for Expanding Navajo Nation Soda Tax

<table>
<thead>
<tr>
<th>Betas</th>
<th>Model 1</th>
<th>Model 2</th>
<th>Model 3</th>
<th>Model 4</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Media Coverage</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Heard of NN Tax(^a)</td>
<td>0.777***</td>
<td>0.754***</td>
<td>0.801***</td>
<td>0.888***</td>
</tr>
<tr>
<td></td>
<td>(0.145)</td>
<td>(0.152)</td>
<td>(0.151)</td>
<td>(0.168)</td>
</tr>
<tr>
<td>Heard of Other ST(^b)</td>
<td>-0.763***</td>
<td>-0.782***</td>
<td>-0.754***</td>
<td>-0.721***</td>
</tr>
<tr>
<td></td>
<td>(0.151)</td>
<td>(0.156)</td>
<td>(0.158)</td>
<td>(0.180)</td>
</tr>
<tr>
<td><strong>Knowledge, Attitude, Behavior</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Poor Diet</td>
<td>-</td>
<td>-0.069</td>
<td>-0.061</td>
<td>-0.074</td>
</tr>
<tr>
<td></td>
<td>(0.137)</td>
<td>(0.138)</td>
<td>(0.147)</td>
<td></td>
</tr>
<tr>
<td>Too Many SSBs</td>
<td>-</td>
<td>0.346*</td>
<td>0.327*</td>
<td>0.729***</td>
</tr>
<tr>
<td></td>
<td>(0.178)</td>
<td>(0.181)</td>
<td>(0.151)</td>
<td></td>
</tr>
<tr>
<td>Obesity Problem</td>
<td>-</td>
<td>0.245</td>
<td>0.296*</td>
<td>0.322*</td>
</tr>
<tr>
<td></td>
<td>(0.172)</td>
<td>(0.175)</td>
<td>(0.189)</td>
<td></td>
</tr>
<tr>
<td>Number of SSBs</td>
<td>-</td>
<td>-0.043***</td>
<td>-0.046***</td>
<td>-0.048**</td>
</tr>
<tr>
<td></td>
<td>(0.016)</td>
<td>(0.016)</td>
<td>(0.019)</td>
<td></td>
</tr>
<tr>
<td><strong>Political View</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Moderate</td>
<td>-</td>
<td>-</td>
<td>-0.301*</td>
<td>-0.329</td>
</tr>
<tr>
<td></td>
<td>(0.174)</td>
<td>(0.202)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Conservative</td>
<td>-</td>
<td>-</td>
<td>-1.041***</td>
<td>-1.163***</td>
</tr>
<tr>
<td></td>
<td>(0.206)</td>
<td>(0.233)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Controls</strong></td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>X</td>
</tr>
<tr>
<td><strong>Gamma 2</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Heard of Other ST(^b)</td>
<td>0.123</td>
<td>0.056</td>
<td>0.041</td>
<td>0.136</td>
</tr>
<tr>
<td></td>
<td>(0.105)</td>
<td>(0.109)</td>
<td>(0.112)</td>
<td>(0.122)</td>
</tr>
<tr>
<td>Too Many SSBs</td>
<td>-</td>
<td>0.272**</td>
<td>0.310**</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>(0.120)</td>
<td>(0.125)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Obesity Problem</td>
<td>-</td>
<td>-0.071</td>
<td>-0.062</td>
<td>0.025</td>
</tr>
<tr>
<td></td>
<td>(0.119)</td>
<td>(0.122)</td>
<td>(0.129)</td>
<td></td>
</tr>
<tr>
<td><strong>Gamma 3</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Heard of Other ST(^b)</td>
<td>0.600***</td>
<td>0.571***</td>
<td>0.573***</td>
<td>0.510***</td>
</tr>
<tr>
<td></td>
<td>(0.149)</td>
<td>(0.148)</td>
<td>(0.153)</td>
<td>(0.163)</td>
</tr>
<tr>
<td>Too Many SSBs</td>
<td>-</td>
<td>0.486***</td>
<td>0.544***</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>(0.173)</td>
<td>(0.180)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Obesity Problem</td>
<td>-</td>
<td>0.486***</td>
<td>0.513***</td>
<td>0.610***</td>
</tr>
<tr>
<td></td>
<td>(0.166)</td>
<td>(0.172)</td>
<td>(0.166)</td>
<td></td>
</tr>
<tr>
<td><strong>Gamma 4</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Heard of Other ST(^b)</td>
<td>0.819***</td>
<td>0.829***</td>
<td>0.838***</td>
<td>0.800***</td>
</tr>
<tr>
<td></td>
<td>(0.189)</td>
<td>(0.200)</td>
<td>(0.207)</td>
<td>(0.216)</td>
</tr>
<tr>
<td>Too Many SSBs</td>
<td>-</td>
<td>0.448**</td>
<td>0.495**</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Obesity Problem</td>
<td>Alpha</td>
<td>Summary Stats</td>
<td></td>
</tr>
<tr>
<td>------------------------</td>
<td>-----------------</td>
<td>----------------</td>
<td>---------------</td>
<td></td>
</tr>
<tr>
<td></td>
<td>-</td>
<td>Cut 1</td>
<td>N</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.228)</td>
<td>1.462***</td>
<td>1456</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.234)</td>
<td>1.379***</td>
<td>1364</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.218)</td>
<td>1.795***</td>
<td>1343</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.224)</td>
<td>2.428***</td>
<td>1170</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.222)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.895***</td>
<td>Cut 2</td>
<td>Log Pseudolikelihood</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.218)</td>
<td>0.668***</td>
<td>-2027.338</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.934***</td>
<td>0.490***</td>
<td>-1799.452</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.929***</td>
<td>0.861***</td>
<td>-1744.983</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.219)</td>
<td>1.773***</td>
<td>-1485.682</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.443)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.224)</td>
<td>Cut 3</td>
<td>AIC</td>
<td></td>
</tr>
<tr>
<td></td>
<td>-0.508***</td>
<td>-0.508***</td>
<td>4072.677</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.086)</td>
<td>-1.109***</td>
<td>3636.903</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.154)</td>
<td>-0.810***</td>
<td>3531.965</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.230)</td>
<td>-0.421</td>
<td>3055.364</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.445)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-1.699***</td>
<td>Cut 4</td>
<td>AIC_C</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.108)</td>
<td>-1.699***</td>
<td>4072.801</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.240)</td>
<td>-2.660***</td>
<td>3637.469</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.290)</td>
<td>-2.406***</td>
<td>3532.665</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.489)</td>
<td></td>
<td>3058.569</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Pseudo R²</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.019)</td>
<td>0.019</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.066)</td>
<td>0.066</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.080)</td>
<td>0.080</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0.113</td>
<td></td>
</tr>
</tbody>
</table>

**Notes:** The dependent variable is level of support for expanding the Navajo Nation soda tax on a scale of 1 to 5, where 1 is strongly disagree and 5 is strongly agree with expansion. All regressions use individual weights and robust standard errors are reported in parentheses *** \( p \leq 0.01 \), ** \( p \leq 0.05 \), * \( p \leq 0.1 \). Controls include female, age, married, phone interview, education, income, race/ethnicity, and metropolitan location. 

*NN stands for Navajo Nation Healthy Diné Nation Act of 2014. \( ^{b} \)ST stands for SSB tax. Source: NLMMS Survey.
Table 3.2.A: Ordered Logit for Level of Support for Expanding Navajo Nation Sugar-Sweetened Beverage Tax

<table>
<thead>
<tr>
<th>Betas</th>
<th>Model 1</th>
<th>Model 2</th>
<th>Model 3</th>
<th>Model 4</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Media Coverage</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Heard of NN Tax(^a)</td>
<td>0.797***</td>
<td>0.776***</td>
<td>0.820***</td>
<td>0.904***</td>
</tr>
<tr>
<td></td>
<td>(0.150)</td>
<td>(0.154)</td>
<td>(0.153)</td>
<td>(0.169)</td>
</tr>
<tr>
<td>Heard of Other ST(^b)</td>
<td>-0.382***</td>
<td>-0.414***</td>
<td>-0.387**</td>
<td>-0.387**</td>
</tr>
<tr>
<td></td>
<td>(0.115)</td>
<td>(0.121)</td>
<td>(0.122)</td>
<td>(0.142)</td>
</tr>
<tr>
<td><strong>Knowledge, Attitude, Behavior</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Poor Diet</td>
<td>-</td>
<td>-0.0840</td>
<td>-0.0811</td>
<td>-0.107</td>
</tr>
<tr>
<td></td>
<td>(0.126)</td>
<td>(0.126)</td>
<td>(0.135)</td>
<td>(0.135)</td>
</tr>
<tr>
<td>Too Much Soda</td>
<td>-</td>
<td>0.646***</td>
<td>0.664***</td>
<td>0.713***</td>
</tr>
<tr>
<td></td>
<td>(0.130)</td>
<td>(0.130)</td>
<td>(0.143)</td>
<td>(0.143)</td>
</tr>
<tr>
<td>Obesity Problem</td>
<td>-</td>
<td>0.598***</td>
<td>0.672***</td>
<td>0.722***</td>
</tr>
<tr>
<td></td>
<td>(0.126)</td>
<td>(0.127)</td>
<td>(0.146)</td>
<td>(0.146)</td>
</tr>
<tr>
<td>Number of SSB</td>
<td>-</td>
<td>-0.0398*</td>
<td>-0.0433**</td>
<td>-0.0433*</td>
</tr>
<tr>
<td></td>
<td>(0.0155)</td>
<td>(0.0159)</td>
<td>(0.0184)</td>
<td>(0.0184)</td>
</tr>
<tr>
<td><strong>Political View</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Moderate</td>
<td>-</td>
<td>-</td>
<td>-0.320</td>
<td>-0.358</td>
</tr>
<tr>
<td></td>
<td>(0.164)</td>
<td>(0.164)</td>
<td>(0.187)</td>
<td>(0.187)</td>
</tr>
<tr>
<td>Conservative</td>
<td>-</td>
<td>-</td>
<td>-1.093***</td>
<td>-1.193***</td>
</tr>
<tr>
<td></td>
<td>(0.204)</td>
<td>(0.204)</td>
<td>(0.228)</td>
<td>(0.228)</td>
</tr>
<tr>
<td><strong>Controls</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>X</td>
</tr>
</tbody>
</table>

| Alpha                                           |                  |                  |                  |                  |
| Cut 1                                           | -1.268***        | -0.956***        | -1.376***        | -1.830***        |
|                                               | (0.0831)         | (0.130)          | (0.204)          | (0.405)          |
| Cut 2                                           | -0.539***        | -0.202           | -0.593**         | -1.034**         |
|                                               | (0.0735)         | (0.124)          | (0.196)          | (0.398)          |
| Cut 3                                           | 0.427***         | 0.807***         | 0.463*           | 0.0589           |
|                                               | (0.0742)         | (0.124)          | (0.193)          | (0.392)          |
| Cut 4                                           | 1.520***         | 1.997***         | 1.689***         | 1.269**          |
|                                               | (0.0882)         | (0.143)          | (0.197)          | (0.389)          |

<p>| Summary Stats                                   |                  |                  |                  |                  |
| N                                               | 1456             | 1364             | 1343             | 1170             |
| Log Pseudolikelihood Upon Convergence           | -2042.879        | -1847.291        | -1793.505        | -1561.485        |
| AIC                                             | 4097.8           | 3714.6           | 3611.0           | 3171.0           |</p>
<table>
<thead>
<tr>
<th>AICc</th>
<th>4097.817</th>
<th>3714.744</th>
<th>3611.244</th>
<th>3172.019</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pseudo R^2</td>
<td>0.011</td>
<td>0.042</td>
<td>0.054</td>
<td>0.067</td>
</tr>
</tbody>
</table>

Notes: The dependent variable is level of support for expanding the Navajo Nation soda tax on a scale of 1 to 5, where 1 is strongly disagree and 5 is strongly agree with expansion. All regressions use individual weights and standard errors are reported in parentheses *** p<0.01, ** p<0.05, * p<0.1. Controls include female, age, married, phone interview, education, income, race/ethnicity, and metropolitan location. A NN stands for Navajo Nation Healthy Diné Nation Act of 2014. B ST stands for SSB tax.

Source: NLMMS Survey.
<table>
<thead>
<tr>
<th>Media Coverage</th>
<th>SD vs. D, N, A, and SA</th>
<th>SD and D vs. N, A, and SA</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Model 1</td>
<td>Model 2</td>
</tr>
<tr>
<td>Heard of NN Tax(^a)</td>
<td>0.563**</td>
<td>0.475*</td>
</tr>
<tr>
<td></td>
<td>(0.194)</td>
<td>(0.201)</td>
</tr>
<tr>
<td>Heard of Other ST(^b)</td>
<td>-0.729***</td>
<td>-0.754***</td>
</tr>
<tr>
<td></td>
<td>(0.151)</td>
<td>(0.155)</td>
</tr>
<tr>
<td>Knowledge, Attitude, Behavior</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Poor Diet</td>
<td>-</td>
<td>-0.0671</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.186)</td>
</tr>
<tr>
<td>Too Much Soda</td>
<td>-</td>
<td>0.316</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.184)</td>
</tr>
<tr>
<td>Obesity Problem</td>
<td>-</td>
<td>0.265</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.172)</td>
</tr>
<tr>
<td>Number of SSB</td>
<td>-</td>
<td>-0.0435**</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.0157)</td>
</tr>
<tr>
<td>Political View</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Moderate</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Conservative</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>1.480***</td>
<td>1.417***</td>
</tr>
<tr>
<td></td>
<td>(0.110)</td>
<td>(0.172)</td>
</tr>
<tr>
<td>Controls</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Summary Stats</td>
<td>1456</td>
<td>1364</td>
</tr>
<tr>
<td>---------------</td>
<td>------</td>
<td>------</td>
</tr>
<tr>
<td>N</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AIC</td>
<td>4075.1</td>
<td>3645.6</td>
</tr>
<tr>
<td>AICc</td>
<td>4075.359</td>
<td>3635.059</td>
</tr>
<tr>
<td>Pseudo R²</td>
<td>0.019</td>
<td>0.069</td>
</tr>
</tbody>
</table>

Notes: The dependent variable is level of support for expanding the Navajo Nation soda tax on a scale of 1 to 5, where 1 is strongly disagree and 5 is strongly agree with expansion. All regressions use individual weights and standard errors are reported in parentheses *** p<0.01, ** p<0.05, * p<0.1. Controls include female, age, married, phone interview, education, income, race/ethnicity, and metropolitan location. aNN stands for Navajo Nation Healthy Diné Nation Act of 2014. bST stands for SSB tax.

Source: NLMMS Survey.
### Table 3.2.B. Continued

<table>
<thead>
<tr>
<th></th>
<th>SD, D, N vs. A, and SA</th>
<th>SD, D, N and A vs. SA</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Model 1</td>
<td>Model 2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Media Coverage</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Heard of NN Tax(^a)</td>
<td>0.784***</td>
<td>0.790***</td>
</tr>
<tr>
<td></td>
<td>(0.163)</td>
<td>(0.166)</td>
</tr>
<tr>
<td>Heard of Other ST(^b)</td>
<td>-0.172</td>
<td>-0.231</td>
</tr>
<tr>
<td></td>
<td>(0.133)</td>
<td>(0.140)</td>
</tr>
<tr>
<td><strong>Knowledge, Attitude, Behavior</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Poor Diet</td>
<td>-</td>
<td>-0.0444</td>
</tr>
<tr>
<td></td>
<td>(0.171)</td>
<td>(0.171)</td>
</tr>
<tr>
<td>Too Much Soda</td>
<td>-</td>
<td>0.828***</td>
</tr>
<tr>
<td></td>
<td>(0.163)</td>
<td>(0.166)</td>
</tr>
<tr>
<td>Obesity Problem</td>
<td>-</td>
<td>0.717***</td>
</tr>
<tr>
<td></td>
<td>(0.148)</td>
<td>(0.150)</td>
</tr>
<tr>
<td>Number of SSB</td>
<td>-</td>
<td>-0.0440</td>
</tr>
<tr>
<td></td>
<td>(0.0256)</td>
<td>(0.0277)</td>
</tr>
<tr>
<td><strong>Political View</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Moderate</td>
<td>-</td>
<td>-0.430*</td>
</tr>
<tr>
<td></td>
<td>(0.198)</td>
<td>(0.223)</td>
</tr>
<tr>
<td>Conservative</td>
<td>-</td>
<td>-1.142***</td>
</tr>
<tr>
<td></td>
<td>(0.226)</td>
<td>(0.257)</td>
</tr>
<tr>
<td><strong>Constant</strong></td>
<td>-0.514***</td>
<td>-1.127***</td>
</tr>
<tr>
<td></td>
<td>(0.0872)</td>
<td>(0.171)</td>
</tr>
<tr>
<td><strong>Controls</strong></td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>
## Summary Stats

<table>
<thead>
<tr>
<th></th>
<th>1456</th>
<th>1364</th>
<th>1343</th>
<th>1170</th>
<th>1456</th>
<th>1364</th>
<th>1343</th>
<th>1170</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>1456</td>
<td>1364</td>
<td>1343</td>
<td>1170</td>
<td>1456</td>
<td>1364</td>
<td>1343</td>
<td>1170</td>
</tr>
<tr>
<td>AIC</td>
<td>4075.1</td>
<td>3645.6</td>
<td>3547.2</td>
<td>3104.5</td>
<td>4075.1</td>
<td>3645.6</td>
<td>3547.2</td>
<td>3104.5</td>
</tr>
<tr>
<td>AICC</td>
<td>4075.359</td>
<td>3635.059</td>
<td>3539.379</td>
<td>3112.733</td>
<td>4075.359</td>
<td>3635.059</td>
<td>3539.379</td>
<td>3112.733</td>
</tr>
<tr>
<td>Pseudo R²</td>
<td>0.019</td>
<td>0.069</td>
<td>0.084</td>
<td>0.123</td>
<td>0.019</td>
<td>0.069</td>
<td>0.084</td>
<td>0.123</td>
</tr>
</tbody>
</table>

*Notes:* The dependent variable is level of support for expanding the Navajo Nation soda tax on a scale of 1 to 5, where 1 is strongly disagree and 5 is strongly agree with expansion. All regressions use individual weights and standard errors are reported in parentheses *** p<0.01, ** p<0.05, * p<0.1. Controls include female, age, married, phone interview, education, income, race/ethnicity, and metropolitan location. aNN stands for Navajo Nation Healthy Diné Nation Act of 2014. bST stands for SSB tax.

*Source:* NLMMS Survey.
Table 3.3: Hypothesis Table with Average Marginal Effects on the Level of Support For PPOM Model 4

<table>
<thead>
<tr>
<th>Hypothesis</th>
<th>Strongly Disagree</th>
<th>Disagree</th>
<th>Neutral</th>
<th>Agree</th>
<th>Strongly Agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hypothesis 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Heard of NN Tax&lt;sup&gt;a&lt;/sup&gt;</td>
<td>-0.136***</td>
<td>-0.041***</td>
<td>0.0004</td>
<td>0.061***</td>
<td>0.115***</td>
</tr>
<tr>
<td></td>
<td>(0.026)</td>
<td>(0.009)</td>
<td>(0.006)</td>
<td>(0.012)</td>
<td>(0.022)</td>
</tr>
<tr>
<td>Heard of Other ST&lt;sup&gt;b&lt;/sup&gt;</td>
<td>0.110***</td>
<td>0.006</td>
<td>-0.075***</td>
<td>-0.052**</td>
<td>0.010</td>
</tr>
<tr>
<td></td>
<td>(0.027)</td>
<td>(0.021)</td>
<td>(0.024)</td>
<td>(0.026)</td>
<td>(0.025)</td>
</tr>
<tr>
<td>Hypothesis 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Too Many SSBs</td>
<td>-0.111***</td>
<td>-0.034***</td>
<td>0.0003</td>
<td>0.050***</td>
<td>0.094***</td>
</tr>
<tr>
<td></td>
<td>(0.023)</td>
<td>(0.008)</td>
<td>(0.005)</td>
<td>(0.011)</td>
<td>(0.020)</td>
</tr>
<tr>
<td>Hypothesis 3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Obesity Problem</td>
<td>-0.049*</td>
<td>-0.020</td>
<td>-0.116***</td>
<td>0.023</td>
<td>0.162***</td>
</tr>
<tr>
<td></td>
<td>(0.029)</td>
<td>(0.022)</td>
<td>(0.022)</td>
<td>(0.026)</td>
<td>(0.025)</td>
</tr>
<tr>
<td>Hypothesis 4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of SSBs</td>
<td>0.007**</td>
<td>0.002**</td>
<td>0.00002</td>
<td>-0.003**</td>
<td>-0.006**</td>
</tr>
<tr>
<td></td>
<td>(0.003)</td>
<td>(0.0009)</td>
<td>(0.0003)</td>
<td>(0.001)</td>
<td>(0.002)</td>
</tr>
<tr>
<td>Hypothesis 5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Conservative</td>
<td>0.178***</td>
<td>0.054***</td>
<td>-0.0005</td>
<td>-0.081***</td>
<td>-0.151***</td>
</tr>
<tr>
<td></td>
<td>(0.035)</td>
<td>(0.012)</td>
<td>(0.008)</td>
<td>(0.018)</td>
<td>(0.030)</td>
</tr>
</tbody>
</table>

Notes: For each hypothesis, the relevant variable's average marginal effect is listed. The regression includes the same controls as Model 4. <sup>a</sup>NN stands for Navajo Nation Healthy Diné Nation Act of 2014. <sup>b</sup>ST stands for SSB tax. All regressions use individual weights and robust standard errors are reported in parentheses. *** p ≤ 0.01, ** p ≤ 0.05, * p ≤ 0.1. Source: NLMMS Survey.
Chapter 4. To Light Up or Snuff Out? Estimating Adolescent Smoking Behavior in Nepal

4.1. Introduction

The smoking epidemic is shifting from the developed to the developing world: nearly 80% of smokers worldwide now live in low- to middle-income countries (WHO n.d.). Smoking rates are not spread evenly throughout developing nations nor are they the same for men and women. Although the smoking rates in the South East Asian Region are below the global average (WHO 2015), women in Nepal have the highest smoking rate in the region: 19% of women are current cigarette smokers and 29% of men are as well (World Bank 2014).

Tobacco use is a well-known cause of non-communicable diseases (NCDs) and this alone accounts for one sixth of all NCD deaths globally (Beaglehole et al. 2011). Developing nations like Nepal are facing a dual burden of communicable diseases and NCDs. Within Nepal, the percentage of deaths attributable to NCDs rose from 46% in 2000 to 65% in 2015 (World Bank n.d.). The loss of human life from NCDs most often occurs during a person’s working life between ages 30-69 in developing nations (Strong et al. 2005). Both the direct and indirect medical costs of NCDs are astonishingly high. One estimate predicted, that if nothing was done to reduce the rates of chronic diseases in 23 low- and middle-income countries (including India and Bangladesh) the estimated economic loss of production between 2006 and 2015 would amount to US$84 billion in GDP from heart disease, diabetes, and stroke alone (Abegunde et al. 2007). One way to
reduce NCDs is through effective tobacco control policies and reduced tobacco use (Glantz and Gonzalez 2012).

Additionally, it is vital that tobacco control policies contain measures that specifically target youth smoking behaviors. Smoking initiation normally begins during adolescence: most smokers begin smoking before the age of 18 (USDHHS 2012), and a fifth of adolescent smokers begin before they are 10 years old (Warren et al. 2000). The addictive nature of cigarettes keeps youths smoking longer which causes increased physical health problems (USDHHS 2012).

The Government of Nepal has implemented significant tobacco control policies, but their efficacy regarding youth smoking remains questionable. A health tax on cigarettes was imposed in the early 1990s and tobacco advertisements were banned on television and radio in 1997/1998 (Karki, Pant, and Pande 2003). Nepal signed and ratified the World Health Organization Framework Convention on Tobacco Control (FCTC) in December 2003 and November 2006, respectively (WHO 2013). The FCTC is a WHO treaty created in response to the spread of the tobacco epidemic and outlines measures to reduce both supply and demand of tobacco (WHO 2005). In 2011, Nepal passed legislation, called the Tobacco Control and Regulatory Bill, that follows the FCTC guidelines and contains statuses on anti-smoking issues including banning smoking in private homes, increasing cigarette package warnings, and prohibiting the sale of tobacco to people under the age of 18 (WHO 2011). Nevertheless, youth smoking in Nepal has increased since the bill was passed. In 2012, 3.1% of adolescents between 13-15 years old were current cigarette smokers, whereas, in 2016, the percentage had risen to 5% (WHO 2013 and 2017). This is still less than youth smoking rates in
neighboring Southeast Asian nations (GYTS Collaborating Group 2003), nevertheless the rapid growth in cigarette consumption in Nepal is worrisome.

Understanding the unique way factors such as proximity to other smokers; formal and informal anti-tobacco education; and social perceptions of smoking and tobacco advertisements influence youth smoking behaviors is critical to designing policies that effectively reduce cigarette use. In this analysis we exclusively focus on cigarette smoking behaviors, the implications of which will be returned to in the discussion. Although there have been studies in Nepal that examine youth smoking behaviors, they often focus on only one smoking status and are from a particular region.14 We contribute to the existing literature on youth smoking in several ways. First, we explore the roles proximity to other smokers, pro-tobacco marketing and social perception, and anti-tobacco awareness have on three smoking statuses: never smoker, former smoker, and level of cigarettes smoked by current smokers. Second, our econometric analysis explicitly examines how these factors impact smoking behaviors differently by gender. Next, we utilize two waves of nationally representative data. Finally, we estimate the smoking status model using a trivariate ordered probit model with independence based up Kasteridis, Munkin and Yen (2010).

The results indicate that being around other smokers greatly increases the odds of engaging in smoking behaviors for both girls and boys, although parental smoking uniquely increases boys’ smoking behaviors. We find gendered variations in the ways that social perceptions of smoking influences smoking statuses. Although formal anti-

14 For example see Pradhan and Marahatta 2016, Aryal et al. 2013 and Aryal, Petzold, and Krettek 2014.
tobacco education generally did not influence smoking behaviors, high quality education reduced the likelihood of low and medium cigarette consumption for boys. Greater exposure to pro-tobacco media increased girl’s level of cigarette consumption. Counterintuitively, we find that youths exposed to anti-tobacco media were more likely to begin smoking and to engage in a higher level of cigarette consumption. We did not find a relationship between weight changes and the likelihood girls would engage in smoking behaviors, but there was a negative association of weight and smoking habits for boys. Finally, knowing that cigarettes are harmful to your health was statistically significantly associated with increased likelihood of higher cigarette consumption for boys.

4.2. Background Literature

The body of literature on causes and correlates of smoking behaviors is extensive in its breath and depth. A sizable portion of this literature focuses on adults in developed nations while its direct applicability to youth in developing nations is questionable. As is noted by Jonathan Gruber (2001), there could be differences in youth and adult risky behaviors depending on what theoretical lens is applied. Traditional economic theory does not find reason to differentiate behaviors between these two groups. However, examining risky behaviors from a developmental psychological perspective can support these differences: youth have been shown to have lower decision-making competence, to be more impacted by social reactions, and less able to understand future risks and consequences of their actions.
Nevertheless, there are certain facets of economic theory that influence youth smoking which should not vary between developed and developing nations such as the impact of taxes and age-restrictions on purchases. The level to which either policy is actually implemented or enforced can vary dramatically across countries. The social acceptability of smoking within a nation is mediated by the nation’s cultural, religious, and patriarchal landscape. These social norms could diminish the impact of economic factors for certain subpopulations.

There are many factors which influence a youth’s decision to smoke including, but not limited to, smoking behavior of friends (Avenevoli and Merikangas 2003), parental smoking (Gilman et al. 2009), pro-tobacco advertisements (Aryal, Petzold, and Krettek 2014), anti-tobacco advertisements (Wakefield et al. 2003), formal anti-tobacco curriculum in schools (Perry et al. 1980), health perceptions of smoking (Song et al. 2009), access to pocket money (Mohan, Sarma, and Thankappan 2005), alcohol use (Pradhan and Karlra 2015), smoking by school personnel (Nikaj and Chaloupka 2015), and cigarette prices and taxes (Nonnemaker and Farrelly 2011). In this paper we focus on the impact proximity to other smokers; pro-tobacco marketing and social perception; and anti-tobacco marketing have on the smoking behaviors of youth in Nepal.

Youth’s exposure to other smokers immensely influences their own smoking status, and their specific relationship to that smoker is of great importance. The impact of parental smoking on child’s smoking is mixed. Gender concordance was seen in parent and offspring cigarette use but not in cross-gender pairs (Loureiro, Sanz-de-Galdeano, and Vuri 2010). However, overall there is weak evidence that having a parent who smokes will increase the likelihood of adolescent smoking (Avenevoli and Merikangas
2003). Consistently the most important personal relationship that impacts a youth’s decision to start smoking is if their peers are smokers (Avenevoli and Merikangas 2003). It is not just the impact of having a close friend who smokes that can increase the likelihood of being a smoker, but merely being around a large number of adolescent smokers can also increase the odds of being a smoker. Powell, Tauras, and Ross (2005) estimated that if a student were to move from a school where no one smoked to a school where a quarter of students smoke that transfer student would be 14.5 percentage points more likely to be a smoker.

Since most smokers begin smoking when they are young, the tobacco industry recruits new smokers in this age group and often does so with tobacco advertisements. The deleterious impact of pro-tobacco marketing has been recognized globally and enshrined within Article 13 of the FCTC, which calls for a comprehensive ban on advertising (WHO 2005). A comprehensive ban is needed in developed nations to statistically significantly reduce smoking rates; however, this is not necessarily the case for developing nations. Blecher’s 2008 study found that for developing nations a limited advertising ban would reduce consumption by 13.6% but a comprehensive ban would reduce consumption by 23.5%.

Adolescent’s social perception of smoking can impact their smoking status. Teens who believed there were strong social benefits associated with smoking were more likely to start smoking (Song et al. 2009). Within Nepal, youths who perceived there to be smoking related benefits – including looking cool, feeling relaxed, becoming popular, and feeling grown-up – were 1.42 times more likely to be susceptible to start smoking (Aryal et al. 2013).
Anti-tobacco awareness, cultivated through either formal or informal education, can impact smoking behaviors. There are several elements of formal anti-tobacco education. Not only is the frequency of formal anti-tobacco lessons important but so too is the quality of this education. The combination of these two factors can help determine if these lessons impact short- and long-term smoking behaviors. A review of randomized controlled school-based intervention trials in the US found that only one of eight studies had a long-term impact and that particular intervention had a high degree of interaction and participation with students for reducing current smoking (Wiehe et al. 2004).

Additionally, education around the dangers of smoking does not have to be limited to experiences in the classroom. Informal education though anti-tobacco advertisements and parental discussions also have the potential to mitigate smoking initiation and level of consumption. There have been mixed findings on the role that anti-tobacco advertising has on smoking initiation and it mainly due to the different ways in which these studies have been implement those advertisements that have been found to reduce smoking uptake rates elicit an emotional response (Wakefield et al. 2003).

Anti-tobacco awareness also includes understanding the health impacts of smoking. These impacts can be broken down into short and long-term consequences. Short-term consequences would include things like bad breath and wheezing, whereas long-term health consequences are more serious and include cancer, reduced fertility in women, and premature death (USDHHS et al. 2014). A study conducted by Aryal et al. in 2013 on smoking initiation of adolescents in Nepal found that long-term health consequences did not have a statistically significant impact on smoking initiation, but
those who perceived there to be short term health consequences were less likely to start smoking.

Moving beyond the determinates of youth smoking there is the matter of how to estimate smoking behaviors. The empirical smoking literature has employed a variety of econometric specifications to estimate smoking behaviors and cigarette demand. One prominent method is to use a double-hurdle approach spearheaded by Jones (1989) that models two interrelated decisions: to become a smoker and the level of cigarettes consumed. Another method mentioned in Jones (1989) and expanded upon by Kasteridis, Munkin, and Yen (2010) is to estimate a trivariate ordered probit model. This model would estimate three smoking decisions: to abstain from smoking, to quit smoking (conditional on having smoked cigarettes before), and the level of cigarettes consumed by current smokers.

Within the empirical youth smoking literature in Nepal, there have been multiple studies that have studied a variety of smoking status among youth: smoking susceptibility (Aryal, Petzold, and Krettek 2014), tried smoking (Pradhan and Kalra 2015), current smoking (Kabir and Goh 2007), and quit smoking attempts (Pradhan and Marahatta 2016). However, most of the aforementioned studies have two things in common: their main analysis focuses on only one smoking status and uses logistic or stepwise logistic regression.\(^{15}\)

Altogether, the previous literature in both developed and developing nations has shown that proximity to other smokers; anti-tobacco awareness; and pro-tobacco

\(^{15}\) Except the study by Aryal et al. 2013 which used principal component analysis to assess risks and benefits of smoking among youth.
marketing and social perception can impact youth’s smoking decisions. Despite the many ways smoking behaviors can be modeled, empirical studies in Nepal have tended to use logistic regressions. We build upon the existing literature by incorporating above listed determinates of youth smoking into a trivariate ordered probit model with independence. This modeling procedure allows us to simultaneously estimate three smoking statuses and the average marginal effects of the smoking statuses take into previous smoking decisions (for instance level of cigarettes consumed depends upon being a current smoker which itself depends up having started smoking in the first place) which a general logit model cannot do.

4.3. Trivariate Ordered Probit Model with Independence

We follow the work of Kasteridis, Munkin and Yen (KMY) (2010) in setting up the trivariate ordered probit model with independence. Let there be total of \( A \) individuals, where each individual \( i \) belongs to one of three mutually exclusive smoking statuses. The first smoking status category represents if the respondent has abstained from before. This is a binary decision where the decision to never take a puff of a cigarette is modeled by \( N_i^* = X_i \alpha + \varepsilon_{1i} \). The latent variable \( N_i^* \) measures the utility from never smoking a cigarette. Further, \( N_i^* \) takes the value of one when the respondent has abstained from smoking and takes the value of zero when they have smoked at least a puff of a cigarette before. This is modeled using a probit function such that \( \Pr(N_i = 1) = \Phi(X_i \alpha) \), where \( \Phi \) is the normal CDF, represents the probability that someone abstains from smoking.

The next smoking status is if the respondent has quit smoking. This is also a binary outcome and is modeled with \( Q_i^* = X_i \beta + \varepsilon_{2i} \). The latent variable \( Q_i^* \) takes the
value of one when the respondent quits smoking and takes the value of zero when the respondent continues to smoke. A probit model is used for this relationship where
\[ \Pr(Q_i = 1) = \Phi(X_i \beta) \]
determines the probability of quitting smoking.

The final smoking status measures the number of cigarettes smoked, \( y_i \): this is the dependent variable of interest and a large proportion of the responses are zeros. The KMY paper establishes that the quantity of cigarettes smoked can be measured using a variety of methods. They tested the efficiency of a type 1 and type 2 negative binomial, a Gaussian model that truncates the error term to ensure non-negativity, an ordered probit, and a sequential probit. Ultimately, they find that ordered probit performs the best, robust to varying category definitions. Additionally, this modeling approach conforms best with our data. Generically, the respondent will belong to smoking level category \( m \) when they cross the threshold \( \tau_{m-1} \) for \( \tau = (\tau_1, \ldots, \tau_{M-1}) \). The ordered probit function is written as
\[ \Pr(y_i = m) = \Phi(\tau_m - X_i \gamma) - \Phi(\tau_{m-1} - X_i \gamma). \]

In these three previous equations, \( X_i \) is a vector of variables related to the factors which influence smoking status; \( \alpha, \beta, \) and \( \gamma \) are conformable vectors of coefficients; and \( \varepsilon_{1i} \) and \( \varepsilon_{2i} \) are i.i.d. error terms.

To set up the likelihood function, we first establish when a non-zero value will come to fruition in \( y_i \). Respondents who have never smoked before \( (N_i = 1) \) or who have quit smoking \( (Q_i = 1 | N_i = 0) \) will take the value of 0 in \( y_i \). Only respondents who are currently smoking, \( (N_i = 0 \text{ and } Q_i = 0) \) will have non-zero values present in \( y_i \). All together this leads to the following likelihood function:
\[
\mathcal{L}(y|\theta) = \prod_{N_i=1} \Pr(N_i = 1)
\]

\[
\times \prod_{N_i=0,Q_i=1} [1 - \Pr(N_i = 1)] \Pr(Q_i = 1|N_i = 0)
\]

\[
\times \prod_{N_i=0,Q_i=0} [1 - \Pr(N_i = 1)] [1 - \Pr(Q_i = 1|N_i = 0)]
\]

\[
\times \prod_{m=1}^{M} [\Pr(y_i = m|N_i = 0, Q_i = 0)]^{d_{im}} \quad (1)
\]

The first product of the likelihood function uses all \( A \) observations to determine who is a never smoker and how the variables in \( X_i \) influence the probability to belong to this smoking category. The contents of the next product are restricted to observations that have quit smoking given that they have started smoking and it estimates the likelihood of being a former smoker. The third product sign establishes the conditions for respondents who are currently smoking: they started smoking and they continue to smoke given that they started. The last product sign estimates the ordered probit for the level of cigarette smoking where \( d_{im} \) takes the value of one iff \( y_i \) belongs to the \( m^{th} \) category, else zero.

For identification purposes, we restrict \( \tau_1 \) to equal zero in the ordered probit.

### 4.4. Data

#### 4.4.1. Global Youth and Tobacco Survey

The data for this project comes from the Global Youth and Tobacco Survey (GYTS). The GYTS has been administer in over 131 nations and was initially created by the WHO and
the Centers for Disease Control and Prevention to collect detailed international youth smoking prevalence data (Warren et al. 2006). Each GYTS uses a two-stage cluster sample to produce nationally representative data for students in the grades associated with 13-15 years old in each nation. The first stage selects the schools where the probability of selection is proportional to the number of students enrolled in those specific grades. The second stage randomly selects classes within the selected school where all students present that day are eligible to participate. The data is then weighted to adjust for non-response and other factors that influence the probability of selection (GYTS Collaborating Group 2003).

For this analysis, we use the 2007 and 2011 waves of the Nepal GYTS which are nationally representative. Although the survey was also administered in 2001, due to political insurgency it was only conducted in one region in Nepal. As such, we do not use the data from 2001 because it does not conform to the other nationally representative waves.

4.4.2. Smoking Status

Three separate variables were used to create the smoking status dependent variable: 1) “Have you ever tried or experimented with cigarette smoking, even one or two puffs?”; 2) “During the past 30 days (one month), on how many days did you smoke cigarettes?”; and 3) “During the past 30 days (one month), on the days you smoked, how many cigarettes did you usually smoke?”. Students responded with a yes or no to the first question. Question two was dichotomized to equal zero for no days in the past month versus any number of days in the past month takes the value one. For the last question,
respondents selected the best category for their level of cigarette smoking: *I did not smoke in the past 30 days, Less than 1 cigarette per day, 1 cigarette per day, 2-5 cigarettes per day, 6-10 cigarettes per day, 11-20 cigarettes per day, or 20+ cigarettes per day*. Due to the low number of observations, the last three categories were combined to be *6+ cigarettes per day*.

With the GYTS data, all questions were asked to all respondents. As such, we had to impose a conditional coding upon the respondents. Figure 4.1 places these three questions into a smoking decision tree. Students who have never had a puff of a cigarette before ($N_t = 1$) are considered never smokers. Conditional upon having a puff of a cigarette before ($N_t = 0$) these respondents were included in the next decision: did they smoked in the past thirty days. If they said no ($Q_t = 1$) they are considered to have quit smoking and are hence a former smoker. This categorization in practice proved to be more problematic and the implications of which are seen later in the results section. On the other hand, if they said yes ($Q_t = 0$) these observations were used to observe the level of cigarettes smoked on the days they smoked, $y_t$.

### 4.4.3. Independent Variables

The first category of independent variables contains two variables capturing proximity to other smokers. The first is a binary indicator equaling one if the respondent has close

---

16 There were five respondents who said that they did not smoke in the past month yet still reported smoking a positive number of cigarettes during that period. They were recoded from former smoker to current smoker.
friends who smoke cigarettes. The latter is another binary variable equaling one if either
or both of the respondent’s parents smoke.

The next main category consists of two groups of variables around the
respondent’s social perception of tobacco and exposure to pro-tobacco marketing. The
first group is centered on social perceptions of tobacco use and was constructed using
principal component analysis (PCA). We took seven variables and transformed them into
a set of three continuous variables measuring the extent that respondents believe that
smoking increases a person’s attractiveness and ease at parties; perceived popularity of
smokers; and perception of adult smokers. Respondent’s exposure to pro-tobacco
messages was also created using PCA where a larger number indicates a greater intensity
of exposure. Five variables are used in this PCA and include topics such as seeing
cigarettes advertised at sporting events or on billboards to seeing actors smoke in movies
or on television. An overview of the PCA is found in the Methods section, and a detailed
explanation can be found in Appendix A.

The last main category of independent variables focuses on a respondent’s
awareness of the dangers of smoking through formal and informal education and
student’s health perceptions of cigarettes. Formal health education is captured by the
quality of anti-smoking education in school. This is an index constructed with three
binary variables about what aspects of smoking were they taught during the current
school year and includes being taught about the dangers of smoking, why people smoke,
and what are the effects of smoking. For each item the respondent said yes to, they
received one point resulting in an index on a scale of 0 to 3, where the higher the number,
the better quality anti-tobacco education they received during that year. Informal
education includes two variables. Exposure to anti-tobacco messages equals one if the respondent was exposed to both anti-smoking media messages during the past thirty days and anti-smoking messages at sporting events and social events. The other informal education variable is a binary indicator equaling one if the respondent’s parents had discussed the dangers of smoking with them. The final group of variables in this category relate to health perceptions of cigarette smoking. It includes if they think that cigarettes make you lose weight, gain weight, or have no impact on weight (base category); and two binary indicators for if they believe that smoking is harmful to their health and if other people’s cigarette smoke is harmful to the respondent, separately.

The following controls are included in the final model sex (female =1), age (13 years or younger [base category], 14-15 years old, and 16 years or older), grade (seventh [base category], eighth, ninth, and tenth), and a year dummy for 2011.

4.4.4. Hypotheses

In this study we formally test three sets of hypotheses on what influences adolescent’s smoking status in Nepal. The first category involves the impact that smoking by people close to the adolescent have on the adolescent’s smoking status. The proximity hypotheses are:

\textit{H1.1: Being around another smoker will increase the odds of engaging in smoking behaviors and with a greater intensity.}

\textit{H1.2: Having a parent who smokes will have a greater impact on boy’s smoking behaviors.}
The development of smoking perceptions is cultivated through many avenues: from discussion with parents, friends, and teachers to exposure to tobacco advertising. The second set of hypotheses centers around how adolescents perceive smoking in terms of social status and their exposure to pro-tobacco media. Formally this is tested with the following two hypotheses:

*H2.1: Adolescents who hold stronger social perceptions of smoking will be more likely to engage in smoking behaviors and with a greater intensity.*

*H2.2: Greater exposure to pro-tobacco media will increase the likelihood of smoking behaviors.*

The final set of hypotheses encompasses factors that are expected to increase resiliency to be a never smoker or promote smoking cessation through anti-tobacco awareness. These factors include the quality of anti-smoking education and media exposure, and lead to the last four hypotheses:

*H3.1: The higher the quality of anti-tobacco education the adolescent received during the current school year the less likely they will engage in smoking behaviors.*

*H3.2: Exposure to informal anti-tobacco education will decrease smoking behaviors.*

*H3.3: Knowing that cigarettes are dangerous to human health will decrease the likelihood of engaging in smoking behaviors.*
H3.4: Knowing that cigarettes can cause weight loss will increase the likelihood girls will engage in smoking behaviors.

4.5. Estimation Strategy

4.5.1. Principal Component Analysis

As was described in the data section, we use principal component analysis (PCA) as a method of data reduction for two measures: social perception of smoking and pro-tobacco exposure. If we were to include all seven variables for social perception and five for pro-tobacco exposure there would be issues of multicollinearity and reduced efficiency in the modeling. PCA distills these various measures into the minimum number of components with the maximum amount of variation. Following Kaiser’s Rule, eigenvalues exceeding unity were kept and a varimax rotation was used to improve factor-loading distribution. This resulted in three components for social perception and one component for pro-tobacco exposure. These components either captured close to or more than half of the total variation from the respective original variables.

4.5.2. Maximum Likelihood Estimation

We estimate Equation 1 using maximum likelihood estimation. In doing so we use all observation which have non-missing values to simultaneously estimate the three smoking statuses: never smoker, former smoker, and level of cigarettes smoked by current smokers. However, by estimating the varying smoking statuses in one equation we are
forfeiting the ability to estimate the covariances between the three smoking categories and hence, we are assuming independence between the error terms.

Conceptually, the quantity of cigarettes smoked can be expressed by the following function:

\[ y_{mi} = f(Proximity, Pro-Tobacco Perception, Anti-Tobacco Awareness, Controls) \]

Where \( y_{mi} \) equals one if individual \( i \) belongs to category \( m \) cigarettes smoked; \( Proximity \) is a vector of two variables related to being around other smokers as defined in the data section; \( Pro Tobacco Perception \) is a vector of variables related to social perception and pro-tobacco message exposure; \( Anti-Tobacco Awareness \) is a vector of formal and informal education, health perception and anti-tobacco message exposure variables; and the socio-demographic controls are in the vector \( Controls \).

4.5.3. Coefficient Interactions

In our preliminary modeling we estimated a multinomial logit, separating boys from girls, to determine what factors influenced the likelihood that a respondent was a never smoker, former smoker, or current smoker (found in Appendix B). Previous literature has found some factors influence smoking behaviors of boys and girls differently (Tsai et al. 2008, Nikaj and Chaloupka 2005, and Nonnemaker et al. 2011). Indeed, this is what our preliminary results indicated. However, when we moved to estimating the trivariate ordered probit model with independence, we were unable to estimate separate regressions for boys and girls. In the preferred model there are 37 female smokers and there is low
variation within the level of cigarettes smoked, three-fourths of girls smoke less than one cigarette a day.

So as to not lose out on this gender variation in results found in the preliminary analyses, we instead interact female with the variables that showed significant gender differences. For the portions of Equation 1 which are include the full sample and the former smokers subsample, we include a full factorial of female with the following variables: parents smoke, each social perception variable, the impact cigarettes have on weight, if you believe cigarettes are harmful to your health, and each formal anti-smoking education quality variable. For the ordered probit portion of Equation 1 the aforementioned list of variables excluding the impact cigarettes have on weight as well as low and medium education quality were included. Those variable interactions were omitted because their inclusion led to perfect predictability for some observations and hence we were unable to estimate the trivariate ordered probit model with independence.

4.5.4. Average Marginal Effects

The raw coefficients from the maximum likelihood estimation might not provide the correct sign nor magnitude of the included variables for three reasons. First, with the inclusion of the variable interactions the impact of one variable will be captured by two coefficients. Second, with the nature of the trivariate ordered probit model, smoking states are dependent upon previous smoking states. For instance, the likelihood of being a quitter depends upon having started smoking initially. Thirdly, with the ordered probit, the signs on the coefficients might not correctly express the direction of the impact of the variable because the cut values or $\tau_m$ also need to be taken into account. As such, to
assess the overall impact the variables have on $y_i$ we estimate the average marginal effect of the conditional mean as is presented in KMY.

\[
AME_L = \frac{1}{A} \sum_{i=1}^{A} \left\{ \sum_{m=1}^{M} \frac{\partial[Pr(y_i = m|N_i = 0, Q_i = 0)] \partial X_{ik}}{[1 - Pr(N_i = 1)][1 - Pr(Q_i = 1)]} \right\}
\]

Or

\[
AME_L = \frac{1}{A} \sum_{i=1}^{A} \left\{ \sum_{m=1}^{M} \bar{y}_m [\Phi(\tau_m - X_i\gamma) - \Phi(\tau_{m-1} - X_i\gamma)] [1 - \Phi(X_i\alpha)][1 - \Phi(X_i\beta)] \right\}
\]

\[
\left. * \frac{\gamma_k [\Phi(\tau_{m-1} - X_i\gamma) - \Phi(\tau_m - X_i\gamma)] - \alpha_k \phi(X_i\alpha)}{\Phi(\tau_m - X_i\gamma) - \Phi(\tau_{m-1} - X_i\gamma)} - \frac{\beta_k \phi(X_i\beta)}{1 - \Phi(X_i\beta)} \right\}
\]

Where $\bar{y}_m$ represents the category mean and it equals zero when $m=0$, when the respondent does not smoke cigarettes and $\phi$ represents the normal PDF. Additionally, we also calculate the average marginal effects for being a quitter and an abstainer from smoking which are shown in equations 3 and 4, respectively.

\[
AME_Q = \frac{1}{A} \sum_{i=1}^{A} \left\{ [1 - \Phi(X_i\alpha)][\Phi(X_i\beta)] * \left[ \frac{\beta \phi(X_i\beta)}{\Phi(X_i\beta)} - \frac{\alpha \phi(X_i\alpha)}{1 - \Phi(X_i\alpha)} \right] \right\}
\]

\[
AME_N = \frac{1}{A} \sum_{i=1}^{A} \left\{ \alpha \phi(X_i\alpha) \right\}
\]

To calculate the change of a discrete variable, we compute a finite change for moving from 0 to 1, all else held equal.
4.5.5. Models

Each vector of independent variables is added to the regression in subsequent models resulting in a total of four models. For brevity we only report the results from Model 3, which contains all of the hypothesis variables but no control variables. The results of the other models were consistent. All regressions use probability weights and are clustered at the school-class level to account for similarity of students from the same classroom and school.

4.6. Results

4.6.1. Descriptive Statistics

Table 4.1 presents the descriptive statistics of the analytical sample for the variables used in the preferred model for both sexes combined and for boys and girls separately. Most boys and girls have never had a puff of a cigarette before, at 84.1% and 95.4% respectively. Additionally, more boys than girls are former smokers. Finally, in total about 5% of the sample is currently smoking and this is driven by male smokers: 8.5% of boys are current smokers whereas only 1.7% of girls are smoking. Furthermore, most current smokers – 75.68% of girls and 61.70% of boys - smoke less than one cigarette a day. Boys have a wider distribution of smoking across the categories. Although a larger percentage of girls report being in the top level of cigarette consumption than boys, because of the small number of girl smokers overall there are fewer girls in this category than boys (3 girls versus 8 boys).
Substantially more boys report that they have close friends who smoke than girls, from over half to just over one quarter, respectively. However, both report similar rates of parental smoking at around 50%.

For the social perception variables, under the pro-tobacco marketing and social perception section, boys and girls on average have a negative perception of smoking being attractive and the sentiment is stronger for boys. Smoking cigarettes is not seen to be a popular act by both genders. Although girls have a negative perception of adult smokers, boys have a positive perception of them. Finally, girls have less exposure to pro-tobacco media than boys.

Within anti-tobacco awareness and knowledge, we find that around forty percent of respondents are exposed to high quality education. There is no indication in the survey if the student attends same sex schools, so it is likely that on average both boys and girls are exposed to the same quality of education. However, the recall of the education differs by sex. Girls report having a higher quality smoking education than boys by 6 percentage points. Around a quarter of both boys and girls are exposed to anti-tobacco media. Additionally, parents talk to their sons and daughters at nearly equal rates about the dangers of smoking at close to 66%. The health perception of cigarettes and smoking remains relatively constant across the sexes: most know that smoking can help you lose weight and that smoking (be it your own smoking or being exposed to other’s smoking) is harmful. Finally, the controls are similar across sex.

4.6.2. Trivariate Ordered Probit Regression Results
The raw regression results for the trivariate ordered probit with independence are found in Table 4.2. Although there are three separate columns indicating the smoking status, each smoking status was estimated simultaneously in the same maximum likelihood estimation.

Being around other smokers had a strong impact on adolescents smoking behaviors. Having close friends who smoke or parents who smoked decreased the likelihood of being a never smoker and it also decreased the likelihood of quitting smoking. Of the two variables, having friends who smoked had a stronger impact than having parents who smoke for engaging and sustaining smoking behaviors. Additionally, being a female increased the likelihood of abstaining from smoking. These variables did not statistically significantly influence the level of cigarettes consumed by current smokers.

All of the social perception of smoker’s variables were interacted with female and with the raw regression results we cannot determine the full impact these variables have on smoking status. The results do show adolescents who hold stronger positive beliefs around smoking increasing popularity were statistically significantly less likely to be a never smoker and the interaction between female and popularity was positive and statistically significant. Adolescents who perceive adult smokers to be cooler at a higher rate are also more likely to start engaging in smoking behaviors. As expected, exposure to pro-tobacco media was negatively correlated with being a never smoker and former smoker and also positively correlated with level of cigarettes consumed. Although these correlations were not statistically significant.
Anti-tobacco awareness and knowledge had mixed impacts on smoking behaviors. On the whole, lower quality education did not influence smoking behaviors, but high quality formal anti-tobacco education in school did make a difference. Those who had a high-quality education, as opposed to no formal anti-tobacco education, were more likely to be never smoker. Further when this was interacted with female, it was positive and statistically significant. Counterintuitively, being exposed to anti-tobacco media, compared to no exposure to this media, was correlated with a decreased likelihood of being a never smoker and lower likelihood of being a former smoker. Family discussions about the dangers of cigarettes did not statistically significantly influence smoking behaviors. Believing that cigarettes could change weight—either increase or decrease weight—was associated with an increase in the chance of being a never smoker or a former smoker. Additionally, the interactions of these weight variables were negative and statistically significant for never smokers indicating that the impact cigarettes have on weight changes could be different for boys and girls. Knowing that cigarettes are harmful for your health was not statistically significant across the smoking statuses. However, the interaction of this variable with female was statistically significant for never smoker. Knowing that the smoke from other’s cigarettes is harmful to your own health did not influence smoking behaviors.

4.6.3. Average Marginal Effects Results

To determine the full impact these factors have on smoking behaviors, we transform the results to incorporate the gender effect from the interaction and the structure of the trivariate ordered probit likelihood function. To do so we calculate the average marginal
effects for each smoking status. The never smoker and former smoker average marginal effects were calculated after the trivariate ordered probit model was bootstrapped 200 times. The delta-method was used to calculate the standard errors.

Table 4.3.A. contains the results for the never smoker average marginal effects. The first column presents the results for boys and girls combined and the second and third column present the results for girls and boys, respectively. Having a close friend who smokes decreases the likelihood of being a never smoker by a statistically significant 10.7 percentage points, all else held equal. The impact is larger for boys than it is for girls: boys who have a close friend who smoke are 13.5 percentage points more likely to be an ever smoker whereas the impact for girls is 7.8 percentage points. Having a parent who smokes increased the likelihood of smoking initiation for boys by 4.5 percentage points. The combined results for boys and girls indicate that perceived popularity of smokers decreases the likelihood of being a never smoker, but when this is broken down by gender we find that this only impacts boys’ reduced likelihood to be a never smoker. Both boys and girls, together and separately, are less likely to be a never smoker as the intensity of beliefs that adult smokers are cool increases. When it comes to anti-tobacco awareness, generally the quality of formal anti-tobacco education does not statistically influence the likelihood of being a never smoker. Being exposed to anti-tobacco media decreased the likelihood of being a never smoker by 4.9 percentage points and the effect was twice as large for boys (at 6.5 percentage points) than it was for girls (at 3.2 percentage points). We do not find a significant relationship between knowing that cigarettes can lead to weight loss and starting smoking for girls. However, we find that boys are more likely to be a never smoker when they believe smoking can lead to any
changes in their weight—be it weight loss or gain—at 12.9 to 12.3 percentage points, respectively.

The former smoker average marginal effects, found in Table 4.3.B., at first glance seem counterintuitive: having a close friend who smokes or boys who have a higher belief that adult smokers are cool statistically significantly increases the likelihood of being a quitter whereas believing that cigarettes are harmful to your health decreases the likelihood of being a former smoker by 4.4 percentage points for boys and girls combined and by 6.2 percentage points for girls alone. These results also have the opposite sign as the raw coefficient results. However, we also find that girls who receive a low-quality formal anti-tobacco education are 3.9 percentage points more likely to be a quitter. Additionally, anti-tobacco exposure increases the likelihood girls will quit smoking by 1.7 percentage points. We postulate the reason for these contradictory and inconsistent results is due to the ambiguity inherent in the former smoker category that is especially apparent in the average marginal effect results. This limitation is explored further in the discussion section.

The final set of average marginal effects are found in Table 4.3.C. for the conditional mean of the level of consumption. These results were calculated by creating a nested set of programs which ran the maximum likelihood estimation of the trivariate ordered probit model, calculated the average marginal effect of the conditional mean for level of cigarette smoked, and bootstrapped this process 200 times. The reported coefficient is the coefficient resulting from before the random sampling with replacement, the shaded cells indicate the coefficient is statistically significant at the 0.05 level, and the bias corrected 95% confidence interval is presented in parentheses below.
Being around other smokers increases access to cigarettes. Having a close friend who smokes statistically significantly increases the likelihood of all levels of smoking. However, the impact of having a parent who smokes only increases the level of cigarette consumption for boys. Different aspects of the social perception of cigarettes increases boys and girls level of cigarette consumption. Girls who believe that smoking makes someone more attractive/makes social situations easier more likely to engage in medium levels of consumption smoking 1 cigarette a day or 2-5 cigarettes a day. Further, girls who think that adult smokers are cool are more likely to smoke 2-5 cigarettes a day. Boys who think that smoking makes someone popular increases their likelihood of smoking less than one cigarette a day (low consumption) or one cigarette a day. Exposure to pro-tobacco media increased the odds of smoking one cigarette a day or 2-5 cigarettes a day for girls.

Formal anti-tobacco education expressed a gradient of impact: low quality education did not statistically impact level of consumption, but a high-quality education deceased the odds girls would smoke less than one cigarette a day and decreased the likelihood boys would smoke 1 cigarette a day or 2-5 cigarettes a day. Exposure to anti-tobacco media increased all levels of smoking for girls and one or few cigarettes a day for boys. Similar to the raw regression results and the average marginal effects for never smokers, only boys who perceived that cigarettes would cause a deviation from their current weight were less likely to smoke at all levels of cigarette consumption. Finally, we find that knowledge that cigarettes is harmful to your health actually statistically significantly increases the likelihood of 1 cigarette or 2-5 cigarettes a day for boys.
4.6.4. Review of Results

Overall, we find evidence to support our proximity to other smokers hypotheses. Having close friends who smoking statistically significantly increases the odds of smoking initiation, more so for boys than for girls. Additionally, boys who have parents who smoke are not only more likely to be a never smoker, but they are also more likely to smoke low to medium levels of cigarettes. Our hypothesis around social perceptions increasing the likelihood of smoking behaviors was supported by the data, and we found a strong gender effect as well. Boys who held stronger beliefs that smoking makes a person popular or that adult smokers are cool were more likely to start smoking, and the former increased the likelihood of low to medium levels of consumption. Whereas for girls, when they held more positive perceptions of adult smokers they were not only more likely to start smoking but it also increased the odds of medium cigarette consumption. Additionally, girls with more intense perceptions that smoking can make a person attractive were more likely engage in medium levels of consumption. We found that increased exposure to pro-tobacco media increasing the likelihood of girls smoking 1 cigarette a day or 2-5 cigarettes a day. We found mixed levels of support for our anti-tobacco awareness and knowledge hypotheses. On the one hand, better quality formal education did indeed reduce smoking behaviors, but informal education had either null impacts (family discussing the dangers of tobacco) or deleterious impacts on youth smoking (for exposure to anti-tobacco media). Respondents who knew about cigarettes being harmful to your health were more likely to smoke a higher level of cigarettes, counter to our hypothesis. Finally, our last hypothesis on girls smoking for weight loss
was not supported. Rather we found that boys were less motivated to engage in smoking behaviors if it would impact their weight.

4.7. Discussion

The smoking epidemic has transitioned to the developing world: tobacco companies sought out new areas of profit and they successfully did so by taking advantage of international trade liberalization and globalization at large (Yach and Bettcher 2000). The youths in developing nations are particularly in trouble because of this shift. Despite the tobacco regulations outlined in the FCTC, many participating nations have not fully implemented the all aspects of the FCTC. A 2017 monitoring report acknowledges the success in expanding tobacco control measures where 63% of the world’s population is covered by at least one of six MPOWER measures17 but only eight nations have four or more measures in place (WHO and Bloomberg Philanthropies 2017). As such, youths across the globe are still exposed to varying degrees to pro-tobacco advertisements, lax enforcement of underage purchasing laws, and low tobacco taxes all of which have been shown to be correlated with increased tobacco use by youths (Lovato, Watts, and Stead 2011; Lantz et al. 2000; International Agency for Research on Cancer 2011).

Nepal has made substantial progress on implementing smoke-free policies, health warnings and mass media campaigns, and advertising bans, however the rate of cigarettes taxes are some of the lowest in the region (WHO and Bloomberg Philanthropies 2017). Despite these policy wins, the rate of youth smoking in Nepal has increased by about 2 percentage points since the Tobacco Control and Regulatory bill was passed in 2011

17 This does not include monitoring or mass media campaigns.
There is a disconnect between the aims of national level polices and the behaviors of the people it covers. As such, more work is needed to understand why youths in Nepal decide to engage or not engage in a variety of smoking behaviors in order to increase the efficacy of anti-tobacco policies.

We contribute to this effort by studying the impacts proximity to other smokers, pro-tobacco marketing and social behaviors; and anti-tobacco awareness and social perception have on youth in Nepal belonging to one of three smoking statuses using a trivariate ordered probit model with independence following the work by KMY. We included interactions between the variable female and other variables of interests to examine the gendered impact of our findings guided by preliminary analyses. To determine the impact covariates had on the level of cigarette consumption by current smokers we calculated the average marginal effect of the conditional mean of ordered probit portion of the model as was detailed by KMY. Furthermore, we extended this previous study by also calculating the average marginal effects of the two other smoking statuses- never smoker and former smoker.

Our results demonstrated the being around other smokers can have a large impact on youth’s smoking behaviors. The raw regression results found that having a friend or parent who smokes increased the likelihood of being an ever smoker and a non-quitter. The average marginal effects provide additional insight on magnitude and gendered nature of this impact. We find that having a close friend who smokes greatly decreases the likelihood of being a never smoker, by 10.7 percentage points, which is nearly four time greater than the impact of having a parent who smokes. This is consistent with the literature that peer influence is more important that parental influence for initiating
smoking behaviors (Avenevoli and Merikangas 2003). We also find that peer influence has a stronger impact on the likelihood boys will be an ever smoker than girls.

Although the raw regression results found that being around other smokers did not influence level of cigarette consumption, the conditional average marginal effects uncovered that these factors are indeed correlated with increased smoking behavior. Simply being around other smokers increases youth’s access to cigarettes. Having a close friend who smoked increased the level of cigarette consumption for both boys and girls, yet have a parent who smokes only increased level of consumption for boys. Furthermore, we find that parental smoking behavior only statistically significantly decreased boys likelihood to be a never smoker. In many developing nations, it is not acceptable for women to smoke, a study among men and women in Bangladesh and Pakistan found that both agreed it is disrespectful for women to smoke and doing so can impact a woman’s chances for marriage (Bush et al. 2003). Yet, today among young adults and women in Nepal, smoking cigarettes is considered “fashionable” (Ministry of Health and Population 2012). The long-held taboo of women smoking can help explain why parents smoking behavior increases access for boys but not girls, yet this belief is eroding among the youth. Altogether, we found evidence in favor of both of our hypotheses around proximity to other smokers.

Social perceptions of cigarette smoking had differential impacts on boy’s and girl’s smoking behaviors. There was a strong positive correlation between believing that cigarettes make people more popular and believing that adult smokers were cool with the likelihood of trying cigarette smoking. These findings are in line with a previous study on youth perceptions of smoking in Nepal (Aryal et al. 2013) and the United States (Song et
al. 2009). However, as for the level of cigarettes consumed, we find girls were statistically influenced by first increased belief that smoking makes a person attractive and second that adult smokers are cool. On the other hand, boys had an increased likelihood of low levels of cigarette consumption when they held higher beliefs that smoking made someone popular.

We find that greater exposure to pro-tobacco media is positively and statistically associated with increased level of cigarettes smoked for girls but not statistically significantly correlated with being a never smoker or former smoker. The GYTS was collected before and around the time that Tobacco Control and Regulatory Act was passed in 2011 and a comprehensive tobacco ban was to be implemented (WHO 2011). A limited advertising ban in developing nations was found to reduce smoking rates, but a comprehensive ban reduced smoking rates further (Blecher 2008). Although we do not have youth smoking rates before 1997, the GYTS stated in 2001 in Nepal, or after 2011, it is reasonable to expect that a comprehensive ban would reduce smoking rates further.

The impact of formal anti-tobacco education on smoking behaviors provided intermittent support for our hypothesis. Better quality anti-tobacco education decreased the odds of smoking low and medium qualities of cigarettes for boys. However, formal education did not statistically significantly impact youth’s smoking initiation. There is limited evidence in the literature of anti-smoking programs in school having long term impacts on remaining smoke free in a given time span (Wiehe et al. 2005). One program that schools in Nepal could explore implementing is the 2001-2012 Health Education and Tobacco Intervention Program (HETIP). Over this program’s lifespan, it was implemented in 494 public schools across two-thirds of Nepal’s districts. The HETIP
combined tobacco education over a two-day health course to all fourth to twelve grade students and students performed a community street play about the dangers of smoking (Kainulainen and Kivelä 2012). HETIP did not have a control group so we cannot say that this program caused lower smoking rates. Nonetheless, this type of program engages heavily with both students and community members and appears to have the attributes to create a lasting impact on both group’s smoking behaviors, potentially creating a positive spillover effect.

Informal education either did not impact youth’s smoking behaviors, as was the case with family discussion about the dangers of smoking, or had a statistically significantly negative correlation with smoking behaviors. These results were counter to our hypothesis. We find that youths who are exposed to anti-tobacco media are less likely to be a never smoker or a former smoker. The results from the average marginal effects find that this exposure reduces the odds of being a never smoker by 6.5 percentage points for boys and 3.2 percentage points for girls. Anti-tobacco media is also correlated with increased likelihood of nearly all levels of cigarette consumption by both boys and girls. Although it might seem counterintuitive that anti-tobacco media increases smoking behaviors, this is a plausible outcome. Previous literature has shown that the efficacy of anti-tobacco media varies based upon the messenger. Anti-tobacco messages presented by public health entities are perceived by youth to be more effective at preventing smoking initiation than messages presented by tobacco companies (DeBon and Klesges 1996). Worldwide, tobacco companies have an incentive to engage in anti-tobacco media because it is in their own self-interest: it provides them a way to diffuse messages from public health officials, mitigate anti-tobacco legislation, maintain their access to youths,
and improve the standing of the tobacco industry with policy makers (Landman, Ling and Glantz 2002). With the GYTS data, we do not know who was sponsoring the anti-tobacco messages the students saw. However, since the results are correlated with increased smoking behaviors this lends credence to idea that these messages might be sponsored by tobacco industries.

The impacts of health-related knowledge on smoking behaviors also went against our hypothesis. Although there have been previous studies which have found that women and girls smoke to lose weight (Tsai et al. 2008 and French et al. 1994), we find that weight motivations do not influence girl’s smoking behaviors. Rather, we find that it has a strong impact on boys. Boys who perceive that cigarettes will either cause them to gain weight or to lose weight are less likely to be a never smoker and it is also associated with reduced cigarette consumption. This finding is not common in the literature. Previous studies have found that weight concerns do not motivate boys smoking habits (French et al. 1994). Although this previous study was conducted in the US in the 1990’s among mostly white youth, the direct applicability of these findings to boys in Nepal or Southeast Asia are questionable. Future studies could do more to investigate this link.

There is not a consistent impact of knowing that cigarettes are harmful to your health and what youths do with that information. On one hand, for girls it increases the likelihood that they will be a never smoker, but it also increases the level of cigarettes consumed for boys (at medium levels of consumption). The field of youth risky behaviors provides some insights on to why this might be the case. Youth who engage in risky behaviors where the risk itself is uncertain might engage in this behavior more in the future because of the reduced marginal cost (Gruber 2001).
The results of this study illuminate several policy avenues that policy makers can address to reduce youth smoking behaviors. The strategic use of both formal and informal education has the potential to make a positive impact on smoking behaviors. First, tighter regulations need to be placed on who is allowed to disseminate anti-tobacco messages. Not only have tobacco companies determined ways to leverage these messages to their advantage, but they have successfully done so in spite of long held cultural norms against youth, specially girls and women’s smoking. Young people and women in urban Nepal see smoking cigarettes as a fashion statement and a symbol of independence (Ministry of Health and Population Government of Nepal 2012) which harkens back to the early 20th century in the United States when cigarettes were marketed to women as tools of liberations, as “torches of freedom” (Amos and Haglund 2000). A consorted and nationwide effort will be need to create and distribute public health sponsored anti-tobacco messages. However, this will be a difficult endeavor for the government to engage because the entire national governmental expenditure on tobacco control in 2016 was less than US$50,000 (WHO 2017). This effort could be supported by increasing the fine tobacco companies incur from illegally advertising their product. The fine currently stands at a maximum of US$919 (WHO 2011). This will be particularly important in reducing cigarette consumption for girls. 

Regardless, it is pivotal these anti-tobacco messages are targeted to counter the narrative that smoking is makes you popular and that adult smokers are cool. Additional attention needs to be paid to increasing the knowledge that smoking is harmful to your health, especially for boys. Youths have been shown to respond differently to long-term and short-term consequences from smoking. Ayral, Petzold, and Krettek (2013)
examined what the perceived risks and benefits of smoking cigarettes was among non-smoking youth in Nepal. They found that youths who perceived smoking to increase the risk of bad breath, bad cough, and trouble breathing were less likely to be susceptible to start smoking. Anti-tobacco messages which incorporate this short-term consequence of smoking could reduce the likelihood that youth will begin smoking. Our findings also indicate that targeting messages to boys that cigarettes will change your weight could also be useful in reducing their smoking behaviors.

Youths who have people around them that smoke have increased access to cigarettes and were found to smoke more cigarettes. As such, the ban on selling tobacco to people under the age of 18 and the prohibition of smoking in private homes enacted a part of the Tobacco Control and Regulatory Bill of 2011 need to be enforced (WHO 2011). Lastly, more schools should provide higher quality anti-tobacco education and this needs to be reimagined to find ways to reduce smoking initiation especially for girls.

This analysis is subject to several limitations. The GYTS data is only representative of students who showed up to school on the day that the survey took place. It does not necessarily represent students who do not attend formal education or who were not present in class that day. This is also cross-sectional data, as such we are unable to assess the causal impact these variables had on youth smokers within this modeling construct. With longitudinal data, we could follow the same person over the course of their adolescence to see how their smoking status changes over time. However, to the best of our knowledge there is no nationally representative data in Nepal like this. The GYTS data does not contain information on household income, caste, or geography. As such we are unable to control for these important cultural factors.
In our results we found that there were few factors that statistically significantly influenced the likelihood of being a former smoker and the average marginal effects of this smoking status produced several counterintuitive results. This could be partially due to the way that this category was constructed. With the GYTS data, the former smoker category is not ideally defined. As it currently stands, a person who had only one puff of a cigarette before and never touched a cigarette again is combined with other youths who were a bona fide former smokers who successfully quit. There is no question in the GYTS which asks about former smoking intensity, which could have been used to separate out this group. As such, this grouping and designation as former smoker, might in practicality not be the most accurate. It could be better to think of this group as at risk to start smoking again in this context especially in regard to the average marginal effects calculation. Regardless, it was vital that this smoking category was included in the trivariate ordered probit with independence model to prevent issues of selectivity.

As was mentioned in the introduction to this chapter, we are only studying the factors that influence cigarette smoking behaviors and not other tobacco products. The survey questions related to other smokeless and smokeable tobacco usage were asked inconsistently between the Nepal GYTS 2007 and 2011 waves. In the full 2007-2001 Nepal GYTS data, 49 girls have smoked cigarettes in the past 30 days, but 269 and 121 girls have used some sort of smokeless or other smokeable tobacco, respectively, during the same time frame. As such our study only explores one facet of girls’ and youths’ tobacco use in Nepal and we encourage both more consistent GYTS survey instruments to be used within a country and for the further exploration of other tobacco products by boys and girls.
4.8. Summary

In this analysis we used two waves of nationally representative data from adolescents in Nepal to determine how peer influence; pro-tobacco marketing and social perception; and anti-tobacco awareness and knowledge have on belonging to one of three smoking statuses: never smoker, former smoker and the level of consumption by current smokers. The results from the average marginal effects of a trivariate ordered probit model with independence show that being around other smokers not only increased the likelihood of smoking initiation but it also increased the likelihood of smoking higher quantities of cigarettes. Parental smoking behavior only had a statistically significant impact on boys. Social perceptions of cigarette smoking impacted boys and girls differently, but it always increased the odds they would engage in smoking behaviors. High quality formal anti-tobacco education reduced the likelihood of smoking low and medium quantities of cigarettes for girls, but it did not statistically impact abstaining from smoking. Exposure to anti-tobacco media greatly decreased the odds of being a never smoker and was positively associated with increased likelihood of smoking low and medium levels of cigarettes. Finally, we find that boys were more likely to not engage in or had reduced smoking intensity when they believed it would change their weight. No impact was found was found among girls for smoking behaviors and weight.

Truly implementing and enforcing the policies outlined in the Tobacco Control and Regulatory Bill will make immense strides towards reducing smoking among the youth. Our study outlines priority areas of attention. The findings demonstrate the important role that both formal and informal education has on smoking behaviors. Policy
makers need to make sure that anti-tobacco media messages promoted in public or in the classroom help students understand the dangers of cigarettes and counter the narratives of smoking being cool. It will also be important to ensure that the messages are tailored to their audience because different messages will resonate better with boys or girls.

Enforcing the ban on selling tobacco to youth needs to be strongly enforced. By taking these steps we can help shape an ideal of gender parity in smoking behaviors in Nepal, one where both boys and girls have low rates of smoking engagement.
Questions used to create smoking tree:

- Ever Had A Puff?
  - “Have you ever tried or experimented with cigarette smoking, even one or two puffs?”

- Did You Quit Smoking?
  - “During the past 30 days (one month), on how many days did you smoke?”
  - Note: This was dichotomized such that zero days equals one else one.
    Additionally, if people reported smoking a non-zero amount of cigarettes in the past 30 days but said that they did not smoke in the past 30 days,
they were recoded as currently smoking. There were only five instances of that in this data.

- How Many Cigarettes Do You Smoke?
  
  o “During the past 30 days (one month), on the days you smoked, how many cigarettes did you usually smoke?”

  o Note: the categories were combined into: (1) I did not smoke in the past 30 days, (2) less than 1 cigarette a day, (3) 2 cigarette per day, (4) 2-5 cigarettes per day, and (5) 6+ cigarettes per day.
Table 4.1: Descriptive Statistics of Smoking Behavior and Model Covariates by Sex of Analytical Model

<table>
<thead>
<tr>
<th></th>
<th>Together Mean(SD)</th>
<th>Girls Mean(SD)</th>
<th>Boys Mean(SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Smoking Status</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Never Smoker ((1=yes))</td>
<td>0.897</td>
<td>0.954</td>
<td>0.841</td>
</tr>
<tr>
<td>Former Smoker ((1=yes))</td>
<td>0.052</td>
<td>0.029</td>
<td>0.074</td>
</tr>
<tr>
<td>Current Smoker ((1=yes))</td>
<td>0.051</td>
<td>0.017</td>
<td>0.085</td>
</tr>
<tr>
<td><strong>Level of Cigarette Consumption of Current Smokers (%)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Less than one cigarette a day</td>
<td>64.00</td>
<td>75.68</td>
<td>61.70</td>
</tr>
<tr>
<td>1 cigarette a day</td>
<td>21.78</td>
<td>16.22</td>
<td>22.87</td>
</tr>
<tr>
<td>2-5 cigarettes a day</td>
<td>9.33</td>
<td>0.00</td>
<td>11.17</td>
</tr>
<tr>
<td>6+ cigarettes a day</td>
<td>4.89</td>
<td>8.11</td>
<td>4.26</td>
</tr>
<tr>
<td><strong>Proximity to Other Smokers</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>((1=yes))</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Friends Smoke</td>
<td>0.400</td>
<td>0.273</td>
<td>0.528</td>
</tr>
<tr>
<td>Parents Smoke</td>
<td>0.487</td>
<td>0.468</td>
<td>0.505</td>
</tr>
<tr>
<td><strong>Pro-Tobacco Marketing &amp; Social Perception</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Social Perception</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Level of Attraction/ Ease at Parties</td>
<td>-0.030</td>
<td>-0.022</td>
<td>-0.039</td>
</tr>
<tr>
<td>((1.254))</td>
<td></td>
<td>(1.258)</td>
<td>(1.25)</td>
</tr>
<tr>
<td>Popularity</td>
<td>-0.010</td>
<td>-0.010</td>
<td>-0.009</td>
</tr>
<tr>
<td>((1.21))</td>
<td></td>
<td>(1.171)</td>
<td>(1.249)</td>
</tr>
<tr>
<td>Perception of Adult Smokers</td>
<td>-0.018</td>
<td>-0.083</td>
<td>0.047</td>
</tr>
<tr>
<td>Media Exposure</td>
<td>((1.08))</td>
<td>(.987)</td>
<td>(1.163)</td>
</tr>
<tr>
<td>Pro-tobacco Exposure</td>
<td>0.014</td>
<td>-0.018</td>
<td>0.047</td>
</tr>
<tr>
<td>((1.426))</td>
<td></td>
<td>(1.418)</td>
<td>(1.434)</td>
</tr>
<tr>
<td><strong>Anti-Tobacco Awareness &amp; Knowledge</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Quality of Smoking Education (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>None(^b)</td>
<td>16.81</td>
<td>15.91</td>
<td>17.72</td>
</tr>
<tr>
<td>Low</td>
<td>15.41</td>
<td>14.15</td>
<td>16.67</td>
</tr>
<tr>
<td>Medium</td>
<td>25.68</td>
<td>24.68</td>
<td>26.69</td>
</tr>
<tr>
<td>High</td>
<td>42.09</td>
<td>45.25</td>
<td>38.92</td>
</tr>
<tr>
<td>Anti-Tobacco Media Exposure ((1=yes))</td>
<td>0.241</td>
<td>0.246</td>
<td>0.236</td>
</tr>
<tr>
<td>Family discussed smoking ((1=yes))</td>
<td>0.672</td>
<td>0.687</td>
<td>0.657</td>
</tr>
<tr>
<td>Cigarettes and Weight (%)</td>
<td>Gain Weight</td>
<td>Lose Weight</td>
<td>No Difference(^b)</td>
</tr>
<tr>
<td>---------------------------</td>
<td>-------------</td>
<td>-------------</td>
<td>-------------------</td>
</tr>
<tr>
<td></td>
<td>6.16</td>
<td>6.01</td>
<td>6.30</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Smoking Cigarettes is Harmful ((1=\text{yes}))</th>
<th>0.899</th>
<th>0.908</th>
<th>0.890</th>
</tr>
</thead>
<tbody>
<tr>
<td>Others Cigarette Smoking Is Harmful to You ((1=\text{yes}))</td>
<td>0.875</td>
<td>0.877</td>
<td>0.872</td>
</tr>
</tbody>
</table>

**Controls\(^c\)**

<table>
<thead>
<tr>
<th>Age (%)</th>
<th>&lt;13 yo(^b)</th>
<th>14-15 yo</th>
<th>16 + yo</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>23.42</td>
<td>44.20</td>
<td>32.38</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Grade (%)</th>
<th>Sevent(^b)</th>
<th>Eighth</th>
<th>Ninth</th>
<th>Tenth</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>23.92</td>
<td>20.95</td>
<td>28.39</td>
<td>26.74</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Female (%)</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.501</td>
<td>--</td>
<td>--</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Year 2011 ((1=\text{yes}))</th>
<th>0.531</th>
<th>0.541</th>
<th>0.520</th>
</tr>
</thead>
<tbody>
<tr>
<td>Observations</td>
<td>4,419</td>
<td>2,212</td>
<td>2,207</td>
</tr>
</tbody>
</table>

*Notes:* \(^a\)Standard deviation only reported for continuous variables. \(^b\)Base category. \(^c\)Controls are not included in analytical model (except female) and have grade and age have a total of 4389 observations. The unweighted sample reported. Source: GYTS 2007 and 2011.
Table 4.2.: Regression Results of Trivariate Ordered Probit with Independence Model

<table>
<thead>
<tr>
<th></th>
<th>Never Smoke</th>
<th>Former Smoker</th>
<th>Level of Cigarettes Consumed</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Proximity to Other Smokers</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Friends Smoke</td>
<td>-0.630***</td>
<td>-0.781***</td>
<td>0.296</td>
</tr>
<tr>
<td></td>
<td>(0.063)</td>
<td>(0.179)</td>
<td>(0.364)</td>
</tr>
<tr>
<td>Parents Smoke</td>
<td>-0.198**</td>
<td>-0.322*</td>
<td>0.106</td>
</tr>
<tr>
<td></td>
<td>(0.081)</td>
<td>(0.174)</td>
<td>(0.281)</td>
</tr>
<tr>
<td>Parents Smoke * Female</td>
<td>0.126</td>
<td>0.193</td>
<td>-0.001</td>
</tr>
<tr>
<td></td>
<td>(0.135)</td>
<td>(0.376)</td>
<td>(0.874)</td>
</tr>
<tr>
<td>Female</td>
<td>0.683*</td>
<td>0.0979</td>
<td>1.571</td>
</tr>
<tr>
<td></td>
<td>(0.393)</td>
<td>(2.899)</td>
<td>(4.391)</td>
</tr>
<tr>
<td><strong>Pro-Tobacco Marketing &amp; Social Perception</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Social Perception</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Level of Attraction/ Ease at Parties</td>
<td>0.0129</td>
<td>0.0775</td>
<td>0.077</td>
</tr>
<tr>
<td></td>
<td>(0.031)</td>
<td>(0.064)</td>
<td>(0.091)</td>
</tr>
<tr>
<td>Level of Attraction/ Ease at Parties * Female</td>
<td>-0.0386</td>
<td>-0.138</td>
<td>0.443</td>
</tr>
<tr>
<td></td>
<td>(0.055)</td>
<td>(0.144)</td>
<td>(0.358)</td>
</tr>
<tr>
<td>Popularity</td>
<td>-0.105***</td>
<td>-0.110</td>
<td>-0.013</td>
</tr>
<tr>
<td></td>
<td>(0.034)</td>
<td>(0.070)</td>
<td>(0.094)</td>
</tr>
<tr>
<td>Popularity * Female</td>
<td>0.166***</td>
<td>0.214</td>
<td>-0.128</td>
</tr>
<tr>
<td></td>
<td>(0.059)</td>
<td>(0.201)</td>
<td>(0.346)</td>
</tr>
<tr>
<td>Perception of Adult Smokers</td>
<td>-0.0815***</td>
<td>0.0199</td>
<td>0.004</td>
</tr>
<tr>
<td></td>
<td>(0.029)</td>
<td>(0.053)</td>
<td>(0.073)</td>
</tr>
<tr>
<td>Perception of Adult Smokers * Female</td>
<td>0.0175</td>
<td>-0.113</td>
<td>0.115</td>
</tr>
<tr>
<td></td>
<td>(0.052)</td>
<td>(0.189)</td>
<td>(0.227)</td>
</tr>
<tr>
<td>Media Exposure</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pro-tobacco Exposure</td>
<td>-0.0301</td>
<td>-0.0655</td>
<td>0.090</td>
</tr>
<tr>
<td></td>
<td>(0.024)</td>
<td>(0.056)</td>
<td>(0.094)</td>
</tr>
<tr>
<td><strong>Anti-Tobacco Awareness &amp; Knowledge</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Quality of Smoking Education</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low</td>
<td>0.129</td>
<td>0.0290</td>
<td>-0.170</td>
</tr>
<tr>
<td></td>
<td>(0.104)</td>
<td>(0.233)</td>
<td>(0.357)</td>
</tr>
<tr>
<td>Low * Female</td>
<td>-0.326*</td>
<td>0.890</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>(0.188)</td>
<td>(0.711)</td>
<td></td>
</tr>
<tr>
<td>Medium</td>
<td>0.0814</td>
<td>0.256</td>
<td>0.259</td>
</tr>
<tr>
<td></td>
<td>(0.118)</td>
<td>(0.244)</td>
<td>(0.300)</td>
</tr>
<tr>
<td>Medium * Female</td>
<td>0.101</td>
<td>0.761</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>(0.185)</td>
<td>(0.65)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Estimate 1</td>
<td>Estimate 2</td>
<td>Estimate 3</td>
</tr>
<tr>
<td>--------------------------------</td>
<td>------------</td>
<td>------------</td>
<td>------------</td>
</tr>
<tr>
<td>High</td>
<td>0.173*</td>
<td>0.259</td>
<td>-0.393</td>
</tr>
<tr>
<td></td>
<td>(0.113)</td>
<td>(0.212)</td>
<td>(0.306)</td>
</tr>
<tr>
<td>High * Female</td>
<td>-0.110</td>
<td>0.454</td>
<td>1.533*</td>
</tr>
<tr>
<td></td>
<td>(0.180)</td>
<td>(0.662)</td>
<td>(0.899)</td>
</tr>
<tr>
<td>Anti-Tobacco Media Exposure</td>
<td>-0.269***</td>
<td>-0.290*</td>
<td>0.048</td>
</tr>
<tr>
<td></td>
<td>(0.068)</td>
<td>(0.159)</td>
<td>(0.265)</td>
</tr>
<tr>
<td>Family discussed smoking</td>
<td>-0.0485</td>
<td>0.0192</td>
<td>0.118</td>
</tr>
<tr>
<td></td>
<td>(0.062)</td>
<td>(0.166)</td>
<td>(0.276)</td>
</tr>
<tr>
<td>Cigarettes and Weight</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gain Weight</td>
<td>0.456**</td>
<td>0.684*</td>
<td>-0.001</td>
</tr>
<tr>
<td></td>
<td>(0.192)</td>
<td>(0.348)</td>
<td>(0.480)</td>
</tr>
<tr>
<td>Gain Weight * Female</td>
<td>-0.866**</td>
<td>-0.198</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>(0.351)</td>
<td>(2.882)</td>
<td></td>
</tr>
<tr>
<td>Lose Weight</td>
<td>0.481***</td>
<td>0.387*</td>
<td>-0.144</td>
</tr>
<tr>
<td></td>
<td>(0.133)</td>
<td>(0.204)</td>
<td>(0.321)</td>
</tr>
<tr>
<td>Lose Weight * Female</td>
<td>-0.631**</td>
<td>0.096</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>(0.270)</td>
<td>(2.757)</td>
<td></td>
</tr>
<tr>
<td>Smoking Cigarettes is Harmful</td>
<td>0.00420</td>
<td>-0.448</td>
<td>0.302</td>
</tr>
<tr>
<td></td>
<td>(0.116)</td>
<td>(0.282)</td>
<td>(0.396)</td>
</tr>
<tr>
<td>Smoking Cigarettes is Harmful * Female</td>
<td>0.387*</td>
<td>-0.594</td>
<td>-2.950</td>
</tr>
<tr>
<td></td>
<td>(0.205)</td>
<td>(1.127)</td>
<td>(4.514)</td>
</tr>
<tr>
<td>Other's Cigarette Smoking Is Harmful to You</td>
<td>-0.0735</td>
<td>0.158</td>
<td>0.098</td>
</tr>
<tr>
<td></td>
<td>(0.1)</td>
<td>(0.246)</td>
<td>(0.333)</td>
</tr>
<tr>
<td>Constant</td>
<td>1.135***</td>
<td>0.778*</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>(0.182)</td>
<td>(0.448)</td>
<td></td>
</tr>
<tr>
<td>Cut 1</td>
<td>--</td>
<td>--</td>
<td>0.993</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.699)</td>
</tr>
<tr>
<td>Cut 2</td>
<td>--</td>
<td>--</td>
<td>1.751***</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.716)</td>
</tr>
<tr>
<td>Cut 3</td>
<td>--</td>
<td>--</td>
<td>2.376***</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.740)</td>
</tr>
</tbody>
</table>

*Notes: Clustered and bootstrapped standard errors are reported. Results were bootstrapped 200 times. *** p<0.01, ** p<0.05, * p<0.1. Source: GYTS 2007 and 2011.*
Table 4.3.A.: Average Marginal Effects for Never Smoking

<table>
<thead>
<tr>
<th>Hypothesis</th>
<th>Together</th>
<th>Girls</th>
<th>Boys</th>
<th>Observations</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Hypothesis 1</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Parents Smoke</td>
<td>-0.027**</td>
<td>-0.008</td>
<td>-0.045**</td>
<td>Only impacts boys.</td>
</tr>
<tr>
<td>(0.011)</td>
<td>(0.012)</td>
<td>(0.018)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Friends Smoke</td>
<td>-0.107***</td>
<td>-0.078***</td>
<td>-0.135***</td>
<td></td>
</tr>
<tr>
<td>(0.011)</td>
<td>(0.0107)</td>
<td>(0.013)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Hypothesis 2</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Level of Attraction/ Ease at Parties</td>
<td>0.0001</td>
<td>-0.003</td>
<td>0.003</td>
<td></td>
</tr>
<tr>
<td>(0.004)</td>
<td>(0.0048)</td>
<td>(0.0069)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Popularity</td>
<td>-0.009*</td>
<td>0.007</td>
<td>-0.024***</td>
<td>Popularity matters to boys.</td>
</tr>
<tr>
<td>(0.005)</td>
<td>(0.0057)</td>
<td>(0.0076)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Perception of Adult Smokers</td>
<td>-0.013***</td>
<td>-0.007*</td>
<td>-0.018***</td>
<td></td>
</tr>
<tr>
<td>(0.004)</td>
<td>(0.0041)</td>
<td>(0.0066)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pro-Tobacco Exposure</td>
<td>-0.005</td>
<td>-0.003</td>
<td>-0.007</td>
<td></td>
</tr>
<tr>
<td>(0.004)</td>
<td>(0.0026)</td>
<td>(0.0055)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Hypothesis 3</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low Quality Education</td>
<td>0.003</td>
<td>-0.026</td>
<td>0.030</td>
<td>Education does not impact smoking initiation.</td>
</tr>
<tr>
<td>(0.018)</td>
<td>(0.023)</td>
<td>(0.025)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Middle Quality Education</td>
<td>0.019</td>
<td>0.018</td>
<td>0.020</td>
<td></td>
</tr>
<tr>
<td>(0.017)</td>
<td>(0.0156)</td>
<td>(0.0286)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>High Quality Education</td>
<td>0.024</td>
<td>0.007</td>
<td>0.040</td>
<td></td>
</tr>
<tr>
<td>(0.016)</td>
<td>(0.0167)</td>
<td>(0.0267)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anti-Tobacco Exposure</td>
<td>-0.049***</td>
<td>-0.032***</td>
<td>-0.065***</td>
<td></td>
</tr>
<tr>
<td>(0.013)</td>
<td>(0.009)</td>
<td>(0.018)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Smoking Cigarettes is Harmful</td>
<td>0.026</td>
<td>0.053*</td>
<td>0.001</td>
<td>Girls less likely to smoke if they know it's harmful.</td>
</tr>
<tr>
<td>(0.021)</td>
<td>(0.031)</td>
<td>(0.026)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other's Cigarette Smoking Is Harmful to You</td>
<td>-0.012</td>
<td>-0.008</td>
<td>-0.016</td>
<td></td>
</tr>
<tr>
<td>(0.016)</td>
<td>(0.0102)</td>
<td>(0.0215)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cigarettes Cause Weight Gain</td>
<td>0.039</td>
<td>-0.048</td>
<td>0.123**</td>
<td>Only boys are motivated by weight to not start smoking.</td>
</tr>
<tr>
<td>(0.030)</td>
<td>(0.0327)</td>
<td>(0.051)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cigarettes Cause Weight Loss</td>
<td>0.059**</td>
<td>-0.014</td>
<td>0.129***</td>
<td></td>
</tr>
<tr>
<td>(0.024)</td>
<td>(0.021)</td>
<td>(0.041)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Family Discussed Smoking</td>
<td>-0.008</td>
<td>-0.005</td>
<td>-0.011</td>
<td></td>
</tr>
<tr>
<td>(0.010)</td>
<td>(0.0066)</td>
<td>(0.0139)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gender</td>
<td>Coefficient 1</td>
<td>Coefficient 2</td>
<td>Coefficient 3</td>
<td></td>
</tr>
<tr>
<td>--------</td>
<td>---------------</td>
<td>---------------</td>
<td>---------------</td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>0.072***</td>
<td>0.060***</td>
<td>0.083***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.011)</td>
<td>(0.01)</td>
<td>(0.012)</td>
<td></td>
</tr>
</tbody>
</table>

*Notes*: Clustered and delta-method standard errors are reported. Results were bootstrapped 200 times. *** $p<0.01$, ** $p<0.05$, * $p<0.1$. Source: GYTS 2007 and 2011.
<table>
<thead>
<tr>
<th>Hypothesis 1</th>
<th>Together</th>
<th>Girls</th>
<th>Boys</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parents Smoke</td>
<td>0.005</td>
<td>0.004</td>
<td>0.007</td>
</tr>
<tr>
<td>(0.009)</td>
<td>(0.01)</td>
<td>(0.014)</td>
<td></td>
</tr>
<tr>
<td>Friends Smoke</td>
<td>0.043***</td>
<td>0.042***</td>
<td>0.044***</td>
</tr>
<tr>
<td>(0.010)</td>
<td>(0.011)</td>
<td>(0.011)</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Hypothesis 2</th>
<th>Together</th>
<th>Girls</th>
<th>Boys</th>
</tr>
</thead>
<tbody>
<tr>
<td>Level of Attraction/ Ease at Parties</td>
<td>0.002</td>
<td>0.001</td>
<td>0.003</td>
</tr>
<tr>
<td>(0.003)</td>
<td>(0.004)</td>
<td>(0.005)</td>
<td></td>
</tr>
<tr>
<td>Popularity</td>
<td>0.002</td>
<td>-0.003</td>
<td>0.007</td>
</tr>
<tr>
<td>(0.004)</td>
<td>(0.005)</td>
<td>(0.006)</td>
<td></td>
</tr>
<tr>
<td>Perception of Adult Smokers</td>
<td>0.008</td>
<td>0.003</td>
<td>0.012**</td>
</tr>
<tr>
<td>(0.003)</td>
<td>(0.004)</td>
<td>(0.005)</td>
<td></td>
</tr>
<tr>
<td>Pro-Tobacco Exposure</td>
<td>0.001</td>
<td>0.001</td>
<td>0.000007</td>
</tr>
<tr>
<td>(0.003)</td>
<td>(0.002)</td>
<td>(0.005)</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Hypothesis 3</th>
<th>Together</th>
<th>Girls</th>
<th>Boys</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low Quality Education</td>
<td>0.012</td>
<td>0.039*</td>
<td>-0.014</td>
</tr>
<tr>
<td>(0.015)</td>
<td>(0.021)</td>
<td>(0.020)</td>
<td></td>
</tr>
<tr>
<td>Middle Quality Education</td>
<td>0.005</td>
<td>0.005</td>
<td>0.006</td>
</tr>
<tr>
<td>(0.013)</td>
<td>(0.014)</td>
<td>(0.022)</td>
<td></td>
</tr>
<tr>
<td>High Quality Education</td>
<td>0.002</td>
<td>0.010</td>
<td>-0.006</td>
</tr>
<tr>
<td>(0.013)</td>
<td>(0.015)</td>
<td>(0.020)</td>
<td></td>
</tr>
<tr>
<td>Anti-Tobacco Exposure</td>
<td>0.017</td>
<td>0.017**</td>
<td>0.016</td>
</tr>
<tr>
<td>(0.012)</td>
<td>(0.009)</td>
<td>(0.017)</td>
<td></td>
</tr>
<tr>
<td>Smoking Cigarettes is Harmful</td>
<td>-0.044*</td>
<td>-0.062*</td>
<td>-0.026</td>
</tr>
<tr>
<td>(0.022)</td>
<td>(0.035)</td>
<td>(0.022)</td>
<td></td>
</tr>
<tr>
<td>Other's Cigarette Smoking Is Harmful to You</td>
<td>0.013</td>
<td>0.008</td>
<td>0.018</td>
</tr>
<tr>
<td>(0.013)</td>
<td>(0.008)</td>
<td>(0.018)</td>
<td></td>
</tr>
<tr>
<td>Cigarettes Cause Weight Gain</td>
<td>0.012</td>
<td>0.042</td>
<td>-0.016</td>
</tr>
<tr>
<td>(0.030)</td>
<td>(0.054)</td>
<td>(0.039)</td>
<td></td>
</tr>
<tr>
<td>Cigarettes Cause Weight Loss</td>
<td>-0.010</td>
<td>0.017</td>
<td>-0.036</td>
</tr>
<tr>
<td>(0.027)</td>
<td>(0.047)</td>
<td>(0.027)</td>
<td></td>
</tr>
<tr>
<td>Family Discussed Smoking</td>
<td>0.006</td>
<td>0.004</td>
<td>0.007</td>
</tr>
<tr>
<td>(0.009)</td>
<td>(0.006)</td>
<td>(0.012)</td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>-0.033***</td>
<td>-0.031</td>
<td>-0.034</td>
</tr>
<tr>
<td>--------</td>
<td>-----------</td>
<td>--------</td>
<td>--------</td>
</tr>
<tr>
<td></td>
<td>(0.01)</td>
<td>(0.009)</td>
<td>(0.012)</td>
</tr>
</tbody>
</table>

*Notes: Clustered and delta-method standard errors are reported. Results were bootstrapped 200 times. *** p<0.01, ** p<0.05, * p<0.1. Source: GYTS 2007 and 2011.*
### Table 4.3.C.: Average Marginal Effects of the Conditional Mean of Level of Cigarettes Smoked

<table>
<thead>
<tr>
<th>Hypothesis 1</th>
<th>Together</th>
<th>Girls</th>
<th>Boys</th>
<th>Observations</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Parents Smoke</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;1 cig/day</td>
<td>0.006</td>
<td>0.001</td>
<td>0.011</td>
<td>Parental smoking only impacts boy's level of cigarette consumption.</td>
</tr>
<tr>
<td></td>
<td>(0.001, 0.012)</td>
<td>(-0.005, 0.005)</td>
<td>(0.002, 0.022)</td>
<td></td>
</tr>
<tr>
<td>1 cig/day</td>
<td>0.005</td>
<td>0.001</td>
<td>0.009</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0006, 0.008)</td>
<td>(-0.001, 0.003)</td>
<td>(0.002, 0.016)</td>
<td></td>
</tr>
<tr>
<td>2-5 cigs/day</td>
<td>0.008</td>
<td>0.001</td>
<td>0.014</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.002, 0.015)</td>
<td>(-0.002, 0.006)</td>
<td>(0.003, 0.029)</td>
<td></td>
</tr>
<tr>
<td>6+ cigs/day</td>
<td>0.014</td>
<td>0.004</td>
<td>0.023</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(-0.002, 0.037)</td>
<td>(-0.010, 0.025)</td>
<td>(-0.002, 0.069)</td>
<td></td>
</tr>
<tr>
<td><strong>Friends Smoke</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;1 cig/day</td>
<td>0.021</td>
<td>0.014</td>
<td>0.028</td>
<td>Having close friends who smoke increases all levels of cigarette consumption and the greatest impact is for 6+ cigs/day.</td>
</tr>
<tr>
<td></td>
<td>(0.015, 0.026)</td>
<td>(0.010, 0.021)</td>
<td>(0.020, 0.037)</td>
<td></td>
</tr>
<tr>
<td>1 cig/day</td>
<td>0.013</td>
<td>0.005</td>
<td>0.021</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.009, 0.018)</td>
<td>(0.003, 0.009)</td>
<td>(0.016, 0.029)</td>
<td></td>
</tr>
<tr>
<td>2-5 cigs/day</td>
<td>0.020</td>
<td>0.009</td>
<td>0.032</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.009, 0.032)</td>
<td>(0.004, 0.017)</td>
<td>(0.015, 0.052)</td>
<td></td>
</tr>
<tr>
<td>6+ cigs/day</td>
<td>0.038</td>
<td>0.028</td>
<td>0.048</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.013, 0.071)</td>
<td>(0.005, 0.067)</td>
<td>(0.019, 0.086)</td>
<td></td>
</tr>
<tr>
<td><strong>Hypothesis 2</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Level of Attraction/ Ease at Parties</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;1 cig/day</td>
<td>-0.002</td>
<td>-0.002</td>
<td>-0.003</td>
<td>We see gendered differences in social perception of smoking and level of consumption.</td>
</tr>
<tr>
<td></td>
<td>(-0.004, 0.0004)</td>
<td>(-0.002, 0.0009)</td>
<td>(-0.006, 0.0002)</td>
<td></td>
</tr>
<tr>
<td>1 cig/day</td>
<td>0.0002</td>
<td>0.001</td>
<td>-0.0004</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(-0.001, 0.001)</td>
<td>(0.000009, 0.002)</td>
<td>(-0.003, 0.002)</td>
<td></td>
</tr>
<tr>
<td>2-5 cigs/day</td>
<td>0.001</td>
<td>0.002</td>
<td>0.001</td>
<td>Popularity influences boy's level of consumption, whereas level of attraction/ease at parties and perception of adults impacts girl's consumption.</td>
</tr>
<tr>
<td></td>
<td>(-0.001, 0.004)</td>
<td>(0.00004, 0.004)</td>
<td>(-0.004, 0.005)</td>
<td></td>
</tr>
<tr>
<td>6+ cigs/day</td>
<td>0.005</td>
<td>0.007</td>
<td>0.003</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(-0.002, 0.015)</td>
<td>(-0.002, 0.029)</td>
<td>(-0.011, 0.015)</td>
<td></td>
</tr>
<tr>
<td><strong>Popularity</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;1 cig/day</td>
<td>0.002</td>
<td>-0.001</td>
<td>0.005</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(-0.00009, 0.006)</td>
<td>(-0.003, 0.0009)</td>
<td>(0.001, 0.010)</td>
<td></td>
</tr>
<tr>
<td>1 cig/day</td>
<td>0.001</td>
<td>-0.001</td>
<td>0.003</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(-0.000005, 0.003)</td>
<td>(-0.002, 0.0004)</td>
<td>(0.00002, 0.005)</td>
<td></td>
</tr>
<tr>
<td>2-5 cigs/day</td>
<td>0.002</td>
<td>-0.001</td>
<td>0.005</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(-0.001, 0.005)</td>
<td>(-0.003, 0.0009)</td>
<td>(-0.001, 0.010)</td>
<td></td>
</tr>
<tr>
<td>6+ cigs/day</td>
<td>0.001</td>
<td>-0.004</td>
<td>0.006</td>
<td></td>
</tr>
</tbody>
</table>
### Perception of Adult Smokers

<table>
<thead>
<tr>
<th>Cigarettes per Day</th>
<th>&lt;1 cig/day</th>
<th>1 cig/day</th>
<th>2-5 cigs/day</th>
<th>6+ cigs/day</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>0.002</td>
<td>0.001</td>
<td>0.002</td>
</tr>
<tr>
<td></td>
<td>0.002</td>
<td>0.001</td>
<td>0.001</td>
<td>0.002</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.002</td>
<td>0.001</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.003</td>
<td>0.004</td>
<td>0.003</td>
</tr>
</tbody>
</table>

### Pro-Tobacco Exposure

<table>
<thead>
<tr>
<th>Cigarettes per Day</th>
<th>&lt;1 cig/day</th>
<th>1 cig/day</th>
<th>2-5 cigs/day</th>
<th>6+ cigs/day</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.001</td>
<td>0.001</td>
<td>0.001</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.001</td>
<td>0.001</td>
<td>0.005</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.003</td>
<td>0.001</td>
<td>0.002</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.003</td>
<td>0.001</td>
<td>0.005</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.006</td>
<td>0.002</td>
<td>0.009</td>
</tr>
</tbody>
</table>

Greater exposure to pro-tobacco media increases girl's level of cigarette consumption.

### Hypothesis 3

#### Low Quality Education

<table>
<thead>
<tr>
<th>Cigarettes per Day</th>
<th>&lt;1 cig/day</th>
<th>1 cig/day</th>
<th>2-5 cigs/day</th>
<th>6+ cigs/day</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>-0.004</td>
<td>-0.005</td>
<td>-0.002</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>-0.006</td>
<td>-0.003</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>-0.011</td>
<td>-0.005</td>
<td>-0.011</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-0.028</td>
<td>-0.009</td>
<td>-0.052</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-0.014</td>
<td>-0.008</td>
<td>-0.020</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-0.056</td>
<td>-0.045</td>
<td>-0.088</td>
</tr>
</tbody>
</table>

Low quality education does not impact cigarette consumption. Medium quality education does reduce boy's and girl's likelihood of consuming less than one cig/day. Further, high quality education reduces the odds boys will smoke one or few cigs/day.

#### Middle Quality Education

<table>
<thead>
<tr>
<th>Cigarettes per Day</th>
<th>&lt;1 cig/day</th>
<th>1 cig/day</th>
<th>2-5 cigs/day</th>
<th>6+ cigs/day</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>-0.020</td>
<td>-0.002</td>
<td>-0.004</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>-0.001</td>
<td>-0.001</td>
<td>-0.003</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-0.001</td>
<td>-0.002</td>
<td>-0.000</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-0.010</td>
<td>-0.008</td>
<td>-0.029</td>
</tr>
</tbody>
</table>

#### High Quality Education

<table>
<thead>
<tr>
<th>Cigarettes per Day</th>
<th>&lt;1 cig/day</th>
<th>1 cig/day</th>
<th>2-5 cigs/day</th>
<th>6+ cigs/day</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>-0.007</td>
<td>-0.008</td>
<td>-0.006</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>-0.007</td>
<td>-0.009</td>
<td>-0.012</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-0.010</td>
<td>-0.008</td>
<td>-0.029</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-0.001</td>
<td>-0.009</td>
<td>-0.038</td>
</tr>
<tr>
<td></td>
<td>&lt;1 cig/day</td>
<td>1 cig/day</td>
<td>2-5 cigs/day</td>
<td>6+ cigs/day</td>
</tr>
<tr>
<td>------------------------</td>
<td>------------</td>
<td>-----------</td>
<td>--------------</td>
<td>-------------</td>
</tr>
<tr>
<td><strong>Anti-Tobacco Exposure</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;1 cig/day</td>
<td>0.010</td>
<td>0.005</td>
<td>0.014</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.005, 0.018)</td>
<td>(0.003, 0.009)</td>
<td>(0.006, 0.027)</td>
<td></td>
</tr>
<tr>
<td>1 cig/day</td>
<td>0.006</td>
<td>0.002</td>
<td>0.011</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.003, 0.012)</td>
<td>(0.0008, 0.004)</td>
<td>(0.002, 0.020)</td>
<td></td>
</tr>
<tr>
<td>2-5 cigs/day</td>
<td>0.010</td>
<td>0.003</td>
<td>0.016</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(-0.0008, 0.0257)</td>
<td>(0.001, 0.007)</td>
<td>(-0.003, 0.046)</td>
<td></td>
</tr>
<tr>
<td>6+ cigs/day</td>
<td>0.018</td>
<td>0.010</td>
<td>0.025</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(-0.005, 0.047)</td>
<td>(0.002, 0.033)</td>
<td>(-0.014, 0.079)</td>
<td></td>
</tr>
<tr>
<td><strong>Smoking Cigarettes is Harmful</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;1 cig/day</td>
<td>0.006</td>
<td>0.006</td>
<td>0.005</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(-0.004, 0.012)</td>
<td>(0.003, 0.010)</td>
<td>(-0.014, 0.018)</td>
<td></td>
</tr>
<tr>
<td>1 cig/day</td>
<td>0.004</td>
<td>0.001</td>
<td>0.008</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0007, 0.008)</td>
<td>(-0.003, 0.003)</td>
<td>(0.001, 0.015)</td>
<td></td>
</tr>
<tr>
<td>2-5 cigs/day</td>
<td>0.006</td>
<td>-0.002</td>
<td>0.014</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(-0.0003, 0.017)</td>
<td>(-0.009, 0.003)</td>
<td>(0.002, 0.032)</td>
<td></td>
</tr>
<tr>
<td>6+ cigs/day</td>
<td>-0.008</td>
<td>-0.040</td>
<td>0.023</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(-0.073, 0.020)</td>
<td>(-0.253, 0.002)</td>
<td>(-0.005, 0.048)</td>
<td></td>
</tr>
<tr>
<td><strong>Other's Cigarette Smoking Is Harmful to You</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;1 cig/day</td>
<td>-0.001</td>
<td>-0.0003</td>
<td>-0.002</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(-0.009, 0.005)</td>
<td>(-0.004, 0.002)</td>
<td>(-0.016, 0.009)</td>
<td></td>
</tr>
<tr>
<td>1 cig/day</td>
<td>0.0004</td>
<td>0.0001</td>
<td>0.001</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(-0.007, 0.004)</td>
<td>(-0.001, 0.0009)</td>
<td>(-0.012, 0.007)</td>
<td></td>
</tr>
<tr>
<td>2-5 cigs/day</td>
<td>0.001</td>
<td>0.0002</td>
<td>0.002</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(-0.008, 0.011)</td>
<td>(-0.002, 0.002)</td>
<td>(-0.014, 0.021)</td>
<td></td>
</tr>
<tr>
<td>6+ cigs/day</td>
<td>0.003</td>
<td>0.0003</td>
<td>0.006</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(-0.026, 0.021)</td>
<td>(-0.011, 0.010)</td>
<td>(-0.041, 0.037)</td>
<td></td>
</tr>
<tr>
<td><strong>Cigarettes Cause Weight Gain</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;1 cig/day</td>
<td>-0.015</td>
<td>0.002</td>
<td>-0.032</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(-0.031, -0.001)</td>
<td>(-0.009, 0.013)</td>
<td>(-0.062, -0.006)</td>
<td></td>
</tr>
<tr>
<td>1 cig/day</td>
<td>-0.011</td>
<td>0.001</td>
<td>-0.023</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(-0.024, -0.004)</td>
<td>(-0.003, 0.005)</td>
<td>(-0.049, -0.009)</td>
<td></td>
</tr>
<tr>
<td>2-5 cigs/day</td>
<td>-0.018</td>
<td>0.001</td>
<td>-0.037</td>
<td></td>
</tr>
</tbody>
</table>

*This increases level of cigarette consumption of all levels for girls, but the magnitude is larger for boys at lower levels of consumption.*
and medium levels of cigarettes.

<table>
<thead>
<tr>
<th>Cigarettes Cause Weight Loss</th>
<th>6+ cigs/day</th>
<th>&lt;1 cig/day</th>
<th>1 cig/day</th>
<th>2-5 cigs/day</th>
<th>6+ cigs/day</th>
</tr>
</thead>
<tbody>
<tr>
<td>6+ cigs/day</td>
<td>-0.027</td>
<td>0.002</td>
<td>-0.056</td>
<td>(-0.131, 0.009)</td>
<td>(-0.039, 0.038)</td>
</tr>
<tr>
<td>&lt;1 cig/day</td>
<td>-0.013</td>
<td>-0.001</td>
<td>-0.026</td>
<td>(-0.028, -0.003)</td>
<td>(-0.009, 0.005)</td>
</tr>
<tr>
<td>1 cig/day</td>
<td>-0.011</td>
<td>0.000</td>
<td>-0.022</td>
<td>(-0.023, -0.003)</td>
<td>(-0.005, 0.002)</td>
</tr>
<tr>
<td>2-5 cigs/day</td>
<td>-0.018</td>
<td>-0.001</td>
<td>-0.036</td>
<td>(-0.051, -0.004)</td>
<td>(-0.011, 0.002)</td>
</tr>
<tr>
<td>6+ cigs/day</td>
<td>-0.032</td>
<td>-0.005</td>
<td>-0.058</td>
<td>(-0.130, 0.003)</td>
<td>(-0.039, 0.014)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Family Discuss Harm of Cigarettes</th>
<th>6+ cigs/day</th>
<th>&lt;1 cig/day</th>
<th>1 cig/day</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;1 cig/day</td>
<td>-0.001</td>
<td>0.0002</td>
<td>-0.0004</td>
</tr>
<tr>
<td>1 cig/day</td>
<td>0.001</td>
<td>0.0003</td>
<td>0.002</td>
</tr>
<tr>
<td>2-5 cigs/day</td>
<td>0.003</td>
<td>0.001</td>
<td>0.005</td>
</tr>
<tr>
<td>6+ cigs/day</td>
<td>0.006</td>
<td>0.002</td>
<td>0.009</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Female</th>
<th>6+ cigs/day</th>
<th>&lt;1 cig/day</th>
<th>1 cig/day</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;1 cig/day</td>
<td>-0.020</td>
<td>-0.018</td>
<td>-0.022</td>
</tr>
<tr>
<td>1 cig/day</td>
<td>-0.005</td>
<td>0.001</td>
<td>-0.011</td>
</tr>
<tr>
<td>2-5 cigs/day</td>
<td>-0.002</td>
<td>0.001</td>
<td>-0.005</td>
</tr>
<tr>
<td>6+ cigs/day</td>
<td>0.039</td>
<td>0.006</td>
<td>0.071</td>
</tr>
</tbody>
</table>

Notes: Standard errors were clustered. The 95% bias-corrected confidence interval is reported in parentheses. Results were bootstrapped 200 times. Source: GYTS 2007 and 2011.
Chapter 5: Conclusion

5.1. Dissertation Summary

The rates of NCDs are growing worldwide and both developing and developed nations are burdened by these diseases. The most common NCDs are cardiovascular disease, cancer, chronic respiratory disease, and diabetes. The causes of the most common NCDs all share the same primary risk factors. Further, these risk factors are behavioral in nature including harmful alcohol consumption, a lack of physical activity, unhealthy diets, and tobacco use (Yach et al. 2004). One of these behaviors, tobacco use, accounts for 1 in 6 NCD related deaths (Beaglehole et al. 2011). As such, the spread of these diseases can be reduced.

Guided by the health economics literature, we investigate the ways traditional economic tools, such as taxation, information, advertising, and education can be used to mitigate risky health behaviors. We also consider that the reasons and motivations behind youth risky behaviors could look different than adult risky health behaviors (Gruber 2001).

We focus in on two of the behavioral causes of NCDs in this dissertation: unhealthy diets and tobacco use. Within the United States, the rate of soft drink consumption has increased by 71% among adults 19-29 years old between 1977 and 1996 (Kim and Kawachi 2006). Soft drink consumption falls under the larger umbrella of SSB consumption. These SSBs are one part of a unhealthy diet which have been shown to be positively correlated with weight gain in adults and children (Brownell et al. 2011).
Obesity itself is associated with NCDs such as diabetes, certain cancers, and cardiovascular disease (Kopelman 2000).

The other risky health behavior we focus on is tobacco smoking by youths in a developing nation. Around 80% of smokers worldwide now reside in developing nations (WHO n.d.). Coupled with most smokers starting smoking before the age of 18 (USDHHS 2012), it is vital that we find ways to mitigate tobacco use of youth in developing nations. Coordinated global efforts are currently underway to reduce the global tobacco epidemic. The WHO FCTC outlines ways to reduce supply and demand of tobacco products and the FCTC has been ratified by 181 parties now covering 90% of the world’s population (WHO 2018). In 2011, Nepal passed the Tobacco Control and Regulatory Bill which is based on the recommendations in the FCTC. However, since its passage, the rates of youth smoking have increased by 2 percentage points from 2012 to 2016 (WHO 2013 and 2017).

In Chapter 2, we used a contingent valuation framework, frequently used in environmental economics, and applied it to a novel problem within health economics: calculating the willingness to pay for a SSB tax in New Mexico. With primary data, we focused on the roles that eating habits; knowledge and awareness about food and related policies; and attitude have on both directly and indirectly increasing preferences for SSB taxes. To address issues of protest responses and endogeneity, we formulated a three-equation system that was simultaneously estimated. From this three-equation system we calculated the individual median willingness to pay following Cameron (1988) using bootstrapping. Overall we found that healthier eating habits are correlated with statistically significantly increased preferences for SSB taxes. Knowing that a poor diet
can cause someone to be obese or overweight also statistically significantly increased preferences for taxes. Our estimated individual median willingness to pay value was less than one-penny-per-ounce, which helps explain why a recent two-penny-per-ounce tax in Santa Fe, New Mexico failed. This study recommended that policy makers implement public health campaigns around the dangers of SSB consumption and about the magnitude of the obesity epidemic in their locality before proposing a SSB tax. Additionally, this study showed that going with the status quo tax value of one-penny-per-ounce does not guarantee success of passage.

The next chapter, Chapter 3, addressed a related but more fundamental question: why someone would support a SSB tax in the first place. We examined the level of support for expanding the Healthy Diné Nation Act of 2014, which in part is a 2% tax on SSBs, to all of New Mexico. With the same primary data and a partial proportional odds model we assessed the impact media coverage; knowledge, attitude and behavior; and political ideology have on the respondents level of support for expansion. Our findings showed that respondents who knew about the passage of the Healthy Diné Nation Act were 13.6 percentage points less likely to be strongly against the expansion and this increased the likelihood of strongly supporting the expansion by 11.5 percentage points. However, previous exposure to non-local SSB taxes was associated with an increase of being strongly against expansion by 11.0 percentage points. These results demonstrated that it is important for policy makers to distance their newly proposed tax from existing SSB taxes. Further we also found that increasing awareness that SSB consumption can cause obesity and that obesity is a major problem in their state are win-win strategies for policy makers, for not only are these factors correlated with an increase in strongly
support expansion of the Healthy Diné Nation Act, but it also reduces the likelihood of being strongly against expansion.

The finally analytical chapter, Chapter 4, transitioned from the United States to Nepal and from SSB consumption to youth tobacco use. We used two waves of nationally representative data to study the impact that proximity to other smokers; anti-tobacco education and knowledge; and pro-tobacco exposure and social perceptions have on the probability that youth will belong to one of three smoking statuses: never smoker, former smoker, and level of consumption of cigarettes by current smokers. This study expanded upon previous youth studies in Nepal by examining more than one smoking status in the same paper and by moving beyond using a logistic regression. We followed the work by Kasteridis, Munkin, and Yen (2010) in estimating a trivariate ordered probit model with independence. To capture the differences our hypothesized factors have on boys and girls, guided by preliminary analyses, we included interactions between independent variables of interest and the variable female. In order to determine the full impact that the covariates had on gender across the smoking status, we calculated the average marginal effects for never smokers and former smokers along with the average marginal effect of the conditional mean of level of cigarette consumption outline in Kasteridis, Munkin, and Yen (2010). Our results provided several points policy makers can leverage to reduce youth smoking. First, anti-tobacco media needs to be controlled by the government and needs to be thoroughly redesigned. As our results show youth who were exposed to anti-tobacco media were not only statistically significantly more likely to begin smoking but it also increased the likelihood of engaging in low and medium levels of cigarette consumption. These anti-tobacco media messages need to counter the
idea that smoking can make someone popular and that adult smokers are cool. Further, we find that boys who think that smoking will lead to deviations from their current weight are more likely to be a never smoker and less likely to smoke low and medium levels of cigarettes. This is another media message that can be targeted specially to boys. Formal anti-tobacco education in schools needs to be revamped. Currently it does nothing to prevent students from initiating in smoking behaviors, although it can help reduce the likelihood of smoking low and medium levels of cigarettes, but only for the highest quality education. Finally, strong enforcement of restriction on the sale of tobacco to underage populations is needed to help counteract the impact having close friends who smoke have on smoking initiation and levels of cigarettes smoked.

5.2. Avenues for Future Research

The SSB analysis could be expanded in several ways. First, we could implement it in different locations across the US to determine the variation in the WTP for a SSB tax and study if the covariates vary across locations or remain the same. This study would help develop more generalizable strategies to increase preferences for SSB taxes. Additionally, another study could be conducted to identify if our finding that perceptions of local and non-local SSB taxes had different impacts on support for expanding SSB taxes is also found in other states. To directly test the impact of targeted messages on SSB consumption, we could create a randomized control trial implementing a public health campaign aimed at reducing SSB consumption assessing the differential impacts of campaigns focus on a) obesity being a major problem in the state, b) the dangers of SSB consumption and c) the status quo control group.
There are several ways that the youth smoking in Nepal chapter could be extended. First and foremost, non-cigarette smoking behavior should be conducted. Although the GYTS data does not allow for a similar modeling procedure with non-cigarettes, a modified approach could be taken to see what factors influence youth’s smokeless and non-cigarette smokable tobacco use by gender. Another important venue of exploration would be to dig deeper into the anti-tobacco advertisements in Nepal. If it is indeed the case that there are anti-tobacco messages from both public health officials and the tobacco industry, we could then collect data on youths’ exposure to these differing messages and assess the different impacts they have on smoking behaviors. With sufficient foresight, we could collect longitudinal data to examine the impact of banning tobacco companies from promoting anti-tobacco media would have on youth smoking behaviors. Finally, we would also explore more into what factors are associated with youth smoking cessation behavior in Nepal.
Appendices

Appendix A: Principal Component Analysis .................................................. 139
Appendix B: Supplemental Tables ................................................................. 149
Appendix C: New Landscapes of a Majority-Minority State ............................. 155
Appendix D: Stata Code ........................................................................... 187
Appendix A: Principal Component Analysis

Often time in survey data, there are multiple questions that are measuring the same concept. Including all these as variables into an empirical model would not only reduce efficiency but would also result in issues of multicollinearity. This could also lead to masking statistical significance by spreading out the underlying concept across different measures. One way to overcome these concerns is by using principal component analysis (PCA). The basic idea of which is to distill these various measures into the minimum number of components with the maximum amount of variation/information. This is a form of data reduction to express multivariate data with fewer dimensions.

Overview:
Mathematically, PCA is finding a way maximize the variance of the elements in $z = xu$ where $u'u = 1$. Where $z$ is a vector of the components that result in the linear combination of $x$, the vector of the original variables and $u$, the unexplained variation. PCA is trying to find a $z$ such that it captures as much of the underlying variation in the original variables as possible. The first component will capture the most possible variation; the second component will then capture the most remaining variation, so on and so forth. Ideally, we seek a situation where the components are uncorrelated. In order to find a solution to the PCA, we perform an eigenvalue decomposition of the correlation matrix to find the principal axes of the ensuing scatterplot. The eigenvalue itself represents the direction of one of the principal axes. This is done by solving the equation $(R - \lambda I)u = 0$. Where $R$ is the correlation matrix of the original variables, $\lambda$ are the
eigenvalues that are associated with the variance of the components of \( z \), \( I \) is the identity matrix, and \( u \) is the eigenvector. Then the factor loadings are the correlation between the components and the original variables written as \( F = corr(z, x) = uD^{0.5} \), where \( D \) is the diagonal covariance matrix of the components. For each of the original \( x \) variables, the proportion of the variance explained by the first \( n \) components is given by the sum of the square of the factor loadings. If all of the components are retained, then all variation will be explained and hence the sum of the square of the factor loadings will be unity.

Although using all of these components will account for all of the variation in these variables, this is a method of data reduction, as so keeping all the components would go against this principal. There is a balance between parsimony (keeping a small number of components) and thoroughness (capturing more of the variation). There are several rules to help guide the factor retention process. The first is Kaiser’s rule. This states that only components whose eigenvalue exceeds unity should be kept. The logic behind this rule is that components should be kept only if it gives more variation than the original variables. Another way to determine how many components should be kept is by creating a scree plot and looking for an elbow in the values that exceed unity. A scree plot places of the eigenvalues of the components on the y-axis and the number of the eigenvalue on the x-axis. If there is a bend, or an elbow, in the eigenvalues above one then those on the right-hand side of the bend are not explaining much variation and we might consider excluding these from inclusion in the model even though they exceed one.
When conducting PCA, we want to ensure that the minimum number of components captures the maximum amount of variation. Furthermore, this concept extends to the factor loadings themselves. Not only does this practically help with naming the component, but it allows for a conceptual understanding of each component. For example, say originally there were 15 variables and the Kaiser rule showed that only three components should be kept. If the first seven variables had high factor loadings on all three components then it would be difficult to explain how each component was distinct. A better situation would be that certain variables that are highly correlated would cluster together into different components. As such, it is often the case that the factor loadings matrix is rotated in order to increase the odds that the above condition comes to fruition. There are two main types of factor rotation. The first is orthogonal rotation that does not allow correlation of the components by maintaining the perpendicularity of the axes. Two common types of orthogonal rotation are varimax rotation and quartimax rotation. The other type of rotation is oblique rotation: this allows for the correlation between the rotated factors. One of the most common types of oblique rotation is promax rotation. Which rotation method is used depends on the discipline. After deciding upon the desired factor rotation, it will be time to name the component. Components are generally named after the set of variables that they are most correlated to.

There are certain conditions in which make PCA more applicable to use. There needs to be sufficient correlation between the original variables. Importantly, the original scale of the variable is critical since PCA is scale sensitive. Before doing PCA, the scales should be standardized across variables. Additionally, because of the scale issue, it is preferred
to work with the correlation matrix instead of the covariance matrix. There are empirical
tests to determined the appropriateness of conducting PCA. The main test is the Kaiser-
Meyer-Olking measure of sampling adequacy. This is a scale between zero and one. The
higher the number the more there is in common amongst the variables, and the more
warned PCA is to conduct. Generally a KMO score above 0.5 is acceptable, however,
there are discipline specific thresholds of adequacy which are higher than 0.5.

Social Perception Principal Component Analysis: An Example:

In the Nepal Global Youth Tobacco Survey (GYTS) 2007 and 2011 contain seven
questions on the social perception of cigarette use. They are as follows:

- [SP_1] Do you think boys who smoke cigarettes have more or less friends? *More friends, less friends, no difference from non-smokers*
- [SP_2] Do you think girls who smoke cigarettes have more or less friends? *More friends, less friends, no difference from non-smokers*
- [SP_3] Do you think smoking cigarettes help people feel more or less comfortable at celebrations, parties, or in social gatherings? *More comfortable, less comfortable, no difference from non-smokers*
- [SP_4] Do you think smoking cigarettes makes boys look more or less attractive? *More attractive, less attractive, no difference from non-smokers*
- [SP_5] Do you think smoking cigarettes makes girls look more or less attractive? *More attractive, less attractive, no difference from non-smokers*
• [SP_6] When you see a man smoking, what do you think of him? *Lacks confidence, stupid, loser, successful, intelligent, macho*

• [SP_7] When you see a woman smoking, what do you think of him? *Lacks confidence, stupid, loser, successful, intelligent, sophisticated*

The first step in conducting the PCA is to standardize the data. The first five questions have the same scale. They were then coded to give negative feelings towards smoking the value of -1, no different was coded as 0, and positive feelings towards smoking was coded as 1. The last two variables were dichotomized into positive and negative feelings towards adult smokers. The positive feelings were coded as 1 and the negative feelings were coded as -1, keeping with the coding scheme of the previous variables.

After the standardization of the variables, we then examine the correlation matrix of the social perception standardized variables.

Table A.1.: Correlation Matrix Between Social Perception Variables

<table>
<thead>
<tr>
<th></th>
<th>SP_1</th>
<th>SP_2</th>
<th>SP_3</th>
<th>SP_4</th>
<th>SP_5</th>
<th>SP_6</th>
<th>SP_7</th>
</tr>
</thead>
<tbody>
<tr>
<td>SP_1</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SP_2</td>
<td>0.435</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SP_3</td>
<td>0.190</td>
<td>0.143</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SP_4</td>
<td>0.261</td>
<td>0.172</td>
<td>0.222</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SP_5</td>
<td>0.141</td>
<td>0.272</td>
<td>0.182</td>
<td>0.467</td>
<td>1.000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SP_6</td>
<td>0.050</td>
<td>0.031</td>
<td>0.023</td>
<td>0.029</td>
<td>0.063</td>
<td>1.000</td>
<td></td>
</tr>
<tr>
<td>SP_7</td>
<td>0.006</td>
<td>0.005</td>
<td>-0.008</td>
<td>-0.022</td>
<td>0.009</td>
<td>0.256</td>
<td>1.000</td>
</tr>
</tbody>
</table>

The correlation between the variables ranges from a high of 0.4252 to a low of -0.0224. Overall there is some evidence of correlation between the variables which can be
interpreted of proceeding with the PCA with caution. The final appropriateness of this PCA will be tested at the end with the KMO test.

The next step is to perform the PCA which we will start with an unrotated factor loadings matrix.

Table A.2.: Social Perception Principal Components/Correlation, Unrotated

<table>
<thead>
<tr>
<th>Component</th>
<th>Eigenvalue</th>
<th>Difference</th>
<th>Proportion</th>
<th>Cumulative</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2.0185</td>
<td>0.7655</td>
<td>0.2884</td>
<td>0.2884</td>
</tr>
<tr>
<td>2</td>
<td>1.2529</td>
<td>0.2206</td>
<td>0.1790</td>
<td>0.4673</td>
</tr>
<tr>
<td>3</td>
<td>1.0324</td>
<td>0.1621</td>
<td>0.1475</td>
<td>0.6148</td>
</tr>
<tr>
<td>4</td>
<td>0.8702</td>
<td>0.1296</td>
<td>0.1243</td>
<td>0.7391</td>
</tr>
<tr>
<td>5</td>
<td>0.7407</td>
<td>0.0918</td>
<td>0.1058</td>
<td>0.8450</td>
</tr>
<tr>
<td>6</td>
<td>0.6488</td>
<td>0.2124</td>
<td>0.0927</td>
<td>0.9376</td>
</tr>
<tr>
<td>7</td>
<td>0.4365</td>
<td></td>
<td>0.0624</td>
<td>1.0000</td>
</tr>
</tbody>
</table>

The far-left column list the components from the PCA, there are a total of seven, which equals the number of original variables. The first component has an eigenvalue over 2 and this component alone explains 28.84% of the variation. The first three components of the PCA have an eigenvalue exceeding unity and by Kaiser’s rule, we should keep these three components. Further, these three components explain a total of 61.48% of the variation (this number is found in the rightmost column). To check for potential bends in the eigenvalues, we then examine the Scree plot.
The Scree plot shows that there is a bend in the eigenvalues after the second component. This indicates that we can proceed with using two or three components and moving forward the first three components will be used.

With this information, we then present the factor loadings for the first three components of the PCA. To aid in naming the components and to clean up the matrix, factor loadings less than the absolute value of 0.3 were suppressed in the matrix.
Table A.3.: Principal Components (Eigenvectors) Unrotated

<table>
<thead>
<tr>
<th>Variable</th>
<th>Comp1</th>
<th>Comp2</th>
<th>Comp3</th>
<th>Unexplained</th>
</tr>
</thead>
<tbody>
<tr>
<td>SP_1</td>
<td>0.453</td>
<td>-0.5496</td>
<td>0.2738</td>
<td></td>
</tr>
<tr>
<td>SP_2</td>
<td>0.4539</td>
<td>-0.5197</td>
<td>0.3052</td>
<td></td>
</tr>
<tr>
<td>SP_3</td>
<td>0.3413</td>
<td></td>
<td></td>
<td>0.7417</td>
</tr>
<tr>
<td>SP_4</td>
<td>0.4873</td>
<td>0.4404</td>
<td></td>
<td>0.3111</td>
</tr>
<tr>
<td>SP_5</td>
<td>0.4747</td>
<td>0.4613</td>
<td></td>
<td>0.325</td>
</tr>
<tr>
<td>SP_6</td>
<td></td>
<td>0.695</td>
<td></td>
<td>0.3738</td>
</tr>
<tr>
<td>SP_7</td>
<td></td>
<td>0.711</td>
<td></td>
<td>0.3657</td>
</tr>
</tbody>
</table>

Note: Blanks are abs(loading)<.3.

With the unrotated factor loadings matrix, several of the variables load onto multiple factors. For instance SP_1, SP_2, SP_3, and SP_4 load onto both Component 1 and Component 3. This makes the interpretation of the components convoluted. As such, we will need to rotate the factor loadings matrix. Before doing so, the final column in the table shows how much of the variation of each variable is unexplained. The first variable’s variation is mostly explained, with 0.2738 unexplained. The reason that there is unexplained variation is because we are only using the first three components, if we were to use all seven components, then all of the variation would have been explained. For the most part, the majority of the variation of the variables is captured in the three components with the exception of SP_3.

To address the issue of the multiple loadings from a single variable, we use a varimax rotation on the factor loadings matrix. Note that a promax rotation was also performed and yielded nearly identical results. The following table shows the factor loadings with the varimax rotation while still maintaining only the first three components. Once again factor loadings less than the absolute value of 0.3 were blanked out.
Table A.4.: Principal Component (Eigenvectors), Orthogonal Varimax Rotation

<table>
<thead>
<tr>
<th>Variable</th>
<th>Comp1 (Level of Attraction and Ease at Parties)</th>
<th>Comp2 (Popularity)</th>
<th>Comp3 (Perception of Adult Smokers)</th>
<th>Unexplained</th>
</tr>
</thead>
<tbody>
<tr>
<td>SP_1</td>
<td>0.7120</td>
<td>0.6901</td>
<td>0.7120</td>
<td>0.2738</td>
</tr>
<tr>
<td>SP_2</td>
<td>0.6901</td>
<td>0.3052</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SP_3</td>
<td>0.3494</td>
<td></td>
<td>0.7417</td>
<td></td>
</tr>
<tr>
<td>SP_4</td>
<td>0.6618</td>
<td></td>
<td>0.3111</td>
<td></td>
</tr>
<tr>
<td>SP_5</td>
<td>0.6603</td>
<td></td>
<td>0.3250</td>
<td></td>
</tr>
<tr>
<td>SP_6</td>
<td></td>
<td>0.7019</td>
<td>0.3738</td>
<td></td>
</tr>
<tr>
<td>SP_7</td>
<td></td>
<td>0.7102</td>
<td>0.3657</td>
<td></td>
</tr>
</tbody>
</table>

Note: Blanks are abs(loading)<.3.

With the varimax rotation each variable only loads onto one of the three components. The groups of similar variables have clustered onto a component. This allows us to easily name and conceptually understand the components. Take component 2 for example. Two variables load onto this component and both are regarding the perception of friendship for smokers compared to non-smokers. Together this component is measuring perception of popularity. By examining what variables are loaded into each component, we can name the components easily since there is a clear clustering of variables.

Finally to test the appropriateness of using PCA on the social perception variables we run the Kaiser Meyer Olking Test. This yields a value of 0.5911, which is over the threshold of 0.5, which indicates PCA is indeed appropriate to use.

Altogether, with PCA we were able to reduce the number of variables needed to capture social perception from seven to three. Even more than just reducing the number of variables, these three components captured nearly 70% of the original underlying
variation in the social perception variables. As such we were able to keep a majority of
the variation without compromising model efficiency or having issues of
multicollinearity.
## Appendix B: Supplemental Tables

### Table B.4.4.: Multinomial Logit Results for Current Smokers of Both Sex Combined and Boys and Girls Separately

<table>
<thead>
<tr>
<th></th>
<th>M1</th>
<th>M2</th>
<th>M3</th>
<th>M4</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Together</td>
<td>Boys</td>
<td>Girls</td>
<td>Together</td>
</tr>
<tr>
<td><strong>Proximity to Other Smokers</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Friends Smoke</td>
<td>2.426***</td>
<td>2.029***</td>
<td>2.582***</td>
<td>2.250***</td>
</tr>
<tr>
<td></td>
<td>(0.233)</td>
<td>(0.279)</td>
<td>(0.448)</td>
<td>(0.224)</td>
</tr>
<tr>
<td>Parents Smoke</td>
<td>0.409**</td>
<td>0.414**</td>
<td>0.186</td>
<td>0.482***</td>
</tr>
<tr>
<td></td>
<td>(0.168)</td>
<td>(0.201)</td>
<td>(0.399)</td>
<td>(0.169)</td>
</tr>
<tr>
<td><strong>Pro-Tobacco Marketing &amp; Social Perception</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Social Perception PCA</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Level of Attraction/ Ease at Parties</td>
<td>-0.0856</td>
<td>-0.0753</td>
<td>-0.0332</td>
<td>-0.0806</td>
</tr>
<tr>
<td></td>
<td>(0.0664)</td>
<td>(0.0906)</td>
<td>(0.137)</td>
<td>(0.0731)</td>
</tr>
<tr>
<td>Popularity</td>
<td>0.202***</td>
<td>0.281***</td>
<td>-0.219</td>
<td>0.192***</td>
</tr>
<tr>
<td></td>
<td>(0.0599)</td>
<td>(0.0721)</td>
<td>(0.167)</td>
<td>(0.0612)</td>
</tr>
<tr>
<td>Perception of Adult Smokers</td>
<td>0.192***</td>
<td>0.167**</td>
<td>0.250**</td>
<td>0.139**</td>
</tr>
<tr>
<td></td>
<td>(0.0584)</td>
<td>(0.0668)</td>
<td>(0.106)</td>
<td>(0.0664)</td>
</tr>
<tr>
<td><strong>Media Exposure</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pro-tobacco Exposure</td>
<td>0.215***</td>
<td>0.178**</td>
<td>0.352**</td>
<td>0.149**</td>
</tr>
<tr>
<td></td>
<td>(0.0742)</td>
<td>(0.0740)</td>
<td>(0.146)</td>
<td>(0.0731)</td>
</tr>
<tr>
<td><strong>Anti-Tobacco Awareness &amp; Knowledge</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Quality of Smoking Education</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low</td>
<td>-0.265</td>
<td>-0.296</td>
<td>-0.206</td>
<td>-0.0106</td>
</tr>
<tr>
<td></td>
<td>(0.226)</td>
<td>(0.269)</td>
<td>(0.574)</td>
<td>(0.226)</td>
</tr>
<tr>
<td>Medium</td>
<td>-0.663**</td>
<td>-0.481</td>
<td>-1.362**</td>
<td>-0.465</td>
</tr>
<tr>
<td></td>
<td>(0.300)</td>
<td>(0.342)</td>
<td>(0.648)</td>
<td>(0.299)</td>
</tr>
<tr>
<td></td>
<td>Coefficient 1</td>
<td>Coefficient 2</td>
<td>Coefficient 3</td>
<td>Coefficient 4</td>
</tr>
<tr>
<td>------------------------</td>
<td>---------------</td>
<td>---------------</td>
<td>---------------</td>
<td>---------------</td>
</tr>
<tr>
<td>High</td>
<td>-0.683***</td>
<td>-0.548**</td>
<td>-0.991**</td>
<td>-0.559**</td>
</tr>
<tr>
<td></td>
<td>(0.210)</td>
<td>(0.246)</td>
<td>(0.479)</td>
<td>(0.224)</td>
</tr>
<tr>
<td>Anti Tobacco Media Exposure</td>
<td>0.612***</td>
<td>0.531**</td>
<td>1.225***</td>
<td>0.677***</td>
</tr>
<tr>
<td></td>
<td>(0.203)</td>
<td>(0.231)</td>
<td>(0.453)</td>
<td>(0.206)</td>
</tr>
<tr>
<td>Parents Discuss Smoking Dangers</td>
<td>0.217</td>
<td>0.206</td>
<td>0.273</td>
<td>0.170</td>
</tr>
<tr>
<td></td>
<td>(0.159)</td>
<td>(0.194)</td>
<td>(0.321)</td>
<td>(0.168)</td>
</tr>
<tr>
<td><strong>Cigarettes and Weight</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gain Weight</td>
<td>-1.165**</td>
<td>-1.253**</td>
<td>-0.289</td>
<td>-1.081**</td>
</tr>
<tr>
<td></td>
<td>(0.478)</td>
<td>(0.565)</td>
<td>(0.856)</td>
<td>(0.494)</td>
</tr>
<tr>
<td>Lose Weight</td>
<td>-1.152***</td>
<td>-1.173***</td>
<td>-0.348</td>
<td>-1.012***</td>
</tr>
<tr>
<td></td>
<td>(0.314)</td>
<td>(0.386)</td>
<td>(0.701)</td>
<td>(0.326)</td>
</tr>
<tr>
<td>Smoking Cigarettes is Harmful</td>
<td>0.468</td>
<td>0.445</td>
<td>0.538</td>
<td>0.320</td>
</tr>
<tr>
<td></td>
<td>(0.345)</td>
<td>(0.370)</td>
<td>(0.589)</td>
<td>(0.349)</td>
</tr>
<tr>
<td>Other's Smoking Harmful</td>
<td>-0.0928</td>
<td>0.0211</td>
<td>-0.620</td>
<td>-0.286</td>
</tr>
<tr>
<td></td>
<td>(0.217)</td>
<td>(0.262)</td>
<td>(0.513)</td>
<td>(0.258)</td>
</tr>
<tr>
<td><strong>Controls</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Age</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14-15 yo</td>
<td>0.504</td>
<td>0.772</td>
<td>-0.409</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.379)</td>
<td>(0.498)</td>
<td>(0.524)</td>
<td></td>
</tr>
<tr>
<td>16 + yo</td>
<td>1.237***</td>
<td>1.458***</td>
<td>0.542</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.391)</td>
<td>(0.534)</td>
<td>(0.592)</td>
<td></td>
</tr>
<tr>
<td><strong>Grade</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Eighth</td>
<td>1.115***</td>
<td>0.815*</td>
<td>3.133***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.426)</td>
<td>(0.484)</td>
<td>(0.886)</td>
<td></td>
</tr>
<tr>
<td>Ninth</td>
<td>0.842*</td>
<td>0.602</td>
<td>2.736***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.440)</td>
<td>(0.463)</td>
<td>(0.924)</td>
<td></td>
</tr>
<tr>
<td>Variable</td>
<td>Coefficient</td>
<td>Standard Error</td>
<td>p-Value</td>
<td></td>
</tr>
<tr>
<td>------------</td>
<td>--------------</td>
<td>----------------</td>
<td>---------</td>
<td></td>
</tr>
<tr>
<td>Tenth</td>
<td>0.907**</td>
<td>0.454</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.775</td>
<td>0.492</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2.296**</td>
<td>1.089</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Year 2011</td>
<td>-0.669***</td>
<td>0.245</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-0.729**</td>
<td>0.285</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-0.578</td>
<td>0.475</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>-1.029***</td>
<td>0.272</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>-4.760***</td>
<td>0.224</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-4.111***</td>
<td>0.272</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-5.523***</td>
<td>0.429</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-4.786***</td>
<td>0.236</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-4.113***</td>
<td>0.286</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-5.677***</td>
<td>0.427</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-3.941***</td>
<td>0.575</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-3.437***</td>
<td>0.615</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-5.129***</td>
<td>1.034</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-4.409***</td>
<td>0.796</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-4.344***</td>
<td>0.886</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>6.760***</td>
<td>1.571</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>5,025</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2,494</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2,441</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>4,727</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2,33</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2,319</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>4,580</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2,265</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2,249</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>4,447</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2,234</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2,213</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AIC</td>
<td>4061334</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2628479</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1217868</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>3610320</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2356069</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1065949</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>3386372</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2237266</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>966093</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>3101716</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2128226</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>906936</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BIC</td>
<td>4061374</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2628514</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1217903</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>3610411</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2356150</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1066030</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>3386577</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2237449</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>966276</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>3102011</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2128477</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>907187</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Notes: The base category for smoking status is never smoker. Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1. Standard errors clustered at the school-classroom level. Source: GYTS Nepal 2007 and 2011.
<table>
<thead>
<tr>
<th></th>
<th>M1</th>
<th>M2</th>
<th>M3</th>
<th>M4</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Proximity to Other Smokers</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Friends Smoke</td>
<td>1.020*** 0.877*** 0.947***</td>
<td>0.957*** 0.751*** 0.946***</td>
<td>1.008*** 0.764*** 0.980***</td>
<td>0.781*** 0.645*** 0.849***</td>
</tr>
<tr>
<td>Parents Smoke</td>
<td>0.0808 0.0738 0.113</td>
<td>0.105 0.125 0.0599</td>
<td>0.137 0.145 0.103</td>
<td>0.127 0.0973 0.186</td>
</tr>
<tr>
<td>Pro-Tobacoo Marketing &amp; Social Perception</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Social Perception PCA</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Level of Attraction/ Ease at Parties</td>
<td>0.0410 0.0332 0.0434</td>
<td>0.0287 0.0282 0.0543</td>
<td>0.0448 0.0203 0.0764</td>
<td></td>
</tr>
<tr>
<td>Popularity</td>
<td>0.0562 0.015 -0.0515</td>
<td>0.0430 0.114 -0.0893</td>
<td>0.0449 0.127* -0.120</td>
<td></td>
</tr>
<tr>
<td>Perception of Adult Smokers</td>
<td>0.176*** 0.167** 0.175**</td>
<td>0.156*** 0.161** 0.0955</td>
<td>0.140*** 0.137** 0.0925</td>
<td></td>
</tr>
<tr>
<td>Anti-Tobacoo Awareness &amp; Knowledge</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Quality of Smoking Education</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low</td>
<td>0.0421 -0.193 0.729</td>
<td>0.0767 -0.233 0.675</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Medium</td>
<td>0.0309 0.0806 0.0241</td>
<td>-0.0281 0.0341 -0.156</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Variable</td>
<td>Estimate 1</td>
<td>Estimate 2</td>
<td>Estimate 3</td>
<td>Estimate 4</td>
</tr>
<tr>
<td>----------------------------------</td>
<td>------------</td>
<td>------------</td>
<td>------------</td>
<td>------------</td>
</tr>
<tr>
<td><strong>High</strong></td>
<td>-0.109</td>
<td>-0.133</td>
<td>0.247</td>
<td>-0.136</td>
</tr>
<tr>
<td><strong>Anti Tobacco Media Exposure</strong></td>
<td>0.306*</td>
<td>0.336</td>
<td>0.338</td>
<td>0.316*</td>
</tr>
<tr>
<td><strong>Parents Discuss Smoking Dangers</strong></td>
<td>0.115</td>
<td>0.180</td>
<td>-0.0933</td>
<td>0.104</td>
</tr>
<tr>
<td><strong>Cigarettes and Weight</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Gain Weight</strong></td>
<td>0.000729</td>
<td>-0.361</td>
<td>1.237</td>
<td>0.118</td>
</tr>
<tr>
<td><strong>Lose Weight</strong></td>
<td>-0.265</td>
<td>-0.485*</td>
<td>0.664</td>
<td>-0.199</td>
</tr>
<tr>
<td><strong>Smoking Cigarettes is Harmful</strong></td>
<td>-0.511**</td>
<td>-0.264</td>
<td>-1.039***</td>
<td>-0.671***</td>
</tr>
<tr>
<td><strong>Other's Smoking Harmful</strong></td>
<td>0.246</td>
<td>0.235</td>
<td>0.364</td>
<td>0.206</td>
</tr>
<tr>
<td><strong>Controls</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Age</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14-15 yo</td>
<td>0.355</td>
<td>0.310</td>
<td>0.492</td>
<td></td>
</tr>
<tr>
<td>16 + yo</td>
<td>0.533*</td>
<td>0.571*</td>
<td>0.374</td>
<td></td>
</tr>
<tr>
<td><strong>Grade</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Eighth</td>
<td>0.119</td>
<td>0.256</td>
<td>-0.00320</td>
<td></td>
</tr>
<tr>
<td>Ninth</td>
<td>0.164</td>
<td>0.499</td>
<td>-0.435</td>
<td></td>
</tr>
<tr>
<td>Tenth</td>
<td>0.366</td>
<td>0.700</td>
<td>-0.132</td>
<td></td>
</tr>
<tr>
<td>Year 2011</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-----------</td>
<td>----------</td>
<td>----------</td>
<td>----------</td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.163)</td>
<td>(0.180)</td>
<td>(0.250)</td>
<td>(0.172)</td>
</tr>
<tr>
<td></td>
<td>(0.172)</td>
<td>(0.191)</td>
<td>(0.269)</td>
<td>(0.341)</td>
</tr>
<tr>
<td></td>
<td>(0.341)</td>
<td>(0.398)</td>
<td>(1.006)</td>
<td>(0.454)</td>
</tr>
<tr>
<td></td>
<td>(0.206)</td>
<td></td>
<td></td>
<td>(0.516)</td>
</tr>
<tr>
<td>Observations</td>
<td>5,025</td>
<td>2,494</td>
<td>2,441</td>
<td>4,727</td>
</tr>
<tr>
<td>AIC</td>
<td>4061334</td>
<td>2628479</td>
<td>1217868</td>
<td>3610320</td>
</tr>
<tr>
<td>BIC</td>
<td>4061374</td>
<td>2628514</td>
<td>1217903</td>
<td>3610411</td>
</tr>
</tbody>
</table>

Notes: The base category for smoking status is never smoker. Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1. Standard errors clustered at the school-classroom level. Source: GYTS Nepal 2007 and 2011.
Appendix C: The New Landscapes of A Majority-Minority State Survey

SCREENING CRITERIA

S1: Are you currently a resident of New Mexico?  
Yes…..1  
No…..0  

[If S1=0, thank them for their time and hang up]

S2: Are you over the age of 18?  
Yes…..1  
No…..0  

[If S2=0, thank them for their time and hang up]

S3: What language do you prefer to speak?  
English…0  
Spanish…1  

[If S3=0, proceed with English version; If S3=1 proceed with Spanish version]

Version of Survey:  
V1. This is based upon what version of question 42 is being asked here.  
A…..1  
B…..2  
C…..3  

[For record only, don’t read V1 to the respondent]

SECTION 1: HEALTH OUTCOMES

Thank you. We will start with questions regarding your current health status.

Physical Health

1. How would you rate your overall health -- excellent, very good, good, fair, or poor? [NOTE TO JESSIE, PLEASE USE “MAS O MENOS” AS SPANISH CATEGORY FOR FAIR IN THE SPANISH VERSION NOT “REGULAR”]

Please Read:  
Excellent ………………. 1  
Very Good …………….. 2  
Good …………………. 3  
Fair …………………….. 4  
Poor ……………………. 5  

Do Not Read:  
Don’t know …………… 888  
Refused ……………….. 999
2. Have you ever been told by a doctor or health professional that you have any of the following conditions?  [Mark only one option for each; Don’t read 888 or 999]

<table>
<thead>
<tr>
<th>Condition</th>
<th>Yes = 1</th>
<th>No = 0</th>
<th>Don’t Know = 888</th>
<th>Refused = 999</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. High Blood Pressure</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>b. High Cholesterol</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c. Diabetes</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>d. Heart Disease</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>e. Cancer</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>f. Asthma</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>g. Obesity</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Mental Health and Wellbeing

3. In the past 12 months, did you think you needed help for emotional or mental health problems, such as feeling sad, anxious, or nervous?

   **Please Read:**
   Yes…..1  
   No…..0  
   **Do Not Read:**
   Don't know…..888  
   Refused…..999

4. Overall, how satisfied are you with life as a whole these days on a scale from one to ten: One means “not at all satisfied” and 10 means “completely satisfied.”

   **Record Answer:**
   ____ (Number 0 to 10)  
   **Do Not Read:**
   Don’t Know…..888  
   Refused…..999

Dental Health

5. When was the last time you were seen by a dentist? Stop me when I reach the correct answer. Was it…

   **Please Read:**
   Within the last 6 months…..1  
   More than 6 months but less than a year…..2  
   More than a year but less than two years…..3  
   More than two years ago…..4  
   Never have gone to a dentist…..5  
   **Do Not Read:**
6. In the past 12 months, was there a time when you wanted to see a dentist but did not?

Please Read:
Yes.....1
No.....0

Do Not Read:
Don't know.....888
Refused.....999

[If 6=1 go to 7 if 6=0, 888, 999 go to 8]

7. What was the primary reason you did not see a dentist? [Open Ended Code to List]

- You could not afford it.....1
- The dentist would not accept your health insurance.....2
- You do not have health insurance.....3
- Your health plan does not cover dental health.....4
- You couldn’t get an appointment soon enough.....5
- You couldn’t get there when the dentist’s office or dental clinic was open.....6
- You were too busy with work or other commitments to take the time.....7
- You didn’t think the problem was serious enough.....8
- Other, please specify______________9

Do Not Read:
Don’t know.....888
Refused.....999

SECTION II: HEALTH BEHAVIORS

Now we move onto learning more about your habits and behaviors when it comes to health.

Exercise Level

8. In the past seven days, how many days did you exercise or participate in physical activity for at least 20 minutes that made you sweat and breathe hard?

Record Answer:
______ (0-7 days)

Do Not Read:
Don’t Know.....888
Refused.....999

Alternative Medicine Use
9. In the past 12 months, did you seek care from any of the following? [Mark only one option for each; Don’t read 888 or 999]

<table>
<thead>
<tr>
<th></th>
<th>Yes = 1</th>
<th>No = 0</th>
<th>Don’t Know = 888</th>
<th>Refused = 999</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. Acupuncture</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>b. Chiropractor</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>e. Curandero</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>f. Shaman</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>g. Spiritual Healer</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>h. Other: List Who</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

10. In the past 12 months, did you take any herbs and supplements including teas or oil(s)?

**Please Read:**
- Yes…..1
- No…..0

**Do Not Read:**
- Don't know…..888
- Refused…..999

*Health Seeking Behaviors*

11. Was there any time during the past 12 months when you put off or postponed getting medical care you thought you needed?

**Please Read:**
- Yes…..1
- No…..0

**Do Not Read:**
- Don't know…..888
- Refused…..999

[If 11 =0, 888, 999 go to 13]

12. What was the primary reason you did not get the care that you needed?

**Please Read:**
- You were worried about the cost…1
- The doctor or hospital wouldn’t accepted your health insurance…2
- Your health plan wouldn’t pay for the treatment…3
- You couldn’t get an appointment soon enough…4
- You couldn't get there when the doctor's office or clinic was open…5
- You were too busy with work or other commitments to take the time…6
- You didn’t think the problem was serious enough…7

**Do Not Read:**
- Don’t know…888
- Refused…999
13. In the past 12 months, how many times did you receive care in a hospital emergency room?

Record Answer:
_______(number of visits)

Do Not Read:
Don't know…..888
Refused…..999

SECTION III: HEALTHCARE SYSTEM

This next set of questions is about your experience with health insurance and with healthcare.

Health Insurance

14. In the past 12 months, has there been any time where you went without health insurance?

Please Read:
Yes…..1
No…..0

Do Not Read:
Don't know…..888
Refused…..999

15. Do you have health insurance right now?

Please Read:
Yes….1
No….0

Do Not Read:
Don’t know….888
Refused….999

[If 15 = 1 then skip to 17] [If 15 = 0, 888, or 999 then ask 16 and skip to 18]

16. What is the main reason you do not currently have health insurance? [Open Ended Code to List and ROTATE Options]

Please Read:

- Turned down by insurance company….1
- The benefits package did not cover the service that you needed….2
- It is too hard to purchase coverage….3
- You cannot afford coverage….4
- You do not know how to purchase coverage….5
- Neither you nor your spouse employer provides coverage….6
- You are healthy and you don not need insurance….7
- Other: Please Specify_______________...8
17. What type of health insurance do you have? [READ and ROTATE Options]

**Please Read:**
- Employer-based insurance through work or job…1
- Insurance through the new health exchange marketplace…2
- Medicare of any type…3
- Medicaid…4
- Health insurance through the military, called TRICARE…5
- Some other insurance you privately purchase…6
- Insurance through parents…7
- Insurance through a spouse…8
- Any other type of health insurance plan …..9

**Do Not Read:**
- Don’t know…..888
- Refused…..999

[If 17=1, 3, 4 or 5 then skip to 19]

18. Have you purchased health insurance from the Be Well New Mexico Health Insurance Marketplace?

**Please Read:**
- Yes……1
- No……0

**Do Not Read:**
- Don’t Know…..888
- Refused…..999

**Health Care Provider**

19. Do you have a primary medical provider you usually go to when you are sick or need healthcare?

**Please Read:**
- Yes…..1
- No…..0

[Vol] Has more than one primary doctor…..2

**Do Not Read:**
- Don’t know .....888
- Refused .....999

[If 19=1 or 2 ask 20a but If 19=0, 888, or 999 ask 20b]

20. (A) What is the race or ethnicity of your primary care provider?

**Please Read:**
- White…..1
- Black or African American…..2
Hispanic or Latino…..3  
Asian…..4  
Native Hawaiian or other Pacific Islander…..5  
American Indian or Alaskan Native…..6  
Other (Please Specify ____________)  

Do Not Read:  
Don’t know……888  
Refused…..999  

(B) For your last encounter with a medical provider, what was their race or ethnicity?  
Please Read:  
White…..1  
Black or African American……2  
Hispanic or Latino…..3  
Asian…..4  
Native Hawaiian or other Pacific Islander…..5  
American Indian or Alaskan Native…..6  
Other (Please Specify ____________)  

Do Not Read:  
Don’t know……888  
Refused…..999  

[If S3 = 1 and 19=1 or 2 ask 21a. If S3 = 1 and 19=0, 888, and 999 ask 21b. If S3=0, skip to 22]  

21. (A) Were you able to talk with your primary care provider in your preferred language?  

Please Read:  
Yes….1  
No….0  

Do Not Read:  
Don’t know…888  
Refused….999  

(B) Were you able to talk in your preferred language with the provider you last saw?  

Please Read:  
Yes….1  
No….0  

Do Not Read:  
Don’t know…888  
Refused….999  

[Ask Everyone 22]  

22. If you could choose, would you prefer to be treated by a medical provider of your own race or ethnic group, another race or ethnic group, or do you have no preference?
23. If you could choose, would you prefer to be treated by a medical provider who speaks your primary language?

Please Read:
Yes…..1
No…..2
No preference…..3
Do Not Read:
Don’t know……888
Refused…..999

[If S3 = 1 ask 23 but if S3=0 skip to 24]

24. (A) Overall, how satisfied are you with the quality of health care you have received from your primary care provider?

Please Read:
Very satisfied…..1
Somewhat satisfied…..2
Neither satisfied nor dissatisfied…..3
Somewhat dissatisfied…..4
Very dissatisfied…..5
Do Not Read:
Don’t know……888
Refused…..999

(B) Overall, how satisfied are you with the quality of health care you have received during your encounter with a provider?

Please Read:
Very satisfied…..1
Somewhat satisfied…..2
Neither satisfied nor dissatisfied…..3
Somewhat dissatisfied…..4
Very dissatisfied…..5
Do Not Read:
Don’t know……888
Refused…..999

25. (A) How much trust do you have with your primary care provider?
Please Read:
Great deal…..1
A fair amount…..2
Not too much…..3
None at all…..4
Do Not Read:
Don’t know…..888
Refused…..999

(B) How much trust did you have with the healthcare provider you last saw?

Please Read:
Great deal…..1
A fair amount…..2
Not too much…..3
None at all…..4
Do Not Read:
Don’t know…..888
Refused…..999

SECTION IV: PERCEIVED DISCRIMINATION

We will now be changing to a different topic; these next few questions are about how you feel you are being treated.

26. There are many reasons why people treat others unfairly, in the past 12 months have you been treated unfairly here in New Mexico?

Please Read:
Yes ……1
No……0
Do Not Read:
Don’t Know…..888
Refused…..999

[If 26=0, 888, or 999 go to 29]

27. If yes, what do you think was the main reason for the experience?

Please Read:
Your Ancestry or National Origins…….1
Your Gender……..2
Your Race/Ethnicity…….3
Your Age……..4
Your Religion……..5
Your Height……..6
Your Weight……..7
Some other Aspect of Your Physical Appearance…….8
Your Sexual Orientation……..9
Your Education or Income Level…….10
A physical disability……11
Your shade of skin color …..12
Your accent…….13
Your tribe …..14

Other (SPECIFY) _____________________

Do Not Read:
Don’t Know…..888
Refused…..999

28. In the most typical incident you experienced, what was the race or ethnicity of
the person/s treating you unfairly?

Please Read:
White…..1
Black…..2
Asian…..3
Latino/Hispanic…..4
Native American/American Indian…..5

Do Not Read:
Don’t know…..888
Refused…..999

SECTION V: SOCIO-POLITICAL ECONOMIC ENVIRONMENT

We are now interested in learning more about the environment in which you live and how
you interact with it.

Neighborhood Factors

29. How long have you lived at your current address? Stop me when I reach the
correct answer. Was it…

Please Read:
Less than 1 year…..1
At least one year but less than two years…….2
At least two years but less than five years…..3
At least five years but less than ten years…..4
Ten years or more…..5

Do Not Read:
Don’t know …..888
Refused…..999

30. What is your primary means of transportation to work, school, or the place
where you spend most of your time outside of home? [Open Ended Code to
List and ROTATE Options]

Please Read:
Drive myself…..1
Get a ride with friends or family…..2
31. The next group of questions is about your neighborhoods, that is, the area around your home that you could walk to in 10 or 15 minutes or that area you considered to be your community. How much do you agree or disagree with each of the following statements about your neighborhood/community? Do you strongly agree, somewhat agree, neither agree nor disagree, somewhat disagree, or strongly disagree?

<table>
<thead>
<tr>
<th>Question Number</th>
<th>Question</th>
<th>Strongly Agree =1</th>
<th>Somewhat Agree =2</th>
<th>Neither Agree nor Disagree =3</th>
<th>Somewhat Disagree =4</th>
<th>Strongly Disagree =5</th>
<th>Don't Know =888</th>
<th>Refused =999</th>
</tr>
</thead>
<tbody>
<tr>
<td>a.</td>
<td>There are many banks or financial places within a reasonable distance.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>b.</td>
<td>There are places to walk or bicycle safely in or around my neighborhood/community.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c.</td>
<td>Residents of this neighborhood/community can obtain suitable employment in this area</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>d.</td>
<td>There are free or low cost public recreational facilities in my neighborhood/community such as parks,</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
playground, public swimming pools etc.

e. There are nearby grocery stores within a reasonable distance from my home.

f. There are health clinics, hospitals, or other places of health care within a reasonable distance from my home.

Development

32. Which of the follow approaches comes closer to your views about how to create more jobs here in New Mexico? [READ and ROTATE option]

Please Read:

We should invest more resources to attract large out of state companies to come to New Mexico….1

We should invest more resources to help more New Mexicans start small business and expand their existing businesses….2

Don’t Read

Don’t know … 888
Refused … 999

33. Which of the following best represent your views on economic development in New Mexico: [READ and ROTATE LIST]

Our economy relies too much on government spending for jobs…..1
Big public institutions such as universities and hospitals should do more to invest in local economic development…..2

We should do more to support locally-owned businesses…..3
I don't really care where the jobs come from as long as there are enough jobs…..4

Don’t Read

Don’t know … 888
Refused … 999

Governance s

34. How much of the time do you think that you can trust the state government to do what is right?

Please Read:

Always…..1
Most of the time…..2
Some of the time…..3
35. How would you describe your ability to influence local-government decision-making?

**Please Read:**
Great influence…..1
Moderate influence…..2
A little influence…..3
No influence at all…..4

**Do Not Read:**
Don’t know…..888
Refused…..999

36. There has been a lot of discussion in our state about political corruption over the past couple of years. Which of the following comes closer to your views on this issue?

[Rotate Options]

New Mexico has a culture of political corruption that cannot be improved by reform efforts so we just have to accept that some of this will be part of our political system..................1

Political corruption is not acceptable here in New Mexico and our political leaders should implement reforms such as an independent ethics commission to improve our political system….2

**Do Not Read:**
Don’t know…..888
Refused…..999

Political Views

37. On a scale of 1 to 7 with 1 being extremely liberal and 7 being extremely conservative, where do you fall on this scale?

**Record Answer:**
_____(1-7)

**Do Not Read:**
Don’t Know…..888
Refused…..999

SECTION VI: ORGANIZATIONS AND EMPOWERMENT

Continuing from the last section, we would like to hear more about programs that you participate in and the level of influence you have over decisions in your life.
Services Enrolled In

38. In the past 12 months, have you or anyone in your family participated in the following government programs…

   a. SNAP (formerly called Food Stamps)

      Please Read:
      Yes…..1
      No…..0

      Do Not Read:
      Don't know…..888
      Refused…..999

   b. WIC (Women, Infants, and Children)

      Please Read:
      Yes…..1
      No…..0

      Do Not Read:
      Don't know…..888
      Refused…..999

   c. TANF (Temporary Assistance for Needy Families)

      Please Read:
      Yes…..1
      No…..0

      Do Not Read:
      Don't know…..88
      Refused…..99

   d. Subsidized rental housing

      Please Read:
      Yes…..1
      No…..0

      Do Not Read:
      Don't know…..88
      Refused…..99

   e. Social Security Disability/Survivor Benefits

      Please Read:
      Yes…..1
      No…..0

      Do Not Read:
      Don't know…..888
      Refused…..999

   f. Unemployment compensation

      Please Read:
      Yes…..1
Empowerment

We are interested in how much influence you think you have in your life and in your community. I am going to read you a list of statements. For each one, please tell me how strongly you agree or disagree.

39. I have control over the decision that affects my life.

Please Read:
- Strongly Disagree…..1
- Disagree…..2
- Neither Agree nor Disagree…..3
- Agree…..4
- Strongly Agree…..5

Do Not Read:
- Don’t Know…..888
- Refused…..999

40. I am satisfied with the amount of influence I have over decisions that affect my community.

Please Read:
- Strongly Disagree…..1
- Disagree…..2
- Neither Agree nor Disagree…..3
- Agree…..4
- Strongly Agree…..5

Do Not Read:
- Don’t Know…..888
- Refused…..999

41. SPLIT SAMPLE QUESTION [ASK HALF OF RESPONDENTS AT RANDOM] Do you think that what happens generally to the Lesbian, Gay, Bisexual, Transgender, Queer, and Intersex (LGBTQI) community in this country will have something to do with what happens in your life? Will it affect you a lot, some, a little or not at all?

Please Read:
SECTION VI: OBESITY AND POLICY

42. Questions on Obesity and Policy:

[NOTE: For Question 42, randomly ask the respondent only ONE version (a, b, or c) of the question and record which version was select. After that, randomly select one of the 5 bid values, and record which bid value was selected]

a. Suppose a referendum will be held next week in New Mexico on a sugar-sweetened beverage tax initiative that is designed to fight obesity. The obesity-targeted polices would be financed by a ___ penny per ounce tax on all sugar-sweetened beverages (for example regular soda, sweetened iced teas, sport drinks, and energy drinks). This means that if a 12-ounce can of soda originally cost one dollar, after the tax the same soda would now cost \{insert bid’s corresponding dollar value here\}. Would you vote for or against this referendum?

[Randomly select option: 0 \{$1.00\}, 1/4 \{$1.03\}, 1/2 \{$1.06\}, 1 \{$1.12\}, and 1.5 \{$1.18\}]}

**Please Read:**

For…..1
Against…..0
Would not vote…..2

**Do Not Read:**

Don’t know…888
Refused…999

b. A growing number of scientific studies are showing the link between drinking soda and health problems, including obesity, diabetes, and heart attack. For example, several studies have found that adults and children who drink at least one soda or sugar sweetened beverage a day were more likely to be obese or overweight than people who did not drink soda daily. Further, people who consume at least one sugary drink a day are at a 26% increased risk of developing type two diabetes than people who drink sugary drinks less often.

Suppose a referendum will be held next week in New Mexico on a sugar-sweetened beverage tax initiative that is designed to fight obesity. The obesity-targeted polices would be financed by a ___ penny per ounce tax on all sugar-sweetened beverages (for example regular soda, sweetened iced teas, sport drinks, and energy drinks). This means that if a 12-ounce can of soda originally cost one
A growing number of scientific studies are showing the link between drinking soda and health problems, including obesity, diabetes, and heart attack. For example, several studies have found that adults and children who drink at least one soda or sugar sweetened beverage a day were more likely to be obese or overweight than people who did not drink soda daily. Further, people who consume at least one sugary drink a day are at a 26% increased risk of developing type two diabetes than people who drink sugary drinks less often.

Suppose a referendum will be held next week in New Mexico on a sugar-sweetened beverage tax initiative that is designed to fight obesity. The obesity-targeted polices would be financed by a ___ penny per ounce tax on all sugar-sweetened beverages (for example regular soda, sweetened iced teas, sport drinks, and energy drinks). This means that if a 12-ounce can of soda originally cost one dollar, after the tax the same soda would now cost {insert bid’s corresponding dollar value here}.

Say that the funds generated from a Soda-Tax initiative here in New Mexico would go towards reducing obesity by creating a mass media campaign to educate New Mexicans about the health risks of obesity and by improving the nutritional quality of school lunches from kindergarten through 12th grade. Would you vote for or against this referendum?

[Randomly select option: 0 {$1.00}, 1/4 {$1.03}, 1/2 {$1.06}, 1 {$1.12}, and 1.5 {$1.18}]

**Please Read:**
For…..1
Against…..0
Would not vote…..2

**Do Not Read:**
Don’t know…888
Refused…999

43. On a scale of 1 to 10, with 1 indicating very uncertain and 10 indicating very certain, how certain are you of your decision about how you would vote?

**Record Answer:**
____ (number of 1 to 10)

**Do Not Read:**
44. In November 2015, to combat the growing problem of diabetes and obesity, the Navajo Nation passed a new law where, in part, sugar-sweetened beverages would have an additional tax of 2 percent. Have you heard of this Navajo Nation law before?

**Please Read:**
- Yes.....1
- No.....0

**Do Not Read:**
- Don't know.....888
- Refused.....999

45. How much do you support enacting a similar law for all of New Mexico?

**Please Read:**
- Strongly Support.....1
- Support.....2
- Neither Support nor Against.....3
- Against.....4
- Strongly Against.....5

**Do Not Read:**
- Don't know.....888
- Refused.....999

46. Have you heard of such sugar sweetened beverage tax ballots initiatives being tried in other states or cities before?

**Please Read:**
- Yes.....1
- No.....0

**Do Not Read:**
- Don't know.....888
- Refused.....999

47. On a scale of 1 - 5, one being strongly disagree and five being strongly agree, how much do you agree or disagree with the following statements

a. Obesity is a major problem in New Mexico.

**Record Answer:**

_____ (number of 1 to 5)

**Do Not Read:**
- Don’t Know.....888
- Refused.....999

b. Childhood obesity is a major problem in New Mexico. Again 1 - 5, one being strongly disagree and five being strongly agree.

**Record Answer:**
c. A poor diet can lead to being overweight or obese.

Record Answer:

____ (number of 1 to 5)

Do Not Read:

Don’t Know…..888

Refused…..999

d. Drinking too much soda or other sugar-sweetened beverages can cause a person to be overweight or obese.

Record Answer:

____ (number of 1 to 5)

Do Not Read:

Don’t Know…..888

Refused…..999

48. Thinking of a regular week, how many times during that week did you….

a. Drink a sugar-sweetened beverage (like regular soda, sweetened iced tea, energy drinks, or sports drinks)

Record Answer:

____

Do Not Read:

Don’t Know…..888

Refused…..999

b. Eat fast food (including things like burgers, fries, chicken nuggets, and soda from restaurants like McDonalds, Taco Bell, and Burger King)

Record Answer:

____

Do Not Read:

Don’t Know…..888

Refused…..999

c. Eat fruits and vegetables

Record Answer:

____

Do Not Read:

Don’t Know…..888

Refused…..999

SECTION VII: DEMOGRAPHICS
Thank you for your time, we are almost through with the survey. For this section we would like to know some basic information about you.

Age

49. What year were you born?

**Record Answer:**

________ (Year)

**Do Not Read:**

Don’t know……888
Refused……999

Weight

50. How much do you weigh?

**Record Answer:**

______ (in pounds)
_____(in kilograms)

**Do Not Read:**

Don’t know…..888
Refused…..999

51. How tall are you?

**Record Answer:**

___ft__ in(in feet and inches)
____cm(in centimeters)

**Do Not Read:**

Don’t know…..888
Refused…..999

Race/Ethnicity

52. What is your race? Are you White, Black, American Indian, Asian, or Native Hawaiian/Pacific Islander? [Allow respondent to mark more than one racial group]

**Please Read:**

White…..1
Black or African American…..2
Native Hawaiian or Other Pacific Islander….3
American Indian or Alaskan Native .....4
Asian…..5
(DON'T ASK) Latino/Hispanic .....6

**Do Not Read:**

Something else / Some other race [Record response provided].....888
Refused……999

53. Do you identify as Hispanic, Latino, Chicano, or Spanish?

**Please Read:**
54. If you were walking down the street, what race/ethnicity do you think other Americans who do not know you personally would automatically assume you were, based on what you look like?

Please Read:
- White……1
- Black……2
- Asian……3
- American Indian…..4
- Hispanic or Latino…..5
- Mexican…..6
- Middle Eastern/Arab…..7

Other Race/Ethnicity, please specify ________________________________

Do Not Read:
- Don’t Know…..888
- Refused…..999

Skin Tone

55. We are interested in how you would describe your complexion. Using a scale from 1 to 5 where 1 represents very light and 5 represents very dark, where would you place your skin tone?

Please Read:
- Very light……1
- Light……2
- Medium……3
- Dark…..4
- Very Dark……5

Do Not Read:
- Don’t Know…..888
- Refused…..999

Religious Identity

56. What is your present religion, if any? [Open Ended Code to List]
Protestant (Baptist, Methodist, Non-denominational, Lutheran, Presbyterian, Pentecostal, Episcopal, Reformed, Church of Christ, etc.)...1
Roman Catholic (Catholic)...2
Native American (Tribal Religious Tradition)...3
Mormon (Church of Jesus Christ of Latter-day Saints/LDS)...4
Orthodox (Greek, Russian, or some other orthodox church)...5
Jewish (Judaism)...6
Muslim (Islam)...7
Buddhist...8
Hindu...9
Atheist (do not believe in God)...10
Agnostic (not sure if there is a God)...11
Nothing in particular...12
Christian...13
Unitarian (Universalist)...14
Jehovah’s Witness...15
Other, please specify...

Do Not Read:
Don't Know...888
Refused...999

57. In the past year, how often do you attend religious services, gatherings, or ceremonies? [Question only asked of those who identify a religious affiliation]

Please Read:
More than once a week...1
Once a week...2
Once or twice a month...3
A few times a year...4
Seldom...5
Never...6

Don’t Read:
Don't know...888
Refused...999

New Mexican Identity

58. Do you feel a sense of familial connection to the land here in New Mexico?

Please Read:
Yes...1
No...0

Do Not Read:
Don't know...888
Refused...999

Citizenship Status
59. Are either of your parents born in the U.S.?

Please Read:
Yes…..1
No…..0

Do Not Read:

Don't know…..888
Refused…..999

60. Were you born in the United States?

Please Read:
Yes…..1
No…..0

Do Not Read:

Don't know…..888
Refused…..999

[If 60=0 then go to 62]

61. Were you born in New Mexico?

Please Read:
Yes…..1
No…..0

Do Not Read:

Don't know…..888
Refused…..999

[If 60=1, 888, 999 then go to 70]

62. What is your current documentation status? [Remind them that it is strictly confidential]

Please Read:
Naturalized Citizen ….1
Legal Permanent Resident ….2
DACA recipient ….3
Temporary visa (student, work) ….5
DACA eligible….6
DACA ineligible…..7
None of these apply to me…..8

Do Not Read:

Don’t know .....888
Refused…..999

63. What was your visa or immigration status at the time of entry to the United States?

Please Read:
Tourist/visitor visa….. 1
Business visa….. 2
64. In what country were you born? [ENTER ONE ONLY]

**Record Answer:**
Open Ended ______________

**Do Not Read:**
Don’t know …..888
Refused…..999

65. How many years have you lived in the United States (excluding Puerto Rico)?

**Record Answer:**

_____________ (Record Years From 1 to 97)  

_____________ (Record number of months if less than 1 year)  

_____________ (Record year that the respond came to US)

**Do Not Read:**
Don’t know…..888
Refused…..999

*Country of Origin Health Experience*

66. Did you have regular access to preventive health care services as a child in [country of origin]?

**Please Read:**
Yes…..1
No…..0

**Do Not Read:**
Don’t know…..888
Refused…..999

67. What was the source of your usual healthcare in [country of origin]?

**Please Read:**
Public healthcare services…..1
NGO’s…..2
Charities…..3
Private healthcare services…..4
Family .....5
68. Have you experienced new physical or mental health problems since migrating to the United States?

**Please Read:**
- Yes, physical…1
- Yes, mental…2
- Yes, both mental and physical…3
- No…4

**Do Not Read:**
- Don’t Know…..888
- Refused…..999

69. Do you ever travel to [country of origin] specifically for healthcare services?

**Please Read:**
- Yes…..1
- No….0

**Do Not Read:**
- Don’t know…..888
- Refused…..999

**Sex and Gender**

70. What was the sex on your original birth certificate? [Open Ended Code to List]

**Please Read:**
- Male….1
- Female…..2
- Intersex…..3

**Do Not Read:**
- Don’t Know…..888
- Refused…..999

71. If you were walking the down the street, how would other Americans who do not know you personally identify your gender? Would you say: [Open Ended Code to List]

**Please Read:**
- Man….1
- Woman…..2
- Transgendered…..3
- Other …..4

**Do Not Read:**
- Don’t Know…..888
- Refused…..999
72. Do you identify as a member of the LGBTQI community?

Please Read:
Yes…..1
No…..0

Do Not Read:
Don't know…..888
Refused…..999

[If 72=0, 888, or 999 skip to 82]

73. What is your current gender? [Open Ended Code to List]

Female…..1
Male…..2
Transgender…..3
Part-time male…..4
Part-time female…..5

Do Not Read:
Don’t know…..888

74. What best describes your gender identity? [Mark only one option for each; Don’t read 888 or 999] [Open Ended Code To List]

<table>
<thead>
<tr>
<th></th>
<th>Yes = 1</th>
<th>No = 0</th>
<th>Don’t Know = 888</th>
<th>Refused =999</th>
</tr>
</thead>
<tbody>
<tr>
<td>a.</td>
<td>Female</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>b.</td>
<td>Male</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c.</td>
<td>Transgender</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>d.</td>
<td>Transmale</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>e.</td>
<td>Transfemale</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>f.</td>
<td>Gender queer</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>g.</td>
<td>Gender non-conforming</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>h.</td>
<td>Intersex</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>i.</td>
<td>Two-spirit</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>j.</td>
<td>Other (specify)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

75. What terms best describe your gender expression? [Mark only one option for each; Don’t read 888 or 999] [Open Ended Code To List]

<table>
<thead>
<tr>
<th></th>
<th>Yes = 1</th>
<th>No = 0</th>
<th>Don’t Know = 888</th>
<th>Refused =9999</th>
</tr>
</thead>
<tbody>
<tr>
<td>a.</td>
<td>Transgender</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>b.</td>
<td>Male to female</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c.</td>
<td>Transexual</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>d.</td>
<td>Gender non-conforming</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>e.</td>
<td>Female to Male</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>f.</td>
<td>Gender queer</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>g.</td>
<td>Two-Spirit</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>h.</td>
<td>Cross-dresser</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>j.</td>
<td>Androgyrous</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>k.</td>
<td>Third Gender</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>l.</td>
<td>Feminine male</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>m.</td>
<td>Masculine female or butch</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>n.</td>
<td>Intersex</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>o.</td>
<td>Drag performer (King/Queen)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>p.</td>
<td>AG or Aggressive</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>q.</td>
<td>Other (specify)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

76. Do you think people can tell you are LGBTQI even you don’t tell them:

**Please Read:**

- Always…..1
- Most of the time……2
- Sometimes…..3
- Occasionally…..4
- Never…..5

**Do Not Read:**

- Don’t Know…..888
- Refused…..999

77. What best describes your sexual orientation?

**Please Read:**

- Heterosexual…..1
- Lesbian…..2
- Bisexual…..3
- Gay…..4
- Pansexual…..5
- Queer…..6
- Questioning…..7
- Asexual…..8
- Other…..9

**Do Not Read:**

- Don’t Know…..888
- Refused…..999

78. Is there an LGBTQI resource center in your community or an organization that provides services for other minority groups including LGBTQI individuals?

**Please Read:**

- Yes…..1
- No…..0

**Do Not Read:**
Don't know…..888
Refused…..999

[If 78=0, 888, 999 go to 82]

79. Have you utilized these services?

Please Read:
Yes…..1
No…..0

Do Not Read:
Don't know…..888
Refused…..999

80. Please indicate the level of agreement or disagreement with the following
statements: I feel connected to my local LGBTQI community.

Please Read:
Strongly Disagree…..1
Disagree…..2
Neither Agree nor Disagree .....3
Agree…..4
Strongly Agree…..5

Do Not Read:
Don't know…..88
Refused…..99

81. How often have you felt uncomfortable in your racial or ethnic community
because of your sexual orientation?

Please Read:
Never…..1
Sometimes……2
Always…..3

Do Not Read:
Don’t Know……888
Refused…..999

Marital Status

82. Are you currently single, married, divorced, separated, living with a partner or
widowed?

Please Read:
Single…..1
Married…..2
Divorced…..3
Separated…..4
Widowed…..5
Living with a partner……6
Other…..7

Do Not Read:
83. Including yourself, how many people live in your household?

**Record Answer:**

_____ (number of people under the age of 18)

**Do Not Read:**

Don’t Know…..888

Refused…..999

**Education**

84. What is the highest level of school you have completed or the highest degree you have received?

**Please Read:**

Less than high school (Grades 1-8 or no formal schooling)…..1

High school incomplete (Grades 9-11 or Grade 12 with NO diploma)…..2

High school graduate (Grade 12 with diploma or GED certificate)…..3

Some college, no degree (includes community college)…..4

Two year associate degree from a college or university…..5

Four year college or university degree/Bachelor’s degree (e.g., BS, BA, AB)….6

Some postgraduate or professional schooling, no postgraduate degree…..7

Postgraduate or professional degree, including master’s, doctorate, medical or law degree (e.g., MA, MS, PhD, MD, JD)….8

**Do Not Read:**

Don’t know…888

Refused…999

**Employment Status**

85. What is your employment status?

**Please Read:**

Currently employed full time … 1

Currently employed part time … 2

Seasonal employment….3

Not employed, but was employed during the past 12 months … 4

Not employed, and was not employed during the past 12 months……5

**Do Not Read:**

Don’t know … 888

Refused … 999

**Dominant Language**

86. What are the primary languages spoken in your home? Mark All That Apply

**Please Read:**

English…..1

Spanish…..2

Tribal Language…..3
Socioeconomic Status

87. What was your total combined household income in 2015 before taxes? I will provide a range of income categories, and just stop me when I read the correct category that best captures your household income. [Probe: This question is completely confidential and just used to help classify the responses, but it is very important for our research. Repeat the question.]

Record Answer:
- Less than $5,000…..0
- Less than $10,000…..1
- $10,000 to $14,999…..2
- $15,000 to $19,999…..3
- $20,000 to $29,999…..4
- $30,000 to $39,999…..5
- $40,000 to $59,999…..6
- $60,000 to $79,999…..7
- $80,000 to $99,999…..8
- $100,000 to $150,000…..9
- More than $150,000…..10

Do Not Read:
Don’t know…..888
Refused…..999

[If 87=888 or 999 go to 88]

88. Do you make more or less than $50,000.

- Less than $50,000…..1
- More than $50,000…..2

89. Would you be interested in being part of a follow up focus group and/or interview?

Please Read:
- Yes…..1
- No…..0

Do Not Read:
Don’t know…..888
Refused…..999

If Yes, what is the best phone number to reach you at? _____________________

Prescription Drugs and Illicit Drugs
Before concluding, we would like to ask you a few more questions about prescription and illicit drugs. Your answers are confidential and there is no identifying information of yours that will be linked to your response.

90. Have you or an immediate family member ever suffered from alcohol and/or drug dependence?

**Please Read:**
- Yes…..1
- No…..0

**Do Not Read:**
- Don't know…..888
- Refused…..999

91. Do you have a close friend that is either an active addict or in recovery from alcohol and/or drug dependence?

**Please Read:**
- Yes…..1
- No…..0

**Do Not Read:**
- Don't know…..888
- Refused…..999

Do you agree or disagree with the following two statements:

92. Drugs are a problem in your community.

**Please Read:**
- Strongly Disagree…..1
- Disagree…..2
- Neither Agree nor Disagree…..3
- Agree…..4
- Strongly Agree…..5

**Do Not Read:**
- Don't know…..888
- Refused…..999

93. Tax dollars should pay for substance use prevention or intervention services.

**Please Read:**
- Strongly Disagree…..1
- Disagree…..2
- Neither Agree nor Disagree…..3
- Agree…..4
- Strongly Agree…..5

**Do Not Read:**
- Don't know…..888
94. Do you know anyone who has died from a prescription or illegal drug overdose?

**Please Read:**
- Yes.....1
- No.....0

**Do Not Read:**
- Don't know.....888
- Refused.....999

Thank you for your participation in this research study. If you have any questions later on you may reach me by email at [INSERT EMAIL ADDRESS] or by phone at [INSERT PHONE NUMBER]
Appendix D: Stata Code

 sentenced text: (Note: this is a natural representation of the code as written in the document, without any additional formatting or annotation.)

REM: Chapter 2 (Willingness to Pay for SSB Tax) Code
* Conditional Mixed Process, Probit, Probit with Selection, and Probit with Selection and Control Function Approach Regressions and Average Marginal Effects
* Bootstrapping Program and WTP Estimates: with and without selection
Kristina N. Piorkowski */

******************** Regressions and AMEs ********************
***** This reproduces the results found in Table 2.2, 2.3, 2.2.A, 2.2.B, and 2.2.C *****
clear all
set more off
capture log close
capture timer clear
cd /Users/kristinapiorkowski/Chp2"

use " cleaned_data_103116"

log using full_regressions_070617, replace

keep if q42a != .
gen YesNo = q42a
recode YesNo (888 = .) (2 = .)

*Soda Tax
gen soda_tax =.
replace soda_tax = 1 if split_bid_value == 4 | split_bid_value == 5
replace soda_tax = 2 if split_bid_value == 6
replace soda_tax = 5 if split_bid_value == 7
replace soda_tax = 10 if split_bid_value == 8
replace soda_tax = 15 if split_bid_value == 9
replace soda_tax = 25 if split_bid_value == 10

gen log_soda_tax = log(soda_tax)

*Values Recoded Based Upon Certainty
gen certain = q43_1
recode certain (888 = .)

gen Yes10 = YesNo
replace Yes10 = 0 if YesNo == 1 & certain<10
gen snav_bi = .
replace snav_bi = 0 if sup_nav == 1
replace snav_bi = 1 if sup_nav == 2 | sup_nav == 3 | sup_nav == 4 | sup_nav == 5

****** Recoded Variables ******

gen snav_bi_rc = snav_bi
recode snav_bi_rc (. = 1)

gen conservative_rc = conservative
recode conservative_rc (. = 0)

gen opin_union_rc = opin_union
recode opin_union_rc (. = 0)

gen log_num_ssb = (num_ssb + 1)
replace log_num_ssb = log(log_num_ssb)

gen log_num_fv = (num_fv + 1)
replace log_num_fv = log(log_num_fv)

gen ssb_bi = num_ssb > 0 if num_ssb < .

gen log_ssb_cen = log(num_ssb) if num_ssb > 0

gen new_Yes10 = Yes10
replace new_Yes10 = . if snav_bi_rc == 0

cmp setup

**************************************************************************
*Define global variables
global endo_vars "know_2 age age2 race_white inc_50_ab"

global selection_vars "heard_nav heard_other_st conservative_rc opin_union_rc"

global m1_rhs "log_soda_tax"

global m2_rhs "log_soda_tax log_num_ssb num_fv obesity know_1 gov_trust_never"

global m3_rhs "log_soda_tax log_num_ssb num_fv obesity know_1 gov_trust_never female age age2 edu_med edu_high race_white phone farm_metro non_metro lc_metro abq_metro inc_50_ab"

************************************************************************** Model 1**************************************************************************
*dropping variables ensures that all the models will have the same number of observations
drop if Yes10 == .
drop if heard_nav == .
drop if heard_other_st == .
drop if log_soda_tax == .

*Probit
probit Yes10 $m1_rhs [iweight=weight], vce(robust)
outreg2 using probit_results.xls, excel ctitle(M1) replace
estat ic
margins, dydx(*)

*Probit w selection
heckprob Yes10 $m1_rhs [iweight=weight],
   select(snav_bi_rc = $selction_vars) vce(robust)
outreg2 using probit_w_selection_results.xls, excel ctitle(M1) replace
estat ic
margins, dydx(*) predict(pcond) noestimcheck

*Probit w selection w cfa
heckprob Yes10 $m1_rhs [iweight=weight],
   select(snav_bi_rc = $selction_vars) vce(robust)
outreg2 using probit_w_selection_cfa_results.xls, excel ctitle(M1) replace
estat ic
margins, dydx(*) predict(pcond) noestimcheck

*3 eq system
cmp (eq2:snav_bi_rc = $selction_vars) ///
   (eq3:Yes10 = $m1_rhs) [iweight=weight], ///
   ind($cmp_probit snav_bi*$cmp_probit) vce(robust)
outreg2 using cmp_results.xls, excel ctitle(M1) replace
estat ic
margins, dydx(*) predict(pr eq(#2) condition(0 ,eq(#1)))

********************* Model 2 *********************

drop if num_ssb == .
drop if num_fv == .
drop if know_2 == .
drop if age == .
drop if age2 == .
drop if race_white == .
drop if obesity == .
drop if know_1 == .
drop if gov_trust_never == .
*Probit
probit Yes10 $m2_rhs [iweight=weight], vce(robust)
outreg2 using probit_results.xls, excel ctitle(M2) append
estat ic
margins, dydx(*)

*Probit w selection
heckprob Yes10 $m2_rhs [iweight=weight], ///
       select(snav_birc = $selection_vars) vce(robust)
outreg2 using probit_w_selection_results.xls, excel ctitle(M2) append
estat ic
margins, dydx(*) predict(pcond) noestimcheck

*Probit w selection w cfa
reg log_num_ssb $endo_vars [iweight=weight]
predict ssb_resid_m1, residuals
outreg2 using probit_w_selection_cfa_results.xls, excel ctitle(Endo) append

heckprob Yes10 $m2_rhs ssb_resid_m1 [iweight=weight], ///
       select(snav_birc = $selection_vars) vce(robust)
outreg2 using probit_w_selection_cfa_results.xls, excel ctitle(M2) append
estat ic
margins, dydx(*) predict(pcond) noestimcheck

*3 eq system
cmp (log_num_ssb = $endo_vars) ///
   (snav_birc = $selection_vars) ///
   (Yes10 = $m2_rhs) [iweight=weight], ///
   ind($cmp_cont $cmp_probit snav_birc*$cmp_probit) vce(robust)
outreg2 using cmp_results.xls, excel ctitle(M2) append
estat ic
margins, dydx(*) predict(pr eq(#3) condition(0 ,eq(#2)))

******************************* Model 3 *********************************
drop if female == .
drop if edu_med == .
drop if edu_high == .

*Probit
probit Yes10 $m3_rhs [iweight=weight], vce(robust)
outreg2 using probit_results.xls, excel ctitle(M3) append
estat ic
margins, dydx(*)
*Probit w selection
heckprob Yes10 $m3_rhs [iweight=weight], ///
    select(snav_bi_rc = $selction_vars) vce(robust)
outreg2 using probit_w_selection_results.xls, excel ctitle(M3) append
estat ic
margins, dydx(*) predict(pcond) noestimcheck

*Probit w selection w cfa
heckprob Yes10 $m3_rhs ssb_resid_m1 [iweight=weight], ///
    select(snav_bi_rc = $selction_vars) vce(robust)
outreg2 using probit_w_selection_cfa_results.xls, excel ctitle(M3) append
estat ic
margins, dydx(*) predict(pcond) noestimcheck

*3 eq system
cmp (log_num_ssb = $endo_vars) ///
    (snav_bi_rc = $selction_vars) ///
    (Yes10 = $m3_rhs) [iweight=weight], ///
    ind($cmp_cont $cmp_probit snav_bi*$cmp_probit) vce(robust)
outreg2 using cmp_results.xls, excel ctitle(M3) append
estat ic
margins, dydx(*) predict(pr eq(#3) condition(0 . ,eq(#2)))

log close

***********************************************************************
******* Bootstrapping and WTP Estimates **************
********** This and the following section reproduces Table 2.4 and Figure 2.1 **********
***********************************************************************

*************** With Selection ***************

******* Set up for file ***************

clear all
set more off
capture log close
capture timer clear
cd /Users/kristinapiorkowski/Chp2"
use " short_data_061917"
log using cmp_boot_wtp_selection_062617_new, replace
cmp setup
global endo_vars "know_2 age age2 race_white inc_50_ab"
global selection_vars "heard_nav heard_other_st conservative_rc opin_union_rc"

194
global m2_rhs        "log_soda_tax log_num_ssb num_fv obesity know_1
gov_trust_never"
capture program drop cmp_prg
program define cmp_prg, rclass
preserve
cmp (log_num_ssb = $endo_vars) ///
         (snav_bi_rc = $selection_vars) ///
         (Yes10 = $m2_rhs) [iweight=weight], ///
         ind($cmp_cont $cmp_probit snav_bi_rc*$cmp_probit) vce(robust)
matrix b_cmp = e(b)
matrix list b_cmp
scalar b1_yes10_con = b_cmp[1,18]
scalar b1_yes10_lns = b_cmp[1,13]
scalar b1_yes10_nfv = b_cmp[1,14]
scalar b1_yes10_obe = b_cmp[1,15]
scalar b1_yes10_kn1 = b_cmp[1,16]
scalar b1_yes10_gov = b_cmp[1,17]
scalar b1_yes10_lst = b_cmp[1,12]
scalar b1_rho       = b_cmp[1,22]
scalar b1_snav_hnn  = b_cmp[1,7]
scalar b1_snav_hot  = b_cmp[1,8]
scalar b1_snav_cns  = b_cmp[1,9]
scalar b1_snav_opn  = b_cmp[1,10]
scalar b1_snav_con  = b_cmp[1,11]
//return values
return scalar b_yes10_con = b1_yes10_con
return scalar b_yes10_lns = b1_yes10_lns
return scalar b_yes10_nfv = b1_yes10_nfv
return scalar b_yes10_obe = b1_yes10_obe
return scalar b_yes10_kn1 = b1_yes10_kn1
return scalar b_yes10_gov = b1_yes10_gov
return scalar b_yes10_lst = b1_yes10_lst
return scalar b_rho       = b1_rho
return scalar b_snav_hnn = b1_snav_hnn
return scalar b_snav_hot  = b1_snav_hot
return scalar b_snav_cns  = b1_snav_cns
return scalar b_snav_opn  = b1_snav_opn
return scalar b_snav_con  = b1_snav_con
end
capture prog drop BootWTP_062617
prog define BootWTP_062617, rclass

tempvar swtp_cmp_kr swtp_cmp_kr_mean swtp_cmp_kr_median
WTP_mean_99_1 ///
    WTP_median_99_1 WTP_mean_95_1 WTP_median_95_1 WTP_mean_90_1 ///
    WTP_median_90_1 WTP_at_mean_1 WTP_98 WTP_95 WTP_mean_98_2_1 ///
    WTP_median_98_2_1 WTP_99 WTP_95 WTP_90

gui cmp_prv

matrix b_cmp = e(b)
matrix list b_cmp

scalar b_yes10_con = b_cmp[1,18]
scalar b_yes10_lns = b_cmp[1,13]
scalar b_yes10_nf = b_cmp[1,14]
scalar b_yes10_obe = b_cmp[1,15]
scalar b_yes10_kn1 = b_cmp[1,16]
scalar b_yes10_gov = b_cmp[1,17]
scalar b_yes10_lst = b_cmp[1,12]
scalar b_rho = b_cmp[1,22]
scalar b_snav_hnn = b_cmp[1,7]
scalar b_snav_hot = b_cmp[1,8]
scalar b_snav_cns = b_cmp[1,9]
scalar b_snav_opn = b_cmp[1,10]
scalar b_snav_con = b_cmp[1,11]

gen `swtp_cmp_kr' = exp((b_yes10_con + b_yes10_lns * log_num_ssb + b_yes10_nf * num_ffv ///
    + b_yes10_obe * obesity + b_yes10_kn1 * know_1 + b_yes10_gov * gov_trust_never) ///
    / -b_yes10_lst)* ///
    exp((exp(2*b_rho)-1)/(exp(b_rho*2)+1) * ///
    (-1/b_yes10_lst) * ///
    (normalden(b_snav_hnn * heard_nav + b_snav_hot * heard_other_st + b_snav_cns * conservative_rc ///
        + b_snav_opn * opin_union_rc + b_snav_con) ///
    normal(b_snav_hnn * heard_nav + b_snav_hot * heard_other_st + b_snav_cns * conservative_rc ///
        + b_snav_opn * opin_union_rc + b_snav_con))

// Untrimmed
qui sum `swtp_cmp_kr', detail
    scalar  `swtp_cmp_kr_mean'   = r(mean)
    scalar  `swtp_cmp_kr_median' = r(p50)
    scalar  `WTP_99'             = r(p99)
    scalar  `WTP_95'             = r(p95)
    scalar  `WTP_90'             = r(p90)

// 99% Trimmed
qui sum `swtp_cmp_kr' if `swtp_cmp_kr' < `WTP_99', detail
    scalar  `WTP_mean_99_1'   = r(mean)
    scalar  `WTP_median_99_1' = r(p50)

// 95% Trimmed
qui sum `swtp_cmp_kr' if `swtp_cmp_kr' < `WTP_95', detail
    scalar  `WTP_mean_95_1'   = r(mean)
    scalar  `WTP_median_95_1' = r(p50)

// 90% Trimmed
qui sum `swtp_cmp_kr' if `swtp_cmp_kr' < `WTP_90', detail
    scalar  `WTP_mean_90_1'   = r(mean)
    scalar  `WTP_median_90_1' = r(p50)

// At mean
qui sum log_num_ssb
    scalar  m1_log_num_ssb = r(mean)
qui sum num_fv
    scalar  m1_num_fv = r(mean)
qui sum obesity
    scalar  m1_obesity = r(mean)
qui sum know_1
    scalar  m1_know_1 = r(mean)
qui sum gov_trust_never
    scalar  m1_gov_trust_never = r(mean)
qui sum heard_nav
    scalar  m1_heard_nav = r(mean)
qui sum heard_other_st
    scalar  m1_heard_other_st = r(mean)
qui sum conservative_rc
    scalar  m1_conservative_rc = r(mean)
qui sum opin_union_rc
    scalar  m1_opin_union_rc = r(mean)

    scalar  `WTP_at_mean_1' = exp((b_yes10_con + b_yes10_lns * m1_log_num_ssb + b_yes10_nf * m1_num_fv) ///
+ b_yes10 obe * m1_obesity + b_yes10 kn1 * m1_know_1 + b_yes10 gov * m1_gov_trust_never) ///
-b_yes10 lst)* ///
/ (exp((exp(2*b_rho) -1)/(exp(b_rho*2)+1) * ///
(-1/b_yes10 lst) * ///
(normalden(b_snav_hnn * m1_heard_nav + b_snav_hot * m1_heard_other_st + b_snav_cns * m1_conservative_rc ///
+ b_snav_opn * m1_opin_union rc + b_snav_con) / ///
(normal(b_snav_hnn * m1_heard_nav + b_snav_hot * m1_heard_other_st + b_snav_cns * m1_conservative_rc ///
+ b_snav_opn * m1_opin_union rc + b_snav_con)))

// Return Values
return scalar WTP_mean      = `swtp_cmp_kr_mean'
return scalar WTP_median    = `swtp_cmp_kr_median'
return scalar WTP_mean_99   = `WTP_mean_99_1'
return scalar WTP_median_99 = `WTP_median_99_1'
return scalar WTP_mean_95   = `WTP_mean_95_1'
return scalar WTP_median_95 = `WTP_median_95_1'
return scalar WTP_mean_90   = `WTP_mean_90_1'
return scalar WTP_median_90 = `WTP_median_90_1'
return scalar WTP_at_mean   = `WTP_at_mean_1'

end

bootstrap WTP_mean      = r(WTP_mean) ///
    WTP_mean_99   = r(WTP_mean_99) ///
    WTP_mean_95   = r(WTP_mean_95) ///
    WTP_mean_90   = r(WTP_mean_90) ///
    WTP_median    = r(WTP_median) ///
    WTP_median_99 = r(WTP_median_99) ///
    WTP_median_95 = r(WTP_median_95) ///
    WTP_median_90 = r(WTP_median_90) ///
    WTP_at_mean   = r(WTP_at_mean) ///
, rep(500) saving(bootVar_selection_062217, replace) ///
    seed(1234) noisily :  BootWTP_062617
estat bootstrap, all

use bootVar_selection_062217, replace
************************* WTP Median Untrimmed *************************

***** 95% CI
sort WTP_median
egen WTP_median_2_5 = pctlile(WTP_median), p(2.5)
egen WTP_median_97_5 = pctlile(WTP_median), p(97.5)
sum WTP_median_2_5 WTP_median_97_5
sum WTP_median if WTP_median < WTP_median_97_5 & WTP_median > WTP_median_2_5, detail

***** Histogram
histogram WTP_median, percent kdensity ytitle(Percent) xtitle(Willingness to Pay (in pennies per ounce)) ///
    title(Histogram of Bootstrapped Median WTP) ///
    xline(0.167) ///
    note(Notes: WTP was bootstrapped 500 times. Median of the bootstrapped WTP reported.)

graph twoway (histogram WTP_median, ytitle(Density) xtitle(Willingness to Pay (in pennies per ounce)) ///
    title(Figure 1: Histogram of Bootstrapped Median WTP) ///
    note(Notes: WTP was bootstrapped 500 times. Median of the bootstrapped WTP reported.)) ///
    (kdensity WTP_median) ///
    (scatteri 0 0.167 3 0.167, c(l) m(i)), ///
    legend(order(2 3) label(2 "Kernel Density") label(3 "Median"))
graph save Median_WTP, replace
graph export Median_WTP.pdf, replace

************************* WTP Median 99 Trimmed *************************

***** 95% CI
egen WTP_median_99_2_5  = pctile(WTP_median_99), p(2.5)
egen WTP_median_99_97_5 = pctile(WTP_median_99), p(97.5)

sum WTP_median_99_2_5 WTP_median_99_97_5
sum WTP_median_99 if WTP_median_99 < WTP_median_99_97_5 & WTP_median_99 > WTP_median_99_2_5, detail

************************* WTP Median 95 Trimmed *************************

***** 95% CI
egen WTP_median_95_2_5  = pctile(WTP_median_95), p(2.5)
egen WTP_median_95_97_5 = pctile(WTP_median_95), p(97.5)

sum WTP_median_95_2_5 WTP_median_95_97_5
sum WTP_median_95 if WTP_median_95 < WTP_median_95_97_5 & WTP_median_95 > WTP_median_95_2_5, detail
************************ WTP Median 90 Trimmed ***************************

***** 95% CI
egen WTP_median_90_2_5  = pctile(WTP_median_90), p(2.5)
egen WTP_median_90_97_5 = pctile(WTP_median_90), p(97.5)

sum WTP_median_90_2_5 WTP_median_90_97_5

sum WTP_median_90 if WTP_median_90 < WTP_median_90_97_5 & WTP_median_90 > WTP_median_90_2_5, detail

************************** WTP Mean Untrimmed ******************************

***** 95% CI
sort WTP_mean
egen WTP_mean_2_5  = pctile(WTP_mean), p(2.5)
egen WTP_mean_97_5 = pctile(WTP_mean), p(97.5)

sum WTP_mean_2_5 WTP_mean_97_5

sum WTP_mean if WTP_mean < WTP_mean_97_5 & WTP_mean > WTP_mean_2_5, detail

************************* WTP Mean 99 Trimmed ***********************

***** 95% CI
egen WTP_mean_99_2_5  = pctile(WTP_mean_99), p(2.5)
egen WTP_mean_99_97_5 = pctile(WTP_mean_99), p(97.5)

sum WTP_mean_99_2_5 WTP_mean_99_97_5

sum WTP_mean_99 if WTP_mean_99 < WTP_mean_99_97_5 & WTP_mean_99 > WTP_mean_99_2_5, detail

************************* WTP Mean 95 Trimmed **************************

***** 95% CI
egen WTP_mean_95_2_5  = pctile(WTP_mean_95), p(2.5)
egen WTP_mean_95_97_5 = pctile(WTP_mean_95), p(97.5)

sum WTP_mean_95_2_5 WTP_mean_95_97_5

sum WTP_mean_95 if WTP_mean_95 < WTP_mean_95_97_5 & WTP_mean_95 > WTP_mean_95_2_5, detail

200
************** WTP Mean 90 Trimmed ***********************

***** 95% CI
egen WTP_mean_90_2_5 = pctile(WTP_mean_90), p(2.5)
egen WTP_mean_90_97_5 = pctile(WTP_mean_90), p(97.5)

sum WTP_mean_90_2_5 WTP_mean_90_97_5

sum WTP_mean_90 if WTP_mean_90 < WTP_mean_90_97_5 & WTP_mean_90 > WTP_mean_90_2_5, detail

log close

********************** Bootstrapping and WTP Estimates ***********************

****************************** Without Selection ******************************

***************************** Set up for file *****************************

clear all
set more off
capture log close
capture timer clear

cd /Users/kristinapiorkowski/Chp2"

use " short_data_061917"

log using cmp_boot_wtp_no_selection_062617_new, replace

cmp

global endo_vars "know_2 age age2 race_white inc_50_ab"
global selction_vars "heard_nav heard_other_st conservative_rc opin_union_rc"
global m2_rhs "log_soda_tax log_num_ssb num_fv obesity know_1 gov_trust_never"
capture program drop cmp_prg
program define cmp_prg, rclass
preserve
cmp (log_num_ssb = $endo_vars) ///
   (snav_bi_rc = $selction_vars) ///
   (Yes10 = $m2_rhs) [iweight=weight], ///
   ind($cmp_cont $cmp_probit snav_bi_rc*$cmp_probit) vce(robust)

matrix b_cmp = e(b)
matrix list b_cmp

scalar b1_lst = b_cmp[1,12]
scalar b1_lns = b_cmp[1,13]
scalar b1_nfv = b_cmp[1,14]
scalar b1_obe = b_cmp[1,15]
scalar b1_kn1 = b_cmp[1,16]
scalar b1_gov = b_cmp[1,17]
scalar b1_con = b_cmp[1,18]

//return values
return scalar b_lst = b1_lst
return scalar b_lns = b1_lns
return scalar b_nfv = b1_nfv
return scalar b_obe = b1_obe
return scalar b_kn1 = b1_kn1
return scalar b_gov = b1_gov
return scalar b_con = b1_con
end

capture prog drop BootWTP_062617
prog define BootWTP_062617, rclass
tempvar swtp_cmp_kr swtp_cmp_kr_mean swtp_cmp_kr_median
WTP_mean_99_1 ///
  WTP_median_99_1 WTP_mean_95_1 WTP_median_95_1 WTP_mean_90_1 ///
  WTP_median_90_1 WTP_at_mean_1 WTP_98 WTP_95 WTP_mean_98_2_1 ///
  WTP_median_98_2_1 WTP_99 WTP_95 WTP_90
qui cmp_prg

matrix b_cmp = e(b)

scalar b_lst = b_cmp[1,12]
scalar b_lns = b_cmp[1,13]
scalar b_nfv = b_cmp[1,14]
scalar b_obe = b_cmp[1,15]
scalar b_kn1 = b_cmp[1,16]
scalar b_gov = b_cmp[1,17]
scalar b_con = b_cmp[1,18]

g`swtp_cmp_kr' = exp(-(b_con + b_lns * log_num_ssb + b_nfv * num_fv +
  b_obe * obesity ///
    + b_kn1 * know_1 + b_gov * gov_trust_never) / b_lst)
// Untrimmed
qui sum `swtp_cmp_kr', detail
scalar `swtp_cmp_kr_mean' = r(mean)
scalar `swtp_cmp_kr_median' = r(p50)
scalar `WTP_99' = r(p99)
scalar `WTP_95' = r(p95)
scalar `WTP_90' = r(p90)

// 99% Trimmed
qui sum `swtp_cmp_kr' if `swtp_cmp_kr' < `WTP_99', detail
scalar `WTP_mean_99_1' = r(mean)
scalar `WTP_median_99_1' = r(p50)

// 95% Trimmed
qui sum `swtp_cmp_kr' if `swtp_cmp_kr' < `WTP_95', detail
scalar `WTP_mean_95_1' = r(mean)
scalar `WTP_median_95_1' = r(p50)

// 90% Trimmed
qui sum `swtp_cmp_kr' if `swtp_cmp_kr' < `WTP_90', detail
scalar `WTP_mean_90_1' = r(mean)
scalar `WTP_median_90_1' = r(p50)

// At mean
qui sum log_num_ssb
scalar m1_log_num_ssb = r(mean)
qui sum num_fv
scalar m1_num_fv = r(mean)
qui sum obesity
scalar m1_obesity = r(mean)
qui sum know_1
scalar m1_know_1 = r(mean)
qui sum gov_trust_never
scalar m1_gov_trust_never = r(mean)

scalar `WTP_at_mean_1' = exp(-b_con + b_lns * m1_log_num_ssb + b_nfv * m1_num_fv + b_obe * m1_obesity + b_kn1 * m1_know_1 + b_gov * m1_gov_trust_never) / b_lst)

return scalar WTP_mean = `swtp_cmp_kr_mean'
return scalar WTP_median = `swtp_cmp_kr_median'
return scalar WTP_mean_99 = `WTP_mean_99_1'
return scalar WTP_median_99 = `WTP_median_99_1'
return scalar WTP_mean_95 = `WTP_mean_95_1'
return scalar WTP_median_95 = `WTP_median_95_1'
return scalar WTP_mean_90 = `WTP_mean_90_1'
return scalar WTP_median_90 = `WTP_median_90_1'
return scalar WTP_at_mean = `WTP_at_mean_1'

end

bootstrap WTP_mean = r(WTP_mean) ///
    WTP_mean_99 = r(WTP_mean_99) ///
    WTP_mean_95 = r(WTP_mean_95) ///
    WTP_mean_90 = r(WTP_mean_90) ///
    WTP_median = r(WTP_median) ///
    WTP_median_99 = r(WTP_median_99) ///
    WTP_median_95 = r(WTP_median_95) ///
    WTP_median_90 = r(WTP_median_90) ///
    WTP_at_mean = r(WTP_at_mean) ///
, rep(500) saving(bootVar_no_selection_062217, replace) ///
    seed(1234) : BootWTP_062617

estat bootstrap, all

use bootVar_no_selection_062217, replace

************************* WTP Median Untrimmed *************************

***** 95% CI

sort WTP_median
egen WTP_median_2.5 = pctl(WTP_median), p(2.5)
egen WTP_median_97.5 = pctl(WTP_median), p(97.5)

sum WTP_median_2.5 WTP_median_97.5

sum WTP_median if WTP_median < WTP_median_97.5 & WTP_median > WTP_median_2.5, detail

***** Histogram

graph twoway (histogram WTP_median, ytitle(Density) xtitle(Willingness to Pay (in pennies per ounce))) ///
    (kdensity WTP_median) ///
    (scatteri 0 0.878 .6 0.878, c(l) m(i)) ///
    legend(order(2 3) label(2 "Kernel Density") label(3 "Median"))

graph save Median_WTP_no_selection, replace
graph export Median_WTP_no_selection.pdf, replace

************************* WTP Median 99 Trimmed *************************

204
***** 95% CI
egen WTP_median_99_2_5 = pctl(WTP_median_99), p(2.5)
egen WTP_median_99_97_5 = pctl(WTP_median_99), p(97.5)

sum WTP_median_99_2_5 WTP_median_99_97_5

sum WTP_median_99 if WTP_median_99 < WTP_median_99_97_5 &
WTP_median_99 > WTP_median_99_2_5, detail

************************ WTP Median 95 Trimmed ************************

***** 95% CI
egen WTP_median_95_2_5 = pctl(WTP_median_95), p(2.5)
egen WTP_median_95_97_5 = pctl(WTP_median_95), p(97.5)

sum WTP_median_95_2_5 WTP_median_95_97_5

sum WTP_median_95 if WTP_median_95 < WTP_median_95_97_5 &
WTP_median_95 > WTP_median_95_2_5, detail

************************ WTP Median 90 Trimmed ************************

***** 95% CI
egen WTP_median_90_2_5 = pctl(WTP_median_90), p(2.5)
egen WTP_median_90_97_5 = pctl(WTP_median_90), p(97.5)

sum WTP_median_90_2_5 WTP_median_90_97_5

sum WTP_median_90 if WTP_median_90 < WTP_median_90_97_5 &
WTP_median_90 > WTP_median_90_2_5, detail

************************ WTP Mean Untrimmed ************************

***** 95% CI
sort WTP_mean
egen WTP_mean_2_5 = pctl(WTP_mean), p(2.5)
egen WTP_mean_97_5 = pctl(WTP_mean), p(97.5)

sum WTP_mean_2_5 WTP_mean_97_5

sum WTP_mean if WTP_mean < WTP_mean_97_5 & WTP_mean > WTP_mean_2_5,
detail

************************ WTP Mean 99 Trimmed ************************
***** 95% CI
egen WTP_mean_99_2_5 = pctile(WTP_mean_99), p(2.5)
egen WTP_mean_99_97_5 = pctile(WTP_mean_99), p(97.5)

sum WTP_mean_99_2_5 WTP_mean_99_97_5

sum WTP_mean_99 if WTP_mean_99 < WTP_mean_99_97_5 & WTP_mean_99 > WTP_mean_99_2_5, detail

************************* WTP Mean 95 Trimmed **************************

***** 95% CI
egen WTP_mean_95_2_5 = pctile(WTP_mean_95), p(2.5)
egen WTP_mean_95_97_5 = pctile(WTP_mean_95), p(97.5)

sum WTP_mean_95_2_5 WTP_mean_95_97_5

sum WTP_mean_95 if WTP_mean_95 < WTP_mean_95_97_5 & WTP_mean_95 > WTP_mean_95_2_5, detail

************************* WTP Mean 90 Trimmed **************************

***** 95% CI
egen WTP_mean_90_2_5 = pctile(WTP_mean_90), p(2.5)
egen WTP_mean_90_97_5 = pctile(WTP_mean_90), p(97.5)

sum WTP_mean_90_2_5 WTP_mean_90_97_5

sum WTP_mean_90 if WTP_mean_90 < WTP_mean_90_97_5 & WTP_mean_90 > WTP_mean_90_2_5, detail

log close

*************************************************************************************************************************************************
clear all
capture log close
eststo clear
cd /Users/kristinapiorkowski/Chp3"

use "cleaned_data_103116"

log using ologit_gologit_ppom_mfx_040517, replace

global m1 heard_nav heard_other_st
global m2 heard_nav heard_other_st know_1 know_2 opin_1 num_ssb
global m3 heard_nav heard_other_st know_1 know_2 opin_1 num_ssb moderate ///
   conservative
global m4 heard_nav heard_other_st know_1 know_2 opin_1 num_ssb moderate ///
   conservative edu_med edu_high phone inc_med inc_high race_white ///
   female age b4.metro

****************************** Ordered Logit ******************************
*Makes results in table 3.2.A
eststo clear

eststo: ologit sup_nav $m1 [iweight=weight], vce(robust)
scalar k_1_ol = e(rank)
scalar n_1_ol = e(N)
scalar ll_1_ol = e(ll)
estat ic
matrix aic_1_ol_mat = r(S)
scalar aic_1_ol = aic_1_ol_mat[1,5]
scalar aic1_ol = -2*ll_1_ol + 2*k_1_ol
scalar aicc1_ol = -2*ll_1_ol + 2*k_1_ol +(2*k_1_ol*(k_1_ol+1))/(n_1_ol-k_1_ol-1)
display aic1_ol
display aicc1_ol

eststo: ologit sup_nav $m2 [iweight=weight], vce(robust)
scalar k_2_ol = e(rank)
scalar n_2_ol = e(N)
scalar ll_2_ol = e(ll)
estat ic
matrix aic_2_ol_mat = r(S)
scalar aic_2_ol = aic_2_ol_mat[1,5]
scalar aic2_ol = -2*ll_2_ol + 2*k_2_ol
scalar aicc2_ol = -2*ll_2_ol + 2*k_2_ol + (2*k_2_ol*(k_2_ol+1))/(n_2_ol-k_2_ol-1)
display aic2_ol
display aicc2_ol

eststo: ologit sup_nav $m3 [iweight=weight], vce(robust)
scalar k_3_ol = e(rank)
scalar n_3_ol = e(N)
scalar ll_3_ol = e(ll)
estat ic
matrix aic_3_ol_mat = r(S)
scalar aic_3_ol = aic_3_ol_mat[1,5]
scalar aic3_ol = -2*ll_3_ol + 2*k_3_ol
scalar aicc3_ol = -2*ll_3_ol + 2*k_3_ol + (2*k_3_ol*(k_3_ol+1))/(n_3_ol-k_3_ol-1)
display aic3_ol
display aicc3_ol

eststo xi: ologit sup_nav $m4 [iweight=weight], vce(robust)
scalar k_4_ol = e(rank)
scalar n_4_ol = e(N)
scalar ll_4_ol = e(ll)
estat ic
matrix aic_4_ol_mat = r(S)
scalar aic_4_ol = aic_4_ol_mat[1,5]
scalar aic4_ol = -2*ll_4_ol + 2*k_4_ol
scalar aicc4_ol = -2*ll_4_ol + 2*k_4_ol + (2*k_4_ol*(k_4_ol+1))/(n_4_ol-k_4_ol-1)
display aic4_ol
display aicc4_ol

esttab using ologit_030317.csv, replace se aic bic

**************************** Generalized Ordered Logit ****************************
*Makes the results in table 3.2.B.
eststo clear

eststo: gologit2 sup_nav $m1 [iweight=weight], vce(robust)
scalar k_1_go = e(rank)
scalar n_1_go = e(N)
scalar ll_1_go = e(ll)
estat ic
matrix aic_1_go_mat = r(S)
scalar aic_1_go = aic_1_go_mat[1,5]
scalar aic1_go = -2*ll_1_go + 2*k_1_go
scalar aicc1_go = -2*ll_1_go + 2*k_1_go + (2*k_1_go*(k_1_go+1))/(n_1_go-k_1_go-1)
display aic1_go
display aicc1_go

eststo: gologit2 sup_nav $m2 [iweight=weight], vce(robust)
scalar k_2_go = e(rank)
scalar n_2_go = e(N)
scalar ll_2_go = e(ll)
estat ic
matrix aic_2_go_mat = r(S)
scalar aic_2_go = aic_2_go_mat[1,5]
scalar aic2_go = -2*ll_2_go + 2*k_2_go
scalar aicc2_go = -2*ll_2_go + 2*k_2_go + (2*k_2_go*(k_2_go+1))/(n_2_go-k_2_go-1)
display aic2_go
display aicc2_go

eststo: gologit2 sup_nav $m3 [iweight=weight], vce(robust)
scalar k_3_go = e(rank)
scalar n_3_go = e(N)
scalar ll_3_go = e(ll)
estat ic
matrix aic_3_go_mat = r(S)
scalar aic_3_go = aic_3_go_mat[1,5]
scalar aic3_go = -2*ll_3_go + 2*k_3_go
scalar aicc3_go = -2*ll_3_go + 2*k_3_go + (2*k_3_go*(k_3_go+1))/(n_3_go-k_3_go-1)
display aic3_go
display aicc3_go

eststo: gologit2 sup_nav $m4 [iweight=weight], vce(robust)
scalar k_4_ol = e(rank)
scalar n_4_ol = e(N)
scalar ll_4_ol = e(ll)
estat ic
matrix aic_4_ol_mat = r(S)
scalar aic_4_ol = aic_4_ol_mat[1,5]
scalar aic4_ol = -2*ll_4_ol + 2*k_4_ol
scalar aicc4_ol = -2*ll_4_ol + 2*k_4_ol + (2*k_4_ol*(k_4_ol+1))/(n_4_ol-k_4_ol-1)
display aic4_ol
display aicc4_ol

esttab using gologit_030317.csv, se replace aic bic ///
title(Factors of Support for Navajo Nation Soda Tax: GOlogit w weights)
Partial Proportional Odds Model

*Makes the results in table 3.1

eststo clear

eststo: gologit2 sup_nav $m1 [iweight=weight], vce(robust) ///
    pl(heard_nav) gamma
scalar k_1_ppom = e(rank)
scalar n_1_ppom = e(N)
scalar ll_1_ppom = e(ll)
estat ic
matrix aic_1_ppom_mat = r(S)
scalar aic_1_ppom = aic_1_ppom_mat[1,5]
scalar aic1_ppom = -2*ll_1_ppom + 2*k_1_ppom
scalar aicc1_ppom = -2*ll_1_ppom + 2*k_1_ppom
+(2*k_1_ppom*(k_1_ppom+1))/(n_1_ppom-k_1_ppom-1)
display aicc1_ppom

eststo: gologit2 sup_nav $m2 [iweight=weight], vce(robust) ///
    pl(know_1 num_ssb heard_nav) gamma
scalar k_2_ppom = e(rank)
scalar n_2_ppom = e(N)
scalar ll_2_ppom = e(ll)
estat ic
matrix aic_2_ppom_mat = r(S)
scalar aic_2_ppom = aic_2_ppom_mat[1,5]
scalar aic2_ppom = -2*ll_2_ppom + 2*k_2_ppom
scalar aicc2_ppom = -2*ll_2_ppom + 2*k_2_ppom
+(2*k_2_ppom*(k_2_ppom+1))/(n_2_ppom-k_2_ppom-1)
display aicc2_ppom

eststo: gologit2 sup_nav $m3 [iweight=weight], vce(robust) ///
    pl(know_1 conservative heard_nav num_ssb moderate) gamma
scalar k_3_ppom = e(rank)
scalar n_3_ppom = e(N)
scalar ll_3_ppom = e(ll)
estat ic
matrix aic_3_ppom_mat = r(S)
scalar aic_3_ppom = aic_3_ppom_mat[1,5]
scalar aic3_ppom = -2*ll_3_ppom + 2*k_3_ppom
scalar aicc3_ppom = -2*ll_3_ppom + 2*k_3_ppom
+(2*k_3_ppom*(k_3_ppom+1))/(n_3_ppom-k_3_ppom-1)
display aicc3_ppom
eststo xi: gologit2 sup_nav $m4 [iweight=weight], vce(robust) ///
    npl(heard_other_st opin_1 edu_high phone female age) gamma
scalar k_4_ol = e(rank)
scalar n_4_ol = e(N)
scalar ll_4_ol = e(ll)
estat ic
matrix aic_4_ol_mat = r(S)
scalar aic_4_ol = aic_4_ol_mat[1,5]
scalar aicc4_ol = -2*ll_4_ol + 2*k_4_ol + (2*k_4_ol*(k_4_ol+1))/(n_4_ol-k_4_ol-1)
display aic4_ol
display aicc4_ol

esttab using ppom_OMP_only_040517.csv, se replace aic bic ///
title(Factors of Support for Navajo Nation Soda Tax: PPOM Result)

************************************************ Marginal Effects Results ************************************************
*Makes the results in table 3.3
xi: gologit2 sup_nav i.heard_nav i.heard_other_st i.know_1 i.know_2 i.opin_1 ///
c.num_ssb i.short_pv i.edu_cat i.phone i.inc_cat ///
i.race_white i.female c.age b4.metro [iweight=weight], vce(robust) ///
npl(i.heard_other_st i.opin_1 _Iedu_cat_3 i.phone i.female c.age)
margins, dydx(_Iheard_nav_1 _Iheard_oth_1 _Iknow_1_1 _Iknow_2_1 _Iopin_1_1 ///
c.num_ssb _Ishort_pv_1 _Ishort_pv_2 _Iedu_cat_2 _Iedu_cat_3 ///
_phone_1 _Iinc_cat_1 _Iinc_cat_2 _Irace_whit_1 _Ifemale_1 age ///
0.metro 1.metro 2.metro 3.metro 4.metro) vce(unconditional) post
outreg2 using full_ppom_mfx.xls, excel replace

************************************************ Falsification Test ************************************************
eststo: gologit2 drug_tax $m1 [iweight=weight], vce(robust) autofit
eststo: gologit2 drug_tax $m2 [iweight=weight], vce(robust) autofit
eststo: gologit2 drug_tax $m3 [iweight=weight], vce(robust) autofit
eststo xi: gologit2 drug_tax $m4 [iweight=weight], vce(robust) autofit
esttab using falsification_030317.csv, se replace aic bic title(Falsification Test Using Drug Tax: PPOM Result)

********************************************************************************************************************
********************************************************************************************************************
/* Chapter 4 (Youth Smoking In Nepal) Code * Preliminary Results Multinomial Logit; Basic Results (Probits and Ordered Probit), MLE Trivariate Ordered Probit and Average Marginal Effects, Average Marginal Effects of Conditional Mean for Level of Cigarette Consumption Kristina N. Piorkowski */

****************** Preliminary Multinomial Logit Regressions ******************

* The dependent variable for this regression is `smoke_3' whose categories are
* 1=never smoke, 2 = current smoker, and 3 = former smoker.

*This reproduces the preliminary results found in Appendix B

clear all
set more off
capture log close
cd /Users/kristinapiorkowski/Chp4"
use “GYTS_0711_var_082917"
global m1 friends_smoke_bi parent_smoke_bi

global m2 friends_smoke_bi parent_smoke_bi sp_pca_1 sp_pca_2 sp_pca_3 pt_pca_1

global m3 friends_smoke_bi parent_smoke_bi sp_pca_1 sp_pca_2 sp_pca_3 ///
   b3.cig_weight cig_harm_bi other_cig_harm_bi pt_pca_1 i.edu_quality ///
   ate_v1 fam_smoke_dis

global m4 friends_smoke_bi parent_smoke_bi sp_pca_1 sp_pca_2 sp_pca_3 ///
   b3.cig_weight cig_harm_bi other_cig_harm_bi pt_pca_1 i.edu_quality ///
   ate_v1 fam_smoke_dis i.age_3 i.grade i.year

***** Model 1
mlogit smoke_3 $m1 [pweight = FinalWgt]               , vce(cluster sch_class) // Together
estat ic
outreg2 using mnl_bg_ate_v1_092217.xls, excel ctitle(M1 Together)  ///
   title("Multinomial Logit Results: Boys and Girls Seperate with Never Smokers Base Category Anti-Tob V1") replace

mlogit smoke_3 $m1 [pweight = FinalWgt] if female == 0, vce(cluster sch_class) // B
estat ic
outreg2 using mnl_bg_ate_v1_092217.xls, excel ctitle(M1 Boys) append
mlogit smoke_3 $m1 [pweight = FinalWgt] if female == 1, vce(cluster sch_class) // G
estat ic
outreg2 using mnl_bg_ate_v1_092217.xls, excel ctitle(M1 Girls) append

***** Model 2

mlogit smoke_3 $m2 [pweight = FinalWgt] , vce(cluster sch_class) // Together
estat ic
outreg2 using mnl_bg_ate_v1_092217.xls, excel ctitle(M2 Together) append

mlogit smoke_3 $m2 [pweight = FinalWgt] if female == 0, vce(cluster sch_class) // B
estat ic
outreg2 using mnl_bg_ate_v1_092217.xls, excel ctitle(M2 Boys) append

mlogit smoke_3 $m2 [pweight = FinalWgt] if female == 1, vce(cluster sch_class) // G
estat ic
outreg2 using mnl_bg_ate_v1_092217.xls, excel ctitle(M2 Girls) append

***** Model 3

mlogit smoke_3 $m3 [pweight = FinalWgt] , vce(cluster sch_class) // Together
estat ic
outreg2 using mnl_bg_ate_v1_092217.xls, excel ctitle(M3 Together) append

mlogit smoke_3 $m3 [pweight = FinalWgt] if female == 0, vce(cluster sch_class) // B
estat ic
outreg2 using mnl_bg_ate_v1_092217.xls, excel ctitle(M3 Boys) append

mlogit smoke_3 $m3 [pweight = FinalWgt] if female == 1, vce(cluster sch_class) // G
estat ic
outreg2 using mnl_bg_ate_v1_092217.xls, excel ctitle(M3 Girls) append

***** Model 4

mlogit smoke_3 $m4 female [pweight = FinalWgt] , vce(cluster sch_class) // Together
estat ic
outreg2 using mnl_bg_ate_v1_092217.xls, excel ctitle(M4 Together) append

mlogit smoke_3 $m4 [pweight = FinalWgt] if female == 0, vce(cluster sch_class) // B
estat ic
outreg2 using mnl_bg_ate_v1_092217.xls, excel ctitle(M4 Boys) append

mlogit smoke_3 $m4 [pweight = FinalWgt] if female == 1, vce(cluster sch_class) // G
estat ic

213
outreg2 using mnl_bg_ate_v1_092217.xls, excel ctitle(M4 Girls) append
log close

**************************************************************************** Basic Models ****************************
/*These models produce the same results as the trivariate ordered probit model when the MLE is run WITHOUT bootstrapping the standard errors */
clear all
set more off
capture log close
cd /Users/kristinapiorkowski/Chp4"
use “GYTS_0711_var_082917”

global m3_op female friends_smoke_bi parent_smoke_bi f_parent sp_pca_1 f_sp1 ///
sp_pca_2 f_sp2 sp_pca_3 f_sp3 b3.cig_weight cig_harm_bi f_cigh ///
other_cig_harm_bi pt_pca_1 i.edu_quality f_edu4 ate_v1 fam_smoke_dis

global m3_ff friends_smoke_bi female##parent_smoke_bi female##c.sp_pca_1 ///
female##c.sp_pca_2 female##c.sp_pca_3 female##b3.cig_weight ///
female##cig_harm_bi other_cig_harm_bi pt_pca_1 female##edu_quality ///
ate_v1 fam_smoke_dis

probit ever_smoke $m3_ff [pweight = FinalWgt], vce(cluster sch_class)
probit cur_smoker $m3_ff [pweight = FinalWgt], vce(cluster sch_class)
oprob num_cigs $m3_op [pweight = FinalWgt] if current_ind == 1, vce(cluster sch_class)
log close

**************************************************************************** Trivariate Ordered Probit MLE ****************************
**************************************************************************** AME for Never and Former Smokers ********************

*This reproduces the results found in Table 4.2., 4.3.A., 4.3.B., and 4.3.C

clear all
set more off
capture log close
cd /Users/kristinapiorkowski/Chp4"
use “GYTS_0711_var_082917”
keep ever_smoke friends_smoke_bi parent_smoke_bi female num_cigs cur_smoker ///
   sch_class sp_pca_1 sp_pca_2 sp_pca_3 cig_weight cig_harm_bi ///
   other_cig_harm_bi pt_pca_1 edu_quality ate_v1 fam_smoke_dis FinalWgt ///

***** Set up variables needed for likelihood function
*Recode number of cigs smoked, all obs
replace num_cigs = 1 if num_cigs == . & ever_smoke != .
recode num_cigs (5/7 = 5)
tab num_cigs, generate(nc_)

*Generate Indicator Variable for Never Smoked (S_i=0)
gen never_ind = 0 if ever_smoke != .
replace never_ind = 1 if ever_smoke == 0

*Generate Indicator Variable for Current Smokers (S_i=1, Q_i=1)
gen current_ind = 0 if ever_smoke != .
replace current_ind = 1 if cur_smoker == 1

*Generate Indicator Variable for Former Smokers (S_i=1, Q_i=0)
gen former_ind = 0 if ever_smoke != .
replace former_ind = 1 if cur_smoker == 0

*Create Interaction Variables
gen edu = edu_quality
gen never_smoke = ever_smoke
recode never_smoke (1 = 2)
recode never_smoke (0 = 1)
recode never_smoke (2 = 0)
gen former_smoker = cur_smoker
recode former_smoker (1 = 2)
recode former_smoker (0 = 1)
recode former_smoker (2 = 0)

*Drop missing and unneeded variables for mle program
drop if ever_smoke == .
drop if friends_smoke_bi == .
drop if parent_smoke_bi == .
drop if female == .
drop if sp_pca_1 == .
drop if sp_pca_2 == .
drop if sp_pca_3 == .
drop if cig_weight == .
drop if cig_harm_bi == .
drop if other_cig_harm_bi==.
drop if pt_pca_1 ==.
drop if edu_quality ==.
drop if ate_v1 ==.
drop if fam_smoke_dis ==.
drop if num_cigs ==.
drop if current_ind ==.
drop if former_ind ==.
drop if never_ind ==.

global vars i.friends_smoke_bi i.female i.parent_smoke_bi c.sp_pca_1 ///
c.sp_pca_2 c.sp_pca_3 b3.cig_weight i.cig_harm_bi i.other_cig_harm_bi ///
c.pt_pca_1 i.edu_quality i.ate_v1 i.fam_smoke_dis

global never i.friends_smoke_bi i.female##i.parent_smoke_bi i.female##c.sp_pca_1 ///
i.female##c.sp_pca_2 i.female##c.sp_pca_3 i.female##b3.cig_weight ///
i.female##i.cig_harm_bi i.other_cig_harm_bi c.pt_pca_1 ///
i.female##edu_quality i.ate_v1 i.fam_smoke_dis

global quitter i.friends_smoke_bi i.female##i.parent_smoke_bi i.female##c.sp_pca_1 ///
i.female##c.sp_pca_2 i.female##c.sp_pca_3 i.female##b3.cig_weight ///
i.female##i.cig_harm_bi i.other_cig_harm_bi c.pt_pca_1 ///
i.female##edu_quality i.ate_v1 i.fam_smoke_dis

global level i.friends_smoke_bi female##parent_smoke_bi female##c.sp_pca_1 ///
female##c.sp_pca_2 female##c.sp_pca_3 b3.cig_weight ///
female##cig_harm_bi i.other_cig_harm_bi c.pt_pca_1 ///
i.edu_quality i(1)bn.female##i(3)bn.edu i.ate_v1 i.fam_smoke_dis

{ capture program drop trivariate_model
  program define trivariate_model
    args lnL xa xb xc t2 t3 t4
    tempvar p_nsmoke p_smoke pa p_fsmoke p_csmoke pb p_count v2 v3 v4 p2 p3 p4 p5

    quietly gen double `pa' = normprob(`xa')
    quietly gen double `p_nsmoke' = never_ind * ln(`pa')
    quietly gen double `p_smoke' = (1 - never_ind) * ln(1 - `pa')

    quietly gen double `pb' = normprob(`xb')
    quietly gen double `p_fsmoke' = former_ind * ln(`pb')
    quietly gen double `p_csmoke' = (1 - former_ind) * ln(1 - `pb')

    quietly gen double `v2' = `t2' - (`xc')
    quietly gen double `v3' = `t3' - (`xc')
quietly gen double `v4' = `t4' - ('xc')
quietly gen double `p2' = normprob(`v2')
quietly gen double `p3' = normprob(`v3') - normprob(`v2')
quietly gen double `p4' = normprob(`v4') - normprob(`v3')
quietly gen double `p5' = 1 - normprob(`v4')
quietly gen double `p_count' = nc_2*ln(`p2') + nc_3*ln(`p3') + nc_4*ln(`p4') +
  nc_5*ln(`p5')
quietly replace `lnL' = never_ind * (`p_nsmoke') + /// //never smokers
  former_ind * (`p_smoke' + `p_fsmoke') + /// //quitters
  current_ind * (`p_smoke' + `p_csmoke' + `p_count') // quantity
end
}
capture program trivariate_model_ame drop
program trivariate_model_ame
ml model lf trivariate_model ///
  (xa:never_smoke = $never if never_ind == 1) ///
  (xb:former_ind = $quitter if former_ind == 1) ///
  (xc:num_cigs = $level, noconstant if current_ind == 1) /cut1 /cut2 /cut3 ///
  [pweight = FinalWgt], vce(cluster sch_class)
ml search
ml maximize, difficult iterate(5000)
end
bootstrap, rep(200) seed(1): trivariate_model_ame
margins , dydx($vars) ///
expression(normal(predict(eq(xa), `xa')))\n
margins , dydx($vars) ///
expression(normal(predict(eq(xa), `xa'))) over(female)
margins , dydx($vars) ///
expression((1 - normal(predict(eq(xa), `xa'))) * (normal(predict(eq(xb), `xb'))))\n
margins , dydx($vars) ///
expression((1 - normal(predict(eq(xa), `xa'))) * (normal(predict(eq(xb), `xb')))) ///
over(female)
log close

************ AME for Conditional Mean for Level of Cigarettes Smoked ************
clear all
set more off
capture log close

cd /Users/kristinapiorkowski/Chp4"

use “GYTS_0711_var_082917”

keep ever_smoke friends_smoke_bi parent_smoke_bi female num_cigs cur_smoker ///
sch_class sp_pca_1 sp_pca_2 sp_pca_3 cig_weight cig_harm_bi ///
other_cig_harm_bi pt_pca_1 edu_quality ate_v1 fam_smoke_dis FinalWgt ///

***** Set up variables needed for likelihood function
*Recode number of cigs smoked, all obs
replace num_cigs = 1 if num_cigs == . & ever_smoke != .
recode num_cigs (5/7 = 5)
tab num_cigs, generate(nc_)

*Generate Indicator Variable for Never Smoked (S_i=0)
gen never_ind = 0 if ever_smoke != .
replace never_ind = 1 if ever_smoke == 0

*Generate Indicator Variable for Current Smokers (S_i=1, Q_i=1)
gen current_ind = 0 if ever_smoke != .
replace current_ind = 1 if cur_smoker == 1

*Generate Indicator Variable for Former Smokers (S_i=1, Q_i=0)
gen former_ind = 0 if ever_smoke != .
replace former_ind = 1 if cur_smoker == 0

*Create Interaction Variables
gen edu = edu_quality

gen never_smoke = ever_smoke
recode never_smoke (1 = 2)
recode never_smoke (0 = 1)
recode never_smoke (2 = 0)

gen former_smoker = cur_smoker
recode former_smoker (1 = 2)
recode former_smoker (0 = 1)
recode former_smoker (2 = 0)

*Create categories for variables
tab cig_weight, generate(cw_)
tab edu_quality, generate(ed_)

*Create Interaction Variables
gen f_parent = female * parent_smoke_bi
gen f_sp1 = female * sp_pca_1
gen f_sp2 = female * sp_pca_2
gen f_sp3 = female * sp_pca_3
gen f_cw1 = female * cw_1
gen f_cw2 = female * cw_2
gen f_cigh = female * cig_harm_bi
gen f_edu2 = female * ed_2
gen f_edu3 = female * ed_3
gen f_edu4 = female * ed_4

*Create y_bar_m
gen cig_cat_mean = .
replace cig_cat_mean = 0 if num_cigs == 1
replace cig_cat_mean = .5 if num_cigs == 2
replace cig_cat_mean = 1 if num_cigs == 3
replace cig_cat_mean = 3.5 if num_cigs == 4
replace cig_cat_mean = 10 if num_cigs == 5

tab cig_cat_mean, gen(ccm_)
replace ccm_1 = 0
replace ccm_2 = 0.5
replace ccm_3 = 1
replace ccm_4 = 3.5
replace ccm_5 = 10

*Drop missing and unneeded variables for mle program
drop if ever_smoke == .
drop if friends_smoke_bi == .
drop if parent_smoke_bi == .
drop if female == .
drop if sp_pca_1 == .
drop if sp_pca_2 == .
drop if sp_pca_3 == .
drop if cig_weight == .
drop if cig_harm_bi == .
drop if other_cig_harm_bi== .
drop if pt_pca_1 == .
drop if edu_quality == .
drop if ate_v1 == .
drop if fam_smoke_dis == .
drop if num_cigs == .
drop if current_ind == .
drop if former_ind == .
drop if never_ind == .

global vars i.friends_smoke_bi i.female i.parent_smoke_bi c.sp_pca_1 ///
c.sp_pca_2 c.sp_pca_3 b3.cig_weight i.cig_harm_bi i.other_cig_harm_bi ///
c.pt_pca_1 i.edu_quality i.ate_v1 i.fam_smoke_dis

global never i.friends_smoke_bi i.female##i.parent_smoke_bi i.female##c.sp_pca_1 ///
i.female##c.sp_pca_2 i.female##c.sp_pca_3 i.female##b3.cig_weight ///
i.female##i.cig_harm_bi i.other_cig_harm_bi c.pt_pca_1 ///
i.female##edu_quality i.ate_v1 i.fam_smoke_dis

global quitter i.friends_smoke_bi i.female##i.parent_smoke_bi i.female##c.sp_pca_1 ///
i.female##c.sp_pca_2 i.female##c.sp_pca_3 i.female##b3.cig_weight ///
i.female##i.cig_harm_bi i.other_cig_harm_bi c.pt_pca_1 ///
i.female##edu_quality i.ate_v1 i.fam_smoke_dis

global level i.friends_smoke_bi i.female##parent_smoke_bi i.female##c.sp_pca_1 ///
female##c.sp_pca_2 female##c.sp_pca_3 b3.cig_weight ///
female##i.cig_harm_bi i.other_cig_harm_bi c.pt_pca_1 ///
i.edu_quality i(1)bn.female##i(3)bn.edu i.ate_v1 i.fam_smoke_dis

*****
{
    capture program drop trivariate_model
    program define trivariate_model
        args lnL xa xb xc t2 t3 t4
        tempvar p_nsmoke p_smoke pa p_fsmoke p_csmoke pb p_count v2 v3 v4 p2 p3 p4 p5
        quietly gen double `pa' = normprob(`xa')
        quietly gen double `p_nsmoke' = never_ind * ln(`pa')
        quietly gen double `p_smoke' = (1 - never_ind) * ln(1 - `pa')

        quietly gen double `pb' = normprob(`xb')
        quietly gen double `p_fsmoke' = former_ind * ln(`pb')
        quietly gen double `p_csmoke' = (1 - former_ind) * ln(1 - `pb')

        quietly gen double `v2' = `t2' - (`xc')
        quietly gen double `v3' = `t3' - (`xc')
        quietly gen double `v4' = `t4' - (`xc')
        quietly gen double `p2' = normprob(`v2')
        quietly gen double `p3' = normprob(`v3') - normprob(`v2')
        quietly gen double `p4' = normprob(`v4') - normprob(`v3')
        quietly gen double `p5' = 1 - normprob(`v4')
}

****
quietly gen double `p_count' = nc_2*ln(`p2') + nc_3*ln(`p3') + nc_4*ln(`p4') +
nc_5*ln(`p5')

quietly replace `lnL' = never_ind   * (`p_nsmoke') + /// //never smokers
    former_ind  * (`p_smoke' + `p fsmoke') + /// //quitters
    current_ind * (`p_smoke' + `p csmoke' + `p count') // quantity

end

******
capture program trivariate_model_est drop
program trivariate_model_est
	ml model lf trivariate_model ///
    (xa:never_smoke = $never if never_ind == 1) ///
    (xb:former_ind = $quitter if former_ind == 1) ///
    (xc:num_cigs = $level, noconstant if current_ind == 1) /cut1 /cut2 /cut3 ///
[pweight = FinalWgt], vce(cluster sch_class) //technique(bhhh bfgs)

ml search
ml maximize, difficult iterate(5000)
end

******
capture prog drop level_con_w_int_marginal
prog define level_con_w_int_marginal, rclass
tempvar xa pa qa xb pb qb xc v1 v2 v3 p1 p2 p3 q1 q2 q3 q4 tog_sp1_a ///
    fem_sp1_a mal_sp1_a tog_sp1_b fem_sp1_b mal_sp1_b tog_sp1_c fem_sp1_c ///
    mal_sp1_c tog_sp2_a fem_sp2_a mal_sp2_a tog_sp2_b fem_sp2_b mal_sp2_b ///
    tog_sp2_c fem_sp2_c mal_sp2_c tog_sp3_a fem_sp3_a mal_sp3_a tog_sp3_b ///
    fem_sp3_b mal_sp3_b tog_sp3_c fem_sp3_c mal_sp3_c me_sp1_1 ///
    me_sp1_1_f me_sp1_1_m me_sp2_1 me_sp2_1_f me_sp2_1_m me_sp3_1 ///
    me_sp3_1_f me_sp3_1_m me_sp1_2 me_sp1_2_f me_sp1_2_m me_sp2_2 ///
    me_sp2_2_f me_sp2_2_m me_sp3_2 me_sp3_2_f me_sp3_2_m me_sp1_3 ///
    me_sp1_3_f me_sp1_3_m me_sp2_3 me_sp2_3_f me_sp2_3_m me_sp3_3 ///
    me_sp3_3_f me_sp3_3_m me_sp1_4 me_sp1_4_f me_sp1_4_m me_sp2_4 ///
    me_sp2_4_f me_sp2_4_m me_sp3_4 me_sp3_4_f me_sp3_4_m
tqui trivariate_model_est

matrix b_tri = e(b)
matrix list b_tri

scalar a0_constant_cons      = b_tri[1,54]
scalar a1_female_cons        = b_tri[1,4]
scalar a2_friends_cons = b_tri[1,2]
scalar a3_parents_cons = b_tri[1,6]
scalar a4_f_parent_cons = b_tri[1,10]
scalar a5_sp1_cons = b_tri[1,11]
scalar a6_f_sp1_cons = b_tri[1,13]
scalar a7_sp2_cons = b_tri[1,14]
scalar a8_f_sp2_cons = b_tri[1,16]
scalar a9_sp3_cons = b_tri[1,17]
scalar a10_f_sp3_cons = b_tri[1,19]
scalar a11_cw1_cons = b_tri[1,20]
scalar a12_f_cw1_cons = b_tri[1,26]
scalar a13_cw2_cons = b_tri[1,21]
scalar a14_f_cw2_cons = b_tri[1,27]
scalar a15_cigharm_cons = b_tri[1,30]
scalar a16_f_cigharm_cons = b_tri[1,34]
scalar a17_othercigharm_cons = b_tri[1,36]
scalar a18_pte1_cons = b_tri[1,37]
scalar a19_edu2_cons = b_tri[1,39]
scalar a20_f_edu2_cons = b_tri[1,47]
scalar a21_edu3_cons = b_tri[1,40]
scalar a22_f_edu3_cons = b_tri[1,48]
scalar a23_edu4_cons = b_tri[1,41]
scalar a24_f_edu_cons = b_tri[1,49]
scalar a25_ate_cons = b_tri[1,51]
scalar a26_famdis_cons = b_tri[1,53]
scalar b0_constant_cons = b_tri[1,108]
scalar b1_female_cons = b_tri[1,58]
scalar b2_friends_cons = b_tri[1,56]
scalar b3_parents_cons = b_tri[1,60]
scalar b4_f_parent_cons = b_tri[1,64]
scalar b5_sp1_cons = b_tri[1,65]
scalar b6_f_sp1_cons = b_tri[1,67]
scalar b7_sp2_cons = b_tri[1,68]
scalar b8_f_sp2_cons = b_tri[1,70]
scalar b9_sp3_cons = b_tri[1,71]
scalar b10_f_sp3_cons = b_tri[1,73]
scalar b11_cw1_cons = b_tri[1,74]
scalar b12_f_cw1_cons = b_tri[1,80]
scalar b13_cw2_cons = b_tri[1,75]
scalar b14_f_cw2_cons = b_tri[1,81]
scalar b15_cigharm_cons = b_tri[1,84]
scalar b16_f_cigharm_cons = b_tri[1,88]
scalar b17_othercigharm_cons = b_tri[1,90]
scalar b18_pte1_cons = b_tri[1,91]
scalar b19_edu2_cons = b_tri[1,93]
scalar b20_f_edu2_cons = b_tri[1,101]
scalar b21_edu3_cons = b_tri[1,94]
scalar b22_f_edu3_cons = b_tri[1,102]
scalar b23_edu4_cons = b_tri[1,95]
scalar b24_f_edu_cons = b_tri[1,103]
scalar b25_ate_cons = b_tri[1,105]
scalar b26_famdis_cons = b_tri[1,107]

scalar c1_female_cons = b_tri[1,111]
scalar c2_friends_cons = b_tri[1,110]
scalar c3_parents_cons = b_tri[1,113]
scalar c4_fparent_cons = b_tri[1,115]
scalar c5_sp1_cons = b_tri[1,116]
scalar c6_fsp1_cons = b_tri[1,117]
scalar c7_sp2_cons = b_tri[1,118]
scalar c8_fsp2_cons = b_tri[1,119]
scalar c9_sp3_con = b_tri[1,120]
scalar c10_fsp3_cons = b_tri[1,121]
scalar c11_cw1_cons = b_tri[1,122]
scalar c13_cw2_cons = b_tri[1,123]
scalar c15_cigharm_cons = b_tri[1,126]
scalar c16_fcigh_cons = b_tri[1,128]
scalar c17_othercigharm_cons = b_tri[1,130]
scalar c18_pte1_cons = b_tri[1,131]
scalar c19_edu2_cons = b_tri[1,133]
scalar c21_edu3_cons = b_tri[1,134]
scalar c23_edu4_cons = b_tri[1,135]
scalar c24_fedu4_cons = b_tri[1,136]
scalar c25_ate_cons = b_tri[1,138]
scalar c26_famdis_cons = b_tri[1,140]
scalar cut1_cons = b_tri[1,141]
scalar cut2_cons = b_tri[1,142]
scalar cut3_cons = b_tri[1,143]

*For Ever Smokers
quietly gen double `xa' = a0_constant_cons ///
  + a1_female_cons * female ///
  + a2_friends_cons * friends_smoke_bi ///
  + a3_parents_cons * parent_smoke_bi ///
  + a4_fparent_cons * f_parent ///
  + a5_sp1_cons * sp_pca_1 ///
  + a6_fsp1_cons * f_sp1 ///
  + a7_sp2_cons * sp_pca_2 ///
  + a8_fsp2_cons * f_sp2 ///
  + a9_sp3_cons * sp_pca_3 ///
  + a10_fsp3_cons * f_sp3 ///
+ a11_cw1_cons * cw_1 ///
+ a12_f_cw1_cons * f_cw1 ///
+ a13_cw2_cons * cw_2 ///
+ a14_f_cw2_cons * f_cw2 ///
+ a15_cigharm_cons * cig_harm_bi ///
+ a16_f_cigharm_cons * f_cigh ///
+ a17_othercigharm_cons * other_cig_harm_bi ///
+ a18_pte1_cons * pt_pca_1 ///
+ a19_edu2_cons * ed_2 ///
+ a20_f_edu2_cons * f_edu2 ///
+ a21_edu3_cons * ed_3 ///
+ a22_f_edu3_cons * f_edu3 ///
+ a23_edu4_cons * ed_4 ///
+ a24_f_edu4_cons * f_edu4 ///
+ a25_ate_cons * ate_v1 ///
+ a26_famdis_cons * fam_smoke_dis

quietly gen double `pa' = normalden(`xa')
quietly gen double `qa' = 1 - normprob(`xa')

*For Current Smoker Portion
quietly gen double `xb' = b0_constant_cons ///
+ b1_female_cons * female ///
+ b2_friends_cons * friends_smoke_bi ///
+ b3_parents_cons * parent_smoke_bi ///
+ b4_f_parent_cons * f_parent ///
+ b5_sp1_cons * sp_pca_1 ///
+ b6_f_sp1_cons * f_sp1 ///
+ b7_sp2_cons * sp_pca_2 ///
+ b8_f_sp2_cons * f_sp2 ///
+ b9_sp3_cons * sp_pca_3 ///
+ b10_f_sp3_cons * f_sp3 ///
+ b11_cw1_cons * cw_1 ///
+ b12_f_cw1_cons * f_cw1 ///
+ b13_cw2_cons * cw_2 ///
+ b14_f_cw2_cons * f_cw2 ///
+ b15_cigharm_cons * cig_harm_bi ///
+ b16_f_cigharm_cons * f_cigh ///
+ b17_othercigharm_cons * other_cig_harm_bi ///
+ b18_pte1_cons * pt_pca_1 ///
+ b19_edu2_cons * ed_2 ///
+ b20_f_edu2_cons * f_edu2 ///
+ b21_edu3_cons * ed_3 ///
+ b22_f_edu3_cons * f_edu3 ///
+ b23_edu4_cons * ed_4 ///
+ b24_f_edu4_cons * f_edu4 ///
+ b25_ate_cons * ate_v1 ///
quietly gen double `pb' = normalden(`xb')
quietly gen double `qb' = 1 - normprob(`xb')

*For Number of Cigarettes Smoked Portion
quietly gen double `xc' = c1_female_cons * female ///
  + c2_friends_cons * friends_smoke_bi ///
  + c3_parents_cons * parent_smoke_bi ///
  + c4_fparent_cons * f_parent ///
  + c5_sp1_cons * sp_pca_1 ///
  + c6_fsp1_cons * f_sp1 ///
  + c7_sp2_cons * sp_pca_2 ///
  + c8_fsp2_cons * f_sp2 ///
  + c9_sp3_cons * sp_pca_3 ///
  + c10_fsp3_cons * f_sp3 ///
  + c11_cw1_cons * cw_1 ///
  + c12_cw2_cons * cw_2 ///
  + c13_cw3_cons * cw_3 ///
  + c14_cw4_cons * cw_4 ///
  + c15_cigharm_cons * cig_harm_bi ///
  + c16_fcigh_cons * f_cigh ///
  + c17_othercigharm_cons * other_cig_harm_bi ///
  + c18_ptel1_cons * pt_pca_1 ///
  + c19_edu2_cons * ed_2 ///
  + c20_edu3_cons * ed_3 ///
  + c21_edu4_cons * ed_4 ///
  + c22_fedu4_cons * f_edu4 ///
  + c23_ate_cons * ate_v1 ///
  + c24_fsp3_cons * f_sp3 ///
  + c25_ate_cons * ate_v1 ///
quietly gen double `v1' = cut1_cons - (`xc')
quietly gen double `v2' = cut2_cons - (`xc')
quietly gen double `v3' = cut3_cons - (`xc')

quietly gen double `p1' = normalden(`v1') * (-1) // this has a negtive trick - see derivation
quietly gen double `p2' = normalden(`v1') - normalden(`v2')
quietly gen double `p3' = normalden(`v2') - normalden(`v3')
quietly gen double `p4' = normalden(`v3')

quietly gen double `q1' = normprob(`v1')
quietly gen double `q2' = normprob(`v2') - normprob(`v1')
quietly gen double `q3' = normprob(`v3') - normprob(`v2')
quietly gen double `q4' = 1 - normprob(`v3')

*SP1
quietly gen `tog_sp1_a' = a5_sp1_cons + a6_f_sp1_cons * female
quietly gen `fem_sp1_a' = a5_sp1_cons + a6_f_sp1_cons * female if female == 1
quietly gen `mal_sp1_a' = a5_sp1_cons + a6_f_sp1_cons * female if female == 0
quietly gen `tog_sp1_b' = b5_sp1_cons + b6_f_sp1_cons * female
quietly gen `fem_sp1_b' = b5_sp1_cons + b6_f_sp1_cons * female if female == 1
quietly gen `mal_sp1_b' = b5_sp1_cons + b6_f_sp1_cons * female if female == 0

quietly gen `tog_sp1_c' = c5_sp1_cons + c6_fsp1_cons * female
quietly gen `fem_sp1_c' = c5_sp1_cons + c6_fsp1_cons * female if female == 1
quietly gen `mal_sp1_c' = c5_sp1_cons + c6_fsp1_cons * female if female == 0

*SP2
quietly gen `tog_sp2_a' = a7_sp2_cons + a8_f_sp2_cons * female
quietly gen `fem_sp2_a' = a7_sp2_cons + a8_f_sp2_cons * female if female == 1
quietly gen `mal_sp2_a' = a7_sp2_cons + a8_f_sp2_cons * female if female == 0

quietly gen `tog_sp2_b' = b7_sp2_cons + b8_f_sp2_cons * female
quietly gen `fem_sp2_b' = b7_sp2_cons + b8_f_sp2_cons * female if female == 1
quietly gen `mal_sp2_b' = b7_sp2_cons + b8_f_sp2_cons * female if female == 0

quietly gen `tog_sp2_c' = c7_sp2_cons + c8_fsp2_cons * female
quietly gen `fem_sp2_c' = c7_sp2_cons + c8_fsp2_cons * female if female == 1
quietly gen `mal_sp2_c' = c7_sp2_cons + c8_fsp2_cons * female if female == 0

*SP3
quietly gen `tog_sp3_a' = a9_sp3_cons + a10_f_sp3_cons * female
quietly gen `fem_sp3_a' = a9_sp3_cons + a10_f_sp3_cons * female if female == 1
quietly gen `mal_sp3_a' = a9_sp3_cons + a10_f_sp3_cons * female if female == 0

quietly gen `tog_sp3_b' = b9_sp3_cons + b10_f_sp3_cons * female
quietly gen `fem_sp3_b' = b9_sp3_cons + b10_f_sp3_cons * female if female == 1
quietly gen `mal_sp3_b' = b9_sp3_cons + b10_f_sp3_cons * female if female == 0

quietly gen `tog_sp3_c' = c9_sp3_cons + c10_fsp3_cons * female
quietly gen `fem_sp3_c' = c9_sp3_cons + c10_fsp3_cons * female if female == 1
quietly gen `mal_sp3_c' = c9_sp3_cons + c10_fsp3_cons * female if female == 0

local Ames1 `me_sp1_1' `me_sp1_1_f' `me_sp1_1_m' `me_sp2_1' `me_sp2_1_f' `me_sp2_1_m' 
local Ames2 `me_sp1_2' `me_sp1_2_f' `me_sp1_2_m' `me_sp2_2' `me_sp2_2_f' `me_sp2_2_m' 
local Ames3 `me_sp1_3' `me_sp1_3_f' `me_sp1_3_m' `me_sp2_3' `me_sp2_3_f' `me_sp2_3_m' 
local Ames4 `me_sp1_4' `me_sp1_4_f' `me_sp1_4_m' `me_sp2_4' `me_sp2_4_f' `me_sp2_4_m'

forvalues i = 1/4{
   foreach x of local Ames'i'{
      quietly gen double `x' = .
}
local i = `i' + 1

*****SP 1:

*Together
quietly replace `me_sp1_4' = ccm_5 * (`q4'*`qa'*`qb')* ((`tog_sp1_c'*`p4')/(`q4')) ///
    - ((`tog_sp1_a'*`pa')/(`qa')) - ((`tog_sp1_b'*`pb')/(`qb'))
quietly sum `me_sp1_4'
scalar mesp14_mean = r(mean)
return scalar ame_sp1_4 = mesp14_mean

quietly replace `me_sp1_3' = ccm_4 * (`q3'*`qa'*`qb')* ((`tog_sp1_c'*`p3')/(`q3')) ///
    - ((`tog_sp1_a'*`pa')/(`qa')) - ((`tog_sp1_b'*`pb')/(`qb'))
quietly sum `me_sp1_3'
scalar mesp13_mean = r(mean)
return scalar ame_sp1_3 = mesp13_mean

quietly replace `me_sp1_2' = ccm_3 * (`q2'*`qa'*`qb')* ((`tog_sp1_c'*`p2')/(`q2')) ///
    - ((`tog_sp1_a'*`pa')/(`qa')) - ((`tog_sp1_b'*`pb')/(`qb'))
quietly sum `me_sp1_2'
scalar mesp12_mean = r(mean)
return scalar ame_sp1_2 = mesp12_mean

quietly replace `me_sp1_1' = ccm_2 * (`q1'*`qa'*`qb')* ((`tog_sp1_c'*`p1')/(`q1')) ///
    - ((`tog_sp1_a'*`pa')/(`qa')) - ((`tog_sp1_b'*`pb')/(`qb'))
quietly sum `me_sp1_1'
scalar mesp11_mean = r(mean)
return scalar ame_sp1_1 = mesp11_mean

*Girls
quietly replace `me_sp1_4_f' = ccm_5 * (`q4'*`qa'*`qb')* ((`fem_sp1_c'*`p4')/(`q4')) ///
    - ((`fem_sp1_a'*`pa')/(`qa')) - ((`fem_sp1_b'*`pb')/(`qb'))
quietly sum `me_sp1_4_f'
scalar mesp14f_mean = r(mean)
return scalar ame_sp1_4_f = mesp14f_mean

quietly replace `me_sp1_3_f' = ccm_4 * (`q3'*`qa'*`qb')* ((`fem_sp1_c'*`p3')/(`q3')) ///
    - ((`fem_sp1_a'*`pa')/(`qa')) - ((`fem_sp1_b'*`pb')/(`qb'))
quietly sum `me_sp1_3_f'
scalar mesp13f_mean = r(mean)
return scalar ame_sp1_3_f = mesp13f_mean

quietly replace `me_sp1_2_f' = ccm_3 * (`q2'*`qa'*`qb')* ((`fem_sp1_c'*`p2')/(`q2')) ///
    - ((`fem_sp1_a'*`pa')/(`qa')) - ((`fem_sp1_b'*`pb')/(`qb'))
quietly sum `me_sp1_2_f'
scalar mesp12f_mean = r(mean)
return scalar ame_sp1_2_f = mesp12f_mean
quietly sum `me_sp1_2_f'
scalar mesp12f_mean = r(mean)
return scalar ame_sp1_2_f = mesp12f_mean

quietly replace `me_sp1_1_f' = ccm_2 * (`q1'*`qa'*`qb')* ((`fem_sp1_c'*`p1')/(`q1') ///
   - (`fem_sp1_a'*`pa')/(`qa')) - ((`fem_sp1_b'*`pb')/(`qb'))) 
quietly sum `me_sp1_1_f'
scalar mesp11f_mean = r(mean)
return scalar ame_sp1_1_f = mesp11f_mean

*Boys
quietly replace `me_sp1_4_m' = ccm_5 * (`q4'*`qa'*`qb')* ((`mal_sp1_c'*`p4')/(`q4') ///
   - (`mal_sp1_a'*`pa')/(`qa')) - ((`mal_sp1_b'*`pb')/(`qb'))) 
quietly sum `me_sp1_4_m'
scalar mesp14m_mean = r(mean)
return scalar ame_sp1_4_m = mesp14m_mean

quietly replace `me_sp1_3_m' = ccm_4 * (`q3'*`qa'*`qb')* ((`mal_sp1_c'*`p3')/(`q3') ///
   - (`mal_sp1_a'*`pa')/(`qa')) - ((`mal_sp1_b'*`pb')/(`qb'))) 
quietly sum `me_sp1_3_m'
scalar mesp13m_mean = r(mean)
return scalar ame_sp1_3_m = mesp13m_mean

quietly replace `me_sp1_2_m' = ccm_3 * (`q2'*`qa'*`qb')* ((`mal_sp1_c'*`p2')/(`q2') ///
   - (`mal_sp1_a'*`pa')/(`qa')) - ((`mal_sp1_b'*`pb')/(`qb'))) 
quietly sum `me_sp1_2_m'
scalar mesp12m_mean = r(mean)
return scalar ame_sp1_2_m = mesp12m_mean

quietly replace `me_sp1_1_m' = ccm_2 * (`q1'*`qa'*`qb')* ((`mal_sp1_c'*`p1')/(`q1') ///
   - (`mal_sp1_a'*`pa')/(`qa')) - ((`mal_sp1_b'*`pb')/(`qb'))) 
quietly sum `me_sp1_1_m'
scalar mesp11m_mean = r(mean)
return scalar ame_sp1_1_m = mesp11m_mean

******SP 2:

*Together
quietly replace `me_sp2_4' = ccm_5 * (`q4'*`qa'*`qb')* ((`tog_sp2_c'*`p4')/(`q4') ///
   - (`tog_sp2_a'*`pa')/(`qa')) - ((`tog_sp2_b'*`pb')/(`qb'))) 
quietly sum `me_sp2_4'
scalar mesp24_mean = r(mean)
return scalar ame_sp2_4 = mesp24_mean

quietly replace `me_sp2_3' = ccm_4 * (`q3'*`qa'*`qb')* ((`tog_sp2_c'*`p3')/(`q3') ///
   - (`tog_sp2_a'*`pa')/(`qa')) - ((`tog_sp2_b'*`pb')/(`qb'))) 

228
quietly sum `me_sp2_3'
scalar mesp23_mean = r(mean)
return scalar ame_sp2_3 = mesp23_mean

quietly replace `me_sp2_2' = ccm_3 * (`q2'*`qa'*`qb')* (('tog_sp2_c'*`p2')/(`q2')) ///
- (('tog_sp2_a'*pa')/(`qa')) - (('tog_sp2_b'*pb')/(`qb'))
quietly sum `me_sp2_2'
scalar mesp22_mean = r(mean)
return scalar ame_sp2_2 = mesp22_mean

quietly replace `me_sp2_1' = ccm_2 * (`q1'*`qa'*`qb')* (('tog_sp2_c'*`p1')/(`q1')) ///
- (('tog_sp2_a'*pa')/(`qa')) - (('tog_sp2_b'*pb')/(`qb'))
quietly sum `me_sp2_1'
scalar mesp21_mean = r(mean)
return scalar ame_sp2_1 = mesp21_mean

*Girls
quietly replace `me_sp2_4_f' = ccm_5 * (`q4'*`qa'*`qb')* (('fem_sp2_c'*`p4')/(`q4')) ///
- (('fem_sp2_a'*pa')/(`qa')) - (('fem_sp2_b'*pb')/(`qb'))
quietly sum `me_sp2_4_f'
scalar mesp24f_mean = r(mean)
return scalar ame_sp2_4_f = mesp24f_mean

quietly replace `me_sp2_3_f' = ccm_4 * (`q3'*`qa'*`qb')* (('fem_sp2_c'*`p3')/(`q3')) ///
- (('fem_sp2_a'*pa')/(`qa')) - (('fem_sp2_b'*pb')/(`qb'))
quietly sum `me_sp2_3_f'
scalar mesp23f_mean = r(mean)
return scalar ame_sp2_3_f = mesp23f_mean

quietly replace `me_sp2_2_f' = ccm_3 * (`q2'*`qa'*`qb')* (('fem_sp2_c'*`p2')/(`q2')) ///
- (('fem_sp2_a'*pa')/(`qa')) - (('fem_sp2_b'*pb')/(`qb'))
quietly sum `me_sp2_2_f'
scalar mesp22f_mean = r(mean)
return scalar ame_sp2_2_f = mesp22f_mean

quietly replace `me_sp2_1_f' = ccm_2 * (`q1'*`qa'*`qb')* (('fem_sp2_c'*`p1')/(`q1')) ///
- (('fem_sp2_a'*pa')/(`qa')) - (('fem_sp2_b'*pb')/(`qb'))
quietly sum `me_sp2_1_f'
scalar mesp21f_mean = r(mean)
return scalar ame_sp2_1_f = mesp21f_mean

*Boys
quietly replace `me_sp2_4_m' = ccm_5 * (`q4'*`qa'*`qb')* (('mal_sp2_c'*`p4')/(`q4')) ///
- (('mal_sp2_a'*pa')/(`qa')) - (('mal_sp2_b'*pb')/(`qb'))
quietly sum `me_sp2_4_m'
scalar mesp24m_mean = r(mean)
return scalar ame_sp2_4_m = mesp24m_mean 

quietly replace `me_sp2_3_m' = ccm_4 * (`q3'*`qa'*`qb')* (((`mal_sp2_c'*`p3')/(`q3')) ///
 - ((`mal_sp2_a*' pa')/(`qa')) - ((`mal_sp2_b*' pb')/(`qb'))) 

quietly sum `me_sp2_3_m'
scalar mesp23m_mean = r(mean)
return scalar ame_sp2_3_m = mesp23m_mean

quietly replace `me_sp2_2_m' = ccm_3 * (`q2'*`qa'*`qb')* (((`mal_sp2_c'*`p2')/(`q2')) ///
 - ((`mal_sp2_a*' pa')/(`qa')) - ((`mal_sp2_b*' pb')/(`qb'))) 

quietly sum `me_sp2_2_m'
scalar mesp22m_mean = r(mean)
return scalar ame_sp2_2_m = mesp22m_mean

quietly replace `me_sp2_1_m' = ccm_2 * (`q1'*`qa'*`qb')* (((`mal_sp2_c'*`p1')/(`q1')) ///
 - ((`mal_sp2_a*' pa')/(`qa')) - ((`mal_sp2_b*' pb')/(`qb'))) 

quietly sum `me_sp2_1_m'
scalar mesp21m_mean = r(mean)
return scalar ame_sp2_1_m = mesp21m_mean

******SP 3:
*Together
quietly replace `me_sp3_4' = ccm_5 * (`q4'*`qa'*`qb')* (((`tog_sp3_c'*`p4')/(`q4')) ///
 - ((`tog_sp3_a*' pa')/(`qa')) - ((`tog_sp3_b*' pb')/(`qb'))) 

quietly sum `me_sp3_4'
scalar mesp34_mean = r(mean)
return scalar ame_sp3_4 = mesp34_mean

quietly replace `me_sp3_3' = ccm_4 * (`q3'*`qa'*`qb')* (((`tog_sp3_c'*`p3')/(`q3')) ///
 - ((`tog_sp3_a*' pa')/(`qa')) - ((`tog_sp3_b*' pb')/(`qb'))) 

quietly sum `me_sp3_3'
scalar mesp33_mean = r(mean)
return scalar ame_sp3_3 = mesp33_mean

quietly replace `me_sp3_2' = ccm_3 * (`q2'*`qa'*`qb')* (((`tog_sp3_c'*`p2')/(`q2')) ///
 - ((`tog_sp3_a*' pa')/(`qa')) - ((`tog_sp3_b*' pb')/(`qb'))) 

quietly sum `me_sp3_2'
scalar mesp32_mean = r(mean)
return scalar ame_sp3_2 = mesp32_mean

quietly replace `me_sp3_1' = ccm_2 * (`q1'*`qa'*`qb')* (((`tog_sp3_c'*`p1')/(`q1')) ///
 - ((`tog_sp3_a*' pa')/(`qa')) - ((`tog_sp3_b*' pb')/(`qb'))) 

quietly sum `me_sp3_1'
scalar mesp31_mean = r(mean)
return scalar ame_sp3_1 = mesp31_mean

*Girls
quietly replace `me_sp3_4_f' = ccm_5 * (`q4'*`qa'*`qb')* (((`fem_sp3_c'*p4')/(`q4')) ///
- (((`fem_sp3_a'*pa')/(`qa')) - (((`fem_sp3_b'*pb')/(`qb'))) )))
quietly sum `me_sp3_4_f'
scalar mesp34f_mean = r(mean)
return scalar ame_sp3_4_f = mesp34f_mean

quietly replace `me_sp3_3_f' = ccm_4 * (`q3'*`qa'*`qb')* (((`fem_sp3_c'*p3')/(`q3')) ///
- (((`fem_sp3_a'*pa')/(`qa')) - (((`fem_sp3_b'*pb')/(`qb'))) )))
quietly sum `me_sp3_3_f'
scalar mesp33f_mean = r(mean)
return scalar ame_sp3_3_f = mesp33f_mean

quietly replace `me_sp3_2_f' = ccm_3 * (`q2'*`qa'*`qb')* (((`fem_sp3_c'*p2')/(`q2')) ///
- (((`fem_sp3_a'*pa')/(`qa')) - (((`fem_sp3_b'*pb')/(`qb'))) )))
quietly sum `me_sp3_2_f'
scalar mesp32f_mean = r(mean)
return scalar ame_sp3_2_f = mesp32f_mean

quietly replace `me_sp3_1_f' = ccm_2 * (`q1'*`qa'*`qb')* (((`fem_sp3_c'*p1')/(`q1')) ///
- (((`fem_sp3_a'*pa')/(`qa')) - (((`fem_sp3_b'*pb')/(`qb'))) )))
quietly sum `me_sp3_1_f'
scalar mesp31f_mean = r(mean)
return scalar ame_sp3_1_f = mesp31f_mean

*Boys
quietly replace `me_sp3_4_m' = ccm_5 * (`q4'*`qa'*`qb')* (((`mal_sp3_c'*p4')/(`q4')) ///
- (((`mal_sp3_a'*pa')/(`qa')) - (((`mal_sp3_b'*pb')/(`qb'))) )))
quietly sum `me_sp3_4_m'
scalar mesp34m_mean = r(mean)
return scalar ame_sp3_4_m = mesp34m_mean

quietly replace `me_sp3_3_m' = ccm_4 * (`q3'*`qa'*`qb')* (((`mal_sp3_c'*p3')/(`q3')) ///
- (((`mal_sp3_a'*pa')/(`qa')) - (((`mal_sp3_b'*pb')/(`qb'))) )))
quietly sum `me_sp3_3_m'
scalar mesp33m_mean = r(mean)
return scalar ame_sp3_3_m = mesp33m_mean

quietly replace `me_sp3_2_m' = ccm_3 * (`q2'*`qa'*`qb')* (((`mal_sp3_c'*p2')/(`q2')) ///
- (((`mal_sp3_a'*pa')/(`qa')) - (((`mal_sp3_b'*pb')/(`qb'))) )))
quietly sum `me_sp3_2_m'
scalar mesp32m_mean = r(mean)
return scalar ame_sp3_2_m = mesp32m_mean
quietly replace `me_sp3_1_m' = ccm_2 * (q1'*qa'*qb')* (((mal_sp3_c'*p1')/(q1')) ///
- ((mal_sp3_a'*pa')/(qa')) - ((mal_sp3_b'*pb')/(qb')))
quietly sum `me_sp3_1_m'
scalar mesp31m_mean = r(mean)
return scalar ame_sp3_1_m = mesp31m_mean

**** Pro-Tobacco Exposure

(tempvar fem_pte_a mal_pte_a fem_pte_b mal_pte_b fem_pte_c mal_pte_c ///
   me_pte_1 me_pte_1_f me_pte_1_m me_pte_2 me_pte_2_f me_pte_2_m ///
   me_pte_3 me_pte_3_f me_pte_3_m me_pte_4 me_pte_4_f me_pte_4_m ///
tog_pte_a tog_pte_b tog_pte_c)
quietly gen `tog_pte_a' = a18_pte1_cons
quietly gen `fem_pte_a' = a18_pte1_cons if female == 1
quietly gen `mal_pte_a' = a18_pte1_cons if female == 0

quietly gen `tog_pte_b' = b18_pte1_cons
quietly gen `fem_pte_b' = b18_pte1_cons if female == 1
quietly gen `mal_pte_b' = b18_pte1_cons if female == 0

quietly gen `tog_pte_c' = c18_pte1_cons
quietly gen `fem_pte_c' = c18_pte1_cons if female == 1
quietly gen `mal_pte_c' = c18_pte1_cons if female == 0

local Bmes1 `me_pte_1' `me_pte_1_f' `me_pte_1_m'
local Bmes2 `me_pte_2' `me_pte_2_f' `me_pte_2_m'
local Bmes3 `me_pte_3' `me_pte_3_f' `me_pte_3_m'
local Bmes4 `me_pte_4' `me_pte_4_f' `me_pte_4_m'

forvalues i = 1/4{
   foreach x of local Bmes' i'{
      quietly gen double 'x' = .
   }
   local i = 'i' +1
}

*Together
quietly replace `me_pte_4' = ccm_5 * (q4'*qa'*qb')* (((tog_pte_c'*p4')/(q4')) ///
- ((tog_pte_a'*pa')/(qa')) - ((tog_pte_b'*pb')/(qb')))
quietly sum `me_pte_4'
scalar mepte4_mean = r(mean)
return scalar ame_pte_4 = mepte4_mean

quietly replace `me_pte_3' = ccm_4 * (q3'*qa'*qb')* (((tog_pte_c'*p3')/(q3')) ///
quietly sum `me_pte_3'
scalar mepte3_mean = r(mean)
return scalar ame_pte_3 = mepte3_mean

quietly replace `me_pte_2' = ccm_3 * (`q2'*`qa'*`qb')* (((`tog_pte_c'*`p2')/(`q2')) - ((`tog_pte_a'*`pa')/(`qa')) - ((`tog_pte_b'*`pb')/(`qb')))  
quietly sum `me_pte_2'
scalar mepte2_mean = r(mean)
return scalar ame_pte_2 = mepte2_mean

quietly replace `me_pte_1' = ccm_2 * (`q1'*`qa'*`qb')* (((`tog_pte_c'*`p1')/(`q1')) - ((`tog_pte_a'*`pa')/(`qa')) - ((`tog_pte_b'*`pb')/(`qb')))  
quietly sum `me_pte_1'
scalar mepte1_mean = r(mean)
return scalar ame_pte_1 = mepte1_mean

*Girls
quietly replace `me_pte_4_f' = ccm_5 * (`q4'*`qa'*`qb')* (((`fem_pte_c'*`p4')/(`q4')) - ((`fem_pte_a'*`pa')/(`qa')) - ((`fem_pte_b'*`pb')/(`qb')))  
quietly sum `me_pte_4_f'
scalar mepte4f_mean = r(mean)
return scalar ame_pte_4_f = mepte4f_mean

quietly replace `me_pte_3_f' = ccm_4 * (`q3'*`qa'*`qb')* (((`fem_pte_c'*`p3')/(`q3')) - ((`fem_pte_a'*`pa')/(`qa')) - ((`fem_pte_b'*`pb')/(`qb')))  
quietly sum `me_pte_3_f'
scalar mepte3f_mean = r(mean)
return scalar ame_pte_3_f = mepte3f_mean

quietly replace `me_pte_2_f' = ccm_3 * (`q2'*`qa'*`qb')* (((`fem_pte_c'*`p2')/(`q2')) - ((`fem_pte_a'*`pa')/(`qa')) - ((`fem_pte_b'*`pb')/(`qb')))  
quietly sum `me_pte_2_f'
scalar mepte2f_mean = r(mean)
return scalar ame_pte_2_f = mepte2f_mean

*Boys
quietly replace `me_pte_4_m' = ccm_5 * (`q4'*`qa'*`qb')* (((`mal_pte_c'*`p4')/(`q4')) - ((`mal_pte_a'*`pa')/(`qa')) - ((`mal_pte_b'*`pb')/(`qb')))  
quietly sum `me_pte_4_m'
scalar mepte4m_mean = r(mean)
return scalar ame_pte_4_m = mepte4m_mean

quietly replace `me_pte_3_m' = ccm_4 * (`q3'*`qa'*`qb')* (((`mal_pte_c'*`p3')/(`q3')) - (`'mal_pte_a'*`pa')/(`qa')) - ((`mal_pte_b'*`pb')/(`qb'))
quietly sum `me_pte_3_m'
scalar mepte3m_mean = r(mean)
return scalar ame_pte_3_m = mepte3m_mean

quietly replace `me_pte_2_m' = ccm_3 * (`q2'*`qa'*`qb')* (((`mal_pte_c'*`p2')/(`q2')) - (`'mal_pte_a'*`pa')/(`qa')) - ((`mal_pte_b'*`pb')/(`qb'))
quietly sum `me_pte_2_m'
scalar mepte2m_mean = r(mean)
return scalar ame_pte_2_m = mepte2m_mean

quietly replace `me_pte_1_m' = ccm_2 * (`q1'*`qa'*`qb')* (((`mal_pte_c'*`p1')/(`q1')) - (`'mal_pte_a'*`pa')/(`qa')) - ((`mal_pte_b'*`pb')/(`qb'))
quietly sum `me_pte_1_m'
scalar mepte1m_mean = r(mean)
return scalar ame_pte_1_m = mepte1m_mean

************* Discrete Variables

*****Friends_smoke_bi

tempvar Friends_xa_1 Friends_xb_1 Friends_xc_1 Friends_xa_0 Friends_xb_0 ///
   Friends_xc_0 Friends_pa_1 Friends_pb_1 Friends_v2_1 Friends_v3_1 ///
   Friends_v4_1 Friends_p2_1 Friends_p3_1 Friends_p4_1 Friends_p5_1 ///
   Friends_pa_0 Friends_pb_0 Friends_v2_0 Friends_v3_0 Friends_v4_0 ///
   Friends_p2_0 Friends_p3_0 Friends_p4_0 Friends_p5_0 ///
   Friends_when_one_5 Friends_when_one_4 Friends_when_one_3 ///
   Friends_when_one_2 Friends_when_zero_5 Friends_when_zero_4 ///
   Friends_when_zero_3 Friends_when_zero_2 Friends_when_one_5_f ///
   Friends_when_one_4_f Friends_when_one_3_f Friends_when_one_2_f ///
   Friends_when_one_5_m Friends_when_one_4_m ///
   Friends_when_one_3_m Friends_when_one_2_m ///
   Friends_when_zero_5_f Friends_when_zero_4_f ///
   Friends_when_zero_3_f Friends_when_zero_2_f ///
   Friends_when_zero_5_m Friends_when_zero_4_m ///
   Friends_when_zero_3_m Friends_when_zero_2_m

****** When Equal to One
*For Ever Smokers
quietly gen double `Friends_xa_1' = a0_constant_cons ///
    + a1_female_cons * female ///
    + a2_friends_cons * 1 ///
+ a3_parents_cons * parent_smoke_bi ///
+ a4_f_parent_cons * f_parent ///
+ a5_sp1_cons * sp_pca_1 ///
+ a6_f_sp1_cons * f_sp1 ///
+ a7_sp2_cons * sp_pca_2 ///
+ a8_f_sp2_cons * f_sp2 ///
+ a9_sp3_cons * sp_pca_3 ///
+ a10_f_sp3_cons * f_sp3 ///
+ a11_cw1_cons * cw_1 ///
+ a12_f_cw1_cons * f_cw1 ///
+ a13_cw2_cons * cw_2 ///
+ a14_f_cw2_cons * f_cw2 ///
+ a15_cigharm_cons * cig_harm_bi ///
+ a16_f_cigharm_cons * f_cigh ///
+ a17_othercigharm_cons * other_cig_harm_bi ///
+ a18_pte1_cons * pt_pca_1 ///
+ a19_edu2_cons * ed_2 ///
+ a20_f_edu2_cons * f_edu2 ///
+ a21_edu3_cons * ed_3 ///
+ a22_f_edu3_cons * f_edu3 ///
+ a23_edu4_cons * ed_4 ///
+ a24_f_edu_cons * f_edu4 ///
+ a25_ate_cons * ate_v1 ///
+ a26_famdis_cons * fam_smoke_dis
quietly gen double `Friends_pa_1' = 1 - normprob(`Friends_xa_1')

*For Current Smoker Portion
quietly gen double `Friends_xb_1' = b0_constant_cons ///
+ b1_female_cons * female ///
+ b2_friends_cons * 1 ///
+ b3_parents_cons * parent_smoke_bi ///
+ b4_f_parent_cons * f_parent ///
+ b5_sp1_cons * sp_pca_1 ///
+ b6_f_sp1_cons * f_sp1 ///
+ b7_sp2_cons * sp_pca_2 ///
+ b8_f_sp2_cons * f_sp2 ///
+ b9_sp3_cons * sp_pca_3 ///
+ b10_f_sp3_cons * f_sp3 ///
+ b11_cw1_cons * cw_1 ///
+ b12_f_cw1_cons * f_cw1 ///
+ b13_cw2_cons * cw_2 ///
+ b14_f_cw2_cons * f_cw2 ///
+ b15_cigharm_cons * cig_harm_bi ///
+ b16_f_cigharm_cons * f_cigh ///
+ b17_othercigharm_cons * other_cig_harm_bi ///
+ b18_pte1_cons * pt_pca_1 ///
quietly gen double `Friends_pb_1' = 1 - normprob(`Friends_xb_1')

*For Number of Cigarettes Smoked Portion
quietly gen double `Friends_xc_1' = c1_female_cons * female ///
+ c2_friends_cons * 1 ///
+ c3_parents_cons * parent_smoke_bi ///
+ c4_fparent_cons * f_parent ///
+ c5_sp1_cons * sp_pca_1 ///
+ c6_fsp1_cons * f_sp1 ///
+ c7_sp2_cons * sp_pca_2 ///
+ c8_fsp2_cons * f_sp2 ///
+ c9_sp3_cons * sp_pca_3 ///
+ c10_fsp3_cons * f_sp3 ///
+ c11 cw1_cons * cw_1 ///
+ c13 cw2_cons * cw_2 ///
+ c15_cigharm_cons * cig_harm_bi ///
+ c16_f_cigharm_cons * f_cigh ///
+ c17_othercigharm_cons * other_cig_harm_bi ///
+ c18_pte1_cons * pt_pca_1 ///
+ c19_edu2_cons * ed_2 ///
+ c21_edu3_cons * ed_3 ///
+ c23_edu4_cons * ed_4 ///
+ c24_fedu4_cons * f_edu4 ///
+ c25_ate_cons * ate_v1 ///
+ c26_famdis_cons * fam_smoke_dis
quietly gen double `Friends_v2_1' = cut1_cons - (`Friends_xc_1')
quietly gen double `Friends_v3_1' = cut2_cons - (`Friends_xc_1')
quietly gen double `Friends_v4_1' = cut3_cons - (`Friends_xc_1')
quietly gen double `Friends_p2_1' = normprob(`Friends_v2_1')
quietly gen double `Friends_p3_1' = normprob(`Friends_v3_1') - normprob(`Friends_v2_1')
quietly gen double `Friends_p4_1' = normprob(`Friends_v4_1') - normprob(`Friends_v3_1')
quietly gen double `Friends_p5_1' = 1 - normprob(`Friends_v4_1')

*Top Category m=5
quietly gen double `Friends_when_one_5' = `Friends_PA_1' * `Friends_pb_1' * `Friends_p5_1'
quietly gen double `Friends_when_one_5_f' = `Friends_pa_1' * `Friends_pb_1' * `Friends_p5_1' if female == 1
quietly gen double `Friends_when_one_5_m' = `Friends_pa_1' * `Friends_pb_1' * `Friends_p5_1' if female == 0

*Middle Category m=4
quietly gen double `Friends_when_one_4' = `Friends_pa_1' * `Friends_pb_1' * `Friends_p4_1'
quietly gen double `Friends_when_one_4_f' = `Friends_pa_1' * `Friends_pb_1' * `Friends_p4_1' if female == 1
quietly gen double `Friends_when_one_4_m' = `Friends_pa_1' * `Friends_pb_1' * `Friends_p4_1' if female == 0

*Middle Category m=3
quietly gen double `Friends_when_one_3' = `Friends_pa_1' * `Friends_pb_1' * `Friends_p3_1'
quietly gen double `Friends_when_one_3_f' = `Friends_pa_1' * `Friends_pb_1' * `Friends_p3_1' if female == 1
quietly gen double `Friends_when_one_3_m' = `Friends_pa_1' * `Friends_pb_1' * `Friends_p3_1' if female == 0

*Bottom Category m=2
quietly gen double `Friends_when_one_2' = `Friends_pa_1' * `Friends_pb_1' * `Friends_p2_1'
quietly gen double `Friends_when_one_2_f' = `Friends_pa_1' * `Friends_pb_1' * `Friends_p2_1' if female == 1
quietly gen double `Friends_when_one_2_m' = `Friends_pa_1' * `Friends_pb_1' * `Friends_p2_1' if female == 0

local when_one `Friends_when_one_5' `Friends_when_one_4' `Friends_when_one_3' `Friends_when_one_2'
local when_one_f `Friends_when_one_5_f' `Friends_when_one_4_f' `Friends_when_one_3_f' `Friends_when_one_2_f'
local when_one_m `Friends_when_one_5_m' `Friends_when_one_4_m' `Friends_when_one_3_m' `Friends_when_one_2_m'

***** When Equal to Zero
*For Ever Smoker Portion
quietly gen double `Friends_xa_0' = a0_constant_cons ///
+ a1_female_cons * female ///
+ a2_friends_cons * 0 ///
+ a3_parents_cons * parent_smoke_bi ///
+ a4_f_parent_cons * f_parent ///
+ a5_sp1_cons * sp_pca_1 ///
+ a6_f_sp1_cons * f_sp1 ///
+ a7_sp2_cons * sp_pca_2 ///
+ a8 f_sp2_cons * f_sp2 ///
+ a9 sp3_cons * sp_pca_3 ///
+ a10 f_sp3_cons * f_sp3 ///
+ a11 cw1_cons * cw_1 ///
+ a12 f cw1_cons * f_cw1 ///
+ a13 cw2_cons * cw_2 ///
+ a14 f cw2_cons * f_cw2 ///
+ a15 cigharm_cons * cig_harm_bi ///
+ a16 f cigharm_cons * f_cigh ///
+ a17 othercigharm_cons * other_cig_harm_bi ///
+ a18 pte1_cons * pt_pca_1 ///
+ a19 edu2_cons * ed_2 ///
+ a20 f edu2_cons * f_edu2 ///
+ a21 edu3_cons * ed_3 ///
+ a22 f edu3_cons * f_edu3 ///
+ a23 edu4_cons * ed_4 ///
+ a24 f edu_cons * f_edu4 ///
+ a25 ate_cons * ate_v1 ///
+ a26 famdis_cons * fam_smoke_dis

generates double `Friends_pa_0' = 1 - normprob(`Friends_xa_0')

*For Current Smoker Portion

generates double `Friends_xb_0' = b0 constant_cons ///
+ b1 female_cons * female ///
+ b2 friends_cons * 0 ///
+ b3 parents_cons * parent_smoke_bi ///
+ b4 f parent_cons * f_parent ///
+ b5 sp1_cons * sp_pca_1 ///
+ b6 f sp1_cons * f_sp1 ///
+ b7 sp2_cons * sp_pca_2 ///
+ b8 f sp2_cons * f_sp2 ///
+ b9 sp3_cons * sp_pca_3 ///
+ b10 f sp3_cons * f_sp3 ///
+ b11 cw1_cons * cw_1 ///
+ b12 f cw1_cons * f_cw1 ///
+ b13 cw2_cons * cw_2 ///
+ b14 f cw2_cons * f_cw2 ///
+ b15 cigharm_cons * cig_harm_bi ///
+ b16 f cigharm_cons * f_cigh ///
+ b17 othercigharm_cons * other_cig_harm_bi ///
+ b18 pte1_cons * pt_pca_1 ///
+ b19 edu2_cons * ed_2 ///
+ b20 f edu2_cons * f_edu2 ///
+ b21 edu3_cons * ed_3 ///
+ b22 f edu3_cons * f_edu3 ///
+ b23 edu4_cons * ed_4 ///
quietly gen double `Friends_pb_0' = 1 - normprob(`Friends_xb_0')

*For Number of Cigarettes Smoked Portion
quietly gen double `Friends_xc_0' = c1_female_cons * female + c2_friends_cons * 0 + c3_parents_cons * parent_smoke_bi + c4_fparent_cons * f_parent + c5_sp1_cons * sp_pca_1 + c6_fsp1_cons * f_sp1 + c7_sp2_cons * sp_pca_2 + c8_fsp2_cons * f_sp2 + c9_sp3_cons * sp_pca_3 + c10_fsp3_cons * f_sp3 + c11_cw1_cons * cw_1 + c13_cw2_cons * cw_2 + c15_cigharm_cons * cig_harm_bi + c16_feigh_cons * f_cigh + c17_othercigharm_cons * other_cig_harm_bi + c18_pte1_cons * pt_pca_1 + c19_edu2_cons * ed_2 + c21_edu3_cons * ed_3 + c23_edu4_cons * ed_4 + c24_fedu4_cons * f_edu4 + c25_ate_cons * ate_v1 + c26_famdis_cons * fam_smoke_dis

quietly gen double `Friends_v2_0' = cut1_cons - (`Friends_xc_0')
quietly gen double `Friends_v3_0' = cut2_cons - (`Friends_xc_0')
quietly gen double `Friends_v4_0' = cut3_cons - (`Friends_xc_0')
quietly gen double `Friends_v5_0' = cut4_cons - (`Friends_xc_0')

quietly gen double `Friends_p2_0' = normprob(`Friends_v2_0')
quietly gen double `Friends_p3_0' = normprob(`Friends_v3_0') - normprob(`Friends_v2_0')
quietly gen double `Friends_p4_0' = normprob(`Friends_v4_0') - normprob(`Friends_v3_0')
quietly gen double `Friends_p5_0' = 1 - normprob(`Friends_v4_0')

*Top Category m=5
quietly gen double `Friends_when_zero_5' = `Friends_pa_0' * `Friends_pb_0' * `Friends_p5_0'
quietly gen double `Friends_when_zero_5_f' = `Friends_pa_0' * `Friends_pb_0' * `Friends_p5_0' if female == 1
quietly gen double `Friends_when_zero_5_m' = `Friends_pa_0' * `Friends_pb_0' * `Friends_p5_0' if female == 0
*Middle Category m=4
quietly gen double `Friends_when_zero_4' = `Friends_pa_0' * `Friends_pb_0' * `Friends_p4_0'
quietly gen double `Friends_when_zero_4_f' = `Friends_pa_0' * `Friends_pb_0' * `Friends_p4_0' if female == 1
quietly gen double `Friends_when_zero_4_m' = `Friends_pa_0' * `Friends_pb_0' * `Friends_p4_0' if female == 0

*Middle Category m=3
quietly gen double `Friends_when_zero_3' = `Friends_pa_0' * `Friends_pb_0' * `Friends_p3_0'
quietly gen double `Friends_when_zero_3_f' = `Friends_pa_0' * `Friends_pb_0' * `Friends_p3_0' if female == 1
quietly gen double `Friends_when_zero_3_m' = `Friends_pa_0' * `Friends_pb_0' * `Friends_p3_0' if female == 0

*Bottom Category m=2
quietly gen double `Friends_when_zero_2' = `Friends_pa_0' * `Friends_pb_0' * `Friends_p2_0'
quietly gen double `Friends_when_zero_2_f' = `Friends_pa_0' * `Friends_pb_0' * `Friends_p2_0' if female == 1
quietly gen double `Friends_when_zero_2_m' = `Friends_pa_0' * `Friends_pb_0' * `Friends_p2_0' if female == 0

local when_zero `Friends_when_zero_5' `Friends_when_zero_4'
 `Friends_when_zero_3' `Friends_when_zero_2'
local when_zero_f `Friends_when_zero_5_f' `Friends_when_zero_4_f'
 `Friends_when_zero_3_f' `Friends_when_zero_2_f'
local when_zero_m `Friends_when_zero_5_m' `Friends_when_zero_4_m'
 `Friends_when_zero_3_m' `Friends_when_zero_2_m'

tempvar me_friends_5 me_friends_4 me_friends_3 me_friends_2 ///
 me_friends_5_f me_friends_4_f me_friends_3_f me_friends_2_f ///
 me_friends_5_m me_friends_4_m me_friends_3_m me_friends_2_m

*Friends_
local Cmes1 `me_friends_5' `me_friends_4' `me_friends_3' `me_friends_2'
local Cmes2 `me_friends_5_f' `me_friends_4_f' `me_friends_3_f' `me_friends_2_f'
local Cmes3 `me_friends_5_m' `me_friends_4_m' `me_friends_3_m' `me_friends_2_m'

local cig_cat ccm_5 ccm_4 ccm_3 ccm_2

forvalues i = 1/3{
 foreach x of local Cmes`i'{
 quietly gen double `x' = .
local i = `i' + 1
*
Together
quietly replace `me_friends_5' = ccm_5 * (`Friends_when_one_5' - `Friends_when_zero_5')
quietly sum `me_friends_5'
scalar mefriends5_mean = r(mean)
return scalar ame_friend_4 = mefriends5_mean
quietly replace `me_friends_4' = ccm_4 * (`Friends_when_one_4' - `Friends_when_zero_4')
quietly sum `me_friends_4'
scalar mefriends4_mean = r(mean)
return scalar ame_friend_3 = mefriends4_mean
quietly replace `me_friends_3' = ccm_3 * (`Friends_when_one_3' - `Friends_when_zero_3')
quietly sum `me_friends_3'
scalar mefriends3_mean = r(mean)
return scalar ame_friend_2 = mefriends3_mean
quietly replace `me_friends_2' = ccm_2 * (`Friends_when_one_2' - `Friends_when_zero_2')
quietly sum `me_friends_2'
scalar mefriends2_mean = r(mean)
return scalar ame_friend_1 = mefriends2_mean
*Girls
quietly replace `me_friends_5_f' = ccm_5 * (`Friends_when_one_5_f' - `Friends_when_zero_5_f')
quietly sum `me_friends_5_f'
scalar mefriends5f_mean = r(mean)
return scalar ame_friend_4_f = mefriends5f_mean
quietly replace `me_friends_4_f' = ccm_4 * (`Friends_when_one_4_f' - `Friends_when_zero_4_f')
quietly sum `me_friends_4_f'
scalar mefriends4f_mean = r(mean)
return scalar ame_friend_3_f = mefriends4f_mean
quietly replace `me_friends_3_f' = ccm_3 * (`Friends_when_one_3_f' - `Friends_when_zero_3_f')
quietly sum `me_friends_3_f'
scalar mefriends3f_mean = r(mean)
return scalar ame_friend_2_f = mefriends3f_mean

quietly replace `me_friends_2_f' = ccm_2 * ('Friends_when_one_2_f' - 'Friends_when_zero_2_f')
quietly sum `me_friends_2_f'
scalar mefriends2f_mean = r(mean)
return scalar ame_friend_1_f = mefriends2f_mean

*Boys
quietly replace `me_friends_5_m' = ccm_5 * ('Friends_when_one_5_m' - 'Friends_when_zero_5_m')
quietly sum `me_friends_5_m'
scalar mefriends5m_mean = r(mean)
return scalar ame_friend_4_m = mefriends5m_mean

quietly replace `me_friends_4_m' = ccm_4 * ('Friends_when_one_4_m' - 'Friends_when_zero_4_m')
quietly sum `me_friends_4_m'
scalar mefriends4m_mean = r(mean)
return scalar ame_friend_3_m = mefriends4m_mean

quietly replace `me_friends_3_m' = ccm_3 * ('Friends_when_one_3_m' - 'Friends_when_zero_3_m')
quietly sum `me_friends_3_m'
scalar mefriends3m_mean = r(mean)
return scalar ame_friend_2_m = mefriends3m_mean

quietly replace `me_friends_2_m' = ccm_2 * ('Friends_when_one_2_m' - 'Friends_when_zero_2_m')
quietly sum `me_friends_2_m'
scalar mefriends2m_mean = r(mean)
return scalar ame_friend_1_m = mefriends2m_mean

******Female_smoke_bi

tempvar FemaleX_xa_1 FemaleX_xb_1 FemaleX_xc_1 FemaleX_xa_0 ///
    FemaleX_xb_0 FemaleX_xc_0 FemaleX_pa_1 ///
    FemaleX_pb_1 FemaleX_v2_1 FemaleX_v3_1 ///
    FemaleX_v4_1 FemaleX_p2_1 FemaleX_p3_1 ///
    FemaleX_p4_1 FemaleX_p5_1 FemaleX_pa_0 ///
    FemaleX_pb_0 FemaleX_v2_0 FemaleX_v3_0 FemaleX_v4_0 ///
    FemaleX_p2_0 FemaleX_p3_0 FemaleX_p4_0 FemaleX_p5_0 ///
    FemaleX_when_one_5 /FemaleX_when_one_4 FemaleX_when_one_3 ///
    FemaleX_when_one_2 FemaleX_when_zero_5 FemaleX_when_zero_4 ///
    FemaleX_when_zero_3 FemaleX_when_zero_2 ///
    FemaleX_when_one_5 f FemaleX_when_one_4 f ///

242
**** When Equal to One
*For Ever Smokers
quietly gen double `FemaleX_xa_1' = a0_constant_cons + a1_female_cons * 1
+ a2_friends_cons * friends_smoke_bi
+ a3_parents_cons * parent_smoke_bi
+ a4_f_parent_cons * f_parent
+ a5_sp1_cons * sp_pca_1
+ a6_f_sp1_cons * f_sp1
+ a7_sp2_cons * sp_pca_2
+ a8_f_sp2_cons * f_sp2
+ a9_sp3_cons * sp_pca_3
+ a10_f_sp3_cons * f_sp3
+ a11_cw1_cons * cw_1
+ a12_f_cw1_cons * f_cw1
+ a13_cw2_cons * cw_2
+ a14_f_cw2_cons * f_cw2
+ a15_cigharm_cons * cig_harm_bi
+ a16_f_cigharm_cons * f_cigh
+ a17_othercigharm_cons * other_cig_harm_bi
+ a18_pte1_cons * pt_pca_1
+ a19_edu2_cons * ed_2
+ a20_f_edu2_cons * f_edu2
+ a21_edu3_cons * ed_3
+ a22_f_edu3_cons * f_edu3
+ a23_edu4_cons * ed_4
+ a24_f_edu4_cons * f_edu4
+ a25_ate_cons * ate_v1
+ a26_famdis_cons * fam_smoke_dis
quietly gen double `FemaleX_pa_1' = 1 - normprob(`FemaleX_xa_1')

*For Current Smoker Portion
quietly gen double `FemaleX_xb_1' = b0_constant_cons + b1_female_cons * 1
+ b2_friends_cons * friends_smoke_bi
+ b3_parents_cons * parent_smoke_bi
+ b4_f_parent_cons * f_parent
+ b5_sp1_cons * sp_pca_1
+ b6_f_sp1_cons * f_sp1
+ b7_sp2_cons * sp_pca_2
+ b8_f_sp2_cons * f_sp2
+ b9_sp3_cons * sp_pca_3
+ b10_f_sp3_cons * f_sp3
+ b11_cw1_cons * cw_1
+ b12_f_cw1_cons * f_cw1
+ b13_cw2_cons * cw_2
+ b14_f_cw2_cons * f_cw2
+ b15_cigharm_cons * cig_harm_bi
+ b16_f_cigharm_cons * f_cigh
+ b17_othercigharm_cons * other_cig_harm_bi
+ b18_pte1_cons * pt_pca_1
+ b19_edu2_cons * ed_2
+ b20_f_edu2_cons * f_edu2
+ b21_edu3_cons * ed_3
+ b22_f_edu3_cons * f_edu3
+ b23_edu4_cons * ed_4
+ b24_f_edu_cons * f_edu4
+ b25_ate_cons * ate_v1
+ b26_famdis_cons * fam_smoke_dis

quietly gen double `FemaleX_pb_1' = 1 - normprob(`FemaleX_xb_1')

*For Number of Cigarettes Smoked Portion
quietly gen double `FemaleX_xc_1' = c1_female_cons * 1
+ c2_friends_cons * friends_smoke_bi
+ c3_parents_cons * parent_smoke_bi
+ c4_fparent_cons * f_parent
+ c5_sp1_cons * sp_pca_1
+ c6_fsp1_cons * f_sp1
+ c7_sp2_cons * sp_pca_2
+ c8_fsp2_cons * f_sp2
+ c9_sp3_cons * sp_pca_3
+ c10_fsp3_cons * f_sp3
+ c11_cw1_cons * cw_1
+ c13_cw2_cons * cw_2
+ c15_cigharm_cons * cig_harm_bi
+ c16_fcigh_cons * f_cigh
+ c17_othercigharm_cons * other_cig_harm_bi
+ c18_pte1_cons * pt_pca_1
+ c19_edu2_cons * ed_2
+ c21_edu3_cons * ed_3
+ c23_edu4_cons * ed_4
+ c24_fedu4_cons * f_edu4
+ c25_ate_cons * ate_v1
+ c26_famdis_cons * fam_smoke_dis
quietly gen double `FemaleX_v2_1' = cut1_cons - (`FemaleX_xc_1')
quietly gen double `FemaleX_v3_1' = cut2_cons - (`FemaleX_xc_1')
quietly gen double `FemaleX_v4_1' = cut3_cons - (`FemaleX_xc_1')
quietly gen double `FemaleX_p2_1' = normprob(`FemaleX_v2_1')
quietly gen double `FemaleX_p3_1' = normprob(`FemaleX_v3_1') - normprob(`FemaleX_v2_1')
quietly gen double `FemaleX_p4_1' = normprob(`FemaleX_v4_1') - normprob(`FemaleX_v2_1')
quietly gen double `FemaleX_p5_1' = 1 - normprob(`FemaleX_v4_1')

*Top Category m=5
quietly gen double `FemaleX_when_one_5' = `FemaleX_pa_1' * `FemaleX_pb_1' * `FemaleX_p5_1'
quietly gen double `FemaleX_when_one_5_f' = `FemaleX_pa_1' * `FemaleX_pb_1' * `FemaleX_p5_1' if female == 1
quietly gen double `FemaleX_when_one_5_m' = `FemaleX_pa_1' * `FemaleX_pb_1' * `FemaleX_p5_1' if female == 0

*Middle Category m=4
quietly gen double `FemaleX_when_one_4' = `FemaleX_pa_1' * `FemaleX_pb_1' * `FemaleX_p4_1'
quietly gen double `FemaleX_when_one_4_f' = `FemaleX_pa_1' * `FemaleX_pb_1' * `FemaleX_p4_1' if female == 1
quietly gen double `FemaleX_when_one_4_m' = `FemaleX_pa_1' * `FemaleX_pb_1' * `FemaleX_p4_1' if female == 0

*Middle Category m=3
quietly gen double `FemaleX_when_one_3' = `FemaleX_pa_1' * `FemaleX_pb_1' * `FemaleX_p3_1'
quietly gen double `FemaleX_when_one_3_f' = `FemaleX_pa_1' * `FemaleX_pb_1' * `FemaleX_p3_1' if female == 1
quietly gen double `FemaleX_when_one_3_m' = `FemaleX_pa_1' * `FemaleX_pb_1' * `FemaleX_p3_1' if female == 0

*Bottom Category m=2
quietly gen double `FemaleX_when_one_2' = `FemaleX_pa_1' * `FemaleX_pb_1' * `FemaleX_p2_1'
quietly gen double `FemaleX_when_one_2_f' = `FemaleX_pa_1' * `FemaleX_pb_1' * `FemaleX_p2_1' if female == 1
quietly gen double `FemaleX_when_one_2_m' = `FemaleX_pa_1' * `FemaleX_pb_1' * `FemaleX_p2_1' if female == 0

local when_one `FemaleX_when_one_5' `FemaleX_when_one_4' `FemaleX_when_one_3' `FemaleX_when_one_2'
local when_one_f `FemaleX_when_one_5_f' `FemaleX_when_one_4_f' `FemaleX_when_one_3_f' `FemaleX_when_one_2_f'
local when_one_m `FemaleX_when_one_5_m' `FemaleX_when_one_4_m'
`FemaleX_when_one_3_m' `FemaleX_when_one_2_m'

***** When Equal to Zero
*For Ever Smoker Portion
quietly gen double `FemaleX_xa_0' = a0_constant_cons ///
+ a1_female_cons * 0 ///
+ a2_friends_cons * friends_smoke_bi ///
+ a3_parents_cons * parent_smoke_bi ///
+ a4_f_parent_cons * f_parent ///
+ a5_sp1_cons * sp_pca_1 ///
+ a6_f_sp1_cons * f_sp1 ///
+ a7_sp2_cons * sp_pca_2 ///
+ a8_f_sp2_cons * f_sp2 ///
+ a9_sp3_cons * sp_pca_3 ///
+ a10_f_sp3_cons * f_sp3 ///
+ a11_cw1_cons * cw_1 ///
+ a12_f_cw1_cons * f_cw1 ///
+ a13_cw2_cons * cw_2 ///
+ a14_f_cw2_cons * f_cw2 ///
+ a15_cigharm_cons * cig_harm_bi ///
+ a16_f_cigharm_cons * f_cigh ///
+ a17_othercigharm_cons * other_cig_harm_bi ///
+ a18_pte1_cons * pt_pca_1 ///
+ a19_edu2_cons * ed_2 ///
+ a20_f_edu2_cons * f_edu2 ///
+ a21_edu3_cons * ed_3 ///
+ a22_f_edu3_cons * f_edu3 ///
+ a23_edu4_cons * ed_4 ///
+ a24_f_edu4_cons * f_edu4 ///
+ a25_ate_cons * ate_v1 ///
+ a26_famdis_cons * fam_smoke_dis
quietly gen double `FemaleX_pa_0' = 1 - normprob(`FemaleX_xa_0')

*For Current Smoker Portion
quietly gen double `FemaleX_xb_0' = b0_constant_cons ///
+ b1_female_cons * 0 ///
+ b2_friends_cons * friends_smoke_bi ///
+ b3_parents_cons * parent_smoke_bi ///
+ b4_f_parent_cons * f_parent ///
+ b5_sp1_cons * sp_pca_1 ///
+ b6_f_sp1_cons * f_sp1 ///
+ b7_sp2_cons * sp_pca_2 ///
+ b8_f_sp2_cons * f_sp2 ///
+ b9_sp3_cons * sp_pca_3 ///
+ b10_f_sp3_cons * f_sp3 ///
+ b11_cw1_cons * cw_1 ///
+ b12_f_cw1_cons * f_cw1 ///
+ b13_cw2_cons * cw_2 ///
+ b14_f_cw2_cons * f_cw2 ///
+ b15_cigharm_cons * cig_harm_bi ///
+ b16_f_cigharm_cons * f_cigh ///
+ b17_othercigharm_cons * other_cig_harm_bi ///
+ b18_pte1_cons * pt_pca_1 ///
+ b19_edu2_cons * ed_2 ///
+ b20_f_edu2_cons * f_edu2 ///
+ b21_edu3_cons * ed_3 ///
+ b22_f_edu3_cons * f_edu3 ///
+ b23_edu4_cons * ed_4 ///
+ b24_f_edu_cons * f_edu4 ///
+ b25_ate_cons * ate_v1 ///
+ b26_famdis_cons * fam_smoke_dis

quietly gen double `FemaleX_pb_0' = 1 - normprob(`FemaleX_xb_0')

*For Number of Cigarettes Smoked Portion
quietly gen double `FemaleX_xc_0' = c1_female_cons * 0 ///
+ c2_friends_cons * friends_smoke_bi ///
+ c3_parents_cons * parent_smoke_bi ///
+ c4_fparent_cons * f_parent ///
+ c5_sp1_cons * sp_pca_1 ///
+ c6fsp1_cons * f_sp1 ///
+ c7_sp2_cons * sp_pca_2 ///
+ c8_sp2_cons * f_sp2 ///
+ c9_sp3_cons * sp_pca_3 ///
+ c10fsp3_cons * f_sp3 ///
+ c11_cw1_cons * cw_1 ///
+ c12_cw2_cons * cw_2 ///
+ c13_cigharm_cons * cig_harm_bi ///
+ c14f_cigharm_cons * f_cigh ///
+ c15_othercigharm_cons * other_cig_harm_bi ///
+ c16_pte1_cons * pt_pca_1 ///
+ c19_edu2_cons * ed_2 ///
+ c21_edu3_cons * ed_3 ///
+ c23_edu4_cons * ed_4 ///
+ c24_fedu4_cons * f_edu4 ///
+ c25_ate_cons * ate_v1 ///
+ c26_famdis_cons * fam_smoke_dis

quietly gen double `FemaleX_v2_0' = cut1_cons - (`FemaleX_xc_0')
quietly gen double `FemaleX_v3_0' = cut2_cons - (`FemaleX_xc_0')
quietly gen double `FemaleX_v4_0' = cut3_cons - (`FemaleX_xc_0')
quietly gen double `FemaleX_p2_0' = normprob(`FemaleX_v2_0')
quietly gen double `FemaleX_p3_0' = normprob(`FemaleX_v3_0') - normprob(`FemaleX_v2_0')
quietly gen double `FemaleX_p4_0' = normprob(`FemaleX_v4_0') - normprob(`FemaleX_v3_0')
quietly gen double `FemaleX_p5_0' = 1 - normprob(`FemaleX_v4_0')

*Top Category m=5
quietly gen double `FemaleX_when_zero_5' = `FemaleX_pa_0' * `FemaleX_pb_0' * `FemaleX_p5_0'
quietly gen double `FemaleX_when_zero_5_f' = `FemaleX_pa_0' * `FemaleX_pb_0' * `FemaleX_p5_0' if female == 1
quietly gen double `FemaleX_when_zero_5_m' = `FemaleX_pa_0' * `FemaleX_pb_0' * `FemaleX_p5_0' if female == 0

*Middle Category m=4
quietly gen double `FemaleX_when_zero_4' = `FemaleX_pa_0' * `FemaleX_pb_0' * `FemaleX_p4_0'
quietly gen double `FemaleX_when_zero_4_f' = `FemaleX_pa_0' * `FemaleX_pb_0' * `FemaleX_p4_0' if female == 1
quietly gen double `FemaleX_when_zero_4_m' = `FemaleX_pa_0' * `FemaleX_pb_0' * `FemaleX_p4_0' if female == 0

*Middle Category m=3
quietly gen double `FemaleX_when_zero_3' = `FemaleX_pa_0' * `FemaleX_pb_0' * `FemaleX_p3_0'
quietly gen double `FemaleX_when_zero_3_f' = `FemaleX_pa_0' * `FemaleX_pb_0' * `FemaleX_p3_0' if female == 1
quietly gen double `FemaleX_when_zero_3_m' = `FemaleX_pa_0' * `FemaleX_pb_0' * `FemaleX_p3_0' if female == 0

*Bottom Category m=2
quietly gen double `FemaleX_when_zero_2' = `FemaleX_pa_0' * `FemaleX_pb_0' * `FemaleX_p2_0'
quietly gen double `FemaleX_when_zero_2_f' = `FemaleX_pa_0' * `FemaleX_pb_0' * `FemaleX_p2_0' if female == 1
quietly gen double `FemaleX_when_zero_2_m' = `FemaleX_pa_0' * `FemaleX_pb_0' * `FemaleX_p2_0' if female == 0

local when_zero `FemaleX_when_zero_5' `FemaleX_when_zero_4'
    `FemaleX_when_zero_3' `FemaleX_when_zero_2'
local when_zero_f `FemaleX_when_zero_5_f' `FemaleX_when_zero_4_f'
    `FemaleX_when_zero_3_f' `FemaleX_when_zero_2_f'
local when_zero_m `FemaleX_when_zero_5_m' `FemaleX_when_zero_4_m'
    `FemaleX_when_zero_3_m' `FemaleX_when_zero_2_m'

tempvar me_FemaleX_5   me_FemaleX_4   me_FemaleX_3   me_FemaleX_2  ///
local Dmes1 `me_FemaleX_5'   `me_FemaleX_4'   `me_FemaleX_3'   `me_FemaleX_2'
local Dmes2 `me_FemaleX_5_f' `me_FemaleX_4_f' `me_FemaleX_3_f'
`me_FemaleX_2_f'
local Dmes3 `me_FemaleX_5_m' `me_FemaleX_4_m' `me_FemaleX_3_m'
`me_FemaleX_2_m'

forvalues i = 1/3{
    foreach x of local Dmes`i'{
        quietly gen double `x' = .
    }
    local i = `i' +1
}

*Togther
quietly replace `me_FemaleX_5' = ccm_5 * (`FemaleX_when_one_5' -
`FemaleX_when_zero_5')
quietly sum `me_FemaleX_5'
scalar meFemaleX5_mean = r(mean)
return scalar ame_female_4 = meFemaleX5_mean

quietly replace `me_FemaleX_4' = ccm_4 * (`FemaleX_when_one_4' -
`FemaleX_when_zero_4')
quietly sum `me_FemaleX_4'
scalar meFemaleX4_mean = r(mean)
return scalar ame_female_3 = meFemaleX4_mean

quietly replace `me_FemaleX_3' = ccm_3 * (`FemaleX_when_one_3' -
`FemaleX_when_zero_3')
quietly sum `me_FemaleX_3'
scalar meFemaleX3_mean = r(mean)
return scalar ame_female_2 = meFemaleX3_mean

quietly replace `me_FemaleX_2' = ccm_2 * (`FemaleX_when_one_2' -
`FemaleX_when_zero_2')
quietly sum `me_FemaleX_2'
scalar meFemaleX2_mean = r(mean)
return scalar ame_female_1 = meFemaleX2_mean

*Girls
quietly replace `me_FemaleX_5_f' = ccm_5 * (`FemaleX_when_one_5_f' -
`FemaleX_when_zero_5_f')
quietly sum `me_FemaleX_5_f'
scalar meFemaleX5f_mean = r(mean)
return scalar ame_female_4_f = meFemaleX5f_mean

quietly replace `me_FemaleX_4_f' = ccm_4 * (`FemaleX_when_one_4_f' - `FemaleX_when_zero_4_f')
quietly sum `me_FemaleX_4_f'
scalar meFemaleX4f_mean = r(mean)
return scalar ame_female_3_f = meFemaleX4f_mean

quietly replace `me_FemaleX_3_f' = ccm_3 * (`FemaleX_when_one_3_f' - `FemaleX_when_zero_3_f')
quietly sum `me_FemaleX_3_f'
scalar meFemaleX3f_mean = r(mean)
return scalar ame_female_2_f = meFemaleX3f_mean

quietly replace `me_FemaleX_2_f' = ccm_2 * (`FemaleX_when_one_2_f' - `FemaleX_when_zero_2_f')
quietly sum `me_FemaleX_2_f'
scalar meFemaleX2f_mean = r(mean)
return scalar ame_female_1_f = meFemaleX2f_mean

*Boys
quietly replace `me_FemaleX_5_m' = ccm_5 * (`FemaleX_when_one_5_m' - `FemaleX_when_zero_5_m')
quietly sum `me_FemaleX_5_m'
scalar meFemaleX5m_mean = r(mean)
return scalar ame_female_4_m = meFemaleX5m_mean

quietly replace `me_FemaleX_4_m' = ccm_4 * (`FemaleX_when_one_4_m' - `FemaleX_when_zero_4_m')
quietly sum `me_FemaleX_4_m'
scalar meFemaleX4m_mean = r(mean)
return scalar ame_female_3_m = meFemaleX4m_mean

quietly replace `me_FemaleX_3_m' = ccm_3 * (`FemaleX_when_one_3_m' - `FemaleX_when_zero_3_m')
quietly sum `me_FemaleX_3_m'
scalar meFemaleX3m_mean = r(mean)
return scalar ame_female_2_m = meFemaleX3m_mean

quietly replace `me_FemaleX_2_m' = ccm_2 * (`FemaleX_when_one_2_m' - `FemaleX_when_zero_2_m')
quietly sum `me_FemaleX_2_m'
scalar meFemaleX2m_mean = r(mean)
return scalar ame_female_1_m = meFemaleX2m_mean

250
tempvar Ocigh_xa_1 Ocigh_xb_1 Ocigh_xc_1 Ocigh_xa_0 Ocigh_xb_0
  Ocigh_v4_1 Ocigh_p2_1 Ocigh_p3_1 Ocigh_p4_1 Ocigh_p5_1
  Ocigh_pa_0 Ocigh_pb_0 Ocigh_v2_0 Ocigh_v3_0 Ocigh_v4_0
  Ocigh_p2_0 Ocigh_p3_0 Ocigh_p4_0 Ocigh_p5_0 Ocigh_when_one_5
  Ocigh_when_one_4 Ocigh_when_one_3 Ocigh_when_one_2
  Ocigh_when_zero_5 Ocigh_when_zero_4 Ocigh_when_zero_3
  Ocigh_when_zero_2 Ocigh_when_one_5 f Ocigh_when_one_4 f
  Ocigh_when_one_3 f Ocigh_when_one_2 f Ocigh_when_one_5 m
  Ocigh_when_one_4 m Ocigh_when_one_3 m Ocigh_when_one_2 m
  Ocigh_when_zero_5 f Ocigh_when_zero_4 f Ocigh_when_zero_3 f
  Ocigh_when_zero_2 f Ocigh_when_zero_5 m Ocigh_when_zero_4 m
  Ocigh_when_zero_3 m Ocigh_when_zero_2 m

***** When Equal to One
*For Ever Smokers

quietly gen double `Ocigh_xa_1'
  = a0 constant_cons
  + a1 female_cons * female
  + a2 friends_cons * friends_smoke_bi
  + a3 parents_cons * parent_smoke_bi
  + a4 f_parent_cons * f_parent
  + a5 sp1_cons * sp_pca_1
  + a6 f_sp1_cons * f_sp1
  + a7 sp2_cons * sp_pca_2
  + a8 f_sp2_cons * f_sp2
  + a9 sp3_cons * sp_pca_3
  + a10 f_sp3_cons * f_sp3
  + a11 cw1_cons * cw_1
  + a12 f_cw1_cons * f_cw1
  + a13 cw2_cons * cw_2
  + a14 f_cw2_cons * f_cw2
  + a15 cigharm_cons * cig_harm_bi
  + a16 f_cigharm_cons * f_cigh
  + a17 othercigharm_cons * 1
  + a18 pte1_cons * pt_pca_1
  + a19 edu2_cons * ed_2
  + a20 f_edu2_cons * f_edu2
  + a21 edu3_cons * ed_3
  + a22 f_edu3_cons * f_edu3
  + a23 edu4_cons * ed_4
  + a24 f_edu4_cons * f_edu4
  + a25 ate_cons * ate_v1
  + a26 famdis_cons * fam_smoke_dis

quietly gen double `Ocigh_pa_1'
  = 1 - normprob(`Ocigh_xa_1')
*For Current Smoker Portion

```
quietly gen double `Ocigh_xb_1' = b0_constant_cons ///
   + b1_female_cons * female ///
   + b2_friends_cons * friends_smoke_bi ///
   + b3_parents_cons * parent_smoke_bi ///
   + b4_f_parent_cons * f_parent ///
   + b5_sp1_cons * sp_pca_1 ///
   + b6_f_sp1_cons * f_sp1 ///
   + b7_sp2_cons * sp_pca_2 ///
   + b8_f_sp2_cons * f_sp2 ///
   + b9_sp3_cons * sp_pca_3 ///
   + b10_f_sp3_cons * f_sp3 ///
   + b11_cw1_cons * cw_1 ///
   + b12_f_cw1_cons * f_cw1 ///
   + b13_cw2_cons * cw_2 ///
   + b14_f_cw2_cons * f_cw2 ///
   + b15_cigharm_cons * cig_harm_bi ///
   + b16_f_cigharm_cons * f_cigh ///
   + b17_othercigharm_cons * 1 ///
   + b18_pte1_cons * pt_pca_1 ///
   + b19_edu2_cons * ed_2 ///
   + b20_f_edu2_cons * f_edu2 ///
   + b21_edu3_cons * ed_3 ///
   + b22_f_edu3_cons * f_edu3 ///
   + b23_edu4_cons * ed_4 ///
   + b24_f_edu_cons * f_edu4 ///
   + b25_ate_cons * ate_v1 ///
   + b26_famdis_cons * fam_smoke_dis
```

quietly gen double `Ocigh_pb_1' = 1 - normprob(`Ocigh_xb_1')

*For Number of Cigarettes Smoked Portion

```
quietly gen double `Ocigh_xc_1' = c1_female_cons * female ///
   + c2_friends_cons * friends_smoke_bi ///
   + c3_parents_cons * parent_smoke_bi ///
   + c4_fparent_cons * f_parent ///
   + c5_sp1_cons * sp_pca_1 ///
   + c6_fsp1_cons * f_sp1 ///
   + c7_sp2_cons * sp_pca_2 ///
   + c8_fsp2_cons * f_sp2 ///
   + c9_sp3_cons * sp_pca_3 ///
   + c10_fsp3_cons * f_sp3 ///
   + c11_cw1_cons * cw_1 ///
   + c13_cw2_cons * cw_2 ///
   + c15_cigharm_cons * cig_harm_bi ///
   + c16_fcigh_cons * f_cigh ///
```
quietly gen double `Ocigh_v2_1' = cut1_cons - ('Ocigh_xc_1')
quietly gen double `Ocigh_v3_1' = cut2_cons - ('Ocigh_xc_1')
quietly gen double `Ocigh_v4_1' = cut3_cons - ('Ocigh_xc_1')
quietly gen double `Ocigh_p2_1' = normprob('Ocigh_v2_1')
quietly gen double `Ocigh_p3_1' = normprob('Ocigh_v3_1') - normprob('Ocigh_v2_1')
quietly gen double `Ocigh_v4_1' = normprob('Ocigh_v4_1') - normprob('Ocigh_v3_1')
quietly gen double `Ocigh_p5_1' = 1 - normprob('Ocigh_v4_1')

*Top Category m=5
quietly gen double `Ocigh_when_one_5' = `Ocigh_pa_1' * `Ocigh_pb_1' * `Ocigh_p5_1'
quietly gen double `Ocigh_when_one_5_f' = `Ocigh_pa_1' * `Ocigh_pb_1' * `Ocigh_p5_1' if female == 1
quietly gen double `Ocigh_when_one_5_m' = `Ocigh_pa_1' * `Ocigh_pb_1' * `Ocigh_p5_1' if female == 0

*Middle Category m=4
quietly gen double `Ocigh_when_one_4' = `Ocigh_pa_1' * `Ocigh_pb_1' * `Ocigh_p4_1'
quietly gen double `Ocigh_when_one_4_f' = `Ocigh_pa_1' * `Ocigh_pb_1' * `Ocigh_p4_1' if female == 1
quietly gen double `Ocigh_when_one_4_m' = `Ocigh_pa_1' * `Ocigh_pb_1' * `Ocigh_p4_1' if female == 0

*Middle Category m=3
quietly gen double `Ocigh_when_one_3' = `Ocigh_pa_1' * `Ocigh_pb_1' * `Ocigh_p3_1'
quietly gen double `Ocigh_when_one_3_f' = `Ocigh_pa_1' * `Ocigh_pb_1' * `Ocigh_p3_1' if female == 1
quietly gen double `Ocigh_when_one_3_m' = `Ocigh_pa_1' * `Ocigh_pb_1' * `Ocigh_p3_1' if female == 0

*Bottom Category m=2
quietly gen double `Ocigh_when_one_2' = `Ocigh_pa_1' * `Ocigh_pb_1' * `Ocigh_p2_1'
quietly gen double `Ocigh_when_one_2_f' = `Ocigh_pa_1' * `Ocigh_pb_1' * `Ocigh_p2_1' if female == 1
quietly gen double `Ocigh_when_one_2_m' = `Ocigh_pa_1' * `Ocigh_pb_1' * `Ocigh_p2_1' if female == 0

local when_one   `Ocigh_when_one_5'   `Ocigh_when_one_4'   `Ocigh_when_one_3'   `Ocigh_when_one_2'
local when_one_f `Ocigh_when_one_5_f' `Ocigh_when_one_4_f' `Ocigh_when_one_3_f' `Ocigh_when_one_2_f'
local when_one_m `Ocigh_when_one_5_m' `Ocigh_when_one_4_m' `Ocigh_when_one_3_m' `Ocigh_when_one_2_m'

***** When Equal to Zero
*For Ever Smoker Portion
quietly gen double `Ocigh_xa_0' = a0_constant_cons ///
+ a1_female_cons * female ///
+ a2_friends_cons * friends_smoke_bi ///
+ a3_parents_cons * parent_smoke_bi ///
+ a4_f_parent_cons * f_parent ///
+ a5_sp1_cons * sp_pca_1 ///
+ a6_f_sp1_cons * f_sp1 ///
+ a7_sp2_cons * sp_pca_2 ///
+ a8_f_sp2_cons * f_sp2 ///
+ a9_sp3_cons * sp_pca_3 ///
+ a10_f_sp3_cons * f_sp3 ///
+ a11_cw1_cons * cw_1 ///
+ a12_f_cw1_cons * f_cw1 ///
+ a13_cw2_cons * cw_2 ///
+ a14_f_cw2_cons * f_cw2 ///
+ a15_cigharm_cons * cig_harm_bi ///
+ a16_f_cigharm_cons * f_cigh ///
+ a17_othercigharm_cons * 0 ///
+ a18_pte1_cons * pt_pca_1 ///
+ a19_edu2_cons * ed_2 ///
+ a20_f_edu2_cons * f_edu2 ///
+ a21_edu3_cons * ed_3 ///
+ a22_f_edu3_cons * f_edu3 ///
+ a23_edu4_cons * ed_4 ///
+ a24_f_edu_cons * f_edu4 ///
+ a25_ate_cons * ate_v1 ///
+ a26_famdis_cons * fam_smoke_dis
quietly gen double `Ocigh_pa_0' = 1 - normprob(`Ocigh_xa_0')

*For Current Smoker Portion
quietly gen double `Ocigh_xb_0' = b0_constant_cons ///
+ b1_female_cons * female ///
quietly gen double `Ocigh_pb_0' = 1 - normprob(`Ocigh_xb_0')

*For Number of Cigarettes Smoked Portion
quietly gen double `Ocigh_xc_0' = c1_female_cons * female ///
+ c2_friends_cons * friends_smoke_bi ///
+ c3_parents_cons * parent_smoke_bi ///
+ c4_fparent_cons * f_parent ///
+ c5_sp1_cons * sp_pca_1 ///
+ c6_fsp1_cons * f_sp1 ///
+ c7_sp2_cons * sp_pca_2 ///
+ c8_fsp2_cons * f_sp2 ///
+ c9_sp3_cons * sp_pca_3 ///
+ c10_fsp3_cons * f_sp3 ///
+ c11_cw1_cons * cw_1 ///
+ c12_fcw1_cons * f_cw1 ///
+ c13_cw2_cons * cw_2 ///
+ c14_fcw2_cons * f_cw2 ///
+ c15_cigharm_cons * cig_harm_bi ///
+ c16_f_cigharm_cons * f_cigh ///
+ c17_othercigharm_cons * 0 ///
+ c18_pte1_cons * pt_pca_1 ///
+ c19_edu2_cons * ed_2 ///
+ c20_f_edu2_cons * f_edu2 ///
+ c21_edu3_cons * ed_3 ///
+ c22_f_edu3_cons * f_edu3 ///
+ c23_edu4_cons * ed_4 ///
+ c24_f_edu_cons * f_edu4 ///
+ c25_ate_cons * ate_v1 ///
+ c26_famdis_cons * fam_smoke_dis
+ c23_edu4_cons * ed_4 ///
  + c24_fedu4_cons * f_edu4 ///
  + c25_ate_cons * ate_v1 ///
  + c26_famdis_cons * fam_smoke_dis

quietly gen double `Ocigh_v2_0' = cut1_cons - (`Ocigh_xc_0')
quietly gen double `Ocigh_v3_0' = cut2_cons - (`Ocigh_xc_0')
quietly gen double `Ocigh_v4_0' = cut3_cons - (`Ocigh_xc_0')
quietly gen double `Ocigh_p2_0' = normprob(`Ocigh_v2_0')
quietly gen double `Ocigh_p3_0' = normprob(`Ocigh_v3_0') - normprob(`Ocigh_v2_0')
quietly gen double `Ocigh_p4_0' = normprob(`Ocigh_v4_0') - normprob(`Ocigh_v3_0')
quietly gen double `Ocigh_p5_0' = 1 - normprob(`Ocigh_v4_0')

*Top Category m=5
quietly gen double `Ocigh_when_zero_5' = `Ocigh_pa_0' * `Ocigh_pb_0' * `Ocigh_p5_0'
quietly gen double `Ocigh_when_zero_5_f' = `Ocigh_pa_0' * `Ocigh_pb_0' * `Ocigh_p5_0' if female == 1
quietly gen double `Ocigh_when_zero_5_m' = `Ocigh_pa_0' * `Ocigh_pb_0' * `Ocigh_p5_0' if female == 0

*Middle Category m=4
quietly gen double `Ocigh_when_zero_4' = `Ocigh_pa_0' * `Ocigh_pb_0' * `Ocigh_p4_0'
quietly gen double `Ocigh_when_zero_4_f' = `Ocigh_pa_0' * `Ocigh_pb_0' * `Ocigh_p4_0' if female == 1
quietly gen double `Ocigh_when_zero_4_m' = `Ocigh_pa_0' * `Ocigh_pb_0' * `Ocigh_p4_0' if female == 0

*Middle Category m=3
quietly gen double `Ocigh_when_zero_3' = `Ocigh_pa_0' * `Ocigh_pb_0' * `Ocigh_p3_0'
quietly gen double `Ocigh_when_zero_3_f' = `Ocigh_pa_0' * `Ocigh_pb_0' * `Ocigh_p3_0' if female == 1
quietly gen double `Ocigh_when_zero_3_m' = `Ocigh_pa_0' * `Ocigh_pb_0' * `Ocigh_p3_0' if female == 0

*Bottom Category m=2
quietly gen double `Ocigh_when_zero_2' = `Ocigh_pa_0' * `Ocigh_pb_0' * `Ocigh_p2_0'
quietly gen double `Ocigh_when_zero_2_f' = `Ocigh_pa_0' * `Ocigh_pb_0' * `Ocigh_p2_0' if female == 1
quietly gen double `Ocigh_when_zero_2_m' = `Ocigh_pa_0' * `Ocigh_pb_0' * `Ocigh_p2_0' if female == 0
local when_zero `Ocigh_when_zero_5' `Ocigh_when_zero_4' `Ocigh_when_zero_3'
`Ocigh_when_zero_2'
local when_zero_f `Ocigh_when_zero_5_f' `Ocigh_when_zero_4_f'
`Ocigh_when_zero_3_f' `Ocigh_when_zero_2_f'
local when_zero_m `Ocigh_when_zero_5_m' `Ocigh_when_zero_4_m'
`Ocigh_when_zero_3_m' `Ocigh_when_zero_2_m'

tempvar me_Ocigh_5  me_Ocigh_4  me_Ocigh_3  me_Ocigh_2  ///
    me_Ocigh_5_f  me_Ocigh_4_f  me_Ocigh_3_f  me_Ocigh_2_f  ///
    me_Ocigh_5_m  me_Ocigh_4_m  me_Ocigh_3_m  me_Ocigh_2_m

local Emes1 `me_Ocigh_5' `me_Ocigh_4' `me_Ocigh_3' `me_Ocigh_2'
local Emes2 `me_Ocigh_5_f' `me_Ocigh_4_f' `me_Ocigh_3_f' `me_Ocigh_2_f'
local Emes3 `me_Ocigh_5_m' `me_Ocigh_4_m' `me_Ocigh_3_m' `me_Ocigh_2_m'

forvalues i = 1/3{
    foreach x of local Emes`i'{
        quietly gen double `x' = .
    }
    local i = `i' +1
}

*Together
quietly replace `me_Ocigh_5' = ccm_5 * (`Ocigh_when_one_5' - `Ocigh_when_zero_5')
quietly sum `me_Ocigh_5'
scalar meOcigh5_mean = r(mean)
return scalar ame_och_4 = meOcigh5_mean

quietly replace `me_Ocigh_4' = ccm_4 * (`Ocigh_when_one_4' - `Ocigh_when_zero_4')
quietly sum `me_Ocigh_4'
scalar meOcigh4_mean = r(mean)
return scalar ame_och_3 = meOcigh4_mean

quietly replace `me_Ocigh_3' = ccm_3 * (`Ocigh_when_one_3' - `Ocigh_when_zero_3')
quietly sum `me_Ocigh_3'
scalar meOcigh3_mean = r(mean)
return scalar ame_och_2 = meOcigh3_mean

quietly replace `me_Ocigh_2' = ccm_2 * (`Ocigh_when_one_2' - `Ocigh_when_zero_2')
quietly sum `me_Ocigh_2'
scalar meOcigh2_mean = r(mean)
return scalar ame_och_1 = meOcigh2_mean

*Girls
quietly replace `me_Ocigh_5_f' = ccm_5 * (`Ocigh_when_one_5_f' - `Ocigh_when_zero_5_f')
quietly sum `me_Ocigh_5_f'
scalar meOcigh5f_mean = r(mean)
return scalar ame_och_4_f = meOcigh5f_mean

quietly replace `me_Ocigh_4_f' = ccm_4 * (`Ocigh_when_one_4_f' - `Ocigh_when_zero_4_f')
quietly sum `me_Ocigh_4_f'
scalar meOcigh4f_mean = r(mean)
return scalar ame_och_3_f = meOcigh4f_mean

quietly replace `me_Ocigh_3_f' = ccm_3 * (`Ocigh_when_one_3_f' - `Ocigh_when_zero_3_f')
quietly sum `me_Ocigh_3_f'
scalar meOcigh3f_mean = r(mean)
return scalar ame_och_2_f = meOcigh3f_mean

quietly replace `me_Ocigh_2_f' = ccm_2 * (`Ocigh_when_one_2_f' - `Ocigh_when_zero_2_f')
quietly sum `me_Ocigh_2_f'
scalar meOcigh2f_mean = r(mean)
return scalar ame_och_1_f = meOcigh2f_mean

*Boys
quietly replace `me_Ocigh_5_m' = ccm_5 * (`Ocigh_when_one_5_m' - `Ocigh_when_zero_5_m')
quietly sum `me_Ocigh_5_m'
scalar meOcigh5m_mean = r(mean)
return scalar ame_och_4_m = meOcigh5m_mean

quietly replace `me_Ocigh_4_m' = ccm_4 * (`Ocigh_when_one_4_m' - `Ocigh_when_zero_4_m')
quietly sum `me_Ocigh_4_m'
scalar meOcigh4m_mean = r(mean)
return scalar ame_och_3_m = meOcigh4m_mean

quietly replace `me_Ocigh_3_m' = ccm_3 * (`Ocigh_when_one_3_m' - `Ocigh_when_zero_3_m')
quietly sum `me_Ocigh_3_m'
scalar meOcigh3m_mean = r(mean)
return scalar ame_och_2_m = meOcigh3m_mean

quietly replace `me_Ocigh_2_m' = ccm_2 * (`Ocigh_when_one_2_m' - `Ocigh_when_zero_2_m')
quietly sum `me_Ocigh_2_m'
scalar meOcigh2m_mean = r(mean)
return scalar ame_och_1_m = meOcigh2m_mean

quietly replace `me_Ocigh_4_f' = ccm_4 * (`Ocigh_when_one_4_f' - `Ocigh_when_zero_4_f')
quietly sum `me_Ocigh_4_f'
scalar meOcigh4f_mean = r(mean)
return scalar ame_och_3_f = meOcigh4f_mean

quietly replace `me_Ocigh_3_f' = ccm_3 * (`Ocigh_when_one_3_f' - `Ocigh_when_zero_3_f')
quietly sum `me_Ocigh_3_f'
scalar meOcigh3f_mean = r(mean)
return scalar ame_och_2_f = meOcigh3f_mean

quietly replace `me_Ocigh_2_f' = ccm_2 * (`Ocigh_when_one_2_f' - `Ocigh_when_zero_2_f')
quietly sum `me_Ocigh_2_f'
scalar meOcigh2f_mean = r(mean)
return scalar ame_och_1_f = meOcigh2f_mean
tempvar ATE_xa_1 ATE_xb_1 ATE_xc_1 ATE_xa_0 ATE_xb_0
ATE_xc_0 ATE_pa_1 ATE_pb_1 ATE_v2_1 ATE_v3_1
ATE_v4_1 ATE_p2_1 ATE_p3_1 ATE_p4_1 ATE_p5_1
ATE_pa_0 ATE_pb_0 ATE_v2_0 ATE_v3_0 ATE_v4_0
ATE_p2_0 ATE_p3_0 ATE_p4_0 ATE_p5_0 ATE_when_one_5
ATE_when_one_4 ATE_when_one_3 ATE_when_one_2
ATE_when_one_5 ATE_when_zero_4 ATE_when_zero_3
ATE_when_zero_2 ATE_when_one_5 ATE_when_one_4_f
ATE_when_one_3 ATE_when_zero_5 ATE_when_one_2_f
ATE_when_one_4 m ATE_when_one_3 m ATE_when_one_2_m
ATE_when_one_5 m ATE_when_zero_4 m ATE_when_zero_3_f
ATE_when_zero_2 m ATE_when_zero_5 m ATE_when_zero_4 m
ATE_when_zero_3 m ATE_when_zero_2_m

***** When Equal to One
*For Ever Smokers
quietly gen double `ATE_xa_1' = a0_constant_cons
+ a1_female_cons * female
+ a2_friends_cons * friends_smoke_bi
+ a3_parents_cons * parent_smoke_bi
+ a4_f_parent_cons * f_parent
+ a5_sp1_cons * sp_pca_1
+ a6_f_sp1_cons * f_sp1
+ a7_sp2_cons * sp_pca_2
+ a8_f_sp2_cons * f_sp2
+ a9_sp3_cons * sp_pca_3
+ a10_f_sp3_cons * f_sp3
+ a11_cw1_cons * cw_1
+ a12_f_cw1_cons * f_cw1
+ a13_cw2_cons * cw_2
+ a14_f_cw2_cons * f_cw2
+ a15_cigharm_cons * cig_harm_bi
+ a16_f_cigharm_cons * f_cigh
+ a17_othercigharm_cons * other_cig_harm_bi
+ a18_pte1_cons * pt_pca_1
+ a19_edu2_cons * ed_2
+ a20_f_edu2_cons * f_edu2
+ a21_edu3_cons * ed_3
+ a22_f_edu3_cons * f_edu3
+ a23_edu4_cons * ed_4
+ a24_f_edu_cons * f_edu4
+ a25_ate_cons * 1

259
quietly gen double `ATE_pa_1' = 1 - normprob(`ATE_xa_1')

*For Current Smoker Portion
quietly gen double `ATE_xb_1'
+ a26_famdis_cons * fam_smoke_dis
= b0_constant_cons ///
+ b1_female_cons * female ///
+ b2_friends_cons * friends_smoke_bi ///
+ b3_parents_cons * parent_smoke_bi ///
+ b4_f_parent_cons * f_parent ///
+ b5_sp1_cons * sp_pca_1 ///
+ b6_f_sp1_cons * f_sp1 ///
+ b7_sp2_cons * sp_pca_2 ///
+ b8_f_sp2_cons * f_sp2 ///
+ b9_sp3_cons * sp_pca_3 ///
+ b10_f_sp3_cons * f_sp3 ///
+ b11_cw1_cons * cw_1 ///
+ b12_f_cw1_cons * f_cw1 ///
+ b13_cw2_cons * cw_2 ///
+ b14_f_cw2_cons * f_cw2 ///
+ b15_cigharm_cons * cig_harm_bi ///
+ b16_f_cigharm_cons * f_cigh ///
+ b17_othercigharm_cons * other_cig_harm_bi ///
+ b18_pte1_cons * pt_pca_1 ///
+ b19_edu2_cons * ed_2 ///
+ b20_f_edu2_cons * f_edu2 ///
+ b21_edu3_cons * ed_3 ///
+ b22_f_edu3_cons * f_edu3 ///
+ b23_edu4_cons * ed_4 ///
+ b24_f_edu4_cons * f_edu4 ///
+ b25_ate_cons * 1 ///
+ b26_famdis_cons * fam_smoke_dis
quietly gen double `ATE_pb_1' = 1 - normprob(`ATE_xb_1')

*For Number of Cigarettes Smoked Portion
quietly gen double `ATE xc_1'
+ a26_famdis_cons * fam_smoke_dis
= c1_female_cons * female ///
+ c2_friends_cons * friends_smoke_bi ///
+ c3_parents_cons * parent_smoke_bi ///
+ c4_fparent_cons * f_parent ///
+ c5_sp1_cons * sp_pca_1 ///
+ c6_fsp1_cons * f_sp1 ///
+ c7_sp2_cons * sp_pca_2 ///
+ c8_fsp2_cons * f_sp2 ///
+ c9_sp3_cons * sp_pca_3 ///
+ c10_fsp3_cons * f_sp3 ///
+ c11_cw1_cons * cw_1 ///
+ c13_cw2_cons * cw_2 ///
+ c14_cw2_cons * cw_2 ///
quietly gen double `ATE_v2_1' = cut1_cons - (`ATE_xc_1')
quietly gen double `ATE_v3_1' = cut2_cons - (`ATE_xc_1')
quietly gen double `ATE_v4_1' = cut3_cons - (`ATE_xc_1')
quietly gen double `ATE_p2_1' = normprob(`ATE_v2_1')
quietly gen double `ATE_p3_1' = normprob(`ATE_v3_1') - normprob(`ATE_v2_1')
quietly gen double `ATE_p4_1' = normprob(`ATE_v4_1') - normprob(`ATE_v3_1')
quietly gen double `ATE_p5_1' = 1 - normprob(`ATE_v4_1')

*Top Category m=5
quietly gen double `ATE_when_one_5' = `ATE_pa_1' * `ATE_pb_1' * `ATE_p5_1'
quietly gen double `ATE_when_one_5_f' = `ATE_pa_1' * `ATE_pb_1' * `ATE_p5_1' if female == 1
quietly gen double `ATE_when_one_5_m' = `ATE_pa_1' * `ATE_pb_1' * `ATE_p5_1' if female == 0

*Middle Category m=4
quietly gen double `ATE_when_one_4' = `ATE_pa_1' * `ATE_pb_1' * `ATE_p4_1'
quietly gen double `ATE_when_one_4_f' = `ATE_pa_1' * `ATE_pb_1' * `ATE_p4_1' if female == 1
quietly gen double `ATE_when_one_4_m' = `ATE_pa_1' * `ATE_pb_1' * `ATE_p4_1' if female == 0

*Middle Category m=3
quietly gen double `ATE_when_one_3' = `ATE_pa_1' * `ATE_pb_1' * `ATE_p3_1'
quietly gen double `ATE_when_one_3_f' = `ATE_pa_1' * `ATE_pb_1' * `ATE_p3_1' if female == 1
quietly gen double `ATE_when_one_3_m' = `ATE_pa_1' * `ATE_pb_1' * `ATE_p3_1' if female == 0

*Bottom Category m=2
quietly gen double `ATE_when_one_2' = `ATE_pa_1' * `ATE_pb_1' * `ATE_p2_1'
quietly gen double `ATE_when_one_2_f' = `ATE_pa_1' * `ATE_pb_1' * `ATE_p2_1' if female == 1
quietly gen double `ATE_when_one_2_m' = `ATE_pa_1' * `ATE_pb_1' * `ATE_p2_1' if female == 0
local when_one `ATE_when_one_5' `ATE_when_one_4' `ATE_when_one_3'
`ATE_when_one_2'
local when_one_f `ATE_when_one_5_f' `ATE_when_one_4_f' `ATE_when_one_3_f'
`ATE_when_one_2_f'
local when_one_m `ATE_when_one_5_m' `ATE_when_one_4_m'
`ATE_when_one_3_m' `ATE_when_one_2_m'

***** When Equal to Zero
*For Ever Smoker Portion
quietly gen double `ATE_xa_0' = a0_constant_cons ///
    + a1_female_cons * female ///
    + a2_friends_cons * friends_smoke_bi ///
    + a3_parents_cons * parent_smoke_bi ///
    + a4_f_parent_cons * f_parent ///
    + a5_sp1_cons * sp_pca_1 ///
    + a6_f_sp1_cons * f_sp1 ///
    + a7_sp2_cons * sp_pca_2 ///
    + a8_f_sp2_cons * f_sp2 ///
    + a9_sp3_cons * sp_pca_3 ///
    + a10_f_sp3_cons * f_sp3 ///
    + a11_cw1_cons * cw_1 ///
    + a12_f_cw1_cons * f_cw1 ///
    + a13_cw2_cons * cw_2 ///
    + a14_f_cw2_cons * f_cw2 ///
    + a15_cigharm_cons * cig_harm_bi ///
    + a16_f_cigharm_cons * f_cigh ///
    + a17_othercigharm_cons * other_cig_harm_bi ///
    + a18_pte1_cons * pt_pca_1 ///
    + a19_edu2_cons * ed_2 ///
    + a20_f_edu2_cons * f_edu2 ///
    + a21_edu3_cons * ed_3 ///
    + a22_f_edu3_cons * f_edu3 ///
    + a23_edu4_cons * ed_4 ///
    + a24_f_edu4_cons * f_edu4 ///
    + a25_ate_cons * 0 ///
    + a26_famdis_cons * fam_smoke_dis
quietly gen double `ATE_pa_0' = 1 - normprob(`ATE_xa_0')

*For Current Smoker Portion
quietly gen double `ATE_xb_0' = b0_constant_cons ///
    + b1_female_cons * female ///
    + b2_friends_cons * friends_smoke_bi ///
    + b3_parents_cons * parent_smoke_bi ///
    + b4_f_parent_cons * f_parent ///
    + b5_sp1_cons * sp_pca_1 ///
    + b6_f_sp1_cons * f_sp1 ///
quietly gen double `ATE_pb_0' = 1 - normprob(`ATE_xb_0')

*For Number of Cigarettes Smoked Portion
quietly gen double `ATE_xc_0' = c1_female_cons * female ///
+ c2_friends_cons * friends_smoke_bi ///
+ c3_parents_cons * parent_smoke_bi ///
+ c4_fparent_cons * f_parent ///
+ c5_sp1_cons * sp_pca_1 ///
+ c6_fsp1_cons * f_sp1 ///
+ c7_sp2_cons * sp_pca_2 ///
+ c8_fsp2_cons * f_sp2 ///
+ c9_sp3_cons * sp_pca_3 ///
+ c10_fsp3_cons * f_sp3 ///
+ c11_cw1_cons * cw_1 ///
+ c13_cw2_cons * cw_2 ///
+ c15_cigharm_cons * cig_harm_bi ///
+ c16_fciharm_cons * f_cigh ///
+ c17_othercigharm_cons * other_cig_harm_bi ///
+ c18_pte1_cons * pt_pca_1 ///
+ c19_edu2_cons * ed_2 ///
+ c21_edu3_cons * ed_3 ///
+ c22_fedu4_cons * ed_4 ///
+ c24_fedu4_cons * f_edu4 ///
+ c25_ate_cons * 0 ///
+ c26_famdis_cons * fam_smoke_dis

quietly gen double `ATE_v2_0' = cut1_cons - (`ATE_xc_0')

263
quietly gen double `ATE_v3_0' = cut2_cons - (`ATE_xc_0')
quietly gen double `ATE_v4_0' = cut3_cons - (`ATE_xc_0')
quietly gen double `ATE_p2_0' = normprob(`ATE_v2_0')
quietly gen double `ATE_p3_0' = normprob(`ATE_v3_0') - normprob(`ATE_v2_0')
quietly gen double `ATE_p4_0' = normprob(`ATE_v4_0') - normprob(`ATE_v3_0')
quietly gen double `ATE_p5_0' = 1 - normprob(`ATE_v4_0')

*Top Category m=5
quietly gen double `ATE_when_zero_5' = `ATE_pa_0' * `ATE_pb_0' * `ATE_p5_0'
quietly gen double `ATE_when_zero_5_f' = `ATE_pa_0' * `ATE_pb_0' * `ATE_p5_0' if female == 1
quietly gen double `ATE_when_zero_5_m' = `ATE_pa_0' * `ATE_pb_0' * `ATE_p5_0'
if female == 0

*Middle Category m=4
quietly gen double `ATE_when_zero_4' = `ATE_pa_0' * `ATE_pb_0' * `ATE_p4_0'
quietly gen double `ATE_when_zero_4_f' = `ATE_pa_0' * `ATE_pb_0' * `ATE_p4_0' if female == 1
quietly gen double `ATE_when_zero_4_m' = `ATE_pa_0' * `ATE_pb_0' * `ATE_p4_0'
if female == 0

*Middle Category m=3
quietly gen double `ATE_when_zero_3' = `ATE_pa_0' * `ATE_pb_0' * `ATE_p3_0'
quietly gen double `ATE_when_zero_3_f' = `ATE_pa_0' * `ATE_pb_0' * `ATE_p3_0' if female == 1
quietly gen double `ATE_when_zero_3_m' = `ATE_pa_0' * `ATE_pb_0' * `ATE_p3_0'
if female == 0

*Bottom Category m=2
quietly gen double `ATE_when_zero_2' = `ATE_pa_0' * `ATE_pb_0' * `ATE_p2_0'
quietly gen double `ATE_when_zero_2_f' = `ATE_pa_0' * `ATE_pb_0' * `ATE_p2_0' if female == 1
quietly gen double `ATE_when_zero_2_m' = `ATE_pa_0' * `ATE_pb_0' * `ATE_p2_0'
if female == 0

local when_zero `ATE_when_zero_5' `ATE_when_zero_4' `ATE_when_zero_3'
`ATE_when_zero_2'
local when_zero_f `ATE_when_zero_5_f' `ATE_when_zero_4_f' `ATE_when_zero_3_f'
`ATE_when_zero_2_f'
local when_zero_m `ATE_when_zero_5_m' `ATE_when_zero_4_m'
`ATE_when_zero_3_m' `ATE_when_zero_2_m'

tempvar me_ATE_5 me_ATE_4 me_ATE_3 me_ATE_2 ///
me_ATE_5_f me_ATE_4_f me_ATE_3_f me_ATE_2_f ///
me_ATE_5_m me_ATE_4_m me_ATE_3_m me_ATE_2_m
local Fmes1 `me_ATE_5' `me_ATE_4' `me_ATE_3' `me_ATE_2'
local Fmes2 `me_ATE_5_f' `me_ATE_4_f' `me_ATE_3_f' `me_ATE_2_f'
local Fmes3 `me_ATE_5_m' `me_ATE_4_m' `me_ATE_3_m' `me_ATE_2_m'

forvalues i = 1/3{
    foreach x of local Fmes`i'{
        quietly gen double `x' = .
    }
    local i = `i' +1
}

*Together
quietly replace `me_ATE_5' = ccm_5 * (`ATE_when_one_5' - `ATE_when_zero_5')
quietly sum `me_ATE_5'
scalar meATE5_mean = r(mean)
return scalar ame_ATE_4 = meATE5_mean

quietly replace `me_ATE_4' = ccm_4 * (`ATE_when_one_4' - `ATE_when_zero_4')
quietly sum `me_ATE_4'
scalar meATE4_mean = r(mean)
return scalar ame_ATE_3 = meATE4_mean

quietly replace `me_ATE_3' = ccm_3 * (`ATE_when_one_3' - `ATE_when_zero_3')
quietly sum `me_ATE_3'
scalar meATE3_mean = r(mean)
return scalar ame_ATE_2 = meATE3_mean

quietly replace `me_ATE_2' = ccm_2 * (`ATE_when_one_2' - `ATE_when_zero_2')
quietly sum `me_ATE_2'
scalar meATE2_mean = r(mean)
return scalar ame_ATE_1 = meATE2_mean

*Girls
quietly replace `me_ATE_5_f' = ccm_5 * (`ATE_when_one_5_f' - `ATE_when_zero_5_f')
quietly sum `me_ATE_5_f'
scalar meATE5f_mean = r(mean)
return scalar ame_ATE_4_f = meATE5f_mean

quietly replace `me_ATE_4_f' = ccm_4 * (`ATE_when_one_4_f' - `ATE_when_zero_4_f')
quietly sum `me_ATE_4_f'
scalar meATE4f_mean = r(mean)
return scalar ame_ATE_3_f = meATE4f_mean

265
quietly replace `me_ATE_3_f' = ccm_3 * (`ATE_when_one_3_f' - `ATE_when_zero_3_f')
quietly sum `me_ATE_3_f'
scalar  meATE3f_mean = r(mean)
return scalar ame_ATE_2_f = meATE3f_mean

quietly replace `me_ATE_2_f' = ccm_2 * (`ATE_when_one_2_f' - `ATE_when_zero_2_f')
quietly sum `me_ATE_2_f'
scalar  meATE2f_mean = r(mean)
return scalar ame_ATE_1_f = meATE2f_mean

*Boys
quietly replace `me_ATE_5_m' = ccm_5 * (`ATE_when_one_5_m' - `ATE_when_zero_5_m')
quietly sum `me_ATE_5_m'
scalar  meATE5m_mean = r(mean)
return scalar ame_ATE_4_m = meATE5m_mean

quietly replace `me_ATE_4_m' = ccm_4 * (`ATE_when_one_4_m' - `ATE_when_zero_4_m')
quietly sum `me_ATE_4_m'
scalar  meATE4m_mean = r(mean)
return scalar ame_ATE_3_m = meATE4m_mean

quietly replace `me_ATE_3_m' = ccm_3 * (`ATE_when_one_3_m' - `ATE_when_zero_3_m')
quietly sum `me_ATE_3_m'
scalar  meATE3m_mean = r(mean)
return scalar ame_ATE_2_m = meATE3m_mean

quietly replace `me_ATE_2_m' = ccm_2 * (`ATE_when_one_2_m' - `ATE_when_zero_2_m')
quietly sum `me_ATE_2_m'
scalar  meATE2m_mean = r(mean)
return scalar ame_ATE_1_m = meATE2m_mean

***** Fam_dis_bi

tempvar famdis_xa_1 famdis_xb_1 famdis_xc_1 famdis_xa_0 famdis_xb_0 ///
famdis_xc_0 famdis_pa_1 famdis_pb_1 famdis_v2_1 famdis_v3_1 ///
famdis_v4_1 famdis_p2_1 famdis_p3_1 famdis_p4_1 famdis_p5_1 ///
famdis_pa_0 famdis_pb_0 famdis_v2_0 famdis_v3_0 famdis_v4_0 ///
famdis_p2_0 famdis_p3_0 famdis_p4_0 famdis_p5_0 ///
famdis_when_one_5 famdis_when_one_4 famdis_when_one_3 ///
famdis_when_one_2  famdis_when_zero_5  famdis_when_zero_4 ///
famdis_when_zero_3  famdis_when_zero_2  famdis_when_one_5_f ///
famdis_when_one_4_f  famdis_when_one_3_f  famdis_when_one_2_f ///
famdis_when_one_5_m  famdis_when_one_4_m ///
famdis_when_one_3_m  famdis_when_one_2_m ///
famdis_when_zero_5_f  famdis_when_zero_4_f  famdis_when_zero_3_f ///
famdis_when_zero_2_f  famdis_when_zero_5_m ///
famdis_when_zero_4_m  famdis_when_zero_3_m ///
famdis_when_zero_2_m

***** When Equal to One
*For Ever Smokers
quietly gen double `famdis_xa_1'       = a0_constant_cons ///
+ a1_female_cons * female ///
+ a2_friends_cons * friends_smoke_bi ///
+ a3_parents_cons * parent_smoke_bi ///
+ a4_f_parent_cons * f_parent ///
+ a5_sp1_cons * sp_pca_1 ///
+ a6_f_sp1_cons * f_sp1 ///
+ a7_sp2_cons * sp_pca_2 ///
+ a8_f_sp2_cons * f_sp2 ///
+ a9_sp3_cons * sp_pca_3 ///
+ a10_f_sp3_cons * f_sp3 ///
+ a11_cw1_cons * cw_1 ///
+ a12_f_cw1_cons * f_cw1 ///
+ a13_cw2_cons * cw_2 ///
+ a14_f_cw2_cons * f_cw2 ///
+ a15_cigharm_cons * cig_harm_bi ///
+ a16_f_cigharm_cons * f_cigh ///
+ a17_othercigharm_cons * other_cig_harm_bi ///
+ a18_pte1_cons * pt_pca_1 ///
+ a19_edu2_cons * ed_2 ///
+ a20_f_edu2_cons * f_edu2 ///
+ a21_edu3_cons * ed_3 ///
+ a22_f_edu3_cons * f_edu3 ///
+ a23_edu4_cons * ed_4 ///
+ a24_f_edu4_cons * f_edu4 ///
+ a25_ate_cons * ate_v1 ///
+ a26_famdis_cons * 1
quietly gen double `famdis_pa_1'       = 1 - normprob(`famdis_xa_1')

*For Current Smoker Portion
quietly gen double `famdis_xb_1'       = b0_constant_cons ///
+ b1_female_cons * female ///
+ b2_friends_cons * friends_smoke_bi ///
+ b3_parents_cons * parent_smoke_bi ///
quietly gen double `famdis_pb_1' = 1 - normprob(`famdis_xb_1')

*For Number of Cigarettes Smoked Portion
quietly gen double `famdis_xc_1' = c1_female_cons * female ///
+ c2_friends_cons * friends_smoke_bi ///
+ c3_parents_cons * parent_smoke_bi ///
+ c4_fparent_cons * f_parent ///
+ c5_sp1_cons * sp_pca_1 ///
+ c6_fsp1_cons * f_sp1 ///
+ c7_sp2_cons * sp_pca_2 ///
+ c8_fsp2_cons * f_sp2 ///
+ c9_sp3_cons * sp_pca_3 ///
+ c10_fsp3_cons * f_sp3 ///
+ c11_cw1_cons * cw_1 ///
+ c13_cw2_cons * cw_2 ///
+ c15_cigharm_cons * cig_harm_bi ///
+ c16_f_cigharm_cons * f_cigh ///
+ c17_othercigharm_cons * other_cig_harm_bi ///
+ c18_pte1_cons * pt_pca_1 ///
+ c19_edu2_cons * ed_2 ///
+ c21_edu3_cons * ed_3 ///
+ c22_f_edu3_cons * f_edu3 ///
+ c23_edu4_cons * ed_4 ///
+ c24_f_edu_cons * f_edu4 ///
+ c25_ate_cons * ate_v1 ///
+ c26_famdis_cons * 1

quietly gen double `famdis_pb_1' = 1 - normprob(`famdis_xb_1')

*For Number of Cigarettes Smoked Portion
quietly gen double `famdis_xc_1' = c1_female_cons * female ///
+ c2_friends_cons * friends_smoke_bi ///
+ c3_parents_cons * parent_smoke_bi ///
+ c4_fparent_cons * f_parent ///
+ c5_sp1_cons * sp_pca_1 ///
+ c6_fsp1_cons * f_sp1 ///
+ c7_sp2_cons * sp_pca_2 ///
+ c8_fsp2_cons * f_sp2 ///
+ c9_sp3_cons * sp_pca_3 ///
+ c10_fsp3_cons * f_sp3 ///
+ c11_cw1_cons * cw_1 ///
+ c13_cw2_cons * cw_2 ///
+ c15_cigharm_cons * cig_harm_bi ///
+ c16_f_cigharm_cons * f_cigh ///
+ c17_othercigharm_cons * other_cig_harm_bi ///
+ c18_pte1_cons * pt_pca_1 ///
+ c19_edu2_cons * ed_2 ///
+ c21_edu3_cons * ed_3 ///
+ c22_f_edu3_cons * f_edu3 ///
+ c23_edu4_cons * ed_4 ///
+ c24_f_edu_cons * f_edu4 ///
+ c25_ate_cons * ate_v1 ///
+ c26_famdis_cons * 1

quietly gen double `famdis_v2_1' = cut1_cons - (famdis_xc_1')
quietly gen double `famdis_v3_1' = cut2_cons - (famdis_xc_1')
quietly gen double `famdis_v4_1' = cut3_cons - (famdis_xc_1')
quietly gen double `famdis_v5_1' = normprob(famdis_v2_1') - normprob(famdis_v3_1')
quietly gen double `famdis_p2_1' = normprob(famdis_v4_1') - normprob(famdis_v5_1')
quietly gen double `famdis_p3_1' = 1 - normprob(famdis_v4_1')

*Top Category m=5
quietly gen double `famdis_when_one_5' = famdis_pa_1' * famdis_pb_1' * famdis_p5_1'
quietly gen double `famdis_when_one_5_f' = famdis_pa_1' * famdis_pb_1' * famdis_p5_1' if female == 1
quietly gen double `famdis_when_one_5_m' = famdis_pa_1' * famdis_pb_1' * famdis_p5_1' if female == 0

*Middle Category m=4
quietly gen double `famdis_when_one_4' = famdis_pa_1' * famdis_pb_1' * famdis_p4_1'
quietly gen double `famdis_when_one_4_f' = famdis_pa_1' * famdis_pb_1' * famdis_p4_1' if female == 1
quietly gen double `famdis_when_one_4_m' = famdis_pa_1' * famdis_pb_1' * famdis_p4_1' if female == 0

*Middle Category m=3
quietly gen double `famdis_when_one_3' = famdis_pa_1' * famdis_pb_1' * famdis_p3_1'
quietly gen double `famdis_when_one_3_f' = famdis_pa_1' * famdis_pb_1' * famdis_p3_1' if female == 1
quietly gen double `famdis_when_one_3_m' = famdis_pa_1' * famdis_pb_1' * famdis_p3_1' if female == 0

*Bottom Category m=2
quietly gen double `famdis_when_one_2' = famdis_pa_1' * famdis_pb_1' * famdis_p2_1'
quietly gen double `famdis_when_one_2_f' = famdis_pa_1' * famdis_pb_1' * famdis_p2_1' if female == 1
quietly gen double `famdis_when_one_2_m' = famdis_pa_1' * famdis_pb_1' * famdis_p2_1' if female == 0

local when_one `famdis_when_one_5' `famdis_when_one_4' `famdis_when_one_3' `famdis_when_one_2'
local when_one_f `famdis_when_one_5_f' `famdis_when_one_4_f'
`famdis_when_one_3_f' `famdis_when_one_2_f'
local when_one_m `famdis_when_one_5_m' `famdis_when_one_4_m'
`famdis_when_one_3_m' `famdis_when_one_2_m'

***** When Equal to Zero
*For Ever Smoker Portion
quietly gen double `famdis_xa_0'         = a0_constant_cons ///
    + a1_female_cons * female ///
    + a2_friends_cons * friends_smoke_bi ///
    + a3_parents_cons * parent_smoke_bi ///
    + a4_f_parent_cons * f_parent ///
    + a5_sp1_cons * sp_pca_1 ///
    + a6_f_sp1_cons * f_sp1 ///
    + a7_sp2_cons * sp_pca_2 ///
    + a8_f_sp2_cons * f_sp2 ///
    + a9_sp3_cons * sp_pca_3 ///
    + a10_f_sp3_cons * f_sp3 ///
    + a11_cw1_cons * cw_1 ///
    + a12_f_cw1_cons * f_cw1 ///
    + a13_cw2_cons * cw_2 ///
    + a14_f_cw2_cons * f_cw2 ///
    + a15_cigharm_cons * cig_harm_bi ///
    + a16_f_cigharm_cons * f_cigh ///
    + a17_othercigharm_cons * other_cig_harm_bi ///
    + a18_pte1_cons * pt_pca_1 ///
    + a19_edu2_cons * ed_2 ///
    + a20_f_edu2_cons * f_edu2 ///
    + a21_edu3_cons * ed_3 ///
    + a22_f_edu3_cons * f_edu3 ///
    + a23_edu4_cons * ed_4 ///
    + a24_f_edu_cons * f_edu4 ///
    + a25_ate_cons * ate_v1 ///
    + a26_famdis_cons * 0
quietly gen double `famdis_pa_0' = 1 - normprob(`famdis_xa_0')

*For Current Smoker Portion
quietly gen double `famdis_xb_0'         = b0_constant_cons ///
    + b1_female_cons * female ///
    + b2_friends_cons * friends_smoke_bi ///
    + b3_parents_cons * parent_smoke_bi ///
    + b4_f_parent_cons * f_parent ///
    + b5_sp1_cons * sp_pca_1 ///
    + b6_f_sp1_cons * f_sp1 ///
    + b7_sp2_cons * sp_pca_2 ///
    + b8_f_sp2_cons * f_sp2 ///
quietly gen double `famdis_pb_0' = 1 - normprob(`famdis_xb_0')

*For Number of Cigarettes Smoked Portion
quietly gen double `famdis_xc_0' = c1_female_cons * female ///
+ c2_friends_cons * friends_smoke_bi ///
+ c3_parents_cons * parent_smoke_bi ///
+ c4_fparent_cons * f_parent ///
+ c5_sp1_cons * sp_pca_1 ///
+ c6fsp1_cons * f_sp1 ///
+ c7_sp2_cons * sp_pca_2 ///
+ c8fsp2_cons * f_sp2 ///
+ c9_sp3_cons * sp_pca_3 ///
+ c10fsp3_cons * f_sp3 ///
+ c11cw1_cons * cw_1 ///
+ c13cw2_cons * cw_2 ///
+ c15cigharm_cons * cig_harm_bi ///
+ c16fcigh_cons * f_cigh ///
+ c17_othercigharm_cons * other_cig_harm_bi ///
+ c18_pte1_cons * pt_pca_1 ///
+ c19_edu2_cons * ed_2 ///
+ c21_edu3_cons * ed_3 ///
+ c23_edu4_cons * ed_4 ///
+ c24_fedu4_cons * f_edu4 ///
+ c25_ate_cons * ate_v1 ///
+ c26_famdis_cons * 0
quietly gen double `famdis_v2_0' = cut1_cons - (`famdis_xc_0')
quietly gen double `famdis_v3_0' = cut2_cons - (`famdis_xc_0')
quietly gen double `famdis_v4_0' = cut3_cons - (`famdis_xc_0')
quietly gen double `famdis_p2_0' = normprob(`famdis_v2_0')
quietly gen double `famdis_p3_0' = normprob(`famdis_v3_0') - normprob(`famdis_v2_0')
quietly gen double `famdis_p4_0' = normprob(`famdis_v4_0') - normprob(`famdis_v3_0')
quietly gen double `famdis_p5_0' = 1 - normprob(`famdis_v4_0')

*Top Category m=5
quietly gen double `famdis_when_zero_5' = `famdis_pa_0' * `famdis_pb_0' * `famdis_p5_0'
quietly gen double `famdis_when_zero_5_f' = `famdis_pa_0' * `famdis_pb_0' * `famdis_p5_0' if female == 1
quietly gen double `famdis_when_zero_5_m' = `famdis_pa_0' * `famdis_pb_0' * `famdis_p5_0' if female == 0

*Middle Category m=4
quietly gen double `famdis_when_zero_4' = `famdis_pa_0' * `famdis_pb_0' * `famdis_p4_0'
quietly gen double `famdis_when_zero_4_f' = `famdis_pa_0' * `famdis_pb_0' * `famdis_p4_0' if female == 1
quietly gen double `famdis_when_zero_4_m' = `famdis_pa_0' * `famdis_pb_0' * `famdis_p4_0' if female == 0

*Middle Category m=3
quietly gen double `famdis_when_zero_3' = `famdis_pa_0' * `famdis_pb_0' * `famdis_p3_0'
quietly gen double `famdis_when_zero_3_f' = `famdis_pa_0' * `famdis_pb_0' * `famdis_p3_0' if female == 1
quietly gen double `famdis_when_zero_3_m' = `famdis_pa_0' * `famdis_pb_0' * `famdis_p3_0' if female == 0

*Bottom Category m=2
quietly gen double `famdis_when_zero_2' = `famdis_pa_0' * `famdis_pb_0' * `famdis_p2_0'
quietly gen double `famdis_when_zero_2_f' = `famdis_pa_0' * `famdis_pb_0' * `famdis_p2_0' if female == 1
quietly gen double `famdis_when_zero_2_m' = `famdis_pa_0' * `famdis_pb_0' * `famdis_p2_0' if female == 0

local when_zero `famdis_when_zero_5' `famdis_when_zero_4'
local when_zero `famdis_when_zero_3' `famdis_when_zero_2'
local when_zero_f `famdis_when_zero_5_f` `famdis_when_zero_4_f'
local when_zero_f `famdis_when_zero_3_f` `famdis_when_zero_2_f'
local when_zero_m `famdis_when_zero_5_m` `famdis_when_zero_4_m`
local when_zero_m `famdis_when_zero_3_m` `famdis_when_zero_2_m`
tempvar me_famdis_5  me_famdis_4  me_famdis_3  me_famdis_2
me_famdis_5_f  me_famdis_4_f  me_famdis_3_f  me_famdis_2_f
me_famdis_5_m  me_famdis_4_m  me_famdis_3_m  me_famdis_2_m

local Gmes1 `me_famdis_5'   `me_famdis_4'   `me_famdis_3'   `me_famdis_2'
local Gmes2 `me_famdis_5_f' `me_famdis_4_f' `me_famdis_3_f' `me_famdis_2_f'
local Gmes3 `me_famdis_5_m' `me_famdis_4_m' `me_famdis_3_m' `me_famdis_2_m'

forvalues i = 1/3{
    foreach x of local Gmes`i'{
        quietly gen double `x' = .
    }
    local i = `i' +1
}

*Together
quietly replace `me_famdis_5' = ccm_5 * (`famdis_when_one_5' - `famdis_when_zero_5')
quietly sum `me_famdis_5'
scalar mefamdis5_mean = r(mean)
return scalar ame_famdis_4 = mefamdis5_mean

quietly replace `me_famdis_4' = ccm_4 * (`famdis_when_one_4' - `famdis_when_zero_4')
quietly sum `me_famdis_4'
scalar mefamdis4_mean = r(mean)
return scalar ame_famdis_3 = mefamdis4_mean

quietly replace `me_famdis_3' = ccm_3 * (`famdis_when_one_3' - `famdis_when_zero_3')
quietly sum `me_famdis_3'
scalar mefamdis3_mean = r(mean)
return scalar ame_famdis_2 = mefamdis3_mean

quietly replace `me_famdis_2' = ccm_2 * (`famdis_when_one_2' - `famdis_when_zero_2')
quietly sum `me_famdis_2'
scalar mefamdis2_mean = r(mean)
return scalar ame_famdis_1 = mefamdis2_mean

*Girls
quietly replace `me_famdis_5_f' = ccm_5 * (`famdis_when_one_5_f' - `famdis_when_zero_5_f')
quietly sum `me_famdis_5_f'
scalar mefamdis5f_mean = r(mean)
return scalar ame_famdis_4_f = mefamdis5f_mean
quietly replace `me_famdis_4_f' = ccm_4 * (`famdis_when_one_4_f' - `famdis_when_zero_4_f')
quietly sum `me_famdis_4_f'
scalar mefamdis4f_mean = r(mean)
return scalar ame_famdis_3_f = mefamdis4f_mean

quietly replace `me_famdis_3_f' = ccm_3 * (`famdis_when_one_3_f' - `famdis_when_zero_3_f')
quietly sum `me_famdis_3_f'
scalar mefamdis3f_mean = r(mean)
return scalar ame_famdis_2_f = mefamdis3f_mean

quietly replace `me_famdis_2_f' = ccm_2 * (`famdis_when_one_2_f' - `famdis_when_zero_2_f')
quietly sum `me_famdis_2_f'
scalar mefamdis2f_mean = r(mean)
return scalar ame_famdis_1_f = mefamdis2f_mean

*Boys
quietly replace `me_famdis_5_m' = ccm_5 * (`famdis_when_one_5_m' - `famdis_when_zero_5_m')
quietly sum `me_famdis_5_m'
scalar mefamdis5m_mean = r(mean)
return scalar ame_famdis_4_m = mefamdis5m_mean

quietly replace `me_famdis_4_m' = ccm_4 * (`famdis_when_one_4_m' - `famdis_when_zero_4_m')
quietly sum `me_famdis_4_m'
scalar mefamdis4m_mean = r(mean)
return scalar ame_famdis_3_m = mefamdis4m_mean

quietly replace `me_famdis_3_m' = ccm_3 * (`famdis_when_one_3_m' - `famdis_when_zero_3_m')
quietly sum `me_famdis_3_m'
scalar mefamdis3m_mean = r(mean)
return scalar ame_famdis_2_m = mefamdis3m_mean

quietly replace `me_famdis_2_m' = ccm_2 * (`famdis_when_one_2_m' - `famdis_when_zero_2_m')
quietly sum `me_famdis_2_m'
scalar mefamdis2m_mean = r(mean)
return scalar ame_famdis_1_m = mefamdis2m_mean

***** Parent_smoke_bi
tempvar parent_xa_1 parent_xb_1 parent_xc_1 parent_xa_0 parent_xb_0 ///
parent_xc_0 parent_pa_1 parent_pb_1 parent_v2_1 parent_v3_1 ///
parent_v4_1 parent_p2_1 parent_p3_1 parent_p4_1 parent_p5_1 ///
parent_pa_0 parent_pb_0 parent_v2_0 parent_v3_0 parent_v4_0 ///
parent_p2_0 parent_p3_0 parent_p4_0 parent_p5_0 ///
parent_when_one_5 parent_when_one_4 ///
parent_when_one_3 parent_when_one_2 ///
parent_when_zero_5 parent_when_zero_4 parent_when_zero_3 ///
parent_when_zero_2 parent_when_zero_1_f parent_when_one_4_f ///
parent_when_one_3_f parent_when_one_2_f parent_when_one_5_m ///
parent_when_zero_5_f parent_when_zero_4_f parent_when_zero_3_f ///
parent_when_zero_2_f parent_when_zero_5_m ///
parent_when_zero_4_m parent_when_zero_3_m parent_when_zero_2_m

***** When Equal to One
*For Ever Smokers
quietly gen double `parent_xa_1'
= a0_constant_cons ///
+ a1_female_cons * female ///
+ a2_friends_cons * friends_smoke_bi ///
+ a3_parents_cons * 1 ///
+ a4_f_parent_cons * female ///
+ a5_sp1_cons * sp_pca_1 ///
+ a6_f_sp1_cons * f_sp1 ///
+ a7_sp2_cons * sp_pca_2 ///
+ a8_f_sp2_cons * f_sp2 ///
+ a9_sp3_cons * sp_pca_3 ///
+ a10_f_sp3_cons * f_sp3 ///
+ a11_cw1_cons * cw_1 ///
+ a12_f_cw1_cons * f_cw1 ///
+ a13_cw2_cons * cw_2 ///
+ a14_f_cw2_cons * f_cw2 ///
+ a15_cigharm_cons * cig_harm_bi ///
+ a16_f_cigharm_cons * f_cigh ///
+ a17_othercigharm_cons * other_cig_harm_bi ///
+ a18_pte1_cons * pt_pca_1 ///
+ a19_edu2_cons * ed_2 ///
+ a20_f_edu2_cons * f_edu2 ///
+ a21_edu3_cons * ed_3 ///
+ a22_f_edu3_cons * f_edu3 ///
+ a23_edu4_cons * ed_4 ///
+ a24_f_edu4_cons * f_edu4 ///
+ a25_ate_cons * ate_v1 ///
+ a26_famdis_cons * fam_smoke_dis
quietly gen double `parent_pa_1'
= 1 - normprob('parent_xa_1')

275
*For Current Smoker Portion*

```
quietly gen double `parent_xb_1' = b0_constant_cons + b1_female_cons * female + b2_friends_cons * friends_smoke_bi + b3_parents_cons * 1 + b4_f_parent_cons * female + b5_sp1_cons * sp_pca_1 + b6_f_sp1_cons * f_sp1 + b7_sp2_cons * sp_pca_2 + b8_f_sp2_cons * f_sp2 + b9_sp3_cons * sp_pca_3 + b10_f_sp3_cons * f_sp3 + b11_cw1_cons * cw_1 + b12_cw1_cons * f_cw1 + b13_cw2_cons * cw_2 + b14_cw2_cons * f_cw2 + b15_cigharm_cons * cig_harm_bi + b16_f_cigharm_cons * f_cigh + b17_othercigharm_cons * other_cig_harm_bi + b18_pte1_cons * pt_pca_1 + b19_edu2_cons * ed_2 + b20_f_edu2_cons * f_edu2 + b21_edu3_cons * ed_3 + b22_f_edu3_cons * f_edu3 + b23_edu4_cons * ed_4 + b24_f_edu_cons * f_edu4 + b25_ate_cons * ate_v1 + b26_famdis_cons * fam_smoke_dis
```

quietly gen double `parent_pb_1' = 1 - normprob(`parent_xb_1')

*For Number of Cigarettes Smoked Portion*

```
quietly gen double `parent_xc_1' = c1_female_cons * female + c2_friends_cons * friends_smoke_bi + c3_parents_cons * 1 + c4_fparent_cons * female + c5_sp1_cons * sp_pca_1 + c6_fsp1_cons * f_sp1 + c7_sp2_cons * sp_pca_2 + c8_fsp2_cons * f_sp2 + c9_sp3_cons * sp_pca_3 + c10_fsp3_cons * f_sp3 + c11_cw1_cons * cw_1 + c12_cw1_cons * f_cw1 + c13_cw2_cons * cw_2 + c15_cigharm_cons * cig_harm_bi + c16_f_cigh_cons * f_cigh + c17_othercigharm_cons * other_cig_harm_bi
```
+ c18_pte1_cons * pt_pca_1 ///
+ c19_edu2_cons * ed_2 ///
+ c21_edu3_cons * ed_3 ///
+ c23_edu4_cons * ed_4 ///
+ c24_fedu4_cons * f_edu4 ///
+ c25_ate_cons * ate_v1 ///
+ c26_famdis_cons * fam_smoke_dis

quietly gen double `parent_v2_1'      = cut1_cons - (`parent_xc_1')
quietly gen double `parent_v3_1'      = cut2_cons - (`parent_xc_1')
quietly gen double `parent_v4_1'      = cut3_cons - (`parent_xc_1')
quietly gen double `parent_p2_1'      = normprob(`parent_v2_1')
quietly gen double `parent_p3_1'      = normprob(`parent_v3_1') - normprob(`parent_v2_1')
quietly gen double `parent_p4_1'      = normprob(`parent_v4_1') - normprob(`parent_v3_1')
quietly gen double `parent_p5_1'      = 1 - normprob(`parent_v4_1')

*Top Category m=5
quietly gen double `parent_when_one_5'   = `parent_pa_1' * `parent_pb_1' * `parent_p5_1'
quietly gen double `parent_when_one_5_f' = `parent_pa_1' * `parent_pb_1' * `parent_p5_1' if female == 1
quietly gen double `parent_when_one_5_m' = `parent_pa_1' * `parent_pb_1' * `parent_p5_1' if female == 0

*Middle Category m=4
quietly gen double `parent_when_one_4'   = `parent_pa_1' * `parent(pb_1' * `parent_p4_1'
quietly gen double `parent_when_one_4_f' = `parent_pa_1' * `parent_pb_1' * `parent_p4_1' if female == 1
quietly gen double `parent_when_one_4_m' = `parent_pa_1' * `parent_pb_1' * `parent_p4_1' if female == 0

*Middle Category m=3
quietly gen double `parent_when_one_3'   = `parent_pa_1' * `parent_pb_1' * `parent_p3_1'
quietly gen double `parent_when_one_3_f' = `parent_pa_1' * `parent_pb_1' * `parent_p3_1' if female == 1
quietly gen double `parent_when_one_3_m' = `parent_pa_1' * `parent_pb_1' * `parent_p3_1' if female == 0

*Bottom Category m=2
quietly gen double `parent_when_one_2'   = `parent_pa_1' * `parent_pb_1' * `parent_p2_1'
quietly gen double `parent_when_one_2_f' = `parent_pa_1' * `parent_pb_1' * `parent_p2_1' if female == 1
quietly gen double `parent_when_one_2_m' = `parent_pa_1' * `parent_pb_1' * `parent_p2_1' if female == 0

local when_one `parent_when_one_5' `parent_when_one_4' `parent_when_one_3' `parent_when_one_2'
local when_one_f `parent_when_one_5_f' `parent_when_one_4_f' `parent_when_one_3_f' `parent_when_one_2_f'
local when_one_m `parent_when_one_5_m' `parent_when_one_4_m' `parent_when_one_3_m' `parent_when_one_2_m'

***** When Equal to Zero
*For Ever Smoker Portion
quietly gen double `parent_xa_0' = a0_constant_cons ///
+ a1_female_cons * female ///
+ a2_friends_cons * friends_smoke_bi ///
+ a3_parents_cons * 0 ///
+ a4_f_parent_cons * 0 ///
+ a5_sp1_cons * sp_pca_1 ///
+ a6_f_sp1_cons * f_sp1 ///
+ a7_sp2_cons * sp_pca_2 ///
+ a8_f_sp2_cons * f_sp2 ///
+ a9_sp3_cons * sp_pca_3 ///
+ a10_f_sp3_cons * f_sp3 ///
+ a11_cw1_cons * cw_1 ///
+ a12_f_cw1_cons * f_cw1 ///
+ a13_cw2_cons * cw_2 ///
+ a14_f_cw2_cons * f_cw2 ///
+ a15_cigharm_cons * cig_harm_bi ///
+ a16_f_cigharm_cons * f_cigh ///
+ a17_othercigharm_cons * other_cig_harm_bi ///
+ a18_pte1_cons * pt_pca_1 ///
+ a19_edu2_cons * ed_2 ///
+ a20_f_edu2_cons * f_edu2 ///
+ a21_edu3_cons * ed_3 ///
+ a22_f_edu3_cons * f_edu3 ///
+ a23_edu4_cons * ed_4 ///
+ a24_f_edu_cons * f_edu4 ///
+ a25_ate_cons * ate_v1 ///
+ a26_famdis_cons * fam_smoke_dis

quietly gen double `parent_pa_0' = 1 - normprob(`parent_xa_0')

*For Current Smoker Portion
quietly gen double `parent_xb_0' = b0_constant_cons ///
+ b1_female_cons * female ///
+ b2_friends_cons * friends_smoke_bi ///
+ b3_parents_cons * 0 ///
quietly gen double `parent_pb_0' = 1 - normprob(`parent_xb_0')

*For Number of Cigarettes Smoked Portion
quietly gen double `parent_xc_0' = c1_female_cons * female //
  + c2_friends_cons * friends_smoke_bi //
  + c3_parents_cons * 0 //
  + c4_fparent_cons * 0 //
  + c5_sp1_cons * sp_pca_1 //
  + c6_fsp1_cons * f_sp1 //
  + c7_sp2_cons * sp_pca_2 //
  + c8_fsp2_cons * f_sp2 //
  + c9_sp3_cons * sp_pca_3 //
  + c10_fsp3_cons * f_sp3 //
  + c11_cw1_cons * cw_1 //
  + c13_cw2_cons * cw_2 //
  + c15_cigharm_cons * cig_harm_bi //
  + c16_f_cigharm_cons * f_cigh //
  + c17_othercigharm_cons * other_cig_harm_bi //
  + c18_pte1_cons * pt_pca_1 //
  + c19_edu2_cons * ed_2 //
  + c20_f_edu2_cons * f_edu2 //
  + c21_edu3_cons * ed_3 //
  + c22_f_edu3_cons * f_edu3 //
  + c23_edu4_cons * ed_4 //
  + c24_f_edu_cons * f_edu4 //
  + c25_ate_cons * ate_v1 //
  + c26_famdis_cons * fam_smoke_dis //

279
quietly gen double `parent_v2_0' = cut1_cons - (`parent_xc_0')
quietly gen double `parent_v3_0' = cut2_cons - (`parent_xc_0')
quietly gen double `parent_v4_0' = cut3_cons - (`parent_xc_0')
quietly gen double `parent_p2_0' = normprob(`parent_v2_0')
quietly gen double `parent_p3_0' = normprob(`parent_v3_0') - normprob(`parent_v2_0')
quietly gen double `parent_p4_0' = normprob(`parent_v4_0') - normprob(`parent_v3_0')
quietly gen double `parent_p5_0' = 1 - normprob(`parent_v4_0')

*Top Category m=5
quietly gen double `parent_when_zero_5' = `parent_pa_0' * `parent_pb_0' * `parent_p5_0' if female == 1
quietly gen double `parent_when_zero_5_m' = `parent_pa_0' * `parent_pb_0' * `parent_p5_0' if female == 0

*Middle Category m=4
quietly gen double `parent_when_zero_4' = `parent_pa_0' * `parent_pb_0' * `parent_p4_0' if female == 1
quietly gen double `parent_when_zero_4_m' = `parent_pa_0' * `parent_pb_0' * `parent_p4_0' if female == 0

*Middle Category m=3
quietly gen double `parent_when_zero_3' = `parent_pa_0' * `parent_pb_0' * `parent_p3_0' if female == 1
quietly gen double `parent_when_zero_3_m' = `parent_pa_0' * `parent_pb_0' * `parent_p3_0' if female == 0

*Bottom Category m=2
quietly gen double `parent_when_zero_2' = `parent_pa_0' * `parent_pb_0' * `parent_p2_0' if female == 1
quietly gen double `parent_when_zero_2_m' = `parent_pa_0' * `parent_pb_0' * `parent_p2_0' if female == 0

local when_zero `parent_when_zero_5' `parent_when_zero_4' `parent_when_zero_3' `parent_when_zero_2'
local when_zero_f `parent_when_zero_5_f` `parent_when_zero_4_f` `parent_when_zero_3_f` `parent_when_zero_2_f` 
local when_zero_m `parent_when_zero_5_m` `parent_when_zero_4_m` `parent_when_zero_3_m` `parent_when_zero_2_m`

tempvar me_parent_5 me_parent_4 me_parent_3 me_parent_2 ///
  me_parent_5_f me_parent_4_f me_parent_3_f me_parent_2_f ///
  me_parent_5_m me_parent_4_m me_parent_3_m me_parent_2_m

local Hmes1 `me_parent_5' `me_parent_4' `me_parent_3' `me_parent_2'
local Hmes2 `me_parent_5_f' `me_parent_4_f' `me_parent_3_f' `me_parent_2_f'
local Hmes3 `me_parent_5_m' `me_parent_4_m' `me_parent_3_m' `me_parent_2_m'

forvalues i = 1/3{
  foreach x of local Hmes`i'{
    quietly gen double `x' = .
  }
  local i = `i' +1
}

*Together
quietly replace `me_parent_5' = ccm_5 * (`parent_when_one_5' - `parent_when_zero_5')
quietly sum `me_parent_5'
scalar meparent5_mean = r(mean)
return scalar ame_parent_4 = meparent5_mean

quietly replace `me_parent_4' = ccm_4 * (`parent_when_one_4' - `parent_when_zero_4')
quietly sum `me_parent_4'
scalar meparent4_mean = r(mean)
return scalar ame_parent_3 = meparent4_mean

quietly replace `me_parent_3' = ccm_3 * (`parent_when_one_3' - `parent_when_zero_3')
quietly sum `me_parent_3'
scalar meparent3_mean = r(mean)
return scalar ame_parent_2 = meparent3_mean

quietly replace `me_parent_2' = ccm_2 * (`parent_when_one_2' - `parent_when_zero_2')
quietly sum `me_parent_2'
scalar meparent2_mean = r(mean)
return scalar ame_parent_1 = meparent2_mean

*Girls
quietly replace `me_parent_5_f' = ccm_5 * (`parent_one_5_f' - `parent_two_5_f')
quietly sum `me_parent_5_f'
scalar meparent5f_mean = r(mean)
return scalar ame_parent_4_f = meparent5f_mean

quietly replace `me_parent_4_f' = ccm_4 * ('parent_when_one_4_f' - 'parent_when_zero_4_f')
quietly sum `me_parent_4_f'
scalar meparent4f_mean = r(mean)
return scalar ame_parent_3_f = meparent4f_mean

quietly replace `me_parent_3_f' = ccm_3 * ('parent_when_one_3_f' - 'parent_when_zero_3_f')
quietly sum `me_parent_3_f'
scalar meparent3f_mean = r(mean)
return scalar ame_parent_2_f = meparent3f_mean

return scalar ame_parent_1_f = meparent2f_mean

*Boys
quietly replace `me_parent_5_m' = ccm_5 * ('parent_when_one_5_m' - 'parent_when_zero_5_m')
quietly sum `me_parent_5_m'
scalar meparent5m_mean = r(mean)
return scalar ame_parent_4_m = meparent5m_mean

quietly replace `me_parent_4_m' = ccm_4 * ('parent_when_one_4_m' - 'parent_when_zero_4_m')
quietly sum `me_parent_4_m'
scalar meparent4m_mean = r(mean)
return scalar ame_parent_3_m = meparent4m_mean

quietly replace `me_parent_3_m' = ccm_3 * ('parent_when_one_3_m' - 'parent_when_zero_3_m')
quietly sum `me_parent_3_m'
scalar meparent3m_mean = r(mean)
return scalar ame_parent_2_m = meparent3m_mean

quietly replace `me_parent_2_m' = ccm_2 * ('parent_when_one_2_m' - 'parent_when_zero_2_m')
quietly sum `me_parent_2_m'
scalar meparent2m_mean = r(mean)
return scalar ame_parent_1_m = meparent2m_mean

***** cw_1
tempvar cw1_xa_1 cw1_xb_1 cw1_xc_1 cw1_xa_0 cw1_xb_0
   cw1_xc_0 cw1_pa_1 cw1_pb_1 cw1_v2_1 cw1_v3_1
   cw1_v4_1 cw1_p2_1 cw1_p3_1 cw1_p4_1 cw1_p5_1
   cw1_pa_0 cw1_pb_0 cw1_v2_0 cw1_v3_0 cw1_v4_0
   cw1_p2_0 cw1_p3_0 cw1_p4_0 cw1_p5_0 cw1_when_one_5
   cw1_when_one_4 cw1_when_one_3 cw1_when_one_2
   cw1_when_zero_5 cw1_when_zero_4 cw1_when_zero_3
   cw1_when_zero_2 cw1_when_one_5_f cw1_when_one_4_f
   cw1_when_one_3_f cw1_when_one_2_f cw1_when_one_5_m
   cw1_when_one_4_m cw1_when_one_3_m cw1_when_one_2_m
   cw1_when_zero_5_f cw1_when_zero_4_f cw1_when_zero_3_f
   cw1_when_zero_2_f cw1_when_zero_5_m cw1_when_zero_4_m
   cw1_when_zero_3_m cw1_when_zero_2_m

***** When Equal to One
*For Ever Smokers
quietly gen double `cw1_xa_1' = a0_constant_cons ///
   + a1_female_cons * female ///
   + a2_friends_cons * friends_smoke_bi ///
   + a3_parents_cons * parent_smoke_bi ///
   + a4_f_parent_cons * f_parent ///
   + a5_sp1_cons * sp_pca_1 ///
   + a6_f_sp1_cons * f_sp1 ///
   + a7_sp2_cons * sp_pca_2 ///
   + a8_f_sp2_cons * f_sp2 ///
   + a9_sp3_cons * sp_pca_3 ///
   + a10_f_sp3_cons * f_sp3 ///
   + a11_cw1_cons * 1 ///
   + a12_f_cw1_cons * female ///
   + a15_cigharm_cons * cig_harm_bi ///
   + a16_f_cigharm_cons * f_cigh ///
   + a17_othercigharm_cons * other_cig_harm_bi ///
   + a18_pte1_cons * pt_pca_1 ///
   + a19_edu2_cons * ed_2 ///
   + a20_f_edu2_cons * f_edu2 ///
   + a21_edu3_cons * ed_3 ///
   + a22_f_edu3_cons * f_edu3 ///
   + a23_edu4_cons * ed_4 ///
   + a24_f_edu_cons * f_edu4 ///
   + a25_ate_cons * ate_v1 ///
   + a26_famdis_cons * fam_smoke_dis
quietly gen double `cw1_pa_1' = 1 - normprob(`cw1_xa_1')

*For Current Smoker Portion
quietly gen double `cw1_xb_1' = b0_constant_cons ///
+ $b_1 \text{female} \cdot \text{female}$ \\
+ $b_2 \text{friends} \cdot \text{friends\_smoke\_bi}$ \\
+ $b_3 \text{parents} \cdot \text{parent\_smoke\_bi}$ \\
+ $b_4 \text{fparent} \cdot \text{f\_parent}$ \\
+ $b_5 \text{sp1} \cdot \text{sp\_pca\_1}$ \\
+ $b_6 \text{fsp1} \cdot \text{f\_sp1}$ \\
+ $b_7 \text{sp2} \cdot \text{sp\_pca\_2}$ \\
+ $b_8 \text{fsp2} \cdot \text{f\_sp2}$ \\
+ $b_9 \text{sp3} \cdot \text{sp\_pca\_3}$ \\
+ $b_10 \text{fsp3} \cdot \text{f\_sp3}$ \\
+ $b_{11} \text{cw1} \cdot 1$ \\
+ $b_{12} \text{fcw1} \cdot \text{female}$ \\
+ $b_{15} \text{cigharm} \cdot \text{cig\_harm\_bi}$ \\
+ $b_{16} \text{fcigharm} \cdot \text{f\_cigh}$ \\
+ $b_{17} \text{othercigharm} \cdot \text{other\_cig\_harm\_bi}$ \\
+ $b_{18} \text{pte1} \cdot \text{pt\_pca\_1}$ \\
+ $b_{19} \text{edu2} \cdot \text{ed\_2}$ \\
+ $b_{20} \text{fedu2} \cdot \text{f\_edu2}$ \\
+ $b_{21} \text{edu3} \cdot \text{ed\_3}$ \\
+ $b_{22} \text{fedu3} \cdot \text{f\_edu3}$ \\
+ $b_{23} \text{edu4} \cdot \text{ed\_4}$ \\
+ $b_{24} \text{fedu4} \cdot \text{f\_edu4}$ \\
+ $b_{25} \text{ate} \cdot \text{ate\_v1}$ \\
+ $b_{26} \text{famdis} \cdot \text{fam\_smoke\_dis}$

quietly gen double `cw1\_pb\_1' = 1 - normprob(`cw1\_xb\_1')

*For Number of Cigarettes Smoked Portion
quietly gen double `cw1\_xc\_1' = c1\_female\_cons * female \\
+ c2\_friends\_cons * friends\_smoke\_bi \\
+ c3\_parents\_cons * parent\_smoke\_bi \\
+ c4\_fparent\_cons * f\_parent \\
+ c5\_sp1\_cons * sp\_pca\_1 \\
+ c6\_fsp1\_cons * f\_sp1 \\
+ c7\_sp2\_cons * sp\_pca\_2 \\
+ c8\_fsp2\_cons * f\_sp2 \\
+ c9\_sp3\_cons * sp\_pca\_3 \\
+ c10\_fsp3\_cons * f\_sp3 \\
+ c11\_cw1\_cons * 1 \\
+ c15\_cigharm\_cons * cig\_harm\_bi \\
+ c16\_fcigh\_cons * f\_cigh \\
+ c17\_othercigharm\_cons * other\_cig\_harm\_bi \\
+ c18\_pte1\_cons * pt\_pca\_1 \\
+ c19\_edu2\_cons * ed\_2 \\
+ c21\_edu3\_cons * ed\_3 \\
+ c23\_edu4\_cons * ed\_4 \\
+ c24\_fedu4\_cons * f\_edu4
quietly gen double `cw1_v2_1' = cut1_cons - (`cw1_xc_1')
quietly gen double `cw1_v3_1' = cut2_cons - (`cw1_xc_1')
quietly gen double `cw1_v4_1' = cut3_cons - (`cw1_xc_1')
quietly gen double `cw1_p2_1' = normprob(`cw1_v2_1')
quietly gen double `cw1_p3_1' = normprob(`cw1_v3_1') - normprob(`cw1_v2_1')
quietly gen double `cw1_p4_1' = normprob(`cw1_v4_1') - normprob(`cw1_v3_1')
quietly gen double `cw1_p5_1' = 1 - normprob(`cw1_v4_1')

*Top Category m=5
quietly gen double `cw1_when_one_5' = `cw1_pa_1' * `cw1_pb_1' * `cw1_p5_1'
quietly gen double `cw1_when_one_5_f' = `cw1_pa_1' * `cw1_pb_1' * `cw1_p5_1' if female == 1
quietly gen double `cw1_when_one_5_m' = `cw1_pa_1' * `cw1_pb_1' * `cw1_p5_1' if female == 0

*Middle Category m=4
quietly gen double `cw1_when_one_4' = `cw1_pa_1' * `cw1_pb_1' * `cw1_p4_1'
quietly gen double `cw1_when_one_4_f' = `cw1_pa_1' * `cw1_pb_1' * `cw1_p4_1' if female == 1
quietly gen double `cw1_when_one_4_m' = `cw1_pa_1' * `cw1_pb_1' * `cw1_p4_1' if female == 0

*Middle Category m=3
quietly gen double `cw1_when_one_3' = `cw1_pa_1' * `cw1_pb_1' * `cw1_p3_1'
quietly gen double `cw1_when_one_3_f' = `cw1_pa_1' * `cw1_pb_1' * `cw1_p3_1' if female == 1
quietly gen double `cw1_when_one_3_m' = `cw1_pa_1' * `cw1_pb_1' * `cw1_p3_1' if female == 0

*Bottom Category m=2
quietly gen double `cw1_when_one_2' = `cw1_pa_1' * `cw1_pb_1' * `cw1_p2_1'
quietly gen double `cw1_when_one_2_f' = `cw1_pa_1' * `cw1_pb_1' * `cw1_p2_1' if female == 1
quietly gen double `cw1_when_one_2_m' = `cw1_pa_1' * `cw1_pb_1' * `cw1_p2_1' if female == 0

local when_one `cw1_when_one_5' `cw1_when_one_4' `cw1_when_one_3' `cw1_when_one_2'
local when_one_f `cw1_when_one_5_f' `cw1_when_one_4_f' `cw1_when_one_3_f' `cw1_when_one_2_f'
local when_one_m `cw1_when_one_5_m' `cw1_when_one_4_m' `cw1_when_one_3_m' `cw1_when_one_2_m'

***** When Equal to Zero
*For Ever Smoker Portion
quietly gen double `cw1_xa_0' = a0_constant_cons ///
+ a1_female_cons * female ///
+ a2_friends_cons * friends_smoke_bi ///
+ a3_parents_cons * parent_smoke_bi ///
+ a4_f_parent_cons * f_parent ///
+ a5_sp1_cons * sp_pca_1 ///
+ a6_f_sp1_cons * f_sp1 ///
+ a7_sp2_cons * sp_pca_2 ///
+ a8_f_sp2_cons * f_sp2 ///
+ a9_sp3_cons * sp_pca_3 ///
+ a10_f_sp3_cons * f_sp3 ///
+ a11_cw1_cons * 0 ///
+ a12_f_cw1_cons * 0 ///
+ a15_cigharm_cons * cig_harm_bi ///
+ a16_f_cigharm_cons * f_cigh ///
+ a17_othercigharm_cons * other_cig_harm_bi ///
+ a18_pte1_cons * pt_pca_1 ///
+ a19_edu2_cons * ed_2 ///
+ a20_f_edu2_cons * f_edu2 ///
+ a21_edu3_cons * ed_3 ///
+ a22_f_edu3_cons * f_edu3 ///
+ a23_edu4_cons * ed_4 ///
+ a24_f_edu4_cons * f_edu4 ///
+ a25_ate_cons * ate_v1 ///
+ a26_famdis_cons * fam_smoke_dis
quietly gen double `cw1_pa_0' = 1 - normprob(`cw1_xa_0')

*For Current Smoker Portion
quietly gen double `cw1_xb_0' = b0_constant_cons ///
+ b1_female_cons * female ///
+ b2_friends_cons * friends_smoke_bi ///
+ b3_parents_cons * parent_smoke_bi ///
+ b4_f_parent_cons * f_parent ///
+ b5_sp1_cons * sp_pca_1 ///
+ b6_f_sp1_cons * f_sp1 ///
+ b7_sp2_cons * sp_pca_2 ///
+ b8_f_sp2_cons * f_sp2 ///
+ b9_sp3_cons * sp_pca_3 ///
+ b10_f_sp3_cons * f_sp3 ///
+ b11_cw1_cons * 0 ///
+ b12_f_cw1_cons * 0 ///
+ b15_cigharm_cons * cig_harm_bi ///
+ b16_f_cigharm_cons * f_cigh ///
+ b17_othercigharm_cons * other_cig_harm_bi ///
+ b18_pte1_cons * pt_pca_1 ///
+ b19_edu2_cons * ed_2 ///
+ b20_f_edu2_cons * f_edu2 ///
+ b21_edu3_cons * ed_3 ///
+ b22_f_edu3_cons * f_edu3 ///
+ b23_edu4_cons * ed_4 ///
+ b24_f_edu4_cons * f_edu4 ///
+ b25_ate_cons * ate_v1 ///
+ b26_famdis_cons * fam_smoke_dis

quietly gen double `cw1_pb_0' = 1 - normprob(`cw1_xb_0')

*For Number of Cigarettes Smoked Portion

quietly gen double `cw1_xc_0' = c1_female_cons * female ///
+ c2_friends_cons * friends_smoke_bi ///
+ c3_parents_cons * parent_smoke_bi ///
+ c4_fparent_cons * f_parent ///
+ c5_sp1_cons * sp_pca_1 ///
+ c6_fsp1_cons * f_sp1 ///
+ c7_sp2_cons * sp_pca_2 ///
+ c8_fsp2_cons * f_sp2 ///
+ c9_sp3_cons * sp_pca_3 ///
+ c10_fsp3_cons * f_sp3 ///
+ c11_cw1_cons * 0 ///
+ c15_cigharm_cons * cig_harm_bi ///
+ c16_fcigh_cons * f_cigh ///
+ c17_othercigharm_cons * other_cig_harm_bi ///
+ c18_pte1_cons * pt_pca_1 ///
+ c19_edu2_cons * ed_2 ///
+ c21_edu3_cons * ed_3 ///
+ c23_edu4_cons * ed_4 ///
+ c24_fedu4_cons * f_edu4 ///
+ c25_ate_cons * ate_v1 ///
+ c26_famdis_cons * fam_smoke_dis

quietly gen double `cw1_v2_0' = cut1_cons - (`cw1_xc_0')
quietly gen double `cw1_v3_0' = cut2_cons - (`cw1_xc_0')
quietly gen double `cw1_v4_0' = cut3_cons - (`cw1_xc_0')
quietly gen double `cw1_v2_0' = normprob(`cw1_v2_0')
quietly gen double `cw1_v3_0' = normprob(`cw1_v3_0') - normprob(`cw1_v2_0')
quietly gen double `cw1_v4_0' = normprob(`cw1_v4_0') - normprob(`cw1_v3_0')
quietly gen double `cw1_v5_0' = 1 - normprob(`cw1_v4_0')

*Top Category m=5
quietly gen double `cw1_when_zero_5' = `cw1_pa_0' * `cw1_pb_0' * `cw1_p5_0'
quietly gen double `cw1_when_zero_5' = `cw1_pa_0' * `cw1_pb_0' * `cw1_p5_0'
if female == 1
quietly gen double `cw1_when_zero_5_m' = `cw1_pa_0' * `cw1_pb_0' * `cw1_p5_0'
if female == 0

287
*Middle Category m=4
quietly gen double `cw1_when_zero_4' = `cw1_pa_0' * `cw1_pb_0' * `cw1_p4_0'
quietly gen double `cw1_when_zero_4_f' = `cw1_pa_0' * `cw1_pb_0' * `cw1_p4_0' if female == 1
quietly gen double `cw1_when_zero_4_m' = `cw1_pa_0' * `cw1_pb_0' * `cw1_p4_0' if female == 0

*Middle Category m=3
quietly gen double `cw1_when_zero_3' = `cw1_pa_0' * `cw1_pb_0' * `cw1_p3_0'
quietly gen double `cw1_when_zero_3_f' = `cw1_pa_0' * `cw1_pb_0' * `cw1_p3_0' if female == 1
quietly gen double `cw1_when_zero_3_m' = `cw1_pa_0' * `cw1_pb_0' * `cw1_p3_0' if female == 0

*Bottom Category m=2
quietly gen double `cw1_when_zero_2' = `cw1_pa_0' * `cw1_pb_0' * `cw1_p2_0'
quietly gen double `cw1_when_zero_2_f' = `cw1_pa_0' * `cw1_pb_0' * `cw1_p2_0' if female == 1
quietly gen double `cw1_when_zero_2_m' = `cw1_pa_0' * `cw1_pb_0' * `cw1_p2_0' if female == 0

local when_zero `cw1_when_zero_5' `cw1_when_zero_4' `cw1_when_zero_3'
local when_zero `cw1_when_zero_5_f' `cw1_when_zero_4_f' `cw1_when_zero_3_f'
local when_zero `cw1_when_zero_5_m' `cw1_when_zero_4_m' `cw1_when_zero_3_m'

tempvar me_cw1_5   me_cw1_4   me_cw1_3   me_cw1_2 ///
   me_cw1_5_f me_cw1_4_f me_cw1_3_f me_cw1_2_f ///
   me_cw1_5_m me_cw1_4_m me_cw1_3_m me_cw1_2_m

local Imes1 `me_cw1_5' `me_cw1_4' `me_cw1_3' `me_cw1_2'
local Imes2 `me_cw1_5_f' `me_cw1_4_f' `me_cw1_3_f' `me_cw1_2_f'
local Imes3 `me_cw1_5_m' `me_cw1_4_m' `me_cw1_3_m' `me_cw1_2_m'

forvalues i = 1/3{
    foreach x of local Imes`i'{
        quietly gen double `x' = .
    }
    local i = `i' +1
}

*Together
quietly replace `me_cw1_5' = ccm_5 * (`cw1_when_one_5' - `cw1_when_zero_5')
quietly sum `me_cw1_5'
scalar mecw15_mean = r(mean)
return scalar ame_cw1_4 = mecw15_mean

quietly replace `me_cw1_4' = ccm_4 * (cw1_when_one_4' - `cw1_when_zero_4')
quietly sum `me_cw1_4'
scalar mecw14_mean = r(mean)
return scalar ame_cw1_3 = mecw14_mean

quietly replace `me_cw1_3' = ccm_3 * (cw1_when_one_3' - `cw1_when_zero_3')
quietly sum `me_cw1_3'
scalar mecw13_mean = r(mean)
return scalar ame_cw1_2 = mecw13_mean

quietly replace `me_cw1_2' = ccm_2 * (cw1_when_one_2' - `cw1_when_zero_2')
quietly sum `me_cw1_2'
scalar mecw12_mean = r(mean)
return scalar ame_cw1_1 = mecw12_mean

*Girls

quietly replace `me_cw1_5_f' = ccm_5 * (cw1_when_one_5_f' - `cw1_when_zero_5_f')
quietly sum `me_cw1_5_f'
scalar mecw15f_mean = r(mean)
return scalar ame_cw1_4_f = mecw15f_mean

quietly replace `me_cw1_4_f' = ccm_4 * (cw1_when_one_4_f' - `cw1_when_zero_4_f')
quietly sum `me_cw1_4_f'
scalar mecw14f_mean = r(mean)
return scalar ame_cw1_3_f = mecw14f_mean

quietly replace `me_cw1_3_f' = ccm_3 * (cw1_when_one_3_f' - `cw1_when_zero_3_f')
quietly sum `me_cw1_3_f'
scalar mecw13f_mean = r(mean)
return scalar ame_cw1_2_f = mecw13f_mean

quietly replace `me_cw1_2_f' = ccm_2 * (cw1_when_one_2_f' - `cw1_when_zero_2_f')
quietly sum `me_cw1_2_f'
scalar mecw12f_mean = r(mean)
return scalar ame_cw1_1_f = mecw12f_mean

*Boys

quietly replace `me_cw1_5_m' = ccm_5*(cw1_when_one_5_m'-`cw1_when_zero_5_m')
quietly sum `me_cw1_5_m'
scalar mecw15m_mean = r(mean)
return scalar ame_cw1_4_m = mecw15m_mean
quietly replace `me_cw1_4_m' = ccm_4*(`cw1_when_one_4_m'-'cw1_when_zero_4_m')
quietly sum `me_cw1_4_m'
scalar mecw14m_mean = r(mean)
return scalar ame_cw1_4_m = mecw14m_mean

quietly replace `me_cw1_3_m' = ccm_3*(`cw1_when_one_3_m'-'cw1_when_zero_3_m')
quietly sum `me_cw1_3_m'
scalar mecw13m_mean = r(mean)
return scalar ame_cw1_3_m = mecw13m_mean

quietly replace `me_cw1_2_m' = ccm_2*(`cw1_when_one_2_m'-'cw1_when_zero_2_m')
quietly sum `me_cw1_2_m'
scalar mecw12m_mean = r(mean)
return scalar ame_cw1_2_m = mecw12m_mean

***** cw_2

tempvar cw2_xa_1 cw2_xb_1 cw2_xc_1 cw2_xa_0 cw2_xb_0 
  cw2_xc_0 cw2_pa_1 cw2_pb_1 cw2_v2_1 cw2_v3_1 
  cw2_v4_1 cw2_p2_1 cw2_p3_1 cw2_p4_1 cw2_p5_1 
  cw2_pa_0 cw2_pb_0 cw2_v2_0 cw2_v3_0 cw2_v4_0 
  cw2_p2_0 cw2_p3_0 cw2_p4_0 cw2_p5_0 cw2_when_one_5 
  cw2_when_one_4 cw2_when_one_3 cw2_when_one_2 
  cw2_when_zero_5 cw2_when_zero_4 cw2_when_zero_3 
  cw2_when_zero_2 cw2_when_one_5_f cw2_when_one_4_f 
  cw2_when_one_3_f cw2_when_one_2_f cw2_when_one_5_m 
  cw2_when_one_4_m cw2_when_one_3_m cw2_when_one_2_m 
  cw2_when_zero_5_f cw2_when_zero_4_f cw2_when_zero_3_f 
  cw2_when_zero_2_f cw2_when_zero_5_m cw2_when_zero_4_m 
  cw2_when_zero_3_m cw2_when_zero_2_m

***** When Equal to One
*For Ever Smokers
quietly gen double `cw2_xa_1' = a0_constant_cons +
  a1_female_cons * female +
  a2_friends_cons * friends_smoke_bi +
  a3_parents_cons * parent_smoke_bi +
  a4_f_parent_cons * f_parent +
  a5_sp1_cons * sp_pca_1 +
  a6_f_sp1_cons * f_sp1 +
  a7_sp2_cons * sp_pca_2 +
  a8_f_sp2_cons * f_sp2 +
  a9_sp3_cons * sp_pca_3 +
  a10_f_sp3_cons * f_sp3 +
  a13_cw2_cons * 1 +
  a14_f_cw2_cons * female

290
+ a15_cigharm_cons * cig_harm_bi ///
+ a16_f_cigharm_cons * f_cigh ///
+ a17_othercigharm_cons * other_cig_harm_bi ///
+ a18_pte1_cons * pt_pca_1 ///
+ a19_edu2_cons * ed_2 ///
+ a20_f_edu2_cons * f_edu2 ///
+ a21_edu3_cons * ed_3 ///
+ a22_f_edu3_cons * f_edu3 ///
+ a23_edu4_cons * ed_4 ///
+ a24_f_edu_cons * f_edu4 ///
+ a25_ate_cons * ate_v1 ///
+ a26_famdis_cons * fam_smoke_dis

quietly gen double `cw2_pa_1' = 1 - normprob(`cw2_xa_1')

*For Current Smoker Portion
quietly gen double `cw2_xb_1' = b0_constant_cons ///
+ b1_female_cons * female ///
+ b2_friends_cons * friends_smoke_bi ///
+ b3_parents_cons * parent_smoke_bi ///
+ b4_f_parent_cons * f_parent ///
+ b5_sp1_cons * sp_pca_1 ///
+ b6_f_sp1_cons * f_sp1 ///
+ b7_sp2_cons * sp_pca_2 ///
+ b8_f_sp2_cons * f_sp2 ///
+ b9_sp3_cons * sp_pca_3 ///
+ b10_f_sp3_cons * f_sp3 ///
+ b13_cw2_cons * 1 ///
+ b14_f_cw2_cons * female ///
+ b15_cigharm_cons * cig_harm_bi ///
+ b16_f_cigharm_cons * f_cigh ///
+ b17_othercigharm_cons * other_cig_harm_bi ///
+ b18_pte1_cons * pt_pca_1 ///
+ b19_edu2_cons * ed_2 ///
+ b20_f_edu2_cons * f_edu2 ///
+ b21_edu3_cons * ed_3 ///
+ b22_f_edu3_cons * f_edu3 ///
+ b23_edu4_cons * ed_4 ///
+ b24_f_edu_cons * f_edu4 ///
+ b25_ate_cons * ate_v1 ///
+ b26_famdis_cons * fam_smoke_dis

quietly gen double `cw2_pb_1' = 1 - normprob(`cw2_xb_1')

*For Number of Cigarettes Smoked Portion
quietly gen double `cw2_xc_1' = c1_female_cons * female ///
+ c2_friends_cons * friends_smoke_bi ///
+ c3_parents_cons * parent_smoke_bi ///
+ c4_fparent_cons * f_parent ///
+ c5_sp1_cons * sp_pca_1 ///
+ c6_fsp1_cons * f_sp1 ///
+ c7_sp2_cons * sp_pca_2 ///
+ c8_fsp2_cons * f_sp2 ///
+ c9_sp3_cons * sp_pca_3 ///
+ c10_fsp3_cons * f_sp3 ///
+ c13_cw2_cons * 1 ///
+ c15_cigharm_cons * cig_harm_bi ///
+ c16_fcigh_cons * f_cigh ///
+ c17_othercigharm_cons * other_cig_harm_bi ///
+ c18_pte1_cons * pt_pca_1 ///
+ c19_edu2_cons * ed_2 ///
+ c21_edu3_cons * ed_3 ///
+ c23_edu4_cons * ed_4 ///
+ c24_fedu4_cons * f_edu4 ///
+ c25_ate_cons * ate_v1 ///
+ c26_famdis_cons * fam_smoke_dis

quietly gen double `cw2_v2_1' = cut1_cons - (`cw2_xc_1')
quietly gen double `cw2_v3_1' = cut2_cons - (`cw2_xc_1')
quietly gen double `cw2_v4_1' = cut3_cons - (`cw2_xc_1')
quietly gen double `cw2_p2_1' = normprob(`cw2_v2_1')
quietly gen double `cw2_p3_1' = normprob(`cw2_v3_1') - normprob(`cw2_v2_1')
quietly gen double `cw2_p4_1' = normprob(`cw2_v4_1') - normprob(`cw2_v3_1')
quietly gen double `cw2_p5_1' = 1 - normprob(`cw2_v4_1')

*Top Category m=5
quietly gen double `cw2_when_one_5' = `cw2_pa_1' * `cw2_pb_1' * `cw2_p5_1'
quietly gen double `cw2_when_one_5_f' = `cw2_pa_1' * `cw2_pb_1' * `cw2_p5_1' if female == 1
quietly gen double `cw2_when_one_5_m' = `cw2_pa_1' * `cw2_pb_1' * `cw2_p5_1' if female == 0

*Middle Category m=4
quietly gen double `cw2_when_one_4' = `cw2_pa_1' * `cw2_pb_1' * `cw2_p4_1'
quietly gen double `cw2_when_one_4_f' = `cw2_pa_1' * `cw2_pb_1' * `cw2_p4_1' if female == 1
quietly gen double `cw2_when_one_4_m' = `cw2_pa_1' * `cw2_pb_1' * `cw2_p4_1' if female == 0

*Middle Category m=3
quietly gen double `cw2_when_one_3' = `cw2_pa_1' * `cw2_pb_1' * `cw2_p3_1'
quietly gen double `cw2_when_one_3_f' = `cw2_pa_1' * `cw2_pb_1' * `cw2_p3_1' if female == 1
quietly gen double `cw2_when_one_3_m' = `cw2_pa_1' * `cw2_pb_1' * `cw2_p3_1' if female == 0
*Bottom Category m=2
quietly gen double `cw2_when_one_2' = `cw2_pa_1' * `cw2_pb_1' * `cw2_p2_1'
quietly gen double `cw2_when_one_2_f' = `cw2_pa_1' * `cw2_pb_1' * `cw2_p2_1' if female == 1
quietly gen double `cw2_when_one_2_m' = `cw2_pa_1' * `cw2_pb_1' * `cw2_p2_1' if female == 0

local when_one `cw2_when_one_5' `cw2_when_one_4' `cw2_when_one_3' `cw2_when_one_2'
local when_one_f `cw2_when_one_5_f' `cw2_when_one_4_f' `cw2_when_one_3_f' `cw2_when_one_2_f'
local when_one_m `cw2_when_one_5_m' `cw2_when_one_4_m' `cw2_when_one_3_m' `cw2_when_one_2_m'

***** When Equal to Zero
*For Ever Smoker Portion
quietly gen double `cw2_xa_0' = a0_constant_cons ///
   + a1_female_cons * female ///
   + a2_friends_cons * friends_smoke_bi ///
   + a3_parents_cons * parent_smoke_bi ///
   + a4_f_parent_cons * f_parent ///
   + a5_sp1_cons * sp_pca_1 ///
   + a6_f_sp1_cons * f_sp1 ///
   + a7_sp2_cons * sp_pca_2 ///
   + a8_f_sp2_cons * f_sp2 ///
   + a9_sp3_cons * sp_pca_3 ///
   + a10_f_sp3_cons * f_sp3 ///
   + a13_cw2_cons * 0 ///
   + a14_f_cw2_cons * 0 ///
   + a15_cigharm_cons * cig_harm_bi ///
   + a16_f_cigharm_cons * f_cigh ///
   + a17_othercigharm_cons * other_cig_harm_bi ///
   + a18_pte1_cons * pt_pca_1 ///
   + a19_edu2_cons * ed_2 ///
   + a20_f_edu2_cons * f_edu2 ///
   + a21_edu3_cons * ed_3 ///
   + a22_f_edu3_cons * f_edu3 ///
   + a23_edu4_cons * ed_4 ///
   + a24_f_edu_cons * f_edu4 ///
   + a25_ate_cons * ate_v1 ///
   + a26_famdis_cons * fam_smoke_dis
quietly gen double `cw2_pa_0' = 1 - normprob(`cw2_xa_0')

*For Current Smoker Portion
quietly gen double `cw2_xb_0' = b0_constant_cons ///
quietly gen double `cw2_pb_0' = 1 - normprob(`cw2_xb_0')

*For Number of Cigarettes Smoked Portion
quietly gen double `cw2_xc_0' = c1_female_cons * female ///
+ c2_friends_cons * friends_smoke_bi ///
+ c3_parents_cons * parent_smoke_bi ///
+ c4_fparent_cons * f_parent ///
+ c5_sp1_cons * sp_pca_1 ///
+ c6_fsp1_cons * f_sp1 ///
+ c7_sp2_cons * sp_pca_2 ///
+ c8_fsp2_cons * f_sp2 ///
+ c9_sp3_cons * sp_pca_3 ///
+ c10_fsp3_cons * f_sp3 ///
+ c13_cw2_cons * 0 ///
+ c15_cigharm_cons * cig_harm_bi ///
+ c16_fcigh_cons * f_cigh ///
+ c17_othercigharm_cons * other_cig_harm_bi ///
+ c18_pte1_cons * pt_pca_1 ///
+ c19_edu2_cons * ed_2 ///
+ c21_edu3_cons * ed_3 ///
+ c23_edu4_cons * ed_4 ///
+ c24_fedu4_cons * f_edu4 ///
quietly gen double `cw2_v2_0' = cut1_cons - (`cw2_xc_0')
quietly gen double `cw2_v3_0' = cut2_cons - (`cw2_xc_0')
quietly gen double `cw2_v4_0' = cut3_cons - (`cw2_xc_0')
quietly gen double `cw2_p2_0' = normprob(`cw2_v2_0')
quietly gen double `cw2_p3_0' = normprob(`cw2_v3_0') - normprob(`cw2_v2_0')
quietly gen double `cw2_p4_0' = normprob(`cw2_v4_0') - normprob(`cw2_v3_0')
quietly gen double `cw2_p5_0' = 1 - normprob(`cw2_v4_0')

*Top Category m=5
quietly gen double `cw2_when_zero_5' = `cw2_pa_0' * `cw2_pb_0' * `cw2_p5_0'
quietly gen double `cw2_when_zero_5_f' = `cw2_pa_0' * `cw2_pb_0' * `cw2_p5_0' if female == 1
quietly gen double `cw2_when_zero_5_m' = `cw2_pa_0' * `cw2_pb_0' * `cw2_p5_0' if female == 0

*Middle Category m=4
quietly gen double `cw2_when_zero_4' = `cw2_pa_0' * `cw2_pb_0' * `cw2_p4_0'
quietly gen double `cw2_when_zero_4_f' = `cw2_pa_0' * `cw2_pb_0' * `cw2_p4_0' if female == 1
quietly gen double `cw2_when_zero_4_m' = `cw2_pa_0' * `cw2_pb_0' * `cw2_p4_0' if female == 0

*Middle Category m=3
quietly gen double `cw2_when_zero_3' = `cw2_pa_0' * `cw2_pb_0' * `cw2_p3_0'
quietly gen double `cw2_when_zero_3_f' = `cw2_pa_0' * `cw2_pb_0' * `cw2_p3_0' if female == 1
quietly gen double `cw2_when_zero_3_m' = `cw2_pa_0' * `cw2_pb_0' * `cw2_p3_0' if female == 0

*Bottom Category m=2
quietly gen double `cw2_when_zero_2' = `cw2_pa_0' * `cw2_pb_0' * `cw2_p2_0'
quietly gen double `cw2_when_zero_2_f' = `cw2_pa_0' * `cw2_pb_0' * `cw2_p2_0' if female == 1
quietly gen double `cw2_when_zero_2_m' = `cw2_pa_0' * `cw2_pb_0' * `cw2_p2_0' if female == 0

local when_zero `cw2_when_zero_5' `cw2_when_zero_4' `cw2_when_zero_3' `cw2_when_zero_2'
local when_zero_f `cw2_when_zero_5_f' `cw2_when_zero_4_f' `cw2_when_zero_3_f' `cw2_when_zero_2_f'
local when_zero_m `cw2_when_zero_5_m' `cw2_when_zero_4_m' `cw2_when_zero_3_m' `cw2_when_zero_2_m'

tempvar me_cw2_5 me_cw2_4 me_cw2_3 me_cw2_2 ///


```
local Jmes1 `me_cw2_5' `me_cw2_4' `me_cw2_3' `me_cw2_2'
local Jmes2 `me_cw2_5_f' `me_cw2_4_f' `me_cw2_3_f' `me_cw2_2_f'
local Jmes3 `me_cw2_5_m' `me_cw2_4_m' `me_cw2_3_m' `me_cw2_2_m'

forvalues i = 1/3{
    foreach x of local Jmes`i'{
        quietly gen double `x' = .
    }
    local i = `i' +1
}

*Together
quietly replace `me_cw2_5' = ccm_5 * (cw2_when_one_5' - cw2_when_zero_5')
quietly sum `me_cw2_5'
scalar mecw25_mean = r(mean)
return scalar ame_cw2_4 = mecw25_mean

quietly replace `me_cw2_4' = ccm_4 * (cw2_when_one_4' - cw2_when_zero_4')
quietly sum `me_cw2_4'
scalar mecw24_mean = r(mean)
return scalar ame_cw2_3 = mecw24_mean

quietly replace `me_cw2_3' = ccm_3 * (cw2_when_one_3' - cw2_when_zero_3')
quietly sum `me_cw2_3'
scalar mecw23_mean = r(mean)
return scalar ame_cw2_2 = mecw23_mean

quietly replace `me_cw2_2' = ccm_2 * (cw2_when_one_2' - cw2_when_zero_2')
quietly sum `me_cw2_2'
scalar mecw22_mean = r(mean)
return scalar ame_cw2_1 = mecw22_mean

*Girls
quietly replace `me_cw2_5_f' = ccm_5 * (cw2_when_one_5_f' - cw2_when_zero_5_f')
quietly sum `me_cw2_5_f'
scalar mecw25f_mean = r(mean)
return scalar ame_cw2_4_f = mecw25f_mean

quietly replace `me_cw2_4_f' = ccm_4 * (cw2_when_one_4_f' - cw2_when_zero_4_f')
quietly sum `me_cw2_4_f'
scalar mecw24f_mean = r(mean)
return scalar ame_cw2_3_f = mecw24f_mean
```

quietly replace `me_cw2_3_f' = ccm_3 * (`cw2_when_one_3_f' - `cw2_when_zero_3_f')
quietly sum `me_cw2_3_f'
scalar mecw23f_mean = r(mean)
return scalar ame_cw2_2_f = mecw23f_mean

quietly replace `me_cw2_2_f' = ccm_2 * (`cw2_when_one_2_f' - `cw2_when_zero_2_f')
quietly sum `me_cw2_2_f'
scalar mecw22f_mean = r(mean)
return scalar ame_cw2_1_f = mecw22f_mean

*Boys
quietly replace `me_cw2_5_m' = ccm_5*(`cw2_when_one_5_m'-`cw2_when_zero_5_m')
quietly sum `me_cw2_5_m'
scalar mecw25m_mean = r(mean)
return scalar ame_cw2_4_m = mecw25m_mean

quietly replace `me_cw2_4_m' = ccm_4 * (`cw2_when_one_4_m' - `cw2_when_zero_4_m')
quietly sum `me_cw2_4_m'
scalar mecw24m_mean = r(mean)
return scalar ame_cw2_3_m = mecw24m_mean

quietly replace `me_cw2_3_m' = ccm_3 * (`cw2_when_one_3_m' - `cw2_when_zero_3_m')
quietly sum `me_cw2_3_m'
scalar mecw23m_mean = r(mean)
return scalar ame_cw2_2_m = mecw23m_mean

quietly replace `me_cw2_2_m' = ccm_2 * (`cw2_when_one_2_m' - `cw2_when_zero_2_m')
quietly sum `me_cw2_2_m'
scalar mecw22m_mean = r(mean)
return scalar ame_cw2_1_m = mecw22m_mean

***** cig_harm_bi

tempvar cigh_xa_1 cigh_xb_1 cigh_xc_1 cigh_xa_0 cigh_xb_0 ///
cigh_xc_0 cigh_pa_1 cigh_pb_1 cigh_v2_1 cigh_v3_1 ///
cigh_v4_1 cigh_p2_1 cigh_p3_1 cigh_p4_1 cigh_p5_1 ///
cigh_pa_0 cigh_pb_0 cigh_v2_0 cigh_v3_0 cigh_v4_0 ///
cigh_p2_0 cigh_p3_0 cigh_p4_0 cigh_p5_0 cught_one_5 ///
cught_one_4 cught_one_3 cught_one_2 ///
cught_zero_5 cught_zero_4 cught_zero_3 ///
cught_zero_2 cught_one_5_f cught_one_4_f ///
cught_one_3_f cught_one_2_f cught_one_5_m ///
***** When Equal to One

*For Ever Smokers

quietly gen double `cigh_xa_1' = a0_constant_cons ///
+ a1_female_cons * female ///
+ a2_friends_cons * friends_smoke_bi ///
+ a3_parents_cons * parent_smoke_bi ///
+ a4_f_parent_cons * f_parent ///
+ a5_sp1_cons * sp_pca_1 ///
+ a6_f_sp1_cons * f_sp1 ///
+ a7_sp2_cons * sp_pca_2 ///
+ a8_f_sp2_cons * f_sp2 ///
+ a9_sp3_cons * sp_pca_3 ///
+ a10_f_sp3_cons * f_sp3 ///
+ a11_cw1_cons * cw_1 ///
+ a12_f_cw1_cons * f_cw1 ///
+ a13_cw2_cons * cw_2 ///
+ a14_f_cw2_cons * f_cw2 ///
+ a15_cigharm_cons * 1 ///
+ a16_f_cigharm_cons * female ///
+ a17_othercigharm_cons * other_cig_harm_bi ///
+ a18_pte1_cons * pt_pca_1 ///
+ a19_edu2_cons * ed_2 ///
+ a20_f_edu2_cons * f_edu2 ///
+ a21_edu3_cons * ed_3 ///
+ a22_f_edu3_cons * f_edu3 ///
+ a23_edu4_cons * ed_4 ///
+ a24_f_edu_cons * f_edu4 ///
+ a25_ate_cons * ate_v1 ///
+ a26_famdis_cons * fam_smoke_dis

quietly gen double `cigh_pa_1' = 1 - normprob(`cigh_xa_1')

*For Current Smoker Portion

quietly gen double `cigh_xb_1' = b0_constant_cons ///
+ b1_female_cons * female ///
+ b2_friends_cons * friends_smoke_bi ///
+ b3_parents_cons * parent_smoke_bi ///
+ b4_f_parent_cons * f_parent ///
+ b5_sp1_cons * sp_pca_1 ///
+ b6_f_sp1_cons * f_sp1 ///
+ b7_sp2_cons * sp_pca_2 ///
+ b8_f_sp2_cons * f_sp2 ///
+ b9_sp3_cons * sp_pca_3 ///
+ b10_f_sp3_cons * f_sp3 ///
+ b11_cw1_cons * cw_1 ///
+ b12_f_cw1_cons * f_cw1 ///
+ b13_cw2_cons * cw_2 ///
+ b14_f_cw2_cons * f_cw2 ///
+ b15_cigharm_cons * 1 ///
+ b16_f_cigharm_cons * female ///
+ b17_othercigharm_cons * other_cig_harm_bi ///
+ b18_pte1_cons * pt_pca_1 ///
+ b19_edu2_cons * ed_2 ///
+ b20_f_edu2_cons * f_edu2 ///
+ b21_edu3_cons * ed_3 ///
+ b22_f_edu3_cons * f_edu3 ///
+ b23_edu4_cons * ed_4 ///
+ b24_f_edu_cons * f_edu4 ///
+ b25_ate_cons * ate_v1 ///
+ b26_famdis_cons * fam_smoke_dis

class `cigh_pb_1'       = normprob(`cigh_xb_1')

quietly gen double `cigh_xc_1'    = c1_female_cons * female ///
    + c2_friends_cons * friends_smoke_bi ///
    + c3_parents_cons * parent_smoke_bi ///
    + c4_fp_cparent_cons * f_parent ///
    + c5_sp1_cons * sp_pca_1 ///
    + c6_fsp1_cons * f_sp1 ///
    + c7_sp2_cons * sp_pca_2 ///
    + c8fsp2_cons * f_sp2 ///
    + c9_sp3_cons * sp_pca_3 ///
    + c10_fsp3_cons * f_sp3 ///
    + c11_cw1_cons * cw_1 ///
    + c13_cw2_cons * cw_2 ///
    + c15_cigharm_cons * 1 ///
    + c16_fc_cigharm_cons * other_cig_harm_bi ///
    + c17_othercigharm_cons * female ///
    + c18_pte1_cons * pt_pca_1 ///
    + c19_edu2_cons * ed_2 ///
    + c21_edu3_cons * ed_3 ///
    + c23_edu4_cons * ed_4 ///
    + c24_fedu4_cons * f_edu4 ///
    + c25_ate_cons * ate_v1 ///
    + c26_famdis_cons * fam_smoke_dis

class `cigh_v2_1'      = cut1_cons - (`cigh_xc_1')

class `cigh_v3_1'      = cut2_cons - (`cigh_xc_1')

class `cigh_v4_1'      = cut3_cons - (`cigh_xc_1')

299
quietly gen double `cigh_p2_1' = normprob(`cigh_v2_1')
quietly gen double `cigh_p3_1' = normprob(`cigh_v3_1') - normprob(`cigh_v2_1')
quietly gen double `cigh_p4_1' = normprob(`cigh_v4_1') - normprob(`cigh_v3_1')
quietly gen double `cigh_p5_1' = 1 - normprob(`cigh_v4_1')

*Top Category m=5
quietly gen double `cigh_when_one_5' = `cigh_pa_1' * `cigh_pb_1' * `cigh_p5_1'
quietly gen double `cigh_when_one_5_f' = `cigh_pa_1' * `cigh_pb_1' * `cigh_p5_1' if female == 1
quietly gen double `cigh_when_one_5_m' = `cigh_pa_1' * `cigh_pb_1' * `cigh_p5_1' if female == 0

*Middle Category m=4
quietly gen double `cigh_when_one_4' = `cigh_pa_1' * `cigh_pb_1' * `cigh_p4_1'
quietly gen double `cigh_when_one_4_f' = `cigh_pa_1' * `cigh_pb_1' * `cigh_p4_1' if female == 1
quietly gen double `cigh_when_one_4_m' = `cigh_pa_1' * `cigh_pb_1' * `cigh_p4_1' if female == 0

*Middle Category m=3
quietly gen double `cigh_when_one_3' = `cigh_pa_1' * `cigh_pb_1' * `cigh_p3_1'
quietly gen double `cigh_when_one_3_f' = `cigh_pa_1' * `cigh_pb_1' * `cigh_p3_1' if female == 1
quietly gen double `cigh_when_one_3_m' = `cigh_pa_1' * `cigh_pb_1' * `cigh_p3_1' if female == 0

*Bottom Category m=2
quietly gen double `cigh_when_one_2' = `cigh_pa_1' * `cigh_pb_1' * `cigh_p2_1'
quietly gen double `cigh_when_one_2_f' = `cigh_pa_1' * `cigh_pb_1' * `cigh_p2_1' if female == 1
quietly gen double `cigh_when_one_2_m' = `cigh_pa_1' * `cigh_pb_1' * `cigh_p2_1' if female == 0

local when_one `cigh_when_one_5' `cigh_when_one_4' `cigh_when_one_3' `cigh_when_one_2'
local when_one_f `cigh_when_one_5_f' `cigh_when_one_4_f' `cigh_when_one_3_f' `cigh_when_one_2_f'
local when_one_m `cigh_when_one_5_m' `cigh_when_one_4_m' `cigh_when_one_3_m' `cigh_when_one_2_m'

***** When Equal to Zero
*For Ever Smoker Portion
quietly gen double `cigh_xa_0' = a0_constant_cons ///
+ a1_female_cons * female ///
+ a2_friends_cons * friends_smoke_bi ///
+ a3_parents_cons * parent_smoke_bi ///
quietly gen double `cigh_pa_0' = 1 - normprob(`cigh_xa_0')

*For Current Smoker Portion
quietly gen double `cigh_xb_0' = b0_constant_cons ///
+ b1_female_cons * female ///
+ b2_friends_cons * friends_smoke_bi ///
+ b3_parents_cons * parent_smoke_bi ///
+ b4_f_parent_cons * f_parent ///
+ b5_sp1_cons * sp_pca_1 ///
+ b6_f_sp1_cons * f_sp1 ///
+ b7_sp2_cons * sp_pca_2 ///
+ b8_f_sp2_cons * f_sp2 ///
+ b9_sp3_cons * sp_pca_3 ///
+ b10_f_sp3_cons * f_sp3 ///
+ b11_cw1_cons * cw_1 ///
+ b12_f_cw1_cons * f_cw1 ///
+ b13_cw2_cons * cw_2 ///
+ b14_f_cw2_cons * f_cw2 ///
+ b15_cigharm_cons * 0 ///
+ b16_cigharm_cons * 0 ///
+ b17_othercigharm_cons * other_cig_harm_bi ///
+ b18_pte1_cons * pt_pca_1 ///
+ b19_edu2_cons * ed_2 ///
+ b20_f_edu2_cons * f_edu2 ///
+ b21_edu3_cons * ed_3 ///
+ b22_f_edu3_cons * f_edu3 ///
+ b23_edu4_cons * ed_4 ///
+ b24_f_edu4_cons * f_edu4 ///
+ b25_ate_cons * ate_v1 ///
+ b26_famdis_cons * fam_smoke_dis
quietly gen double `cigh_pb_0' = 1 - normprob(`cigh_xb_0')

*For Number of Cigarettes Smoked Portion
quietly gen double `cigh_xc_0' = c1_female_cons * female ///
+ c2_friends_cons * friends_smoke_bi ///
+ c3_parents_cons * parent_smoke_bi ///
+ c4_fparent_cons * f_parent ///
+ c5_sp1_cons * sp_pca_1 ///
+ c6_fsp1_cons * f_sp1 ///
+ c7_sp2_cons * sp_pca_2 ///
+ c8_fsp2_cons * f_sp2 ///
+ c9_sp3_cons * sp_pca_3 ///
+ c10_fsp3_cons * f_sp3 ///
+ c11_cw1_cons * cw_1 ///
+ c13_cw2_cons * cw_2 ///
+ c15_cigharm_cons * 0 ///
+ c16_fcigh_cons * 0 ///
+ c17_othercigharm_cons * other_cig_harm_bi ///
+ c18_pte1_cons * pt_pca_1 ///
+ c19_edu2_cons * ed_2 ///
+ c21_edu3_cons * ed_3 ///
+ c23_edu4_cons * ed_4 ///
+ c24_fedu4_cons * f_edu4 ///
+ c25_ate_cons * ate_v1 ///
+ c26_famdis_cons * fam_smoke_dis

quietly gen double `cigh_v2_0' = cut1_cons - (`cigh_xc_0')
quietly gen double `cigh_v3_0' = cut2_cons - (`cigh_xc_0')
quietly gen double `cigh_v4_0' = cut3_cons - (`cigh_xc_0')
quietly gen double `cigh_p2_0' = normprob(`cigh_v2_0')
quietly gen double `cigh_p3_0' = normprob(`cigh_v3_0') - normprob(`cigh_v2_0')
quietly gen double `cigh_p4_0' = normprob(`cigh_v4_0') - normprob(`cigh_v3_0')
quietly gen double `cigh_p5_0' = 1 - normprob(`cigh_v4_0')

*Top Category m=5
quietly gen double `cigh_when_zero_5' = `cigh_pa_0' * `cigh_pb_0' * `cigh_p5_0'
quietly gen double `cigh_when_zero_5_f' = `cigh_pa_0' * `cigh_pb_0' * `cigh_p5_0' if female == 1
quietly gen double `cigh_when_zero_5_m' = `cigh_pa_0' * `cigh_pb_0' * `cigh_p5_0' if female == 0
*Middle Category m=4
quietly gen double `cigh_when_zero_4' = `cigh_pa_0' * `cigh_pb_0' * `cigh_p4_0'
quietly gen double `cigh_when_zero_4_f' = `cigh_pa_0' * `cigh_pb_0' * `cigh_p4_0' if female == 1
quietly gen double `cigh_when_zero_4_m' = `cigh_pa_0' * `cigh_pb_0' * `cigh_p4_0' if female == 0

*Middle Category m=3
quietly gen double `cigh_when_zero_3' = `cigh_pa_0' * `cigh_pb_0' * `cigh_p3_0'
quietly gen double `cigh_when_zero_3_f' = `cigh_pa_0' * `cigh_pb_0' * `cigh_p3_0' if female == 1
quietly gen double `cigh_when_zero_3_m' = `cigh_pa_0' * `cigh_pb_0' * `cigh_p3_0' if female == 0

*Bottom Category m=2
quietly gen double `cigh_when_zero_2' = `cigh_pa_0' * `cigh_pb_0' * `cigh_p2_0'
quietly gen double `cigh_when_zero_2_f' = `cigh_pa_0' * `cigh_pb_0' * `cigh_p2_0' if female == 1
quietly gen double `cigh_when_zero_2_m' = `cigh_pa_0' * `cigh_pb_0' * `cigh_p2_0' if female == 0

local when_zero `cigh_when_zero_5' `cigh_when_zero_4' `cigh_when_zero_3' `cigh_when_zero_2'
local when_zero_f `cigh_when_zero_5_f' `cigh_when_zero_4_f' `cigh_when_zero_3_f' `cigh_when_zero_2_f'
local when_zero_m `cigh_when_zero_5_m' `cigh_when_zero_4_m' `cigh_when_zero_3_m' `cigh_when_zero_2_m'

tempvar me_cigh_5   me_cigh_4   me_cigh_3   me_cigh_2 ///
   me_cigh_5_f me_cigh_4_f me_cigh_3_f me_cigh_2_f ///
   me_cigh_5_m me_cigh_4_m me_cigh_3_m me_cigh_2_m

local Kmes1 `me_cigh_5' `me_cigh_4' `me_cigh_3' `me_cigh_2'
local Kmes2 `me_cigh_5_f' `me_cigh_4_f' `me_cigh_3_f' `me_cigh_2_f'
local Kmes3 `me_cigh_5_m' `me_cigh_4_m' `me_cigh_3_m' `me_cigh_2_m'

forvalues i = 1/3{
    foreach x of local Kmes'i'{
        quietly gen double `x' = .
    }
    local i = `i' +1
}

*Together
quietly replace `me_cigh_5' = ccm_5 * (`cigh_when_one_5' - `cigh_when_zero_5')
quietly sum `me_cigh_5'
scalar mecigh5_mean = r(mean)
return scalar ame_cigh_4 = mecigh5_mean

quietly replace `me_cigh_4' = ccm_4 * (`cigh_when_one_4' - `cigh_when_zero_4')
quietly sum `me_cigh_4'
scalar mecigh4_mean = r(mean)
return scalar ame_cigh_3 = mecigh4_mean

quietly replace `me_cigh_3' = ccm_3 * (`cigh_when_one_3' - `cigh_when_zero_3')
quietly sum `me_cigh_3'
scalar mecigh3_mean = r(mean)
return scalar ame_cigh_2 = mecigh3_mean

quietly replace `me_cigh_2' = ccm_2 * (`cigh_when_one_2' - `cigh_when_zero_2')
quietly sum `me_cigh_2'
scalar mecigh2_mean = r(mean)
return scalar ame_cigh_1 = mecigh2_mean

*Girls
quietly replace `me_cigh_5_f' = ccm_5 * (`cigh_when_one_5_f' - `cigh_when_zero_5_f')
quietly sum `me_cigh_5_f'
scalar mecigh5f_mean = r(mean)
return scalar ame_cigh_4_f = mecigh5f_mean

quietly replace `me_cigh_4_f' = ccm_4 * (`cigh_when_one_4_f' - `cigh_when_zero_4_f')
quietly sum `me_cigh_4_f'
scalar mecigh4f_mean = r(mean)
return scalar ame_cigh_3_f = mecigh4f_mean

quietly replace `me_cigh_3_f' = ccm_3 * (`cigh_when_one_3_f' - `cigh_when_zero_3_f')
quietly sum `me_cigh_3_f'
scalar mecigh3f_mean = r(mean)
return scalar ame_cigh_2_f = mecigh3f_mean

quietly replace `me_cigh_2_f' = ccm_2 * (`cigh_when_one_2_f' - `cigh_when_zero_2_f')
quietly sum `me_cigh_2_f'
scalar mecigh2f_mean = r(mean)
return scalar ame_cigh_1_f = mecigh2f_mean

*Boys
quietly replace `me_cigh_5_m' = ccm_5 * (`cigh_when_one_5_m' - `cigh_when_zero_5_m')
quietly sum `me_cigh_5_m'
scalar mecigh5m_mean = r(mean)
return scalar ame_cigh_4_m = mecigh5m_mean
quietly replace `me_cigh_4_m' = ccm_4 * (`cigh_when_one_4_m' - `cigh_when_zero_4_m')
quietly sum `me_cigh_4_m'
scalar mecigh4m_mean = r(mean)
return scalar ame_cigh_3_m = mecigh4m_mean

quietly replace `me_cigh_3_m' = ccm_3 * (`cigh_when_one_3_m' - `cigh_when_zero_3_m')
quietly sum `me_cigh_3_m'
scalar mecigh3m_mean = r(mean)
return scalar ame_cigh_2_m = mecigh3m_mean

quietly replace `me_cigh_2_m' = ccm_2 * (`cigh_when_one_2_m' - `cigh_when_zero_2_m')
quietly sum `me_cigh_2_m'
scalar mecigh2m_mean = r(mean)
return scalar ame_cigh_1_m = mecigh2m_mean

end

************************************************************************

bootstrap ame_sp1_4_out = r(ame_sp1_4) ///
    ame_sp1_3_out = r(ame_sp1_3) ///
    ame_sp1_2_out = r(ame_sp1_2) ///
    ame_sp1_1_out = r(ame_sp1_1) ///
    ame_sp1_4_f_out = r(ame_sp1_4_f) ///
    ame_sp1_3_f_out = r(ame_sp1_3_f) ///
    ame_sp1_2_f_out = r(ame_sp1_2_f) ///
    ame_sp1_1_f_out = r(ame_sp1_1_f) ///
    ame_sp1_4_m_out = r(ame_sp1_4_m) ///
    ame_sp1_3_m_out = r(ame_sp1_3_m) ///
    ame_sp1_2_m_out = r(ame_sp1_2_m) ///
    ame_sp1_1_m_out = r(ame_sp1_1_m) ///
    ame_sp2_4_out = r(ame_sp2_4) ///
    ame_sp2_3_out = r(ame_sp2_3) ///
    ame_sp2_2_out = r(ame_sp2_2) ///
    ame_sp2_1_out = r(ame_sp2_1) ///
    ame_sp2_4_f_out = r(ame_sp2_4_f) ///
    ame_sp2_3_f_out = r(ame_sp2_3_f) ///
    ame_sp2_2_f_out = r(ame_sp2_2_f) ///
    ame_sp2_1_f_out = r(ame_sp2_1_f) ///
    ame_sp2_4_m_out = r(ame_sp2_4_m) ///
    ame_sp2_3_m_out = r(ame_sp2_3_m) ///
    ame_sp2_2_m_out = r(ame_sp2_2_m) ///
ame_sp2_1_m_out = r(ame_sp2_1_m) ///
amen_sp3_4_out = r(ame_sp3_4) ///
amen_sp3_3_out = r(ame_sp3_3) ///
amen_sp3_2_out = r(ame_sp3_2) ///
amen_sp3_1_out = r(ame_sp3_1) ///
amen_sp3_4_f_out = r(ame_sp3_4_f) ///
amen_sp3_3_f_out = r(ame_sp3_3_f) ///
amen_sp3_2_f_out = r(ame_sp3_2_f) ///
amen_sp3_1_f_out = r(ame_sp3_1_f) ///
amen_sp3_4_m_out = r(ame_sp3_4_m) ///
amen_sp3_3_m_out = r(ame_sp3_3_m) ///
amen_sp3_2_m_out = r(ame_sp3_2_m) ///
amen_sp3_1_m_out = r(ame_sp3_1_m) ///
amen_poe_4_out = r(ame_poe_4) ///
amen_poe_3_out = r(ame_poe_3) ///
amen_poe_2_out = r(ame_poe_2) ///
amen_poe_1_out = r(ame_poe_1) ///
amen_poe_4_f_out = r(ame_poe_4_f) ///
amen_poe_3_f_out = r(ame_poe_3_f) ///
amen_poe_2_f_out = r(ame_poe_2_f) ///
amen_poe_1_f_out = r(ame_poe_1_f) ///
amen_poe_4_m_out = r(ame_poe_4_m) ///
amen_poe_3_m_out = r(ame_poe_3_m) ///
amen_poe_2_m_out = r(ame_poe_2_m) ///
amen_poe_1_m_out = r(ame_poe_1_m) ///
amen_friend_4_out = r(ame_friend_4) ///
amen_friend_3_out = r(ame_friend_3) ///
amen_friend_2_out = r(ame_friend_2) ///
amen_friend_1_out = r(ame_friend_1) ///
amen_friend_4_f_out = r(ame_friend_4_f) ///
amen_friend_3_f_out = r(ame_friend_3_f) ///
amen_friend_2_f_out = r(ame_friend_2_f) ///
amen_friend_1_f_out = r(ame_friend_1_f) ///
amen_friend_4_m_out = r(ame_friend_4_m) ///
amen_friend_3_m_out = r(ame_friend_3_m) ///
amen_friend_2_m_out = r(ame_friend_2_m) ///
amen_friend_1_m_out = r(ame_friend_1_m) ///
amen_female_4_out = r(ame_female_4) ///
amen_female_3_out = r(ame_female_3) ///
amen_female_2_out = r(ame_female_2) ///
amen_female_1_out = r(ame_female_1) ///
amen_female_4_f_out = r(ame_female_4_f) ///
amen_female_3_f_out = r(ame_female_3_f) ///
amen_female_2_f_out = r(ame_female_2_f) ///
amen_female_1_f_out = r(ame_female_1_f) ///
amen_female_4_m_out = r(ame_female_4_m) ///
ame_female_3_m_out = r(ame_female_3_m) ///
ame_female_2_m_out = r(ame_female_2_m) ///
name_female_1_m_out = r(ame_female_1_m) ///
name_och_4_out = r(ame_och_4) ///
name_och_3_out = r(ame_och_3) ///
name_och_2_out = r(ame_och_2) ///
name_och_1_out = r(ame_och_1) ///
name_och_4_f_out = r(ame_och_4_f) ///
name_och_3_f_out = r(ame_och_3_f) ///
name_och_2_f_out = r(ame_och_2_f) ///
name_och_1_f_out = r(ame_och_1_f) ///
name_och_4_m_out = r(ame_och_4_m) ///
name_och_3_m_out = r(ame_och_3_m) ///
name_och_2_m_out = r(ame_och_2_m) ///
name_och_1_m_out = r(ame_och_1_m) ///
name_ATE_4_out = r(ame_ATE_4) ///
name_ATE_3_out = r(ame_ATE_3) ///
name_ATE_2_out = r(ame_ATE_2) ///
name_ATE_1_out = r(ame_ATE_1) ///
name_ATE_4_f_out = r(ame_ATE_4_f) ///
name_ATE_3_f_out = r(ame_ATE_3_f) ///
name_ATE_2_f_out = r(ame_ATE_2_f) ///
name_ATE_1_f_out = r(ame_ATE_1_f) ///
name_ATE_4_m_out = r(ame_ATE_4_m) ///
name_ATE_3_m_out = r(ame_ATE_3_m) ///
name_ATE_2_m_out = r(ame_ATE_2_m) ///
name_ATE_1_m_out = r(ame_ATE_1_m) ///
name_famdis_4_out = r(ame_famdis_4) ///
name_famdis_3_out = r(ame_famdis_3) ///
name_famdis_2_out = r(ame_famdis_2) ///
name_famdis_1_out = r(ame_famdis_1) ///
name_famdis_4_f_out = r(ame_famdis_4_f) ///
name_famdis_3_f_out = r(ame_famdis_3_f) ///
name_famdis_2_f_out = r(ame_famdis_2_f) ///
name_famdis_1_f_out = r(ame_famdis_1_f) ///
name_famdis_4_m_out = r(ame_famdis_4_m) ///
name_famdis_3_m_out = r(ame_famdis_3_m) ///
name_famdis_2_m_out = r(ame_famdis_2_m) ///
name_famdis_1_m_out = r(ame_famdis_1_m) ///
name_parent_4_out = r(ame_parent_4) ///
name_parent_3_out = r(ame_parent_3) ///
name_parent_2_out = r(ame_parent_2) ///
name_parent_1_out = r(ame_parent_1) ///
name_parent_4_f_out = r(ame_parent_4_f) ///
name_parent_3_f_out = r(ame_parent_3_f) ///
name_parent_2_f_out = r(ame_parent_2_f) ///
ame_parent_1_f_out = r(ame_parent_1_f) ///
ame_parent_4_m_out = r(ame_parent_4_m) ///
name_parent_3_m_out = r(ame_parent_3_m) ///
name_parent_2_m_out = r(ame_parent_2_m) ///
name_parent_1_m_out = r(ame_parent_1_m) ///
name_cw1_4_out = r(ame_cw1_4) ///
name_cw1_3_out = r(ame_cw1_3) ///
name_cw1_2_out = r(ame_cw1_2) ///
name_cw1_1_out = r(ame_cw1_1) ///
name_cw1_4_f_out = r(ame_cw1_4_f) ///
name_cw1_3_f_out = r(ame_cw1_3_f) ///
name_cw1_2_f_out = r(ame_cw1_2_f) ///
name_cw1_1_f_out = r(ame_cw1_1_f) ///
name_cw1_4_m_out = r(ame_cw1_4_m) ///
name_cw1_3_m_out = r(ame_cw1_3_m) ///
name_cw1_2_m_out = r(ame_cw1_2_m) ///
name_cw1_1_m_out = r(ame_cw1_1_m) ///
name_cw2_4_out = r(ame_cw2_4) ///
name_cw2_3_out = r(ame_cw2_3) ///
name_cw2_2_out = r(ame_cw2_2) ///
name_cw2_1_out = r(ame_cw2_1) ///
name_cw2_4_f_out = r(ame_cw2_4_f) ///
name_cw2_3_f_out = r(ame_cw2_3_f) ///
name_cw2_2_f_out = r(ame_cw2_2_f) ///
name_cw2_1_f_out = r(ame_cw2_1_f) ///
name_cw2_4_m_out = r(ame_cw2_4_m) ///
name_cw2_3_m_out = r(ame_cw2_3_m) ///
name_cw2_2_m_out = r(ame_cw2_2_m) ///
name_cw2_1_m_out = r(ame_cw2_1_m) ///
name_cigh_4_out = r(ame_cigh_4) ///
name_cigh_3_out = r(ame_cigh_3) ///
name_cigh_2_out = r(ame_cigh_2) ///
name_cigh_1_out = r(ame_cigh_1) ///
name_cigh_4_f_out = r(ame_cigh_4_f) ///
name_cigh_3_f_out = r(ame_cigh_3_f) ///
name_cigh_2_f_out = r(ame_cigh_2_f) ///
name_cigh_1_f_out = r(ame_cigh_1_f) ///
name_cigh_4_m_out = r(ame_cigh_4_m) ///
name_cigh_3_m_out = r(ame_cigh_3_m) ///
name_cigh_2_m_out = r(ame_cigh_2_m) ///
name_cigh_1_m_out = r(ame_cigh_1_m) ///
, rep(200) saving(boot_results_a_level_ame, replace) ///
seed(1): level_con_w_int_marginal

.estat bootstrap, all

.log close
clear all
capture log close

cd /Users/kristinapiorkowski/Chp4"

use “GYTS_0711_var_082917”

keep ever_smoke friends_smoke_bi parent_smoke_bi female num_cigs cur_smoker ///
    sch_class sp_pca_1 sp_pca_2 sp_pca_3 cig_weight cig_harm_bi ///
    other_cig_harm_bi pt_pca_1 edu_quality ate_v1 fam_smoke_dis FinalWgt ///

***** Set up variables needed for likelihood function
*Recode number of cigs smoked, all obs
replace num_cigs = 1 if num_cigs == . & ever_smoke != .
recode num_cigs (5/7 = 5)
tab num_cigs, generate(nc_)

*Generate Indicator Variable for Never Smoked (S_i=0)
gen never_ind = 0 if ever_smoke != .
replace never_ind = 1 if ever_smoke == 0

*Generate Indicator Variable for Current Smokers (S_i=1, Q_i=1)
gen current_ind = 0 if ever_smoke != .
replace current_ind = 1 if cur_smoker == 1

*Generate Indicator Variable for Former Smokers (S_i=1, Q_i=0)
gen former_ind = 0 if ever_smoke != .
replace former_ind = 1 if cur_smoker == 0

*Create Interaction Variables
gen edu = edu_quality

gen never_smoke = ever_smoke
recode never_smoke (1 = 2)
recode never_smoke (0 = 1)
recode never_smoke (2 = 0)

gen former_smoker = cur_smoker
recode former_smoker (1 = 2)
recode former_smoker (0 = 1)
recode former_smoker (2 = 0)
*Create categories for variables
  tab cig_weight, generate(cw_)
  tab edu_quality, generate(ed_)

*Create Interaction Variables
  gen f_parent = female * parent_smoke_bi
  gen f_sp1   = female * sp_pca_1
  gen f_sp2   = female * sp_pca_2
  gen f_sp3   = female * sp_pca_3
  gen f_cw1   = female * cw_1
  gen f_cw2   = female * cw_2
  gen f_cigh  = female * cig_harm_bi
  gen f_edu2  = female * ed_2
  gen f_edu3  = female * ed_3
  gen f_edu4  = female * ed_4

*Create y bar _m
  gen cig_cat_mean = .
  replace cig_cat_mean = 0   if num_cigs == 1
  replace cig_cat_mean = .5  if num_cigs == 2
  replace cig_cat_mean = 1   if num_cigs == 3
  replace cig_cat_mean = 3.5 if num_cigs == 4
  replace cig_cat_mean = 10  if num_cigs == 5

  tab cig_cat_mean, gen(ccm_)
  replace ccm_1 = 0
  replace ccm_2 = 0.5
  replace ccm_3 = 1
  replace ccm_4 = 3.5
  replace ccm_5 = 10

*Drop missing and unneeded variables for mle program
  drop if ever_smoke   == .
  drop if friends_smoke_bi == .
  drop if parent_smoke_bi == .
  drop if female       == .
  drop if sp_pca_1     == .
  drop if sp_pca_2     == .
  drop if sp_pca_3     == .
  drop if cig_weight   == .
  drop if cig_harm_bi  == .
  drop if other_cig_harm_bi== .
  drop if pt_pca_1     == .
  drop if edu_quality  == .
  drop if ate_v1       == .
  drop if fam_smoke_dis == .
drop if num_cigs == .
drop if current_ind == .
drop if former_ind == .
drop if never_ind == .

global vars i.friends_smoke_bi i.female i.parent_smoke_bi c.sp_pca_1 ///
  c.sp_pca_2 c.sp_pca_3 b3.cig_weight i.cig_harm_bi i.other_cig_harm_bi ///
c.pt_pca_1 i.edu_quality i.ate_v1 i.fam_smoke_dis

global never i.friends_smoke_bi i.female##i.parent_smoke_bi i.female##c.sp_pca_1 ///
  i.female##c.sp_pca_2 i.female##c.sp_pca_3 i.female##b3.cig_weight ///
  i.female##i.cig_harm_bi i.other_cig_harm_bi c.pt_pca_1 ///
i.female##edu_quality i.ate_v1 i.fam_smoke_dis

global quitter i.friends_smoke_bi i.female##i.parent_smoke_bi i.female##c.sp_pca_1 ///
  i.female##c.sp_pca_2 i.female##c.sp_pca_3 i.female##b3.cig_weight ///
  i.female##i.cig_harm_bi i.other_cig_harm_bi c.pt_pca_1 ///
i.female##edu_quality i(1)bn.female##i(3)bn.edu i.ate_v1 i.fam_smoke_dis

global level i.friends_smoke_bi female##parent_smoke_bi female##c.sp_pca_1 ///
female##c.sp_pca_2 female##c.sp_pca_3 b3.cig_weight ///
female##cig_harm_bi i.other_cig_harm_bi c.pt_pca_1 ///
i.edu_quality i(1)bn.female##(3)bn.edu i.ate_v1 i.fam_smoke_dis

****
{
capture program drop trivariate_model
program define trivariate_model
args lnL xa xb xc t2 t3 t4

tempvar p_nsmoke p_smoke pa p_fsmoke p_csmoke pb p_count v2 v3 v4 p2 p3 p4 p5

  quietly gen double `pa' = normprob(`xa')
  quietly gen double `p_nsmoke' = never_ind * ln(`pa')
  quietly gen double `p_smoke' = (1 - never_ind) * ln(1 - `pa')

  quietly gen double `pb' = normprob(`xb')
  quietly gen double `p_fsmoke' = former_ind * ln(`pb')
  quietly gen double `p_csmoke' = (1 - former_ind) * ln(1 - `pb')

  quietly gen double `v2' = `t2' - (`xc')
  quietly gen double `v3' = `t3' - (`xc')
  quietly gen double `v4' = `t4' - (`xc')
  quietly gen double `p2' = normprob(`v2')
  quietly gen double `p3' = normprob(`v3') - normprob(`v2')
  quietly gen double `p4' = normprob(`v4') - normprob(`v3')

311
quietly gen double `p5' = 1 - normprob(`v4')
quietly gen double `p_count' = \text{nc}_2 \ln(`p2') + \text{nc}_3 \ln(`p3') + \text{nc}_4 \ln(`p4') + \text{nc}_5 \ln(`p5')

quietly replace `lnL' = never_ind * (`p_nsmoke') + /// //never smokers
former_ind * (`p_smoke' + `p fsmoke') + /// //quitters
current_ind * (`p_smoke' + `p csmoke' + `p_count') // quantity
end
}

******
capture program trivariate_model_est drop
program trivariate_model_est
ml model lf trivariate_model ///
(xa:never_smoke = $never if never_ind == 1) ///
(xb:former_ind = $quitter if former_ind == 1) ///
(xc:num_cigs = $level, noconstant if current_ind == 1) /cut1 /cut2 /cut3 ///
[pweight = FinalWgt], vce(cluster sch_class) //technique(bhhh bfgs)

ml search
ml maximize, difficult iterate(5000)
end

*****
capture prog drop level_con_w_int_marginal
prog define level_con_w_int_marginal, rclass
tempvar xa pa qa xb pb xc v1 v2 v3 p1 p2 p3 p4 q1 q2 q3 q4 tog_sp1_a ///
   fem_sp1_a mal_sp1_a tog_sp1_b fem_sp1_b mal_sp1_b tog_sp1_c fem_sp1_c ///
   mal_sp1_c tog_sp2_a fem_sp2_a mal_sp2_a tog_sp2_b fem_sp2_b mal_sp2_b ///
tog_sp2_c fem_sp2_c mal_sp2_c tog_sp3_a fem_sp3_a mal_sp3_a tog_sp3_b ///
   fem_sp3_b mal_sp3_b tog_sp3_c fem_sp3_c mal_sp3_c me_sp1_1 ///
   me_sp1_1_f me_sp1_1_m me_sp2_1_f me_sp2_1_m ///
   me_sp3_1_f me_sp3_1_m me_sp1_2_f me_sp1_2_m me_sp1_2_f ///
   me_sp1_2_m me_sp2_2_f me_sp2_2_m me_sp2_2_m me_sp3_2 ///
   me_sp3_2_f me_sp3_2_m me_sp1_3_f me_sp1_3_m me_sp1_3_f me_sp1_3_m me_sp2_3 ///
   me_sp2_3_f me_sp2_3_m me_sp3_3_f me_sp3_3_m me_sp3_3_m ///
   me_sp1_4_f me_sp1_4_m me_sp2_4_m me_sp2_4_f ///
   me_sp2_4_m me_sp3_4_m me_sp3_4_f me_sp3_4_m
qui trivariate_model_est
matrix b_tri = e(b)
matrix list b_tri

312
scalar a0_constant_cons = b_tri[1,54]
scalar a1_female_cons = b_tri[1,4]
scalar a2_friends_cons = b_tri[1,2]
scalar a3_parents_cons = b_tri[1,6]
scalar a4_f_parent_cons = b_tri[1,10]
scalar a5_sp1_cons = b_tri[1,11]
scalar a6_f_sp1_cons = b_tri[1,13]
scalar a7_sp2_cons = b_tri[1,14]
scalar a8_f_sp2_cons = b_tri[1,16]
scalar a9_sp3_cons = b_tri[1,17]
scalar a10_f_sp3_cons = b_tri[1,19]
scalar a11_cw1_cons = b_tri[1,20]
scalar a12_f_cw1_cons = b_tri[1,26]
scalar a13_cw2_cons = b_tri[1,21]
scalar a14_f_cw2_cons = b_tri[1,27]
scalar a15_cigharm_cons = b_tri[1,30]
scalar a16_f_cigharm_cons = b_tri[1,34]
scalar a17_othercigharm_cons = b_tri[1,36]
scalar a18_pte1_cons = b_tri[1,37]
scalar a19_edu2_cons = b_tri[1,39]
scalar a20_f_edu2_cons = b_tri[1,47]
scalar a21_edu3_cons = b_tri[1,40]
scalar a22_f_edu3_cons = b_tri[1,48]
scalar a23_edu4_cons = b_tri[1,41]
scalar a24_f_edu4_cons = b_tri[1,49]
scalar a25_ate_cons = b_tri[1,51]
scalar a26_famdis_cons = b_tri[1,53]

scalar b0_constant_cons = b_tri[1,108]
scalar b1_female_cons = b_tri[1,58]
scalar b2_friends_cons = b_tri[1,56]
scalar b3_parents_cons = b_tri[1,60]
scalar b4_f_parent_cons = b_tri[1,64]
scalar b5_sp1_cons = b_tri[1,65]
scalar b6_f_sp1_cons = b_tri[1,67]
scalar b7_sp2_cons = b_tri[1,68]
scalar b8_f_sp2_cons = b_tri[1,70]
scalar b9_sp3_cons = b_tri[1,71]
scalar b10_f_sp3_cons = b_tri[1,73]
scalar b11_cw1_cons = b_tri[1,74]
scalar b12_f_cw1_cons = b_tri[1,80]
scalar b13_cw2_cons = b_tri[1,75]
scalar b14_f_cw2_cons = b_tri[1,81]
scalar b15_cigharm_cons = b_tri[1,84]
scalar b16_f_cigharm_cons = b_tri[1,88]
scalar b17_othercigharm_cons = b_tri[1,90]
scalar b18_pte1_cons = b_tri[1,91]
scalar b19_edu2_cons = b_tri[1,93]
scalar b20_f_edu2_cons = b_tri[1,101]
scalar b21_edu3_cons = b_tri[1,94]
scalar b22_f_edu3_cons = b_tri[1,102]
scalar b23_edu4_cons = b_tri[1,95]
scalar b24_f_edu4_cons = b_tri[1,103]
scalar b25_ate_cons = b_tri[1,105]
scalar b26_famdis_cons = b_tri[1,107]

scalar c1_female_cons = b_tri[1,111]
scalar c2_friends_cons = b_tri[1,110]
scalar c3_parents_cons = b_tri[1,113]
scalar c4_fparent_cons = b_tri[1,115]
scalar c5_sp1_cons = b_tri[1,116]
scalar c6_fsp1_cons = b_tri[1,117]
scalar c7_sp2_cons = b_tri[1,118]
scalar c8_fsp2_cons = b_tri[1,119]
scalar c9_sp3_cons = b_tri[1,120]
scalar c10_fsp3_cons = b_tri[1,121]
scalar c11_cw1_cons = b_tri[1,122]
scalar c13_cw2_cons = b_tri[1,123]
scalar c15_cigharm_cons = b_tri[1,126]
scalar c16_fcigh_cons = b_tri[1,128]
scalar c17_othercigharm_cons = b_tri[1,130]
scalar c18_pte1_cons = b_tri[1,131]
scalar c19_edu2_cons = b_tri[1,133]
scalar c21_edu3_cons = b_tri[1,134]
scalar c23_edu4_cons = b_tri[1,135]
scalar c24_fedu4_cons = b_tri[1,136]
scalar c25_ate_cons = b_tri[1,138]
scalar c26_famdis_cons = b_tri[1,140]
scalar cut1_cons = b_tri[1,141]
scalar cut2_cons = b_tri[1,142]
scalar cut3_cons = b_tri[1,143]

***** edu_2

tempvar edu2_xa_1 edu2_xb_1 edu2_xc_1 edu2_xa_0 edu2_xb_0 although
   edu2_xc_0 edu2_pa_1 edu2_pb_1 edu2_v2_1 edu2_v3_1 although
   edu2_v4_1 edu2_p2_1 edu2_p3_1 edu2_p4_1 edu2_p5_1 although
   edu2_pa_0 edu2_pb_0 edu2_v2_0 edu2_v3_0 edu2_v4_0 although
   edu2_p2_0 edu2_p3_0 edu2_p4_0 edu2_p5_0 edu2_when_one_5 although
   edu2_when_one_4 edu2_when_one_3 edu2_when_one_2 although
   edu2_when_zero_5 edu2_when_zero_4 edu2_when_zero_3 although
   edu2_when_zero_2 edu2_when_one_5_f edu2_when_one_4_f although
edu2_when_one_3_f edu2_when_one_2_f edu2_when_one_5_m ///
edu2_when_one_4_m edu2_when_one_3_m edu2_when_one_2_m ///
edu2_when_zero_5_f edu2_when_zero_4_f edu2_when_zero_3_f ///
edu2_when_zero_2_f edu2_when_zero_5_m edu2_when_zero_4_m ///
edu2_when_zero_3_m edu2_when_zero_2_m

**** When Equal to One
*For Ever Smokers
quietly gen double `edu2_xa_1' = a0_constant_cons ///
 + a1_female_cons * female ///
 + a2_friends_cons * friends_smoke_bi ///
 + a3_parents_cons * parent_smoke_bi ///
 + a4_f_parent_cons * f_parent ///
 + a5_sp1_cons * sp_pca_1 ///
 + a6_f_sp1_cons * f_sp1 ///
 + a7_sp2_cons * sp_pca_2 ///
 + a8_f_sp2_cons * f_sp2 ///
 + a9_sp3_cons * sp_pca_3 ///
 + a10_f_sp3_cons * f_sp3 ///
 + a11_cw1_cons * cw_1 ///
 + a12_f_cw1_cons * f_cw1 ///
 + a13_cw2_cons * cw_2 ///
 + a14_f_cw2_cons * f_cw2 ///
 + a15_cigharm_cons * cig_harm_bi ///
 + a16_f_cigharm_cons * f_cigh ///
 + a17_othercigharm_cons * other_cig_harm_bi ///
 + a18_pte1_cons * pt_pca_1 ///
 + a19_edu2_cons * 1 ///
 + a20_f_edu2_cons * female ///
 + a25_ate_cons * ate_v1 ///
 + a26_famdis_cons * fam_smoke_dis
quietly gen double `edu2_pa_1' = 1 - normprob(`edu2_xa_1')

*For Current Smoker Portion
quietly gen double `edu2_xb_1' = b0_constant_cons ///
 + b1_female_cons * female ///
 + b2_friends_cons * friends_smoke_bi ///
 + b3_parents_cons * parent_smoke_bi ///
 + b4_f_parent_cons * f_parent ///
 + b5_sp1_cons * sp_pca_1 ///
 + b6_f_sp1_cons * f_sp1 ///
 + b7_sp2_cons * sp_pca_2 ///
 + b8_f_sp2_cons * f_sp2 ///
 + b9_sp3_cons * sp_pca_3 ///
 + b10_f_sp3_cons * f_sp3 ///
 + b11_cw1_cons * cw_1 ///
+ b12_f_cw1_cons * f_cw1
+ b13_cw2_cons * cw_2
+ b14_f_cw2_cons * f_cw2
+ b15_cigharm_cons * cig_harm_bi
+ b16_f_cigharm_cons * f_cigh
+ b17_othercigharm_cons * other_cig_harm_bi
+ b18_ptel1_cons * pt_pca_1
+ b19_edu2_cons * 1
+ b20_f_edu2_cons * female
+ b25_ate_cons * ate_v1
+ b26_famdis_cons * fam_smoke_dis

quietly gen double `edu2_pb_1' = 1 - normprob(`edu2_xb_1')

*For Number of Cigarettes Smoked Portion
quietly gen double `edu2_xc_1' = c1_female_cons * female
+ c2_friends_cons * friends_smoke_bi
+ c3_parents_cons * parent_smoke_bi
+ c4_fparent_cons * f_parent
+ c5_sp1_cons * sp_pca_1
+ c6_sp1_cons * f_sp1
+ c7_sp2_cons * sp_pca_2
+ c8_sp2_cons * f_sp2
+ c9_sp3_cons * sp_pca_3
+ c10_sp3_cons * f_sp3
+ c11_cw1_cons * cw_1
+ c13_cw2_cons * cw_2
+ c15_cigharm_cons * cig_harm_bi
+ c16_cigharm_cons * f_cigh
+ c17_othercigharm_cons * other_cig_harm_bi
+ c18_ptel1_cons * pt_pca_1
+ c19_edu2_cons * 1
+ c25_ate_cons * ate_v1
+ c26_famdis_cons * fam_smoke_dis

quietly gen double `edu2_v2_1' = cut1_cons - (`edu2_xc_1')
quietly gen double `edu2_v3_1' = cut2_cons - (`edu2_xc_1')
quietly gen double `edu2_v4_1' = cut3_cons - (`edu2_xc_1')
quietly gen double `edu2_v5_1' = normprob(`edu2_v2_1')
quietly gen double `edu2_v6_1' = normprob(`edu2_v3_1') - normprob(`edu2_v2_1')
quietly gen double `edu2_v7_1' = normprob(`edu2_v4_1') - normprob(`edu2_v3_1')
quietly gen double `edu2_v8_1' = 1 - normprob(`edu2_v4_1')

*Top Category m=5
quietly gen double `edu2_when_one_5' = `edu2_pa_1' * `edu2_pb_1' * `edu2_p5_1'
quietly gen double `edu2_when_one_5_f' = `edu2_pa_1' * `edu2_pb_1' * `edu2_p5_1' if female == 1
quietly gen double `edu2_when_one_5_m' = `edu2_pa_1' * `edu2_pb_1' * `edu2_p5_1' if female == 0

*Middle Category m=4
quietly gen double `edu2_when_one_4' = `edu2_pa_1' * `edu2_pb_1' * `edu2_p4_1'
quietly gen double `edu2_when_one_4_f' = `edu2_pa_1' * `edu2_pb_1' * `edu2_p4_1' if female == 1
quietly gen double `edu2_when_one_4_m' = `edu2_pa_1' * `edu2_pb_1' * `edu2_p4_1' if female == 0

*Middle Category m=3
quietly gen double `edu2_when_one_3' = `edu2_pa_1' * `edu2_pb_1' * `edu2_p3_1'
quietly gen double `edu2_when_one_3_f' = `edu2_pa_1' * `edu2_pb_1' * `edu2_p3_1' if female == 1
quietly gen double `edu2_when_one_3_m' = `edu2_pa_1' * `edu2_pb_1' * `edu2_p3_1' if female == 0

*Bottom Category m=2
quietly gen double `edu2_when_one_2' = `edu2_pa_1' * `edu2_pb_1' * `edu2_p2_1'
quietly gen double `edu2_when_one_2_f' = `edu2_pa_1' * `edu2_pb_1' * `edu2_p2_1' if female == 1
quietly gen double `edu2_when_one_2_m' = `edu2_pa_1' * `edu2_pb_1' * `edu2_p2_1' if female == 0

local when_one `edu2_when_one_5' `edu2_when_one_4' `edu2_when_one_3' `edu2_when_one_2'
local when_one_f `edu2_when_one_5_f' `edu2_when_one_4_f' `edu2_when_one_3_f' `edu2_when_one_2_f'
local when_one_m `edu2_when_one_5_m' `edu2_when_one_4_m' `edu2_when_one_3_m' `edu2_when_one_2_m'

***** When Equal to Zero
*For Ever Smoker Portion
quietly gen double `edu2_xa_0' = a0_constant_cons ///
    + a1_female_cons * female ///
    + a2_friends_cons * friends_smoke_bi ///
    + a3_parents_cons * parent_smoke_bi ///
    + a4_f_parent_cons * f_parent ///
    + a5_sp1_cons * sp_pca_1 ///
    + a6_f_sp1_cons * f_sp1 ///
    + a7_sp2_cons * sp_pca_2 ///
    + a8_f_sp2_cons * f_sp2 ///
    + a9_sp3_cons * sp_pca_3 ///
    + a10_f_sp3_cons * f_sp3 ///
    + a11_cw1_cons * cw_1 ///
    + a12_f_cw1_cons * f_cw1 ///
quietly gen double `edu2_pa_0' = 1 - normprob(`edu2_xa_0')

*For Current Smoker Portion
quietly gen double `edu2_xb_0' = b0_constant_cons
+ b1_female_cons * female
+ b2_friends_cons * friends_smoke_bi
+ b3_parents_cons * parent_smoke_bi
+ b4_f_parent_cons * f_parent
+ b5_sp1_cons * sp_pca_1
+ b6_f_sp1_cons * f_sp1
+ b7_sp2_cons * sp_pca_2
+ b8_f_sp2_cons * f_sp2
+ b9_sp3_cons * sp_pca_3
+ b10_f_sp3_cons * f_sp3
+ b11_cw1_cons * cw_1
+ b12_f_cw1_cons * f_cw1
+ b13_cw2_cons * cw_2
+ b14_f_cw2_cons * f_cw2
+ b15_cig_harm_cons * cig_harm_bi
+ b16_f_cig_harm_cons * f_cigh
+ b17_other_cig_harm_cons * other_cig_harm_bi
+ b18_ptel1_cons * pt_pca_1
+ b19_edu2_cons * 0
+ b20_f_edu2_cons * 0
+ b25_ate_cons * ate_v1
+ b26_famdis_cons * fam_smoke_dis
quietly gen double `edu2_pb_0' = 1 - normprob(`edu2_xb_0')

*For Number of Cigarettes Smoked Portion
quietly gen double `edu2_xc_0' = c1_female_cons * female
+ c2_friends_cons * friends_smoke_bi
+ c3_parents_cons * parent_smoke_bi
+ c4_fparent_cons * f_parent
+ c5_sp1_cons * sp_pca_1
+ c6_fsp1_cons * f_sp1
+ c7_sp2_cons * sp_pca_2
+ c8_fsp2_cons * f_sp2
+ c9_sp3_cons * sp_pca_3
+ c10_fsp3_cons * f_sp3
+ c11_cw1_cons * cw_1
+ c12_f_cw1_cons * f_cw1
+ c13_cw2_cons * cw_2
+ c14_f_cw2_cons * f_cw2
+ c15_cig_harm_cons * cig_harm_bi
+ c16_f_cig_harm_cons * f_cigh
+ c17_other_cig_harm_cons * other_cig_harm_bi
+ c18_ptel1_cons * pt_pca_1
+ c19_edu2_cons * 0
+ c20_f_edu2_cons * 0
+ c25_ate_cons * ate_v1
+ c26_famdis_cons * fam_smoke_dis
quietly gen double `edu2_pc_0' = 1 - normprob(`edu2_xc_0')
quietly gen double `edu2_v2_0' = cut1_cons - (`edu2_xc_0')
quietly gen double `edu2_v3_0' = cut2_cons - (`edu2_xc_0')
quietly gen double `edu2_v4_0' = cut3_cons - (`edu2_xc_0')
quietly gen double `edu2_p2_0' = normprob(`edu2_v2_0')
quietly gen double `edu2_p3_0' = normprob(`edu2_v3_0') - normprob(`edu2_v2_0')
quietly gen double `edu2_p4_0' = normprob(`edu2_v4_0') - normprob(`edu2_v3_0')
quietly gen double `edu2_p5_0' = 1 - normprob(`edu2_v4_0')

*Top Category m=5
quietly gen double `edu2_when_zero_5' = `edu2_pa_0' * `edu2_pb_0' * `edu2_p5_0'
quietly gen double `edu2_when_zero_5_f' = `edu2_pa_0' * `edu2_pb_0' * `edu2_p5_0' if female == 1
quietly gen double `edu2_when_zero_5_m' = `edu2_pa_0' * `edu2_pb_0' * `edu2_p5_0' if female == 0

*Middle Category m=4
quietly gen double `edu2_when_zero_4' = `edu2_pa_0' * `edu2_pb_0' * `edu2_p4_0'
quietly gen double `edu2_when_zero_4_f' = `edu2_pa_0' * `edu2_pb_0' * `edu2_p4_0' if female == 1
quietly gen double `edu2_when_zero_4_m' = `edu2_pa_0' * `edu2_pb_0' * `edu2_p4_0' if female == 0

*Middle Category m=3
quietly gen double `edu2_when_zero_3' = `edu2_pa_0' * `edu2_pb_0' * `edu2_p3_0'
quietly gen double `edu2_when_zero_3_f' = `edu2_pa_0' * `edu2_pb_0' * `edu2_p3_0' if female == 1
quietly gen double `edu2_when_zero_3_m' = `edu2_pa_0' * `edu2_pb_0' * `edu2_p3_0' if female == 0

*Bottom Category m=2
quietly gen double `edu2_when_zero_2' = `edu2_pa_0' * `edu2_pb_0' * `edu2_p2_0'
quietly gen double `edu2_when_zero_2_f' = `edu2_pa_0' * `edu2_pb_0' * `edu2_p2_0' if female == 1
quietly gen double `edu2_when_zero_2_m' = `edu2_pa_0' * `edu2_pb_0' * `edu2_p2_0'
if female == 0

local when_zero `edu2_when_zero_5' `edu2_when_zero_4' `edu2_when_zero_3'
`edu2_when_zero_2'
local when_zero_f `edu2_when_zero_5_f' `edu2_when_zero_4_f' `edu2_when_zero_3_f'
`edu2_when_zero_2_f'
local when_zero_m `edu2_when_zero_5_m' `edu2_when_zero_4_m'
`edu2_when_zero_3_m' `edu2_when_zero_2_m'

tempvar me_edu2_5 me_edu2_4 me_edu2_3 me_edu2_2 ///
me_edu2_5_f me_edu2_4_f me_edu2_3_f me_edu2_2_f ///
me_edu2_5_m me_edu2_4_m me_edu2_3_m me_edu2_2_m

local Lmes1 `me_edu2_5' `me_edu2_4' `me_edu2_3' `me_edu2_2'
local Lmes2 `me_edu2_5_f' `me_edu2_4_f' `me_edu2_3_f' `me_edu2_2_f'
local Lmes3 `me_edu2_5_m' `me_edu2_4_m' `me_edu2_3_m' `me_edu2_2_m'

forvalues i = 1/3{
 foreach x of local Lmes`i'{
 quietly gen double `x' =.
 }
 local i = `i' +1
}
*Together
quietly replace `me_edu2_5' = ccm_5 * (`edu2_when_one_5' - `edu2_when_zero_5')
quietly sum `me_edu2_5'
scalar meedu25_mean = r(mean)
return scalar ame_edu2_4 = meedu25_mean

quietly replace `me_edu2_4' = ccm_4 * (`edu2_when_one_4' - `edu2_when_zero_4')
quietly sum `me_edu2_4'
scalar meedu24_mean = r(mean)
return scalar ame_edu2_3 = meedu24_mean

quietly replace `me_edu2_3' = ccm_3 * (`edu2_when_one_3' - `edu2_when_zero_3')
quietly sum `me_edu2_3'
scalar meedu23_mean = r(mean)
return scalar ame_edu2_2 = meedu23_mean

quietly replace `me_edu2_2' = ccm_2 * (`edu2_when_one_2' - `edu2_when_zero_2')
quietly sum `me_edu2_2'
scalar meedu22_mean = r(mean)
return scalar ame_edu2_1 = meedu22_mean
*Girls
quietly replace `me_edu2_5_f' = ccm_5 * (`edu2_when_one_5_f' - `edu2_when_zero_5_f')
quietly sum `me_edu2_5_f'
scalar meedu25f_mean = r(mean)
return scalar ame_edu2_4_f = meedu25f_mean

quietly replace `me_edu2_4_f' = ccm_4 * (`edu2_when_one_4_f' - `edu2_when_zero_4_f')
quietly sum `me_edu2_4_f'
scalar meedu24f_mean = r(mean)
return scalar ame_edu2_3_f = meedu24f_mean

quietly replace `me_edu2_3_f' = ccm_3 * (`edu2_when_one_3_f' - `edu2_when_zero_3_f')
quietly sum `me_edu2_3_f'
scalar meedu23f_mean = r(mean)
return scalar ame_edu2_2_f = meedu23f_mean

quietly replace `me_edu2_2_f' = ccm_2 * (`edu2_when_one_2_f' - `edu2_when_zero_2_f')
quietly sum `me_edu2_2_f'
scalar meedu22f_mean = r(mean)
return scalar ame_edu2_1_f = meedu22f_mean

*Boys
quietly replace `me_edu2_5_m' = ccm_5 * (`edu2_when_one_5_m' - `edu2_when_zero_5_m')
quietly sum `me_edu2_5_m'
scalar meedu25m_mean = r(mean)
return scalar ame_edu2_4_m = meedu25m_mean

quietly replace `me_edu2_4_m' = ccm_4 * (`edu2_when_one_4_m' - `edu2_when_zero_4_m')
quietly sum `me_edu2_4_m'
scalar meedu24m_mean = r(mean)
return scalar ame_edu2_3_m = meedu24m_mean

quietly replace `me_edu2_3_m' = ccm_3 * (`edu2_when_one_3_m' - `edu2_when_zero_3_m')
quietly sum `me_edu2_3_m'
scalar meedu23m_mean = r(mean)
return scalar ame_edu2_2_m = meedu23m_mean

quietly replace `me_edu2_2_m' = ccm_2 * (`edu2_when_one_2_m' - `edu2_when_zero_2_m')
quietly sum `me_edu2_2_m'
scalar meedu22m_mean = r(mean)
return scalar ame_edu2_1_m = meedu22m_mean

***** edu_3

tempvar edu3_xa_1 edu3_xb_1 edu3_xc_1 edu3_xa_0 edu3_xb_0 ///
    edu3_xc_0 edu3_pa_1 edu3_pb_1 edu3_v2_1 edu3_v3_1 ///
    edu3_v4_1 edu3_p2_1 edu3_p3_1 edu3_p4_1 edu3_p5_1 ///
    edu3_pa_0 edu3_pb_0 edu3_v2_0 edu3_v3_0 edu3_v4_0 ///
    edu3_p2_0 edu3_p3_0 edu3_p4_0 edu3_p5_0 edu3_when_one_5 ///
    edu3_when_one_4 edu3_when_one_3 edu3_when_one_2 ///
    edu3_when_zero_5 edu3_when_zero_4 edu3_when_zero_3 ///
    edu3_when_zero_2 edu3_when_one_5 f edu3_when_one_4 f ///
    edu3_when_one_3 f edu3_when_one_2 f edu3_when_one_5 m ///
    edu3_when_one_4 m edu3_when_one_3 m edu3_when_one_2 m ///
    edu3_when_zero_5 f edu3_when_zero_4 f edu3_when_zero_3 f ///
    edu3_when_zero_2 f edu3_when_zero_5 m edu3_when_zero_4 m ///
    edu3_when_zero_3 m edu3_when_zero_2 m

***** When Equal to One
*For Ever Smokers
quietly gen double `edu3_xa_1' = a0_constant_cons ///
    + a1_female_cons * female ///
    + a2_friends_cons * friends_smoke_bi ///
    + a3_parents_cons * parent_smoke_bi ///
    + a4_f_parent_cons * f_parent ///
    + a5_sp1_cons * sp_pca_1 ///
    + a6_f_sp1_cons * f_sp1 ///
    + a7_sp2_cons * sp_pca_2 ///
    + a8_f_sp2_cons * f_sp2 ///
    + a9_sp3_cons * sp_pca_3 ///
    + a10_f_sp3_cons * f_sp3 ///
    + a11_cw1_cons * cw_1 ///
    + a12_f_cw1_cons * f_cw1 ///
    + a13_cw2_cons * cw_2 ///
    + a14_f_cw2_cons * f_cw2 ///
    + a15_cigharm_cons * cig_harm_bi ///
    + a16_f_cigharm_cons * f_cigh ///
    + a17_othercigharm_cons * other_cig_harm_bi ///
    + a18_pte1_cons * pt_pca_1 ///
    + a21_edu3_cons * 1 ///
    + a22_f_edu3_cons * female ///
    + a25_ate_cons * ate_v1 ///
    + a26_famdis_cons * fam_smoke_dis

322
quietly gen double `edu3_pa_1' = 1 - normprob(`edu3_xa_1')

*For Current Smoker Portion
quietly gen double `edu3_xb_1' = b0_constant_cons ///
+ b1_female_cons * female ///
+ b2_friends_cons * friends_smoke_bi ///
+ b3_parents_cons * parent_smoke_bi ///
+ b4_f_parent_cons * f_parent ///
+ b5_sp1_cons * sp_pca_1 ///
+ b6_f_sp1_cons * f_sp1 ///
+ b7_sp2_cons * sp_pca_2 ///
+ b8_f_sp2_cons * f_sp2 ///
+ b9_sp3_cons * sp_pca_3 ///
+ b10_f_sp3_cons * f_sp3 ///
+ b11 cw1_cons * cw_1 ///
+ b12_f cw1_cons * f_cw1 ///
+ b13 cw2_cons * cw_2 ///
+ b14_f cw2_cons * f_cw2 ///
+ b15 cigharm_cons * cig_harm_bi ///
+ b16_f cigharm_cons * f_cigh ///
+ b17 othercigharm_cons * other_cig_harm_bi ///
+ b18_pte1_cons * pt_pca_1 ///
+ b21 edu3_cons * 1 ///
+ b22_f edu3_cons * female ///
+ b25_ate_cons * ate_v1 ///
+ b26_famdis_cons * fam_smoke_dis

quietly gen double `edu3_pb_1' = 1 - normprob(`edu3_xb_1')

*For Number of Cigarettes Smoked Portion
quietly gen double `edu3_xc_1' = c1_female_cons * female ///
+ c2_friends_cons * friends_smoke_bi ///
+ c3_parents_cons * parent_smoke_bi ///
+ c4_fparent_cons * f_parent ///
+ c5_sp1_cons * sp_pca_1 ///
+ c6_fsp1_cons * f_sp1 ///
+ c7_sp2_cons * sp_pca_2 ///
+ c8_fsp2_cons * f_sp2 ///
+ c9_sp3_cons * sp_pca_3 ///
+ c10_fsp3_cons * f_sp3 ///
+ c11 cw1_cons * cw_1 ///
+ c13 cw2_cons * cw_2 ///
+ c15 cigharm_cons * cig_harm_bi ///
+ c16 fcigh_cons * f_cigh ///
+ c17 othercigharm_cons * other_cig_harm_bi ///
+ c18_pte1_cons * pt_pca_1 ///
+ c21 edu3_cons * 1 ///
+ c25_ate_cons * ate_v1 //
+ c26_famdis_cons * fam_smoke_dis

quietly gen double `edu3_v2_1' = cut1_cons - (`edu3_xc_1')
quietly gen double `edu3_v3_1' = cut2_cons - (`edu3_xc_1')
quietly gen double `edu3_v4_1' = cut3_cons - (`edu3_xc_1')
quietly gen double `edu3_p2_1' = normprob(`edu3_v2_1')
quietly gen double `edu3_p3_1' = normprob(`edu3_v3_1') - normprob(`edu3_v2_1')
quietly gen double `edu3_p4_1' = normprob(`edu3_v4_1') - normprob(`edu3_v3_1')
quietly gen double `edu3_p5_1' = 1 - normprob(`edu3_v4_1')

*Top Category m=5
quietly gen double `edu3_when_one_5' = `edu3_pa_1' * `edu3_pb_1' * `edu3_p5_1'
quietly gen double `edu3_when_one_5_f' = `edu3_pa_1' * `edu3_pb_1' * `edu3_p5_1' if female == 1
quietly gen double `edu3_when_one_5_m' = `edu3_pa_1' * `edu3_pb_1' * `edu3_p5_1' if female == 0

*Middle Category m=4
quietly gen double `edu3_when_one_4' = `edu3_pa_1' * `edu3_pb_1' * `edu3_p4_1'
quietly gen double `edu3_when_one_4_f' = `edu3_pa_1' * `edu3_pb_1' * `edu3_p4_1' if female == 1
quietly gen double `edu3_when_one_4_m' = `edu3_pa_1' * `edu3_pb_1' * `edu3_p4_1' if female == 0

*Middle Category m=3
quietly gen double `edu3_when_one_3' = `edu3_pa_1' * `edu3_pb_1' * `edu3_p3_1'
quietly gen double `edu3_when_one_3_f' = `edu3_pa_1' * `edu3_pb_1' * `edu3_p3_1' if female == 1
quietly gen double `edu3_when_one_3_m' = `edu3_pa_1' * `edu3_pb_1' * `edu3_p3_1' if female == 0

*Bottom Category m=2
quietly gen double `edu3_when_one_2' = `edu3_pa_1' * `edu3_pb_1' * `edu3_p2_1'
quietly gen double `edu3_when_one_2_f' = `edu3_pa_1' * `edu3_pb_1' * `edu3_p2_1' if female == 1
quietly gen double `edu3_when_one_2_m' = `edu3_pa_1' * `edu3_pb_1' * `edu3_p2_1' if female == 0

local when_one `edu3_when_one_5' `edu3_when_one_4' `edu3_when_one_3' `edu3_when_one_2'
local when_one_f `edu3_when_one_5_f' `edu3_when_one_4_f' `edu3_when_one_3_f' `edu3_when_one_2_f'
local when_one_m `edu3_when_one_5_m' `edu3_when_one_4_m' `edu3_when_one_3_m' `edu3_when_one_2_m'
When Equal to Zero

*For Ever Smoker Portion

quietly gen double `edu3_xa_0' = a0_constant_cons ///
+ a1_female_cons * female ///
+ a2_friends_cons * friends_smoke_bi ///
+ a3_parents_cons * parent_smoke_bi ///
+ a4_f_parent_cons * f_parent ///
+ a5_sp1_cons * sp_pca_1 ///
+ a6_f_sp1_cons * f_sp1 ///
+ a7_sp2_cons * sp_pca_2 ///
+ a8_f_sp2_cons * f_sp2 ///
+ a9_sp3_cons * sp_pca_3 ///
+ a10_f_sp3_cons * f_sp3 ///
+ a11_cw1_cons * cw_1 ///
+ a12_f_cw1_cons * f_cw1 ///
+ a13_cw2_cons * cw_2 ///
+ a14_f_cw2_cons * f_cw2 ///
+ a15_cigharm_cons * cig_harm_bi ///
+ a16_f_cigharm_cons * f_cigh ///
+ a17_othercigharm_cons * other_cig_harm_bi ///
+ a18_pte1_cons * pt_pca_1 ///
+ a21_edu3_cons * 0 ///
+ a22_f_edu3_cons * 0 ///
+ a25_ate_cons * ate_v1 ///
+ a26_famdis_cons * fam_smoke_dis

quietly gen double `edu3_pa_0' = 1 - normprob(`edu3_xa_0')

*For Current Smoker Portion

quietly gen double `edu3_xb_0' = b0_constant_cons ///
+ b1_female_cons * female ///
+ b2_friends_cons * friends_smoke_bi ///
+ b3_parents_cons * parent_smoke_bi ///
+ b4_f_parent_cons * f_parent ///
+ b5_sp1_cons * sp_pca_1 ///
+ b6_f_sp1_cons * f_sp1 ///
+ b7_sp2_cons * sp_pca_2 ///
+ b8_f_sp2_cons * f_sp2 ///
+ b9_sp3_cons * sp_pca_3 ///
+ b10_f_sp3_cons * f_sp3 ///
+ b11_cw1_cons * cw_1 ///
+ b12_f_cw1_cons * f_cw1 ///
+ b13_cw2_cons * cw_2 ///
+ b14_f_cw2_cons * f_cw2 ///
+ b15_cigharm_cons * cig_harm_bi ///
+ b16_f_cigharm_cons * f_cigh ///
+ b17_othercigharm_cons * other_cig_harm_bi ///
quietly gen double `edu3_pb_0' = 1 - normprob(`edu3_xb_0')

*For Number of Cigarettes Smoked Portion
quietly gen double `edu3_xc_0' = c1_female_cons * female ///
 + c2_friends_cons * friends_smoke_bi ///
 + c3_parents_cons * parent_smoke_bi ///
 + c4_fparent_cons * f_parent ///
 + c5_sp1_cons * sp_pca_1 ///
 + c6fsp1_cons * f_sp1 ///
 + c7_sp2_cons * sp_pca_2 ///
 + c8fsp2_cons * f_sp2 ///
 + c9_sp3_cons * sp_pca_3 ///
 + c10fsp3_cons * f_sp3 ///
 + c11cw1_cons * cw_1 ///
 + c13cw2_cons * cw_2 ///
 + c15_cigharm_cons * cig_harm_bi ///
 + c16_fcigh_cons * f_cigh ///
 + c17_othercigharm_cons * other_cig_harm_bi ///
 + c18_pte1_cons * pt_pca_1 ///
 + c21_edu3_cons * 0 ///
 + c25_ate_cons * ate_v1 ///
 + c26_famdis_cons * fam_smoke_dis

quietly gen double `edu3_v2_0' = cut1_cons - (`edu3_xc_0')
quietly gen double `edu3_v3_0' = cut2_cons - (`edu3_xc_0')
quietly gen double `edu3_v4_0' = cut3_cons - (`edu3_xc_0')
quietly gen double `edu3_v5_0' = cut5_cons - (`edu3_xc_0')
quietly gen double `edu3_p2_0' = normprob(`edu3_v2_0')
quietly gen double `edu3_p3_0' = normprob(`edu3_v3_0') - normprob(`edu3_v2_0')
quietly gen double `edu3_p4_0' = normprob(`edu3_v4_0') - normprob(`edu3_v3_0')
quietly gen double `edu3_p5_0' = 1 - normprob(`edu3_v4_0')

*Top Category m=5
quietly gen double `edu3_when_zero_5' = `edu3_pa_0' * `edu3_pb_0' * `edu3_p5_0'
quietly gen double `edu3_when_zero_5' = `edu3_pa_0' * `edu3_pb_0' * `edu3_p5_0' if female == 1
quietly gen double `edu3_when_zero_5_m' = `edu3_pa_0' * `edu3_pb_0' * `edu3_p5_0' if female == 0

*Middle Category m=4
quietly gen double `edu3_when_zero_4' = `edu3_pa_0' * `edu3_pb_0' * `edu3_p4_0'
quietly gen double `edu3_when_zero_4' = `edu3_pa_0' * `edu3_pb_0' * `edu3_p4_0' if female == 1
quietly gen double `edu3_when_zero_4_m' = `edu3_pa_0' * `edu3_pb_0' * `edu3_p4_0'
if female == 0

*Middle Category m=3
quietly gen double `edu3_when_zero_3' = `edu3_pa_0' * `edu3_pb_0' * `edu3_p3_0'
quietly gen double `edu3_when_zero_3_f' = `edu3_pa_0' * `edu3_pb_0' * `edu3_p3_0' if female == 1
quietly gen double `edu3_when_zero_3_m' = `edu3_pa_0' * `edu3_pb_0' * `edu3_p3_0'
if female == 0

*Bottom Category m=2
quietly gen double `edu3_when_zero_2' = `edu3_pa_0' * `edu3_pb_0' * `edu3_p2_0'
quietly gen double `edu3_when_zero_2_f' = `edu3_pa_0' * `edu3_pb_0' * `edu3_p2_0' if female == 1
quietly gen double `edu3_when_zero_2_m' = `edu3_pa_0' * `edu3_pb_0' * `edu3_p2_0'
if female == 0

local when_zero `edu3_when_zero_5' `edu3_when_zero_4' `edu3_when_zero_3' `edu3_when_zero_2'
local when_zero_f `edu3_when_zero_5_f' `edu3_when_zero_4_f' `edu3_when_zero_3_f' `edu3_when_zero_2_f'
local when_zero_m `edu3_when_zero_5_m' `edu3_when_zero_4_m' `edu3_when_zero_3_m' `edu3_when_zero_2_m'

tempvar me_edu3_5 me_edu3_4 me_edu3_3 me_edu3_2 ///
    me_edu3_5_f me_edu3_4_f me_edu3_3_f me_edu3_2_f ///
    me_edu3_5_m me_edu3_4_m me_edu3_3_m me_edu3_2_m

local Mmes1 `me_edu3_5' `me_edu3_4' `me_edu3_3' `me_edu3_2'
local Mmes2 `me_edu3_5_f' `me_edu3_4_f' `me_edu3_3_f' `me_edu3_2_f'
local Mmes3 `me_edu3_5_m' `me_edu3_4_m' `me_edu3_3_m' `me_edu3_2_m'

forvalues i = 1/3{
    foreach x of local Mmes`i'{
        quietly gen double `x' = .
    }
    local i = `i' +1
}

*Together
quietly replace `me_edu3_5' = ccm_5 * (`edu3_when_one_5' - `edu3_when_zero_5')
quietly sum `me_edu3_5'
scalar meedu35_mean = r(mean)
return scalar ame_edu3_4 = meedu35_mean

quietly replace `me_edu3_4' = ccm_4 * (`edu3_when_one_4' - `edu3_when_zero_4')
quietly sum `me_edu3_4'
scalar meedu34_mean = r(mean)
return scalar ame_edu3_3 = meedu34_mean

quietly replace `me_edu3_3' = ccm_3 * (`edu3_when_one_3' - `edu3_when_zero_3')
quietly sum `me_edu3_3'
scalar meedu33_mean = r(mean)
return scalar ame_edu3_2 = meedu33_mean

quietly replace `me_edu3_2' = ccm_2 * (`edu3_when_one_2' - `edu3_when_zero_2')
quietly sum `me_edu3_2'
scalar meedu32_mean = r(mean)
return scalar ame_edu3_1 = meedu32_mean

*Girls
quietly replace `me_edu3_5_f' = ccm_5 * (`edu3_when_one_5_f' - `edu3_when_zero_5_f')
quietly sum `me_edu3_5_f'
scalar meedu35f_mean = r(mean)
return scalar ame_edu3_4_f = meedu35f_mean

quietly replace `me_edu3_4_f' = ccm_4 * (`edu3_when_one_4_f' - `edu3_when_zero_4_f')
quietly sum `me_edu3_4_f'
scalar meedu34f_mean = r(mean)
return scalar ame_edu3_3_f = meedu34f_mean

quietly replace `me_edu3_3_f' = ccm_3 * (`edu3_when_one_3_f' - `edu3_when_zero_3_f')
quietly sum `me_edu3_3_f'
scalar meedu33f_mean = r(mean)
return scalar ame_edu3_2_f = meedu33f_mean

quietly replace `me_edu3_2_f' = ccm_2 * (`edu3_when_one_2_f' - `edu3_when_zero_2_f')
quietly sum `me_edu3_2_f'
scalar meedu32f_mean = r(mean)
return scalar ame_edu3_1_f = meedu32f_mean

*Boys
quietly replace `me_edu3_5_m' = ccm_5 * (`edu3_when_one_5_m' - `edu3_when_zero_5_m')
quietly sum `me_edu3_5_m'
scalar meedu35m_mean = r(mean)
return scalar ame_edu3_4_m = meedu35m_mean
quietly replace `me_edu3_4_m' = ccm_4 * (`edu3_when_one_4_m' - `edu3_when_zero_4_m')
quietly sum `me_edu3_4_m'
scalar meedu34m_mean = r(mean)
return scalar ame_edu3_3_m = meedu34m_mean

quietly replace `me_edu3_3_m' = ccm_3 * (`edu3_when_one_3_m' - `edu3_when_zero_3_m')
quietly sum `me_edu3_3_m'
scalar meedu33m_mean = r(mean)
return scalar ame_edu3_2_m = meedu33m_mean

quietly replace `me_edu3_2_m' = ccm_2 * (`edu3_when_one_2_m' - `edu3_when_zero_2_m')
quietly sum `me_edu3_2_m'
scalar meedu32m_mean = r(mean)
return scalar ame_edu3_1_m = meedu32m_mean

***** edu_4

tempvar edu4_xa_1 edu4_xb_1 edu4_xc_1 edu4_xa_0 edu4_xb_0 ///
    edu4_xc_0 edu4_pa_1 edu4_pb_1 edu4_v2_1 edu4_v3_1 ///
    edu4_v4_1 edu4_p2_1 edu4_p3_1 edu4_p4_1 edu4_p5_1 ///
    edu4_pa_0 edu4_pb_0 edu4_v2_0 edu4_v3_0 edu4_v4_0 ///
    edu4_p2_0 edu4_p3_0 edu4_p4_0 edu4_p5_0 edu4_when_one_5 ///
    edu4_when_one_4 edu4_when_one_3 edu4.when_one_2 ///
    edu4.when_one_5 edu4.when_zero_4 edu4.when_zero_3 ///
    edu4.when_zero_2 edu4.when_one_5_f edu4.when_one_4_f ///
    edu4.when_one_3_f edu4.when_one_2_f edu4.when_one_5_m ///
    edu4.when_one_4_m edu4.when_one_3_m edu4.when_one_2_m ///
    edu4.when_zero_5_f edu4.when_zero_4_f edu4.when_zero_3_f ///
    edu4.when_zero_2_f edu4.when_zero_5_m edu4.when_zero_4_m ///
    edu4.when_zero_3_m edu4.when_zero_2_m

***** When Equal to One
*For Ever Smokers
quietly gen double `edu4_xa_1' = a0_constant_cons ///
    + a1_female_cons * female ///
    + a2_friends_cons * friends_smoke_bi ///
    + a3_parents_cons * parent_smoke_bi ///
    + a4_f_parent_cons * f_parent ///
    + a5_sp1_cons * sp_pca_1 ///
    + a6_f_sp1_cons * f_sp1 ///
    + a7_sp2_cons * sp_pca_2 ///
    + a8_f_sp2_cons * f_sp2 ///
    + a9_sp3_cons * sp_pca_3 ///
quietly gen double `edu4_pa_1' = 1 - normprob(`edu4_xa_1')

*For Current Smoker Portion
quietly gen double `edu4_xb_1' = b0\_constant\_cons
+ b1\_female\_cons * female
+ b2\_friends\_cons * friends\_smoke\_bi
+ b3\_parents\_cons * parent\_smoke\_bi
+ b4\_f\_parent\_cons * f\_parent
+ b5\_sp1\_cons * sp\_pca\_1
+ b6\_f\_sp1\_cons * f\_sp1
+ b7\_sp2\_cons * sp\_pca\_2
+ b8\_f\_sp2\_cons * f\_sp2
+ b9\_sp3\_cons * sp\_pca\_3
+ b10\_f\_sp3\_cons * f\_sp3
+ b11\_cw1\_cons * cw\_1
+ b12\_f\_cw1\_cons * f\_cw1
+ b13\_cw2\_cons * cw\_2
+ b14\_f\_cw2\_cons * f\_cw2
+ b15\_cigham\_cons * cig\_harm\_bi
+ b16\_f\_cigham\_cons * f\_cigh
+ b17\_othercigham\_cons * other\_cig\_harm\_bi
+ b18\_ptel1\_cons * pt\_pca\_1
+ b23\_edu4\_cons * 1
+ b24\_f\_edu\_cons * female
+ b25\_ate\_cons * ate\_v1
+ b26\_famdis\_cons * fam\_smoke\_dis
quietly gen double `edu4_pb_1' = 1 - normprob(`edu4_xb_1')

*For Number of Cigarettes Smoked Portion
quietly gen double `edu4_xc_1' = c1\_female\_cons * female
+ c2\_friends\_cons * friends\_smoke\_bi
+ c3\_parents\_cons * parent\_smoke\_bi
+ c4\_fparent\_cons * f\_parent
+ c5_sp1_cons * sp_pca_1  ///
+ c6_fsp1_cons * f_sp1  ///
+ c7_sp2_cons * sp_pca_2  ///
+ c8_fsp2_cons * f_sp2  ///
+ c9_sp3_cons * sp_pca_3  ///
+ c10_fsp3_cons * f_sp3  ///
+ c11_cw1_cons * cw_1  ///
+ c13_cw2_cons * cw_2  ///
+ c15_cigharm_cons * cig_harm_bi  ///
+ c16_fcigh_cons * f_cigh  ///
+ c17_othercigharm_cons * other_cig_harm_bi  ///
+ c18_pte1_cons * pt_pca_1  ///
+ c23_edu4_cons * 1  ///
+ c24_fedu4_cons * female  ///
+ c25_ate_cons * ate_v1  ///
+ c26_famdis_cons * fam_smoke_dis

quietly gen double `edu4_v2_1' = cut1_cons - (`edu4_xc_1')
quietly gen double `edu4_v3_1' = cut2_cons - (`edu4_xc_1')
quietly gen double `edu4_v4_1' = cut3_cons - (`edu4_xc_1')
quietly gen double `edu4_v5_1' = normprob(`edu4_v2_1')
quietly gen double `edu4_v6_1' = normprob(`edu4_v3_1') - normprob(`edu4_v2_1')
quietly gen double `edu4_v7_1' = normprob(`edu4_v4_1') - normprob(`edu4_v3_1')
quietly gen double `edu4_p5_1' = 1 - normprob(`edu4_v4_1')

*Top Category m=5
quietly gen double `edu4_when_one_5' = `edu4_pa_1' * `edu4_pb_1' * `edu4_p5_1'
quietly gen double `edu4_when_one_5_f' = `edu4_pa_1' * `edu4_pb_1' * `edu4_p5_1' if female == 1
quietly gen double `edu4_when_one_5_m' = `edu4_pa_1' * `edu4_pb_1' * `edu4_p5_1' if female == 0

*Middle Category m=4
quietly gen double `edu4_when_one_4' = `edu4_pa_1' * `edu4_pb_1' * `edu4_p4_1'
quietly gen double `edu4_when_one_4_f' = `edu4_pa_1' * `edu4_pb_1' * `edu4_p4_1' if female == 1
quietly gen double `edu4_when_one_4_m' = `edu4_pa_1' * `edu4_pb_1' * `edu4_p4_1' if female == 0

*Middle Category m=3
quietly gen double `edu4_when_one_3' = `edu4_pa_1' * `edu4_pb_1' * `edu4_p3_1'
quietly gen double `edu4_when_one_3_f' = `edu4_pa_1' * `edu4_pb_1' * `edu4_p3_1' if female == 1
quietly gen double `edu4_when_one_3_m' = `edu4_pa_1' * `edu4_pb_1' * `edu4_p3_1' if female == 0

*Bottom Category m=2
quietly gen double `edu4_when_one_2' = `edu4_pa_1' * `edu4_pb_1' * `edu4_p2_1'
quietly gen double `edu4_when_one_2_f' = `edu4_pa_1' * `edu4_pb_1' * `edu4_p2_1' if female == 1
quietly gen double `edu4_when_one_2_m' = `edu4_pa_1' * `edu4_pb_1' * `edu4_p2_1' if female == 0

local when_one `edu4_when_one_5' `edu4_when_one_4' `edu4_when_one_3' `edu4_when_one_2'
local when_one_f `edu4_when_one_5_f' `edu4_when_one_4_f' `edu4_when_one_3_f' `edu4_when_one_2_f'
local when_one_m `edu4_when_one_5_m' `edu4_when_one_4_m' `edu4_when_one_3_m' `edu4_when_one_2_m'

***** When Equal to Zero
*For Ever Smoker Portion
quietly gen double `edu4_xa_0' = a0_constant_cons ///
+ a1_female_cons * female ///
+ a2_friends_cons * friends_smoke_bi ///
+ a3_parents_cons * parent_smoke_bi ///
+ a4_f_parent_cons * f_parent ///
+ a5_sp1_cons * sp_pca_1 ///
+ a6_f_sp1_cons * f_sp1 ///
+ a7_sp2_cons * sp_pca_2 ///
+ a8_f_sp2_cons * f_sp2 ///
+ a9_sp3_cons * sp_pca_3 ///
+ a10_f_sp3_cons * f_sp3 ///
+ a11_cw1_cons * cw_1 ///
+ a12_f_cw1_cons * f_cw1 ///
+ a13_cw2_cons * cw_2 ///
+ a14_f_cw2_cons * f_cw2 ///
+ a15_cigharm_cons * cig_harm_bi ///
+ a16_f_cigharm_cons * f_cigh ///
+ a17_othercigharm_cons * other_cig_harm_bi ///
+ a18_pte1_cons * pt_pca_1 ///
+ a23_edu4_cons * 0 ///
+ a24_f_edu_cons * 0 ///
+ a25_ate_cons * ate_v1 ///
+ a26_famdis_cons * fam_smoke_dis
quietly gen double `edu4_pa_0' = 1 - normprob(`edu4_xa_0')

*For Current Smoker Portion
quietly gen double `edu4_xb_0' = b0_constant_cons ///
+ b1_female_cons * female ///
+ b2_friends_cons * friends_smoke_bi ///
+ b3_parents_cons * parent_smoke_bi ///
+ b4_f_parent_cons * f_parent ///
quietly gen double `edu4_pb_0' = 1 - normprob(`edu4_xb_0')

*For Number of Cigarettes Smoked Portion
quietly gen double `edu4_xc_0' = c1_female_cons * female ///
  + c2_friends_cons * friends_smoke_bi ///
  + c3_parents_cons * parent_smoke_bi ///
  + c4_fparent_cons * f_parent ///
  + c5_sp1_cons * sp_pca_1 ///
  + c6_fsp1_cons * f_sp1 ///
  + c7_sp2_cons * sp_pca_2 ///
  + c8_fsp2_cons * f_sp2 ///
  + c9_sp3_cons * sp_pca_3 ///
  + c10_fsp3_cons * f_sp3 ///
  + c11_cw1_cons * cw_1 ///
  + c13_cw2_cons * cw_2 ///
  + c15_cigharm_cons * cig_harm_bi ///
  + c16_fcigh_cons * f_cigh ///
  + c17_othercigharm_cons * other_cig_harm_bi ///
  + c18_pte1_cons * pt_pca_1 ///
  + c23_edu4_cons * 0 ///
  + c24_fedu4_cons * 0 ///
  + c25_ate_cons * ate_v1 ///
  + c26_famdis_cons * fam_smoke_dis
quietly gen double `edu4_v2_0' = cut1_cons - (`edu4_xc_0')
quietly gen double `edu4_v3_0' = cut2_cons - (`edu4_xc_0')
quietly gen double `edu4_v4_0' = cut3_cons - (`edu4_xc_0')
quietly gen double `edu4_p2_0' = normprob(`edu4_v2_0')
quietly gen double `edu4_p3_0' = normprob(`edu4_v3_0') - normprob(`edu4_v2_0')
quietly gen double `edu4_p4_0' = normprob(`edu4_v4_0') - normprob(`edu4_v3_0')
quietly gen double `edu4_p5_0' = 1 - normprob(`edu4_v4_0')

*Top Category m=5
quietly gen double `edu4_when_zero_5' = `edu4_pa_0' * `edu4_pb_0' * `edu4_p5_0'
quietly gen double `edu4_when_zero_5_f' = `edu4_pa_0' * `edu4_pb_0' * `edu4_p5_0' if female == 1
quietly gen double `edu4_when_zero_5_m' = `edu4_pa_0' * `edu4_pb_0' * `edu4_p5_0' if female == 0

*Middle Category m=4
quietly gen double `edu4_when_zero_4' = `edu4_pa_0' * `edu4_pb_0' * `edu4_p4_0'
quietly gen double `edu4_when_zero_4_f' = `edu4_pa_0' * `edu4_pb_0' * `edu4_p4_0' if female == 1
quietly gen double `edu4_when_zero_4_m' = `edu4_pa_0' * `edu4_pb_0' * `edu4_p4_0' if female == 0

*Middle Category m=3
quietly gen double `edu4_when_zero_3' = `edu4_pa_0' * `edu4_pb_0' * `edu4_p3_0'
quietly gen double `edu4_when_zero_3_f' = `edu4_pa_0' * `edu4_pb_0' * `edu4_p3_0' if female == 1
quietly gen double `edu4_when_zero_3_m' = `edu4_pa_0' * `edu4_pb_0' * `edu4_p3_0' if female == 0

*Bottom Category m=2
quietly gen double `edu4_when_zero_2' = `edu4_pa_0' * `edu4_pb_0' * `edu4_p2_0'
quietly gen double `edu4_when_zero_2_f' = `edu4_pa_0' * `edu4_pb_0' * `edu4_p2_0' if female == 1
quietly gen double `edu4_when_zero_2_m' = `edu4_pa_0' * `edu4_pb_0' * `edu4_p2_0' if female == 0

local when_zero `edu4_when_zero_5' `edu4_when_zero_4' `edu4_when_zero_3'
local when_zero_f `edu4_when_zero_5_f' `edu4_when_zero_4_f' `edu4_when_zero_3_f'
local when_zero_m `edu4_when_zero_5_m' `edu4_when_zero_4_m'
local when_zero_m `edu4_when_zero_3_m'

tempvar me_edu4_5 me_edu4_4 me_edu4_3 me_edu4_2 ///
me_edu4_5_f me_edu4_4_f me_edu4_3_f me_edu4_2_f ///
me_edu4_5_m me_edu4_4_m me_edu4_3_m me_edu4_2_m

local Nmes1 `me_edu4_5' `me_edu4_4' `me_edu4_3' `me_edu4_2'
local Nmes2 `me_edu4_5_f' `me_edu4_4_f' `me_edu4_3_f' `me_edu4_2_f'
local Nmes3 `me_edu4_5_m' `me_edu4_4_m' `me_edu4_3_m' `me_edu4_2_m'
forvalues i = 1/3{
    foreach x of local Nmes`i'{
        quietly gen double `x' = .
    }
    local i = `i' +1
}

*Together
quietly replace `me_edu4_5' = ccm_5 * (`edu4_when_one_5' - `edu4_when_zero_5')
quietly sum `me_edu4_5'
scalar meedu45_mean = r(mean)
return scalar ame_edu4_4 = meedu45_mean
quietly replace `me_edu4_4' = ccm_4 * (`edu4_when_one_4' - `edu4_when_zero_4')
quietly sum `me_edu4_4'
scalar meedu44_mean = r(mean)
return scalar ame_edu4_3 = meedu44_mean
quietly replace `me_edu4_3' = ccm_3 * (`edu4_when_one_3' - `edu4_when_zero_3')
quietly sum `me_edu4_3'
scalar meedu43_mean = r(mean)
return scalar ame_edu4_2 = meedu43_mean
quietly replace `me_edu4_2' = ccm_2 * (`edu4_when_one_2' - `edu4_when_zero_2')
quietly sum `me_edu4_2'
scalar meedu42_mean = r(mean)
return scalar ame_edu4_1 = meedu42_mean

*Girls
quietly replace `me_edu4_5_f' = ccm_5 * (`edu4_when_one_5_f' - `edu4_when_zero_5_f')
quietly sum `me_edu4_5_f'
scalar meedu45f_mean = r(mean)
return scalar ame_edu4_4_f = meedu45f_mean
quietly replace `me_edu4_4_f' = ccm_4 * (`edu4_when_one_4_f' - `edu4_when_zero_4_f')
quietly sum `me_edu4_4_f'
scalar meedu44f_mean = r(mean)
return scalar ame_edu4_3_f = meedu44f_mean
quietly replace `me_edu4_3_f' = ccm_3 * (`edu4_when_one_3_f' - `edu4_when_zero_3_f')
quietly sum `me_edu4_3_f'
scalar meedu43f_mean = r(mean)
return scalar ame_edu4_2_f = meedu43f_mean
quietly replace `me_edu4_2_f' = ccm_2 * (`edu4_when_one_2_f' - `edu4_when_zero_2_f')
quietly sum `me_edu4_2_f'
scalar meedu42f_mean = r(mean)
return scalar ame_edu4_1_f = meedu42f_mean

*Boys
quietly replace `me_edu4_5_m' = ccm_5 * (`edu4_when_one_5_m' - `edu4_when_zero_5_m')
quietly sum `me_edu4_5_m'
scalar meedu45m_mean = r(mean)
return scalar ame_edu4_4_m = meedu45m_mean

quietly replace `me_edu4_4_m' = ccm_4 * (`edu4_when_one_4_m' - `edu4_when_zero_4_m')
quietly sum `me_edu4_4_m'
scalar meedu44m_mean = r(mean)
return scalar ame_edu4_3_m = meedu44m_mean

quietly replace `me_edu4_3_m' = ccm_3 * (`edu4_when_one_3_m' - `edu4_when_zero_3_m')
quietly sum `me_edu4_3_m'
scalar meedu43m_mean = r(mean)
return scalar ame_edu4_2_m = meedu43m_mean

quietly replace `me_edu4_2_m' = ccm_2 * (`edu4_when_one_2_m' - `edu4_when_zero_2_m')
quietly sum `me_edu4_2_m'
scalar meedu42m_mean = r(mean)
return scalar ame_edu4_1_m = meedu42m_mean

end

bootstrap ame_edu2_4_out = r(ame_edu2_4) ///
    ame_edu2_3_out = r(ame_edu2_3) ///
    ame_edu2_2_out = r(ame_edu2_2) ///
    ame_edu2_1_out = r(ame_edu2_1) ///
    ame_edu2_4_f_out = r(ame_edu2_4_f) ///
    ame_edu2_3_f_out = r(ame_edu2_3_f) ///
    ame_edu2_2_f_out = r(ame_edu2_2_f) ///
    ame_edu2_1_f_out = r(ame_edu2_1_f) ///
    ame_edu2_4_m_out = r(ame_edu2_4_m) ///
    ame_edu2_3_m_out = r(ame_edu2_3_m) ///
    ame_edu2_2_m_out = r(ame_edu2_2_m) ///
    ame_edu2_1_m_out = r(ame_edu2_1_m) ///
ame_edu3_4_out = r(ame_edu3_4)
ame_edu3_3_out = r(ame_edu3_3)
name_edu3_2_out = r(ame_edu3_2)
name_edu3_1_out = r(ame_edu3_1)
name_edu3_4_f_out = r(ame_edu3_4_f)
name_edu3_3_f_out = r(ame_edu3_3_f)
name_edu3_2_f_out = r(ame_edu3_2_f)
name_edu3_1_f_out = r(ame_edu3_1_f)
name_edu3_4_m_out = r(ame_edu3_4_m)
name_edu3_3_m_out = r(ame_edu3_3_m)
name_edu3_2_m_out = r(ame_edu3_2_m)
name_edu3_1_m_out = r(ame_edu3_1_m)
name_edu4_4_out = r(ame_edu4_4)
name_edu4_3_out = r(ame_edu4_3)
name_edu4_2_out = r(ame_edu4_2)
name_edu4_1_out = r(ame_edu4_1)
name_edu4_4_f_out = r(ame_edu4_4_f)
name_edu4_3_f_out = r(ame_edu4_3_f)
name_edu4_2_f_out = r(ame_edu4_2_f)
name_edu4_1_f_out = r(ame_edu4_1_f)
name_edu4_4_m_out = r(ame_edu4_4_m)
name_edu4_3_m_out = r(ame_edu4_3_m)
name_edu4_2_m_out = r(ame_edu4_2_m)
name_edu4_1_m_out = r(ame_edu4_1_m)

, rep(200) saving(boot_results_b_level_ame, replace)

estat bootstrap, all
log close
References


[https://doi.org/10.1016/j.ypmed.2011.03.013](https://doi.org/10.1016/j.ypmed.2011.03.013).


Aryal, Umesh R., Max Petzold, Göran Bondjers, and Alexandra Krettek. 2014.


[https://doi.org/10.3402/gha.v7i1.24488](https://doi.org/10.3402/gha.v7i1.24488).


http://hdl.handle.net/10986/13750.


https://doi.org/10.1002/14651858.CD003439.pub2.


Schwartz, Marlene B., Glenn E. Schneider, Yoon-Young Choi, Xun Li, Jennifer Harris, Tatiana Andreyeva, Maia Hyary, Nicolette Highsmith Vernick, and Lawrence J. Appel. 2017. “Association of a Community Campaign for Better Beverage


State Senate of New Mexico. *Sweetened Beverage Excise Tax Act*. By Gerald Ortiz y Pino. 2010. (Senate Bill 243)


   http://www.who.int/tobacco/surveillance/policy/country_profile/npl.pdf?ua=1.


