THE ROLE OF SMOKELESS TOBACCO IN SMOKING CESSATION AND INITIATION: AN INVESTIGATION OF AMERICAN YOUTH

by

JASON D. MILLER

(Under the Direction of Teresa Mauldin)

ABSTRACT

While the overall rate of smoking in the United States has been declining, there exists an unmet societal demand for improved smoking cessation tools and smoking prevention methods. This study built on recent work in the empirical literature on this topic by using longitudinal data from the National Longitudinal Study of Adolescent Health (Add Heath) to analyze the efficacy of smokeless tobacco products as cessation and anti-initiation tools for young Americans. The study applied models adapted from prior literature on smokeless tobacco in cessation. In addition the study surpassed the design phase of prior smokeless tobacco literature by contributing the use of propensity-score matching to "promote honesty" (Rubin, 2001). Furthermore, mixed model and OLS propensity-score regression results were compared for consistency. It was determined that smokeless tobacco is likely of use to dual-users and may be of use in general cessation and initiation programs among youth, although further research is needed.

INDEX WORDS: Smokeless tobacco, Smoking, Prevention, Youth, Cessation, Tobacco, Propensity score, Dual-users, DST, MST, Snus, Add Health, Longitudinal data analysis

THE ROLE OF SMOKELESS TOBACCO IN SMOKING CESSATION AND INITIATION: AN INVESTIGATION OF AMERICAN YOUTH

by

JASON D. MILLER

B.S.F.C.S., The University of Georgia, 2006

A Thesis Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree

MASTER OF SCIENCE

ATHENS, GEORGIA

2009

THE ROLE OF SMOKELESS TOBACCO IN SMOKING CESSATION AND INITIATION: AN INVESTIGATION OF AMERICAN YOUTH

by

JASON D. MILLER

Major Professor: Teresa Mauldin

Committee: Brenda Cude

Joseph Sabia

Electronic Version Approved:

Maureen Grasso

Dean of the Graduate School

The University of Georgia

December 2009

DEDICATION

In loving memory of William H. Miller, Sr.

I know there's a lot of talk going around today
About cigarette smoking whittling your life away
I've seen it and I've heard it so many times
That finally it… it just started to prey on my mind
I guess it scared me a little bit
That's why I decided I was gonna quit
So while I was sitting here forming my battle plan
I took another puff and turned on the fan…

I wish I could think of something bad to say about cigarettes
Boo on cigarettes…
Don't smoke, don't smoke, don't smoke
You quit smoking that'll leave more for me!
I love it, I love it! No I don't love cigarettes ya know
Don't misunderstand me, I hate cigarettes…

Reed, J. (2006). "Another Puff" on Thank You For Smoking (Original Motion Picture Soundtrack). New York City: RCA. (1971)

Jerry Reed died of emphysema September 2nd, 2008.

ACKNOWLEDGEMENTS

This thesis would not have been possible without the direct and indirect support of the following individuals, in alphabetical order: Brenda Cude, Francisco Diaz, Ania Dudziak, Heather Hjorth, M. Blaire Long, Teresa Mauldin, Katie Parks, Uncle Phil, Leslie Green-Pimentel, Bill & Reta Miller, Inez Miller, James Mwai, Joseph J. Sabia, Anne Sweaney, Ratapol Teratanavat and Melissa Wilmarth. In particular, I would like to recognize Professors Teresa Mauldin, Brenda Cude, and Joseph J. Sabia for the knowledge and experience they provided during my undergraduate and graduate programs, as well as for their help in developing this thesis.

TABLE OF CONTENTS

Page

ACKNOWLEDGEMENTS v

LIST OF TABLES viii

CHAPTER

1 Introduction 1

Overview 2

Background 4

2 Literature Review 7

Demand for Smokeless Tobacco 7

Snus and the Swedish Experience 11

Economic Analysis of Smokeless Tobacco in Smoking Prevention 12

3 Methodology 16

Data 16

Hypotheses 20

Theoretical Model 21

Propensity Score Matching 23

4 Results 26

Application of Smokeless Tobacco Models to Youth Institutions 26

Adaptation of the Ault et al. (2004) Cessation Model for Estimating Cessation & Initiation Effects using Longitudinal Data 29

Propensity Score Approach 32

5 Discussion 66

Overview of the Role of Smokeless Tobacco in Youth Smoking Cessation and Initiation 69

Discussion of Findings with a Subgroup of Interest - Dual Users 70

Synthesis of Implications for Consumers, Policymakers and Future Researchers 71

REFERENCES 76

APPENDICES 81

A AMERICAN SNUS TOBACCO PRODUCTS 81

B NON-SNUS SMOKELESS TOBACCO PRODUCTS WITH YOUTH AND DUAL-USER APPEAL 84

LIST OF TABLES

Page

Table 1: Variables of Key Interest in Main (Probit) Models. 37

Table 2: Ex-post Estimates of Quitting Smoking Primary Probit Cessation Models. 40

Table 3: Longitudinal Estimates: Cessation and Initiation (Pr > $\chi 2$ in parentheses) 43

Table 4: Additional Variables Used to Define Propensity Scores. 46

Table 5: Cessation: Key Observables in Control and Treatment Assignment Before Matching. . 48

Table 6: Initiation: Key Observables in Control and Treatment Assignment Before Matching. .. 50

Table 7: Cessation: Improvement in Balance of Initial Groups. 52

Table 8: Cessation: Weighted Probit Regression Matched on Propensity Scores 55

Table 9: Initiation: Weighted Probit Regression Matched on Propensity Scores. 62

Table 10: Cessation: Comparison of OLS and Mixed-Model Propensity Score Regressions. 64

Table 11: Initiation: Comparison of OLS and Mixed Model Propensity Score Regressions…… 65

CHAPTER 1

INTRODUCTION

Prior to the 20th century, consumers in the United States demonstrated a preference for smokeless tobacco over cigarettes and other forms of burned tobacco. Over the course of the 1900s, this domestic trend reversed itself. Flammable tobacco came to dominate the market for tobacco products throughout the world, with the exception of some Scandinavian nations (Rogozinski, 1990, p.1). Growing health concerns related to smoking, combined with increased cigarette regulation and litigation, have deterred many consumers from smoking. Due to the addictive[1] nature of tobacco, demand grew for products to aid in smoking cessation. For instance, there are now gums and mints containing nicotine that are sold over-the-counter as forms of Nicotine Replacement Therapy (NRT) (Chaloupka & Tauras, 2004). Although addictive and high priced, these tools have been accepted as aides in smoking cessation. Only a small number of researchers have examined what role smokeless tobacco may play in consumers' attempts to substitute away from cigarette usage toward a less offending alternative. A related question, the impact of smokeless tobacco on youth smoking initiation, has not been examined at all in the economics literature (although it is informed by many past studies involving youth and smoking policy). Aside from providing tools and insights for the medical community, contributions to the youth smoking cessation and (anti-) initiation literature also provide direction for policymakers.

---

[1] Smoking is addictive in a medical and economic sense (Becker, Grossman, & Murphy, 1994; Benowitz, 1988). Addiction as a product characteristic does not necessarily imply irrational consumption, however.

Overview

When reviewing tobacco consumption patterns in the United States, many investigators begin their discussions with the infamous 1964 "Smoking and Health" report to the Surgeon General by his Advisory Committee. The report credibly and visibly indicted cigarette smoking for its hazardous health effects and served as a major turning point in America's cultural attitudes toward cigarette consumption (Borio, 2008; U.S. Public Health Service, 1964). Since then, increases in the financial costs of cigarettes stemming from increased taxes and legal costs for producers (e.g., settlement payments) have been complemented by increases in the internalization of social costs (such as concern over secondhand smoke effects), increased internalization of health costs by the consumer (i.e., better information) and even legal costs to consumers. The increased costs of smoking have contributed to the reduction of cigarette usage from approximately 42.4% adult smoking participation in 1965 to what the Centers for Disease Control (CDC) refers to as the so-called stable or "hardcore" rate of smoking participation – about 25% of American adults in the 1990s (CDC, 2007). Despite the implied stability of this figure, the smoking rate had declined to 20.8% in 2006 (Borio, 2008; Rock, Malarcher, Kahende, Asman, Husten and Carabollo, 2006; Tauras, Powell, Chaloupka, and Ross, 2007). Expressed in different terms, annual per capita cigarette consumption peaked at 4,345 in 1963 and declined to 1,814 by 2004 (CDC Data Table, 2008).

About two decades after the major anti-smoking catalysts of the 1960s, a 1986 report by the Surgeon General's Advisory Committee linked the use of snuff (a particular form of smokeless tobacco) to oral cancer (National Institutes of Health, 1986). In the months following the release of this report, Congress enacted the Comprehensive Smokeless Tobacco Health Education Act of 1986 (CSTHEA), which, among other things, banned the advertisement of

smokeless tobacco products on television and required certain warning labels to be placed on all smokeless tobacco goods. Ostensibly, the medical findings and related legal action against smokeless tobacco were very similar to that of cigarettes and it correlated with a reduction in sales of many forms of smokeless tobacco (such as dry snuff and loose leaf) over the next two decades (Tauras et al., 2007). While other forms of tobacco suffered a reduction in sales, however, the prevalence of moist snuff tobacco (MST) increased steadily and rapidly. Tauras et al. (2007) noted that MST sales rose from less than 40 million pounds in 1986 to about 65 million in 2001. By 1998 MST was sold in greater quantity than all other forms of smokeless tobacco combined. Interestingly, Sweden began to experience resurgence in the consumption of *snus*[2] during this time due to a change in consumer demand, which has been empirically linked to improvements in public health and is reviewed further in Chapter 2.

The primary aim of this study is to determine how consumers and policymakers should regard smokeless tobacco within the scope of American youth tobacco use reduction, not necessarily to explain the success of American MST. To be more precise, the aims of this study are (1) to evaluate smokeless tobacco as a cessation and initiation tool through adaptation of existing empirical models, now applied to youth (2) to investigate an overlooked subgroup – dual users and (3) to improve on past methodologies, by use of longitudinal data and propensity-scores. A wealth of clinical studies demonstrates that smokeless tobacco is less harmful than smoking in terms of health. Could smokeless tobacco be utilized to aide in smoking cessation programs for some youth groups? What information can the data provide consumers and policymakers to inform the government's ban on sales of smokeless tobacco to youth?

---

[2] Snus is term used for pouch tobacco in Sweden and the non-spit pouches recently inspired in America.

These questions are at the heart of this study. Some can be answered by medical or chemical studies and others are informed by existing econometric literature. The empirical investigation of smokeless tobacco, however, is sparse. This study was an attempt to better answer these questions by examining a large, representative dataset that had yet to be investigated in the smokeless tobacco debate. Additionally, a more thorough set of identification strategies from previous studies that focused on initial design and balance between relevant groups was proposed and is detailed in Chapter 3. To understand the context of this study, a review of the background of the debate on tobacco proves useful.

Empirical Background

A large body of medical evidence has demonstrated that all forms of tobacco pose serious health and addiction risks. CDC fact sheets, which summarize many individual CDC reports, note that smoking causes cancers of the bladder, oral cavity, pharynx, larynx, esophagus, cervix, kidney, lung, pancreas and stomach, as well as acute myeloid leukemia (CDC, 2008). Smokeless tobacco also contains carcinogens and has been linked to cancer of the oral cavity as well as recession of the gums (CDC, 2007). While it is clear that all forms of tobacco are harmful to health in some way, it is equally clear that significant differences exist in the health effects of tobacco's many forms. Smoking is clearly the most harmful to oneself and in terms of health externalities (such as second hand smoke).

The vast majority of harm caused by smoking results not from nicotine or raw tobacco per se, but from the multitude of smoke constituents which are inhaled by smokers (Gilljam & Galanti, 2003). The vast array of particles found in cigarette smoke damages one's lungs and other vital organs, as would the inhalation of smoke arising from any other source (wood, for instance). This is the primary reason that unburned forms of tobacco are considered relatively

safer. This point was illustrated by the National Academy of Science and Institute of Medicine in their report *Clearing the Smoke*, which found that smokeless tobacco may serve as a "valid substitute" for smoking since it (like nicotine replacement products) lowered the morbidity and mortality related to tobacco (Stratton, 2001, pp. 1-14; Ault, Ekelund, Jackson & Saba, 2004).

Another key difference between the health impact of smoked and smokeless tobacco is in the level of Tobacco Specific Nitrosamines (TSNAs)[3], which cause cancer. These are delivered at a much lower level in smokeless tobacco, although it is important to note that differences exist within smokeless groups[4] particularly by curing method. For instance, in North America the blends of moist snuff products are generally high in fire cured tobacco, which is comparatively high in TSNAs as opposed to the air and sun curing methods often used in the manufacture of Swedish snus (Foulds, Ramstrom, Burk, & Fagerstrom, 2003). Finally, one should note that smokeless tobacco products tend to provide at least as much nicotine delivery as cigarettes and often much more (Foulds et al., 2003; Gilljam & Galanti, 2003). The delivery of such nicotine levels indicates both that smokeless tobacco is addictive and that it may be able to satiate nicotine cravings without the presence of smoke constituents. Moreover, smokeless tobacco delivers nicotine at a quicker rate than regulated cessation products (Ault et al., 2004; Stratton, 2001). These differences provide a basis for *a priori* reasoning that smokeless tobacco may be of use in cessation, since it delivers both nicotine and, as a tobacco product, may also satisfy characteristics of the tobacco use experience, which non-tobacco nicotine products cannot. The

---

[3] For a detailed explanation of TSNA levels, see Hoffmann, Brunnemann, Prokopczyk, & Djordjevic, 1994.

[4] Here "groups" could refer to any number of differences between smokeless tobacco products. Examples include tobacco form, brand, cut size (if any), production run, storage conditions, environmental conditions during consumption and so on.

validity of this hypothesis cannot be determined *a priori,* however; it is possible that smokeless tobacco may reinforce smoking behavior if the introduction of smokeless tobacco to smokers causes a reduction or no change in the probability of successful cessation.

Tobacco has been known throughout history as a potent herb – used in a gradient of forms and contexts; in this thesis the author sought to determine how the more natural (smokeless) tobacco-based consumer products would fare as tools to reduce the incidence of youth smoking. While no use of tobacco is virtually always optimal for consumers' health, the addictive and typically long-term consumption of the good persists – and in modern times there is practically universal agreement that the decision to smoke is most myopically (irrationally) made among adolescents. Often referred to as time-inconsistent preferences, these poor decisions impact the user through "negative internalities" far more than it may also burden others with negative externalities (Gruber & Koszegi, 2004). When the decision is rational however, econometric valuation of negative externalities reveal that the utility gained by smokers may be outweighed by the negative externalities, such as secondhand smoke and so on (Sabia & Rees, 2008). Furthermore, youth have shown greater responsiveness to intervention tools such as taxes, relative to other groups (DeCicca, Kenkel, and Mathios, 2008). It is for these reasons that American governments (federal, state, local) have intervened in the market, most notably in the form of an all-out ban on sale to minors.

CHAPTER 2

LITERATURE REVIEW

In light of the large quantity of econometric studies concerning cigarette consumption, the body of economic literature that has examined the demand for smokeless tobacco in the United States is comparatively small. At this time, only five econometric studies of macroeconomic smokeless tobacco's demand have been published (Chaloupka, Tauras & Grossman, 1997; Ohsfeldt & Boyle, 1994; Ohsfeldt, Boyle, & Capilouto, 1997; Ohsfeldt, Boyle, & Capilouto, 1999; Tauras et al., 2007). In recent years, however, smokeless tobacco has begun to receive a substantially larger amount of attention from researchers in many fields. There are two apparent motivations for this increase. One is that a high incidence of smokeless use in Sweden has seemingly correlated with a decreased level of cancer mortality and other health outcomes. This is a phenomenon commonly referred to as "the Swedish experience." The other motivating factor is that American consumption of moist snuff tobacco has been steadily increasing since the late 1980s, whereas cigarette consumption has decreased by about 200 billion cigarettes per year during the period from 1986 to 2001 (Tauras et al., 2007).

Demand for Smokeless Tobacco

Throughout the 1990s, studies by Ohsfeldt and others examined smokeless tobacco's demand primarily as it relates to changes in taxation and policy. The Ohsfeldt and Boyle (1994) analysis estimated smokeless tobacco participation equations for adults (aged 16 and older) by using state-level aggregates in a cross-section constructed from the 1985 Current Population

Survey (CPS). Their findings estimated an own-tax elasticity of demand of about -0.55 using state-level smokeless tobacco excise tax rates and a cross-tax elasticity of demand with respect to cigarettes of about 0.49 using state level cigarette taxes, holding smokeless tobacco tax rates constant. Variation came from state-level differences, such as price, measured by state smokeless tobacco tax, holding constant differences in the general price of tobacco by state. The investigators reported that a 10% increase in the price of smokeless (via tax) is associated with a 5.5% reduction in demand. These results imply that smokeless tobacco consumption is negatively related to increases in final price, but is slightly inelastic to price changes (as one would expect of an addictive good). The positive cross-tax elasticity implied that an increase in the cost of cigarettes is associated with an increase in the prevalence of smokeless tobacco use.

As such, this initial study of smokeless tobacco demand provided some limited evidence that cigarettes and smokeless tobacco may act as economic substitutes (with tax rates effectively representing changes in consumers' real income or purchasing power). This investigation of the 1985 CPS data also found that public and private smoking bans were insignificant determinants of smokeless tobacco use (Ohsfeldt & Boyle, 1994). Cultural changes and increases in smoking regulation (bans) over the past two decades may provide reason to seek further research before concluding that smoking bans have no effect on smokeless tobacco usage today. In addition, the number of states with restrictions on public smoking in place (of which there are currently 17) caused the sample to be restrictively small in Osfeldt & Boyle's study. Due to this small, non-random sample and multicolinnearity problems, the study used cluster analysis to determine that agglomerative hierarchical clustering (AHC) could be employed in developing a binary variable representing "restrictive" states in Ohsfeldt's empirical model. The researchers were unable to obtain data on the (often more restrictive) local level ordinances.

Ohsfeldt, Boyle, and Capilouto (1997) expanded the original Ohsfeldt study by using individual level data from the September 1985 CPS. In this second study the authors again found a negative and significant own-tax elasticity of demand and positive and significant cross-tax elasticity with respect to cigarettes. The magnitudes of the estimates were much smaller than before, however: -.15 and .10, respectively. A central problem with both the 1997 and 1994 Ohsfeldt studies, though, is that they make use of data predating the Comprehensive Smokeless Tobacco Health Education Act of 1986 (CSTHE) and the Synar Amendment (which restricted youth access to tobacco and provided a catalyst for the creation of many additional state level regulations). In this way, the studies based on data from the early 1980s are akin to the study of cigarette consumption prior to 1964. The behavior observed in these studies describes consumption that may have been significantly altered due to the increase in consumer information stemming from CSTHE and the increased financial and social costs following this act.

In 1999 Ohsfeldt and Melkersson published a significantly revised and updated version of their study in the National Bureau of Economic Research (NBER) Conference Report Series. The authors used September 1992, January 1993, and May 1993 CPS data. In this report they estimated prevalence equations for cigarettes and moist snuff tobacco (MST), controlling for both smokeless and cigarette taxes in each equation. The key findings were that MST taxes do not affect the demand for cigarettes and their effect on MST demand was small in magnitude and statistically insignificant when cigarette taxes were treated as exogenous. They also found that higher cigarette taxes reduce smoking but increase snuff use. This latter finding again implies that the two forms of tobacco are substitutes, but the fact that a higher MST tax did not affect cigarette demand may cast doubt on the statistical power of this study. In another study,

however, Ault et al. (2005) point out that in some instances the cross-price effects of two goods might not necessarily be symmetric.

Concurrent with the second Ohsfeldt study, another group of economists became the first to study the demand for smokeless tobacco by youth under the age of 16 (Chaloupka, Tauras, & Grossman, 1997). These authors made use of 1992, 1994, and 1996 data from the Monitoring the Future Surveys to estimate ordered probit models of the impact of price and legislation (such as regulation of cigarette vending machines and public smoking bans) on the frequency of smokeless tobacco use by young males. They estimated a price-elasticity of demand for adolescent males of -0.43 and concluded that increases in smokeless tobacco taxes would lead to significant reductions in the incidence of use but not on the average quantity of use by users.

A decade later, this study was expanded to examine the effect of several policies and offered methodological improvements in estimating conditional smokeless tobacco demand (Tauras et al., 2007). Data were extracted from the 1995, 1997, 1999, and 2001 National School-Based Youth Risk Behaviors Surveys (YRBS), which are conducted by the CDC. The policies examined included 'purchase, possession, and use' legislation, state bans on smoking in high schools, and state regulation of tobacco vending machines. Their findings on the effect of smokeless tobacco taxes and cigarette prices were consistent with past studies, indicating that smokeless tobacco is a good that may act as a substitute for cigarettes. They found little or no effect of other tobacco control policies on smokeless tobacco prevalence or frequency.

Taken as a whole, these traditional demand studies of smokeless tobacco offered evidence that regulators can deter (or encourage) consumption of smokeless tobacco by altering excise taxes on these products. Simultaneously, a fair amount of evidence was presented which

indicates that deterring smoking through taxes may have the consequence of inducing smokeless usage. A more in-depth analysis of this relationship between smokeless tobacco and cigarettes as it relates to cessation and harm-reduction was provided in a recent landmark study (Ault et al, 2004) motivated by the Swedish experience with smokeless tobacco.

Snus and "The Swedish Experience"

An international and interdisciplinary debate over the public health value of smokeless tobacco erupted, following a resurgence in the popularity of *snus* in Sweden among males coupled with seemingly improved aggregate health outcomes. The term *snus* is used generally in Sweden to refer to smokeless tobacco, the most common form being an oral pouch which is typically placed under the upper lip and does not necessitate the spitting of tobacco juice. An overview of this debate about "the Swedish experience," with a focus on medical and sociological journals, is provided by Foulds et al. (2003). *Snus* clearly contains harmful chemicals such as Tobacco Specific Nitrosamines (TSNAs) which can cause cancer.

Several chemical analyses were presented by Foulds et al.(2003) which indicated that snus delivers far fewer TSNAs and harmful substances than cigarettes, although the level of delivered nicotine is commonly equal or greater. These results are consistent with research conducted by the National Academy of Sciences (Stratton, 2001) on smokeless tobacco in general, except that they indicate the unique manufacture of *snus* may make it even less harmful than the traditional forms of American smokeless tobacco (such as lower-lip snuff products). If true, the latter point may mean that positive health results (reductions in mortality, morbidity, and healthcare spending) found in Swedish consumption of smokeless tobacco may not necessarily be robust to American consumers. Even if it were internationally robust, it still

certainly could not be assumed to be so for the youth population in particular. Within the context of Sweden, though, the authors found sufficient evidence to conclude that *snus* usage has contributed to the unusually low rates of smoking among males and to a greater level of overall public health, with some drawbacks such as increased hypertension (Bolinder, Alfredsson, Englund, & de Faire, 1994).

This study and others, which supported the conclusion that Swedish *snus* was associated with increased public health and not linked to cancer in Sweden, generated interest in examining the extent to which *snus* had been used as a tool for smoking cessation. Members of the Swedish government's Karolinska Institute conducted observational research wherein they funded a national telephone survey of 2,000 current and former smokers (Gilljam & Galanti, 2003). The results of their cross-sectional study suggested that by using *snus*, Swedish male smokers increased their overall probability of cigarette abstinence (although most men who quit smoking did so by other means).

Although compelling, theoretically prescribed models did not back these studies on the "Swedish experience" and the observational cessation research conducted may not be credibly generalized to populations that primarily consume non-Swedish smokeless tobacco. Fortunately, these studies have motivated some initial economic analyses of smokeless tobacco's value in smoking cessation.

Economic Analysis of Smokeless Tobacco in Smoking Cessation

Only recently have economists begun to empirically analyze the use of smokeless tobacco in cessation. Aside from general theories of demand and elasticity, this topic lends itself to Becker's rational addiction theory – and to related theoretical concerns of time-inconsistent

preferences. This became a theme in the first economic study on this topic, which was a manuscript by Melkersson (2000). This manuscript was cited by Ault et al. (2004) for its theoretical model of smoking behavior, which was an adaptation of Suranovic, Goldfarb and Leonard's (1999) model for use in smokeless tobacco studies. The model from this study was reproduced in Ault's study and the content and findings of the study are similar to "Rational Addiction When There Are Two Addictive Goods: Cigarettes and Smokeless Tobacco" which was published later that year by Umea University in Stockholm (Bask & Melkersson, 2000) in the Umea Economic Studies. The authors later published essentially the same study, with a shorter discussion of its model, under the title "Should one use smokeless tobacco in smoking cessation programs? A rational addiction approach" in the European Journal of Health Economics (Bask & Melkersson, 2003).

In the Bask and Melkerson studies the authors attempted to adapt Becker and Murphy's (1988) rational addiction theory to answer the question of cessation by making use of aggregated annual time series data from a sample of Swedes (1964-1997). They estimated Ordinary Least Squares (OLS) and Generalized Method of Moments (GMM) models constructed based on the utility maximization of a representative consumer, as according to Greene (2000). Contrary to the direction of most other literature, Bask Melkersson (2003) found that the cross-price elasticity of demand between cigarettes and snus indicated that "…snus contributes to increased smoking. Thus even if snus taking is less harmful…it is not advisable to encourage its use in smoking cessation programs." (p. 267).

A 2005 article by Ault, Beard, Jackson, and Saba in the same journal discussed serious doubts about the credibility of Bask and Melkersson's studies. The doubt primarily concerned problems identified with the statistical techniques used and with a major flaw in interpretation of

their results. Such problems included that their models omitted controls for real income, used questionable instrumental variables without providing their statistical significance in first-stage equations, and aggregated data across genders (whereas smokeless tobacco users are almost entirely male). The instruments used varied without any explanation – most notably there were unexplained shifts between current, lead, and lagged real income. The authors have been criticized for implying that current, lead, and lagged prices (thus real income) are exogenous while lagged consumption is endogenous.

Above all, Bask and Melkersson were criticized because of a "fundamental flaw" in the concept underlying their study. The "flaw" is the fact that interpretation of the cross-price effects is dependent on the reason for the change in price – supply or demand. Bask and Melkersson did not address this point in either iteration of their research and Ault et al. (2005) suggested that the authors' own descriptive statistics may imply that the estimates were interpreted opposite of what would be correct.

Ault and several other researchers carried out their own analysis of smokeless tobacco, cessation, and harm reduction using data from a large, nationally representative sample of adult smokers (ages 16 to 65) who participated in the National Health and Nutrition Examination Survey (NHANES) wave III (Ault et al., 2004). Building on the theoretical models of Suranovic et al. (1999) and Melkersson (2000), they modeled smoking behavior as a function of income, personal characteristics, stock of smoking habits, and the amount of smokeless tobacco used. Unlike Melkersson, these authors modeled smoking behavior as a probit estimation problem.

The results of the Ault et al. (2004) study indicated that if the average male U.S. male smoker used smokeless tobacco his probability of smoking cessation would increase of 10

percentage points to 14%. The authors went on to calculate the life extension and health savings implications of these results. They determined that a conservative estimate is that 2.16 million life years and $3 billion dollars per year in healthcare costs could be saved if smokers were induced to use smokeless tobacco.

The Ault et al. (2004) findings are impressive, but have not yet been confirmed by application of the probit model to other target populations such as youth, or special subgroups such as dual users of cigarettes and smokeless tobacco. This study applied the probit approach to the Add Health dataset and developed the analyses by inclusion of propensity scores matching and mixed model regression (in addition to the direct application of Ault's identification strategy). Thus this study addresses major problems in the prior literature and examines special samples (i.e. American youth and dual users). Furthermore, this study uses pre-analysis sample correction to remove observable heterogeneity between "treatment" and "control" groups, as has been recommended for the econometric analysis of tobacco (Rosenbaum & Rubin, 1983; Rubin, 1997, 2001). The propensity scores are used to promote credibility – the technique removes heterogeneity without reference to an outcome variable, as opposed to simple application of regression (as in the existing literature on youth smokeless tobacco and cessation).

CHAPTER 3

METHODOLOGY

In addition to the subject matter contributions of this study (i.e., investigation of youth by tobacco usage status) there were important contributions in methodology: the use of empirical methods that were appropriate to investigate if findings in the existing literature are robust to another population and to improve on the problems created by self-selection into smokeless tobacco use (in other words, its endogenous nature). The former contribution reflects the first two primary aims of this study as presented in Chapter 1 (p. 3). This was addressed by employing the use of a large, national school-based survey on a special population – American adolescent males – that has not yet been examined in the smokeless tobacco literature. The latter contribution is addressed by the use of propensity matching scores[5] in the identification strategy described in this chapter, which corresponds to the third primary aim of this study – overall enrichment of the identification strategy's credibility.

Data

The data used in this study were drawn from the National Longitudinal Study of Adolescent Health (Add Health). The Add Health dataset is a nationally representative set of surveys of youth, which was mandated by Congress to collect data that measures the impact of the social environment on adolescent health (Udry, 1998; 2003). Funding for data collection was

[5] After introducing the propensity score, some alternate forms of the prediction equation – such as a mixed model – were also introduced to further examine the validity of this study's results.

provided by the National Institute of Child Health and Human Development (NICHD). The data were deposited to the American Family Data Archive (AFDA) of Sociometrics Corporation for public distribution by J.R. Udry of the Carolina Population Center, University of North Carolina at Chapel Hill (Udry, 1998; 2003).

The Add Health was administered in three waves and includes both in-school and in-home surveys. The sample of Wave I in-school respondents numbered over 90,000 adolescents. Respondents in this sample were students in grades 7 to 12 who took the survey between September 1994 and April 1995. All students who took the in-school survey were eligible for the in-home survey, as were students on the schools' rosters who did not attend the day of the in-school survey administration. Students were stratified by grade and sex and about 17 students from each stratum were randomly chosen to take the in-home survey so that about 200 adolescents were selected from each of the 80 pairs of school. A total core sample of 12,105 respondents was interviewed at home in Wave I. In addition a parent, usually the resident mother, of each adolescent was asked to complete a questionnaire in Wave I.

Wave II provided a second data point by re-administering in-home surveys from April to August of 1996. The sample for Wave II was composed of respondents to the Wave I in-home interview. As in Wave I, in-home survey topics included demographic background, health status, substance abuse (including tobacco use), behavior, nutrition, peer networks, family composition, educational performance and aspirations, sexual and romantic relationships, and criminal activity. Questions about personal attributes, which do not vary over time such as one's race, were not repeated in Wave II. The substance abuse and behavior topics contained several questions related to smoking and additional basic measures of smokeless tobacco use.

Wave III took place between August 2001 and April 2002. By administering follow-up interviews to respondents from the original Wave I sample, the researchers were able to capture data concerning the transition between youth and young adulthood. The core respondents were ages 18 to 26 at the time of Wave III. The total sample in Wave III numbered 15,197 with some respondents not eligible due to overseas military service commitments. The survey questionnaire changed somewhat, with the motivation of capturing variables specific to the transition into young adulthood (e.g., questions were added regarding labor force participation, parenting, higher education and so on). The number of questions related to smoking increased noticeably, although only one useful smokeless tobacco question appeared (the number of days in the past month that one has used smokeless tobacco).

It should be noted that the Add Health datasets are available in two forms – public-use and contractual release. For the purposes of this study the public-use form is acceptable, since personally-identifiable information is not critical to the topic at hand. The public use file has a somewhat smaller sample. In Waves I and II the public use data contains 5,800 variables for 6,504 respondents. Of these, 4,769 cases (73%) have all three forms of data (in-school, in-home, and parent). The Wave III public use version of Add Health contains 4,882 individuals, all of who were members of the 6,504 public use samples from Waves I and II. Approximately 1,734 of these observations were on male tobacco users (Figure 1). While all waves of this dataset included an over-sample of some politically important groups, this study made use of the included sample weights such that the data are reflective of the overall youth population.

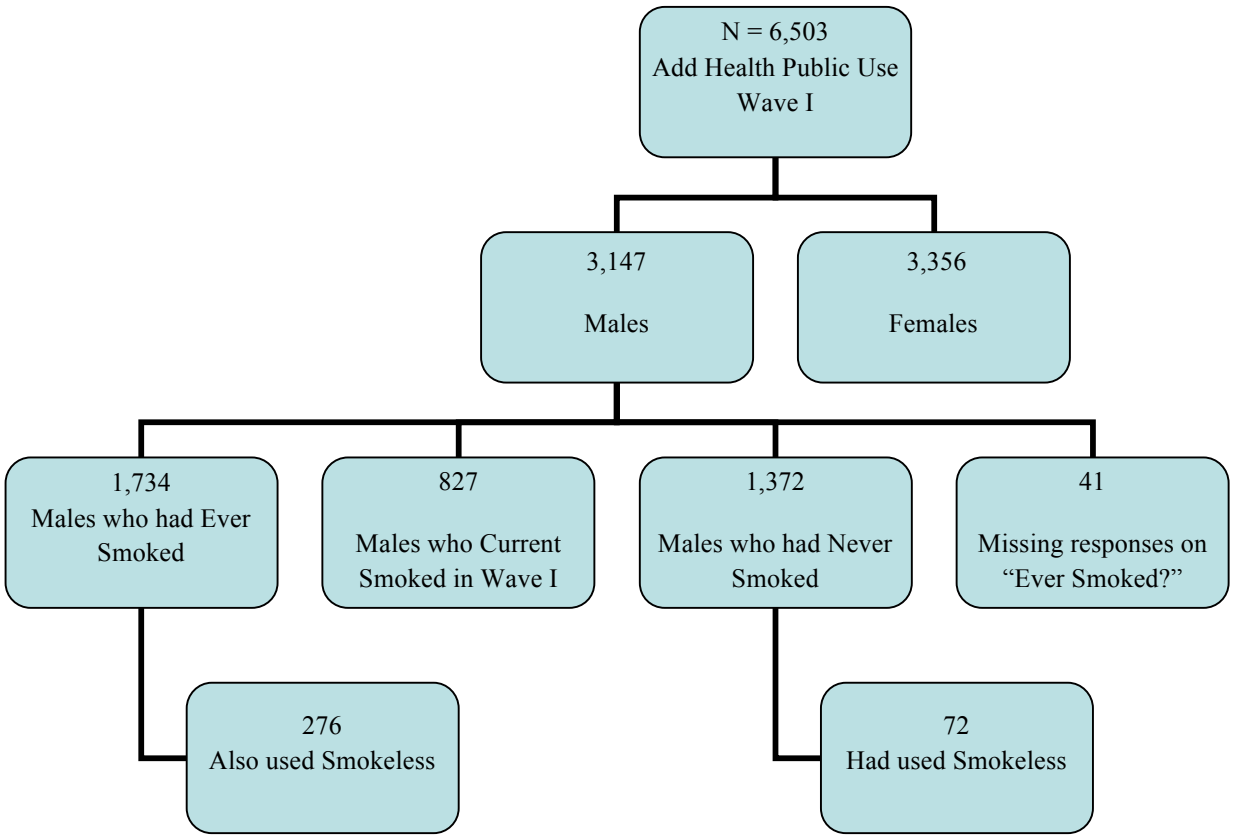

Figure 1: Add Health Public-use Sample Sizes by Group

This particular dataset was chosen for this study for several reasons. The large, representative national sample helps to improve the statistical power over self-collected data and the school-based nature provided a unique mechanism for identifying and surveying adolescents. In this way the data are not entirely unlike other national datasets, such as the National

Longitudinal Survey of Youth. Yet the fact that the Add Health captures data on potentially confounding health and behavior variables so well made this source of data the most attractive. In addition, the broad age range enables this study to focus on a unique, politically important, and hard-to-capture population – adolescents – and to retain the ability to target a more general population by continuing these analyses in future research, since additional Waves of data are still being collected.

Hypotheses

The research hypotheses in this study relate to the research questions posed in Chapter 1. To assess the role of smokeless tobacco in smoking cessation and initiation, this study tests the following hypotheses:

RH1: Adolescent male smokers who dually use smokeless tobacco, or who begin to use it after smoking initiation, will be less likely to discontinue or reduce smoking*, ceteris paribus.*

RH2: Young adult males who use smokeless tobacco are more likely to initiate smoking relative to those who do not, *ceteris paribus*.

These two hypotheses reflect a fundamental purpose of this study – to determine if smokeless tobacco consumption aids, hinders, or does not affect cigarette cessation and initiation. Additional interests of this study included improved data (longitudinal in nature, better questions, higher sample size), improvement of identification (use of propensity-scores) and providing thought to special subgroups, such as dual-users.

The hypotheses relate only to males, because smokeless tobacco consumption is almost an absolutely male phenomenon (Ault et al., 2005). A critical error in a past study by Bask and

Melkersson (2003) was that aggregating data across genders biased the estimated effect of smokeless tobacco usage towards zero.

Theoretical Model

This study follows Suranovic et al. (1999), Melkersson (2000) and Ault et al. (2004) in applying a theoretical model where male smoking behavior is determined by income (Y), a vector of individual characteristics (X), stock of smoking habits (HS), and the frequency of smokeless tobacco use (S). This model was developed from Becker and Murphy's (1988) rational addiction theory, wherein a consumer can be economically rational and addicted to harmful goods. Becker and Murphy (1988) argued that a rationally addicted smoker will choose to quit "cold turkey" when doing so maximizes the present discounted value of current and future utility (see also Keeler, Marciniak and Hu (1999)). Bask and Melkerson (2000) extended this model to include *two* addictive goods, although their empirical application of the model had several flaws (as was discussed in Chapter 2).

The models used in this study, as with many cessation studies, involve dichotomous dependent variables. Thus the probit, which is used in this study and others, is a discreet regression model and estimates probability ratios, which Maddala (1983) details in terms of mathematical derivation. In the primary empirical application of the smoking behavior model this study follows the probit approach of Ault et al. (2004). Ideally one would like to model preference for smoking cigarettes P(C) such that:

$$P(C_i) = \sum_{j=1}^{k} \beta j Xji + \varepsilon i \qquad (1)$$

where preference is assumed to be a linear function of some independent variables X_{ji}, β_j and k are unknown parameters to be estimated, and the error term is independently and identically distributed with a mean of zero and constant variance. Estimation of this model by OLS would require that preferences be measured cardinally however, which is not possible.

Ault et al. (2004) notes that although preferences cannot be estimated directly, the probability of smoking can be estimated. Since preferences are continuous it can be said that there is a critical value, P*, of P(C) – above which a person will choose to smoke. The probability of choosing to smoke is given by the area under a normal distribution whose mean is $\sum_{j=1}^{k} \beta j Xji + \varepsilon i$ to the left of the critical value (P*). This area can be more readily computed if the probability distribution is converted to a standard normal distribution. This requires subtracting the mean and dividing by the standard deviation:

$$\text{Prob(Smoke)} = 1\text{-F}\left[\frac{P^* - \sum \beta j - Xji}{\sigma}\right] \qquad (2)$$

However, since no information on the scale of smoking preferences P(C) is available, one cannot estimate σ. To account for this the equation can be normalized by setting $\sigma^2 = 1$.

Finally, this study invested a greater amount of sample preparation in the design phase relative to the existing empirical research by implementing Rosenbaum and Rubin's (1983) propensity score matching technique. An underlying problem with trying to estimate the effect of smokeless tobacco usage on smoking behavior is that, as a form of treatment (in a quasi-experimental sense), smokeless tobacco usage is not randomly assigned. As such selection bias is

very difficult to eliminate. Propensity score matching is a way to "correct" estimated treatment effects by controlling for factors that confound the effect of a treatment such as smokeless tobacco usage through selection (Becker & Icino, 2002). Furthermore, matching tends to generally provide the benefit of creating tighter confidence intervals around the probability (or odds) ratios; the disadvantages of matching being the costs of time and potential reduction of the sample size (Kleinbaum, 1994, pp. 233-236). Finally, this study expanded the conceptual credibility of American cessation research, by not only estimating a cross-section as in Ault (2004) but by use of longitudinal data as well. This provides additional evidence which will be valuable when consistency is shown between the cross-section and longitudinal data; although the data is limited to only two waves, which minimizes the effectiveness of longitudinal data usage relative to the use of three or move waves of data (Singer & Willet, 2003, pp. 3-10).

Propensity Score Matching

The propensity score is defined as the conditional probability of receiving a treatment given pre-treatment characteristics:

$$P(X) \equiv \text{Prob}\{D = 1 | X\} = E\{D|X\} \quad (3)$$

where D is a dummy variable indicating if treatment was received and X is a vector of pre-treatment characteristics. In other words, a propensity score is simply a predicted probability of assignment to (e.g., membership with) a control or treatment group, conditional on whatever observables are identified. A Harvard statistician and author of the seminal literature on propensity scores, Donald Rubin, identified the lack of outcome variables to be a "tremendous stimulus for honesty" which improved the credibility of his designs thereby boosting public acceptance of his empirical research on tobacco (Rubin, 2001). The latter point is of particular

interest because Rubin has been publically involved in tobacco litigation, yet continues to publish on the topic as an unbiased investigator. [6]

Generally speaking there are *three methods* of using the propensity score to help design observational studies – *regression adjustment, sub-classification, and matching* (Rubin, 2001). Propensity score matching involves predicting the propensity score from a logistic regression based on some observables, then creating one-to-one or one-to-many pairs of control and case matches. Once this has been done, matched-pair techniques may be used to analyze the effect of an independent variable without it being a result of the co-varying factors used to define the propensity scores. Matching is generally the most time consuming, but most credible, of the methods of applying a propensity score.

The other propensity score techniques involve either using the inverse propensity score as a regression variable or its inverse as a regression weight (regression adjustment) or to create five or six subclasses of observations based on equidistant propensity score ranges. Sub-classification was considered as a methodology for this paper, but the sample size seemed to have been just large enough to carry out individual level one-to-many matching. The primary technique in this study is propensity score matching; however, regression adjustment was used with simple linear regression (OLS and mixed model) to determine how replicable the findings were between specifications after all other analyses were computed and reported (Table 10).

---

[6] Similarly, as the author has worked for Philip Morris USA, his intention was to make use of the propensity score to obtain the least bias sample examined in the smokeless tobacco literature to date. Other econometricians involved in the duality of providing credible tobacco empirical analyses whilst having been in the service of that industry include many well-published researchers such as W. Kip Viscusi (Viscusi, 1992).

In these study cases, those individuals who received treatment were matched to controls based on their 5-digit propensity scores. There were more controls than cases and if a case matched multiple controls at the fifth decimal place, then cases were assigned multiple controls. If a case could not be matched based on the first five digits, and then after 5-digit matching had been done, cases were considered for matching based on the first 4-digit propensity score. Theoretically this process would have continued through until matching was being done on only one digit; however, both cessation and initiation datasets were perfectly matched by the fourth decimal place.

CHAPTER 4

RESULTS

The results of this study fall into three broad methodological categories. The first set of results pertains to the direct application of the Ault et al. (2004) model among youth. The second are those resulting from an adaptation of Ault's model to incorporate longitudinal data – where this study not only measured smoking cessation, but initiation as well. Third and finally, results from an alternate identification strategy, namely propensity score matching, are presented with the intention of improving the initial balance and thereby overall credibility. The results from these analyses are presented in Tables 2, 3, 5, 6, 7, 8 and 9. Table 1 provides a list of the variables used to emulate the Ault models as well as definitions and descriptive statistics. Table 4 provides a similar list describing some of the additional variables used in propensity score matching. Table 10 illustrates the stability of the propensity score matched design, facilitating comparison of the past models against simple OLS and mixed-model regressions.

Application of Smokeless Tobacco Cessation Models to Youth Initiation

The first and primary focus of this study was on applying an accepted theoretical model of tobacco consumption behavior (for both smokeless and lit-end products) to youth. The accepted underlying theoretical model, detailed in Chapter 2, describes tobacco behavior (smoking, conditional on smokeless usage) as a function of income, personal characteristics, stock of smoking habits, and the amount of smokeless tobacco used. As in the Ault et al. (2004) study, three probit regression models were specified to examine the impact of smokeless

tobacco on smoking. These models used an *ex post* design such that that cross-sectional data (Wave I of the Add Health, in the case of this study) were used as if they were longitudinal. Specifically, the three models presented in Table 2 are designed to be highly analogous to Ault et al.'s (2004) specifications – which correspond to the Suranovic et al. (1999) and Melkersson (2000) theoretical models of smokeless tobacco use. In this and the Ault et al. study, the first model estimated both a binary smokeless use term and a measure of the degree of smokeless usage (the "weekly use" variable). The second model is the same, save that it only estimates the binary smokeless tobacco term. Likewise, the third model includes the same controls as the first, but omits the "weekly use" term.

Again, these models are binary outcome probit regressions. In general, probit models take the form:

$$\text{Prob}(Y = 1) = \Phi(\beta' x) \tag{4}$$

Here Y is any valid probit response variable (which may be binary, ordered, etc) and where the function Φ(.) indicates the commonly used standard normal distribution (Greene, 2000, pp. 811 – 814). As such, the theoretical model – derived from prior research on smoking, as described in Chapter 3 – gives rise to an empirical model of cessation based on the ex-post cross-sectional data such that:

$$\text{Prob (Quit Smoking} = 1) = \Phi(\hat{\alpha}_0 + \hat{\beta}_{1\ldots n} S + \hat{\chi}_{1\ldots n} X + \hat{\delta}_{1\ldots n} I + \hat{\pi}_{1\ldots n} HS + \varepsilon) \tag{5}$$

where smokeless tobacco measures are represented by vector S, individual characteristics by vector X, (household) income by vector I, and finally, the habit stock by the vector HS. Estimates of this empirical model are shown in Table 2.

The main terms of interest – "Smokeless User" and "Weekly Smokeless Use" – yielded interesting results (Table 2). In the first model, where both terms were estimated, both terms were highly significant but the two had opposing signs – "Smokeless User" was a positive predictor of current smoking whereas "Weekly Smokeless Use" appeared to reduce the probability of smoking. In the second and third models the terms were positive and significant, but had smaller parameter estimates. In addition, an interaction term was calculated but was not reprinted here. The interaction effect was significant at $\alpha = .01$ and had a positive sign of the coefficient, whose parameter estimate was of a magnitude in-between that of the binary-use and Weekly Smokeless Use variables. This demonstrated a consistent statistical story – two significant effects were observed, one that increased the probability of cessation and one that reduced it. Although both are small, the effect that was positively correlated with quitting smoking dominated the overall impact of smokeless tobacco on smoking cessation among youth. The implications of these results are discussed in the next chapter.

Furthermore, it should be noted that the binary nature of both the "Weekly Smokeless Use" and general "Smokeless User" terms in Table 2 result in a special interpretation of the usage effects. Since all smokeless users fall into either "Weekly Use" as a "1" (indicating relatively frequent usage) or a "0" (indicating relatively infrequent usage), the effect of heavy use is captured by the "Weekly Smokeless Use" term and the effect of infrequent use by the overall "Smokeless User" variable.

This said, it is notable in Table 2 that the coefficient of "Smokeless User" is significant and positive whereas that of "Weekly Smokeless User" is significant and

negative. The interpretation of these results, then, implies that infrequent smokeless usage represents a risk to successful cessation whereas heavy smokeless tobacco users are less likely to smoke. This makes theoretical sense, if those who use smokeless tobacco infrequently (1) have passed the 'gateway' into tobacco usage, without being dedicated to a particular product and (2) have not experienced the ultra-high level of nicotine addiction associated with heavy smokeless tobacco usage, which cannot be matched by reasonable cigarette consumption.

Furthermore, consistencies were found with the work of past researches in some cases but not others, in terms of the specific items composing the vectors of personal characteristics and of income (Bask & Melkersson, 2000; Ault et. al, 2004). Youth share many of the significant personal and social correlates – such as the correlation of other risky behaviors (alcohol use, etc.) with smoking; yet family income did not have a meaningful impact on youth in this study – whereas income has been shown to be a significant factor among adults (Bask & Melkersson, 2000; Ault et al., 2004; DeCicca et al., 2008).

Adaptation of the Ault et al. (2004) Cessation Model for Estimating Cessation & Initiation Effects using Longitudinal Data

The next step of this study was to take advantage of the longitudinal nature of the Add Health in terms of using data on the same respondents to measure an observed change in smoking behavior. This has two distinct benefits. The first is that it eliminates some issues associated with using cross-sectional data, such as recall bias. The second is that it enables us to apply essentially the same model used by Ault et al. (2004), but to estimate

the effect of smokeless tobacco use on initiation as well as cessation. While cessation could be estimated *ex-post*, initiation cannot be – it requires at least two waves of repeated measures. This may be a reason that smoking initiation has largely been neglected by the smokeless tobacco literature.[7]

Given the multiple Waves of Add Health data, fortunately, this study was able to estimate not only the empirical model of cessation by longitudinal probit analysis:

$$\text{Prob (Wave 2 Smoker = 1 | Wave 1 Current Smoker = 1)} = \Phi(\hat{\alpha}_0 + \hat{\beta}_{1\ldots n} S + \hat{\chi}_{1\ldots n} X + \hat{\delta}_{1\ldots n} I + \hat{\pi}_{1\ldots n} HS + \varepsilon) \quad (6)$$

but also initiation:

$$\text{Prob (Wave 2 Smoker = 1 | Wave 1 Current Smoker = 0)} = \Phi(\hat{\alpha}_0 + \hat{\beta}_{1\ldots n} S + \hat{\chi}_{1\ldots n} X + \hat{\delta}_{1\ldots n} I + \hat{\pi}_{1\ldots n} HS + \varepsilon) \quad (7)$$

The estimates for both cessation (6) and initiation (7) longitudinal probit models are presented in Table 3.

For cessation, the universe was restricted to not only those males who had ever smoked at least one full cigarette, but further to those who had smoked within a month of the interview in Wave I. The cessation outcome variable was a binary term ("stopped

---

[7] As special thanks is extended to Joseph Sabia for his suggestion to examine initiation effects.

smoking"), which was defined to be equal to one if someone in this restricted sample had not smoked within the past month at the time of interview in Wave II.

For initiation, the universe was restricted to those males who had never smoked as of Wave I. A binary outcome variable ("started smoking") was defined to equal one if a non-smoker from Wave I indicated they had ever used a cigarette by Wave II. The smokeless tobacco use terms correspond to smokeless tobacco consumption in Wave I.

When interpreting the results it is important to keep in mind the probability being estimated. In the "cessation" column of Table 3, the probability of stopping smoking was modeled such that a successful "quit" would be equal to one. As such, the terms with significant and positive coefficients should be interpreted as predictors of successful cessation (i.e., positive terms are "good"). In the "initiation" column it is the probability of picking up smoking that is being modeled. As such the implications are opposite; significant and positive terms are predicting an unwanted event (i.e., positive terms are "bad").

Some interesting comparisons can be made between the cessation and initiation results within this table (Table 3). Most importantly, there seemed to be a conflicting story that, overall, smokeless tobacco was not performing well as a tool to prevent initiation ("Smokeless User" was positive and significant) although both "Smokeless User" and the "Weekly Smokeless Use" variables were positively and significantly associated with cessation. In addition, the fixed effects for age, grade, and other elements of personal characteristics seemed to behave differently between groups – surprisingly age and grade were more often successful predictors of cessation than initiation, despite the habit stock variable present in cessation which one might have expected to have multicolinnearity

issues with age and grade. Still, while these results are interesting, the longitudinal probit suffers from the same design inadequacies as in the ex-post analysis presented prior to it. To address this issue, this study returned to the development of propensity score analyses.

Propensity Score Approaches

Here the same data were used to estimate the same effects, but in the context of a very different design. The motivations for using the propensity score methodology were discussed in Chapter 3; also in that chapter, the calculation of the propensity score was explained (see Equation 3). It is important to reiterate here, however, that these propensity score models were to better the design of smokeless tobacco effect studies and to "promote honesty". In the spirit of creating more robust empirical models and making use of relevant data unique to the Add Health, the models were not constrained to the controls identified by Ault et al. (2004) or any single past study. A number of additional controls, such as log(weight) and log(height), were adapted from Rubin (2001) and others; for instance, various questions addressing "attitudes" and feelings towards parents were specified *a priori*. Table 4 details the specific additional observables included in the propensity score models[8]. These were chosen on the basis of theoretical differences between samples and correlation as per Klien (1995) and Ault et al. (2005) and a cultural based exploration of adolescent tobacco use by Piko, Luszczynska, Gibbons & Tekozel (2005). In addition, the sample means were observed for variation, with interest to characteristics that were shown to vary across age differences (which were significant between the "control" and "treatment" groups of respondents – terms used colloquially in this

[8] Some effects were applied to only one of the two models (initiation or cessation), depending on the amount of variance in the data.

quasi-experimental design) according to a well-published 2005 examination of state-level data (Goel & Nelson, 2005).

The propensity score matching was estimated by use of these variables with Add Health data and an adaptation of the SAS Greedy 5 → 1Digit Match Macro published by Parsons (2001). The results of matching can be observed in the improvements to balance between the initial control and treatment groups for smoking cessation (Table 5) or smoking initiation (Table 6) with the post-matching balance of these two sets of treatment and control groups (Table 7).

It should be noted that all differences between the quasi-experimental "control" and "treatment groups" could not be completely eliminated in the models of either smoking cessation or smoking initiation. Balance was, however, greatly improved by matching on the propensity scores. Approximately 92%[9] of the pre-matching cessation and 83% of the pre-matching initiation terms were significantly different. After matching only 14% and 26% of cessation and initiation observables (respectively) were significantly different between control and treatment groups.

After matching reduced selection bias, the initial sets of controls and cases were analyzed by least-squares regression analysis. The SAS procedure PROC SURVEYREG was used to weight variables by their appropriate statistical weight (Sociometrics' 'gswgt2' variable) and

---

[9] Proportions are presented as percents, since the denominator term – the number of variables in the analysis – differ between pre- and post-matching. This is a result of computational differences between the PROC FREQ and PROC SURVEYFREQ commands in SAS v.9.1. The latter procedure was used to stratify analyses by the values of their matching variables, which broke some ordered terms into binary variables indicating their levels (to be analyzed by the Rao-Scott Chi-square test).

stratified by their match-identification numbers, which were created by the Greedy[10] matching method for propensity scores (Parsons, 2001). Table 8 presents the result for smoking cessation and Table 9 addresses smoking initiation. The latter had less identifying variation and thus could support fewer effects; however, variables were retained to describe each conceptual element of the model.

Whereas these matched regression results in Tables 8 and 9 represent a straightforward application of the propensity-score methodology, there existed a large amount of confounding correlation between the propensity scores and regression factors retained in the models. Results were (not surprisingly) insignificant and likely represent a confounding amount of bias. To combat this, this study made use of the propensity-score variable as a regression term – with simple least-squares and maximum-likelihood models examined with the "started smoking" ballot item (initiation) as the outcome variable and only treatment (control v. case group membership) as an additional explanatory factor. In the case of least-squares regression, this model is represented by the simple OLS equation:

$$\hat{Y} = \hat{\alpha}_0 + \hat{\beta}_1 S + \hat{\kappa} P + \varepsilon \qquad (8)$$

Here, the response variable Y represented the initiation and (in a separate instance) cessation outcomes. Results are found in Tables 10 and 11. P represents the propensity-score, as it is now treated as a regression parameter (Rubin, 2001). These results were compared side-by-side for

[10] "Greedy" is a term which identifies the particular selection mechanism in Parsons' article and accompanying SAS macro (Parsons, 2001).

consistency with a mixed model specification that made use of maximum-likelihood identification and improved on past models by including both fixed and random effects:

$$\hat{Y} = X\beta + Zu + \varepsilon \quad (9)$$

Where X and Z are respectively *n* x *p* and *n* x *q* incidence matrices (X is also called the design matrix), and e is the *n*x1 column vector of residual deviations assumed to be distributed independently of the random individual effects. All of the elements of the incidence matrices are equal to 0 or 1, depending upon whether the relevant effect contributes to the individual's smoking status (Verbeke & Molenberghs, 2000). The findings in Tables 10 and 11 demonstrate a surprising level of consistency within the models displayed.

This provided additional evidence to support both the consistency of these findings and the potential usefulness of propensity scores in promoting parsimony. Importantly, a final piece of supporting evidence provided in the SAS PROC MIXED output was a comparison of adjusted least-squares means between "treatment" and "control" for both cessation and initiation. In both cessation and initiation the means were highly significantly different from zero (< .0001).

The initiation treatment and control groups were highly significantly different from one another (the alpha level being less than .01), although for cessation the treatment groups were only marginally significantly different (an adjusted p-value of 0.059). The minimal confidence in difference between least-square means for cessation implies that it is likely that both the balance of groups (design) and regression estimates may be more precise under the greater size and

robustness of restricted use data – an issue ripe for future investigators to address in the literature.

Table 1

Variables of Key Interest in Main (Probit) Models

Variable Name	Mean (Weighted Mean)	Definition
Current Smoker	.476 (.483)	1 if respondent smoked one or more days in the past month; 0 otherwise
Current Smoker (Alternate)	.344 (.361)	1 if respondent smoked five or more days in the past month; 0 otherwise
Smokeless User	.159 (.161)	1 if respondent used smokeless tobacco one or more days in the past month; 0 otherwise
Smokeless User (Alternate)	.091 (.094)	1 if respondent used smokeless tobacco five or more days in the past month; 0 otherwise
Degree of Smokeless Use	1.94 (2.1)	Number of days in the past month that a respondent used smokeless tobacco
Income	47.49 (46.62)	Family income, reported by a parent, in thousands of dollars
7^{th} Grade F.E.	Reference Group	1 if respondent was in the 7^{th} grade; 0 otherwise
8^{th} Grade F.E.	.137 (.146)	1 if respondent was in the 8^{th} grade; 0 otherwise
9^{th} Grade F.E.	.168 (.161)	1 if respondent was in the 9^{th} grade; 0 otherwise
10^{th} Grade F.E.	.184 (.176)	1 if respondent was in the 10^{th} grade; 0 otherwise

Variable Name	Mean (Weighted Mean)	Definition
11th Grade F.E.	.213 (.178)	1 if respondent was in the 11th grade; 0 otherwise
12th Grade F.E.	.174 (.196)	1 if respondent was in the 12th grade; 0 otherwise
Urban Area F.E.	Reference Group	1 if a respondent lived in a mainly urban area; 0 otherwise
Rural Area F.E.	.284	1 if a respondent lived in a mainly rural area; 0 otherwise
Suburban Area F.E.	.376	1 if a respondent lived in a mainly suburban area; 0 otherwise
"Other" Area F.E.	.007	1 if a respondent lived in an "other" area category; 0 otherwise
Student Employment	9.47 (9.91)	Hours of work (by student) per week; does not include Summer-only jobs
Alcohol Never	(Reference Group)	1 if the respondent reported that they never or almost never drank alcohol; 0 otherwise
Alcohol 1 or 2 per Year	.196 (.197)	1 if the respondent reported "1 or 2 drinks per year" as the alcohol consumption level which best described their behavior; 0 otherwise
Alcohol 1 per Month	.153 (.153)	1 if the respondent reported one drink per month as the alcohol consumption level which best described their

Variable Name	Mean (Weighted Mean)	Definition
		behavior; 0 otherwise
Alcohol 2-3 per Month	.119 (.115)	1 if the respondent reported 2-3 drinks per month as the alcohol consumption level which best described their behavior; 0 otherwise
Alcohol 1-2 per Week	.116 (.118)	1 if the respondent reported 1-2 drinks per week as the alcohol consumption level which best described their behavior; 0 otherwise
Alcohol 3-5 per Week	.053 (.059)	1 if the respondent reported 3-5 drinks per week as the alcohol consumption level which best described their behavior; 0 otherwise
Alcohol Every Day	.018 (.020)	1 if the respondent reported "every day" as the alcohol consumption level which best described their behavior; 0 otherwise
Habit Stock	3.96 (3.45)	Stock of habit measured in years of smoking
HS*Grade	Means vary – there are six interaction terms corresponding to the six grade levels	Interaction of habit stock with grade-level fixed effects (7^{th} grade is the reference group)

Note: Alternate smoking and smokeless tobacco measures were estimated, however results did not vary significantly. As such, the results for the alternate measures are note reproduced here.

Table 2

Ex-post Estimates of Quitting Smoking Primary Probit Models (Pr > χ^2 in parentheses)

Vector	Variable	Model 1	Model 2	Model 3
N/A	Intercept	-2.7029	-2.7129	-2.6223
		(<.0001)	(<.0001)	(<.0001)
Smokeless Tobacco Use	Smokeless User	0.6466	N/A	0.3877
		(<.0001)		(<.0001)
	Weekly Smokeless Use	-0.4081	0.1908	N/A
		(<.0001)	(<.0001)	
Income	Family Income	-0.00066	-0.0055	-0.0006
		(.0098)	(0.0474)	(0.0274)
Individual Characteristics	8th Grade F.E.	-0.2354	-0.2489	-0.2759
		(<.0001)	(<.0001)	(.0003)
	9th Grade F.E.	-0.3074	-0.3446	-0.3405
		(.0008)	(<.0001)	(<.0001)
	10th Grade F.E.	-0.5846	-0.6156	-0.6499
		(<.0001)	(<.0001)	(<.0001)
	11th Grade F.E.	-0.4028	-0.4377	-0.4329
		(<.0001)	(<.0001)	(<.0001)
	12th Grade F.E.	-0.6111	-0.6657	-0.6260
		(<.0001)	(<.0001)	(<.0001)
	Rural Area F.E.	0.2118	0.2330	0.2027

Vector	Variable	Model 1	Model 2	Model 3
		(<.0001)	(<.0001)	(<.0001)
	Suburban Area F.E.	0.0708	0.0707	0.0709
		(.3746)	(.3611)	(.3841)
	"Other" Area F.E.	-0.338	-0.6156	-0.4329
		(.71)	(.5459)	(.3060)
	Student Employment	0.0056	0.0055	0.0055
		(<.0001)	(<.0001)	(<.0001)
	Alcohol 1 or 2 per Year	0.1977	0.2025	0.2072
		(<.0001)	(<.0001)	(<.0001)
	Alcohol 1 per Month	0.4591	0.4695	0.4519
		(<.0001)	(<.0001)	(<.0001)
	Alcohol 2-3 per Month	0.7288	0.7274	0.7399
		(<.0001)	(<.0001)	(<.0001)
	Alcohol 1-2 per Week	0.6625	0.6554	0.6356
		(<.0001)	(<.0001)	(<.0001)
	Alcohol 3-5 per Week	0.8510	0.8998	0.8717
		(<.0001)	(<.0001)	(<.0001)
	Alcohol Every Day	0.9893	1.0070	0.9814
		(<.0001)	(<.0001)	(<.0001)
Smoking – Stock of Habit	Age First Smoked	0.2011	0.2054	0.1969
		(<.0001)	(<.0001)	(<.0001)
	Habit Stock	0.0359	0.0367	0.0273
		(.2625)	(.2707)	(.4090)
	HS * 8^{th}	0.0517	0.0542	0.0595

Vector	Variable	Model 1	Model 2	Model 3
	Grade	(.0019)	(.0004)	(<.0001)
	HS * 9th Grade	0.0655	0.0739	0.0729
		(.0005)	(<.0001)	(<.0001)
	HS * 10th Grade	0.0723	0.0758	0.0797
		(<.0001)	(<.0001)	(<.0001)
	HS * 11th Grade	0.0430	0.0460	0.0476
		(.0088)	(.0042)	(<.0001)
	HS * 12th Grade	0.0356	0.0416	0.0410
		(<.0001)	(.0233)	(.0238)
Respondents		1,734	1,734	1,734
Observations		8,670	8,670	8,670

Note: The numbers of responses are fewer than the number of observations as a result of retaining respondents who may not have answered every question, by approximating missing data in SAS with Bayesian simulation (Jackman, 2000; Schafer & Graham, 2002). There are 8,670 observations with an N = 1,734 in Table 2 for example, because the Bayesian simulation makes 5 copies of the spreadsheet/dataset and estimates the missing cells 5 times.

Table 3

Longitudinal Estimates: Cessation and Initiation (Pr > χ^2 in parentheses)

Vector	Variable	Cessation	Initiation
N/A	Intercept	-1.6643	-0.9006
		(<.0001)	(<.0001)
Smokeless Tobacco Use	Smokeless User	0.0804	0.5034
		(<.0001)	(<.0001)
Smokeless Tobacco Use	Weekly Smokeless Use	0.1957	-0.1158
		(<.0001)	(0.5147)
Income	Family Income	0.000917	-0.0003
		(<.0001)	(0.395)
Individual Characteristics	8th Grade F.E.	-0.3605	0.1935
		(<.0001)	(0.0015)
	9th Grade F.E.	-0.0147	-0.0589
		(<.0001)	(0.352)
	10th Grade F.E.	-0.0340	-0.101
		(<.0001)	(0.1202)
	11th Grade F.E.	-0.4049	-0.2566
		(<.0001)	(0.0004)
	12th Grade F.E.	-0.3261	-0.6126
		(<.0001)	(0.0108)
	Rural Area F.E.	-0.1932	-0.0147
		(<.0001)	(0.7797)
	Suburban Area F.E.	-0.3166	-0.1321
		(<.0001)	(0.0074)

Vector	Variable	Cessation	Initiation
	"Other" Area F.E.	-4.1861	-0.4033
		0.0002	(0.2474)
	Student Employment	0.00679	0.00036
		(<.0001)	(0.8778)
	Alcohol 1 or 2 per Year	-0.0946	0.366
		0.1760	(<.0001)
	Alcohol 1 per Month	-0.4305	0.0302
		(<.0001)	(0.782)
	Alcohol 2-3 per Month	-0.7488	0.9978
		(<.0001)	(<.0001)
	Alcohol 1-2 per Week	-0.00208	0.2808
		(<.0001)	(0.03)
	Alcohol 3-5 per Week	-0.1416	1.4179
		(<.0001)	(<.0001)
	Alcohol Every Day	-1.4793	1.3366
		(<.0001)	(<.0001)
Smoking – Stock of Habit	Age First Smoked	0.0756	N/A
		(<.0001)	
	Habit Stock	0.0853	N/A
		(<.0001)	
	HS * 8th Grade	0.1302	N/A
		(<.0001)	
	HS * 9th Grade	-0.0817	N/A
		(<.0001)	

Vector	Variable	Cessation	Initiation
	HS * 10th Grade	-0.0814	N/A
		(<.0001)	
	HS * 11th Grade	-0.1414	N/A
		(<.0001)	
	HS * 12th Grade	-0.00991	N/A
		(<.0001)	
	Fit – AIC[a]	2431.401	4823.743
	N Respondents	574	1,021
	N Observations[b]	2,870	5,105

[a]This is the Akaike Information Criterion. It is calculated as AIC = -2 Log L + 2((k-1) + s), where k is the number of levels of the outcome variable and s is the number of predictors in the model. The model with the smallest AIC is considered the best.

[b]The number of responses is fewer than the number of observations as a result of retaining respondents who may not have answered every question, by approximating missing data in SAS with Bayesian simulation (Jackman, 2000; Schafer & Graham, 2002).

Table 4.

Additional Variables Used to Define Propensity Scores

Type	Additional Variable	Description	Scale
Attitudes	Assertive	Self-identification as being assertive	Likert (ordinal)
Attitudes	Drives while High on Drugs	Substance and responsibility measure	Binary
Attitudes	Embarrassed	Embarrassed about tobacco, alcohol, and drugs (interviewer assessed)	Binary
Attitudes	Emotional	Self-identification as being emotional	Likert (ordinal)
Attitudes	Expectations 21 Years	Self-assessed probability of survival	Likert (ordinal)
Attitudes	Expectations 35 Years	Self-assessed probability of survival	Likert (ordinal)
Attitudes	Expectations Income	Self-assessed probability of wealth	Likert (ordinal)
Attitudes	Birth Control	Respondents consider birth control "too bothersome"	Likert (ordinal)
Behavior	Baseball Participation	Active or intended team participation	Binary
Behavior	Basketball Participation	Active or intended team participation	Binary
Behavior	Expelled	Respondent's expulsion history	Binary
Behavior	Football	Active or intended team participation	Binary
Behavior	Hangs Out	Times/week hangs out with friends	Continuous
Behavior	No Extra Curricular	Respondent does not participate in any club or team; does not intend to	Binary
Behavior	Seatbelt	Caution; frequency of seatbelt use	Likert (ordinal)

Type	Additional Variable	Description	Scale
Demographic	Birth Month	Respondent's month of birth	Ordinal
Demographic	Birth Year	Respondent's year of birth	Ordinal
Demographic	Ethnicity: African American	Respondent ethnicity	Binary
Demographic	Ethnicity: Asian	Respondent ethnicity	Binary
Demographic	Ethnicity: Caucasian	Respondent ethnicity	Binary
Demographic	Ethnicity: Native American	Respondent ethnicity	Binary
Demographic	Ethnicity: Other	Respondent ethnicity	Binary
Demographic	Ever Married	Marriage of the respondent (youth)	Binary
Demographic	Household Size	Number of individuals in the HH	Continuous
Demographic	Log Height	Log of height	Continuous
Demographic	Log Weight	Log of weight	Continuous
Demographic	Parents Religion	Vector of variables indicating parent's religion	Binary
Demographic	Religion	Vector of variables indicating the respondent's religion	Binary
Demographic	School Enrollment	Indicates current enrollment in school	Binary
Demographic	Sexual Assault	Indicates sexual victimization	Binary
Family	Loving Father	Respondent has a loving Father	Likert (ordinal)
Family	Loving Mother	Respondent has a loving Mother	Likert (ordinal)
Health	HIV/AIDS	Health measure; has HIV or AIDS	Binary
Intelligence	AH-PVT	Peabody Picture Vocabulary Test - rough IQ measure	Pseudo-continuous

Note. All variables used in the application of Ault et al. (2004) are described in Table 1 and as such are not repeated here. The variables in both tables, other than the outcome variables, were used in calculating propensity scores.

Table 5

Cessation: Key Observables in Control and Treatment Assignment Before Matching

	Control N (%)	Treatment N (%)	p-value
Total Subjects	**705**	**200**	
African American	115 (16.31%)	13 (6.50%)	< .0001
Age	N/A -	Continuous	< .0001
AH-PVT	N/A -	Continuous	0.0025
Drives while High on Drugs	N/A -	Categorical	< .0001
Embarrassed	N/A -	Categorical	0.1027
Expectations 35 Years	N/A -	Categorical	0.0043
Expectations Income	N/A -	Categorical	0.1004
Expelled	88 (12.48)	13 (6.50%)	0.0202
Hangs Out	N/A -	Categorical	< .0001
Log Height	N/A -	Continuous	< .0001
Log Weight	N/A -	Continuous	< .0001
School Enrollment	683 (96.88%)	200 (100.00%)	0.0003
8^{th} Grade F.E.	73 (10.35%)	37 (18.50%)	< .0001
9^{th} Grade F.E.	112 (15.89%)	46 (23.00%)	< .0001

11th Grade F.E.	154 (21.18%)	60 (30.00%)	< .0001
12th Grade F.E.	158 (22.41%)	4 (2.00%)	< .0001
Fixed Drinking 1 /mo	115 (16.31%)	40 (20.00%)	0.0063
Fixed Drinking 1 or 2 /year	118 (16.74%)	23 (11.5%)	< .0001
Fixed Drinking 1 to 2 /wk	109 (15.46)	37 (5.25%)	0.0211
Fixed Drinking 2 to 3 /mo	58 (8.23%)	10 (5.00%)	0.0006
Fixed Drinking 3 to 5 /wk	22 (3.12%)	4 (2.00%)	0.0612
Attitudes – Birth Control	N/A -	Categorical	< .0001
Rural Area F.E.	201 (28.51)	91 (45.50%)	< .0001
Suburban Area F.E.	265 (37.59%)	59 (8.37%)	< .0001
Virgin	233 (33.05%)	96 (48.00%)	< .0001

Note. 10th Grade F.E. were omitted from this table due to lack of statistical significance.

Table 6

Initiation: Key Observables in Control and Treatment Assignment Before Matching

	Control N (%)	Treatment N (%)	p-value
Total Subjects	**265**	**265**	
African American	55 (10.38%)	70 (13.21%)	< .0001
Age	N/A -	Continuous	< .0001
AH-PVT	N/A -	Continuous	< .0001
Asian	16 (6.04%)	5 (1.89%)	0.0143
Assertive	N/A -	Categorical	0.0026
Attitudes – Birth Control	N/A -	Categorical	0.0462
Baseball	64 (24.15%)	80 (30.19%)	0.1182
Drives while High on Drugs	30 (11.32%)	40 (15.09%)	0.1995
Emotional	N/A -	Categorical	< .0001
Expectations 35 Years	N/A -	Categorical	0.0043
Expectations Income	N/A -	Categorical	0.1004
Expelled	88 (12.48)	13 (6.50%)	0.0202
Hangs Out	N/A -	Categorical	< .0001
Log Height	N/A -	Continuous	< .0001
Log Weight	N/A -	Continuous	< .0001
School Enrollment	265 (100.00%)	255 (96.23%)	0.0014
8^{th} Grade F.E.	40 (15.09%)	30 (11.32%)	0.1995

	Control N (%)	Treatment N (%)	p-value
9th Grade F.E.	35 (13.21)	60 (22.64%)	0.0046
10th Grade F.E.	31 (11.70%)	50 (18.87%)	0.0218
11th Grade F.E.	41 (15.47%)	30 (11.32%)	0.1607
Fixed Drinking 1 /mo	21 (7.92%)	10 (3.77%)	0.0417
Fixed Drinking 1 or 2 /year	24 (9.06%)	40 (15.09%)	0.0329
Fixed Drinking 1 to 2 /wk	5 (1.89%)	15 (5.66%)	0.0226
Fixed Drinking 2 to 3 /mo	2 (0.75%)	10 (3.77%)	0.0195
Fixed Drinking 3 to 5 /wk	2 (0.75%)	0 (0.00%)	0.2495
Fixed Drinking Everyday	0 (0.00%)	10 (3.77%)	0.0014
Attitudes – Birth Control	N/A -	Categorical	0.0462
Rural Area F.E.	66 (24.91%)	110 (41.51%)	< .0001
Suburban Area F.E.	127 (47.92%)	65 (24.53%)	< .0001
Virgin	43 (16.23%)	85 (32.08%)	< .0001

Table 7

Cessation: Improvement in Balance of Initial Groups from Propensity Score Matching

Effect	Cessation χ^2 p-value	Initiation χ^2 p-value
Age First Smoked	0.7819	N/A
AHPVT	0.5541	0.8103
Asian	0.9791	0.1003
Assertive	0.5079	0.9297
Attitudes BC	0.1807	0.1470
Baseball	0.0596	0.9980
Basketball	0.6346	0.9838
Birth Month	0.2540	0.0021
Birth Year	0.3990	<.0001
Black	1.0000	0.0180
Current Age	0.3673	Birth year used[a]
Diabetes	1.0000	No variation
Drive High	0.4530	0.7062
Eighth Grade	0.5220	0.0003
Eleventh Grade	0.1666	0.1528
Emotional	0.7080	0.0715

[a] *Note*: Birth year was reported for all respondents in Wave I. Given that a sizable fraction of initiation subjects had missing data for their current age, it seemed best to simply make use of birth year (and grade-level fixed effects) rather than to impute age or delete many observations.

Effect	Cessation χ^2 p-value	Initiation χ^2 p-value
Expectations 21	0.4286	0.2467
Expectations 35	0.8531	0.0043
Expectations Income	0.4444	0.0887
Expelled	0.0536	0.4350
Fixed Drinking 1 /mo	0.0824	0.0314
Fixed Drinking 1 or 2 /year	0.0699	0.3524
Fixed Drinking 1 to 2 /week	1.0000	0.3342
Fixed Drinking 2 to 3 /mo	1.0000	0.8975
Fixed Drinking 3 to 5 /week	0.0837	0.7983
Fixed Drinking Everyday	0.1523	0.2527
Fixed Peer One	0.2592	0.1407
Fixed Peer Three	0.1283	0.0045
Fixed Peer Two	0.5382	0.1821
Fixed Area Rural	1.0000	0.0566
Fixed Area Suburban	0.6324	0.6253
Football	0.3311	0.4017
Hangs Out	0.6761	0.0353
Height in Feet	0.0176	0.6550
HIV AIDS	0.9782	No variation
Household Size	0.7314	0.0476
Income	0.6525	0.1572
Log (Height)	0.0401	0.2670
Log (Weight)	0.1577	0.1644

Effect	Cessation χ^2 p-value	Initiation χ^2 p-value
Loving Dad	0.8945	0.1708
Loving Mom	0.2056	0.1941
Native	0.1131	0.0010
Ninth Grade	0.3423	< .0001
No Extra Curricular Activities	0.5858	0.6990
Non-summer Employment	0.6130	0.1385
Other Ethnicity	0.8264	No variation (no 'other')
School Enrollment	0.7954	No variation (all enrolled)
Seatbelt	0.0630	0.8102
Seventh Grade	0.6238	0.0705
Shy	0.5295	0.8815
Tenth Grade	0.0075	0.0202
Twelfth Grade	1.0000	Exactly equal proportions
U.S. Born	0.1523	0.8972
Virgin	0.5748	< .0001
Weight Change	0.9561	0.0667
White	0.7248	0.0499

Table 8

Cessation: Weighted Probit Regression Matched on Propensity Scores

Effect	Estimate	Standard Error	p-value
Intercept	1.1786	9.2304	0.8986
Age First Smoked	-0.0056	0.0303	0.8528
AHPVT	-0.0062	0.0071	0.3860
Asian	-0.4857	0.4378	0.2695
Assertive 1	1.1068	0.7279	0.1310
Assertive 2	0.8197	0.6651	0.2202
Assertive 3	0.9322	0.7510	0.2169
Assertive 4	0.8756	0.7133	0.2220
Assertive 5	Reference	Level	
Attitudes BC 1	-0.0322	0.2619	0.9024
Attitudes BC 2	0.0351	0.2333	0.8808
Attitudes BC 3	0.1909	0.2362	0.4204
Attitudes BC 4	0.3401	0.2956	0.2522
Attitudes BC 5	Reference	Level	
Baseball	-0.0250	0.1777	0.8882

Effect	Estimate	Standard Error	p-value
Basketball	0.1488	0.1599	0.3537
Birth Month April	-0.1216	0.3084	0.6941
Birth Month August	-0.2036	0.2874	0.4799
Birth Month December	-0.0182	0.2816	0.9485
Birth Month February	-0.2355	0.6009	0.6958
Birth Month January	-0.1595	0.2516	0.5274
Birth Month July	-0.1141	0.4011	0.7765
Birth Month June	0.0757	0.3353	0.8219
Birth Month March	-0.1295	0.2700	0.6322
Birth Month May	-0.2911	0.2665	0.2769
Birth Month November	-0.2433	0.2871	0.3984
Birth Month October	0.0087	0.3139	0.9779
Birth Month September	Reference	Level	
Birth Year 1976	-0.1580	0.7727	0.8383
Birth Year 1977	0.1252	0.6322	0.8433
Birth Year 1978	0.0998	0.5751	0.8625
Birth Year 1979	0.1465	0.5270	0.7815
Birth Year 1980	0.1035	0.4720	0.8267

Effect	Estimate	Standard Error	p-value
Birth Year 1981	0.1445	0.4012	0.7193
Birth Year 1982	Reference	Level	
Black	-0.0667	0.4709	0.8876
Diabetes	0.3159	0.6109	0.6060
Drives High	-0.0089	0.1303	0.9458
Eleventh Grade	0.0054	0.4909	0.9912
Emotional 1	0.2170	0.3941	0.5830
Emotional 2	0.3333	0.4257	0.4351
Emotional 3	0.4050	0.3897	0.3008
Emotional 4	0.3863	0.4300	0.3707
Emotional 5	Reference	Level	
Expectations 21 - 0	-0.1574	0.4271	0.7130
Expectations 21 - 1	-0.2205	0.4678	0.6382
Expectations 21 - 2	-0.6742	0.9697	0.4882
Expectations 21 - 3	-0.2383	0.4662	0.6102
Expectations 21 - 4	-0.3046	0.7469	0.6842
Expectations 21 - 5	Reference	Level	
Expectations 35 - 1	-0.2141	0.4475	0.6332

Effect	Estimate	Standard Error	p-value
Expectations 35 - 2	-0.0304	0.4053	0.9403
Expectations 35 - 3	-0.0781	0.2675	0.7708
Expectations 35 - 4	-0.1735	0.2088	0.4077
Expectations 35 - 5	Reference	Level	
Expectations Income 1	0.1922	0.1999	0.3380
Expectations Income 2	-0.0740	0.3312	0.8236
Expectations Income 3	-0.0064	0.1700	0.9701
Expectations Income 4	0.0584	0.1725	0.7355
Expectations Income 5	Reference	Level	
Expelled	-0.0256	0.3417	0.9404
Fixed Drinking 1 /Month	0.1645	0.2171	0.4502
Fixed Drinking1-2/Year	0.1200	0.2511	0.6336
Fixed Drinking1 to 2/wk	0.3494	0.2514	0.1672
Fixed Drinking 2 to 3/mo	0.1776	0.2033	0.3842
Fixed Drinking 3 to 5/wk	0.3216	0.4098	0.4342
Fixed Drinking Everyday	0.0989	0.5392	0.8547
Fixed Area Rural	-0.0216	0.1742	0.9017
Fixed Area Suburban	-0.1516	0.2160	0.4842

Effect	Estimate	Standard Error	p-value
Football	-0.2104	0.1726	0.2252
Hang Out 0	-0.4215	0.3002	0.1629
Hang Out 1	-0.1724	0.2296	0.4542
Hang Out 2	0.0390	0.1554	0.8022
Hang Out 3	Reference	Level	
Height Feet	0.0992	0.1679	0.5556
HIV/AIDS	-0.2136	0.8495	0.8019
Household Size	-0.0751	0.0462	0.1069
Income	-0.0004	0.0011	0.7150
Log (Height)	-0.4652	2.4719	0.8511
Log (Weight)	-0.0077	0.3939	0.9845
Loving Dad 1	-0.1340	0.4898	0.9130
Loving Dad 2	-0.0121	0.5426	0.1153
Loving Dad 3	0.0306	0.5831	0.6284
Loving Dad 4	0.1001	0.6730	0.8321
Loving Dad 5	Reference	Level	
Loving Mom 1	0.1468	0.8077	0.8561
Loving Mom 2	0.1385	0.8465	0.8672

Effect	Estimate	Standard Error	p-value
Loving Mom 3	0.1293	0.9260	0.8892
Loving Mom 4	-0.1527	1.0402	0.9933
Loving Mom 5	Reference	Level	
Native	0.0467	0.3973	0.9066
Ninth Grade	0.0173	0.3849	0.9642
No Extra Curricular	-0.0311	0.1700	0.8550
Non-summer Employment	-0.0012	0.0052	0.8170
Other	0.0077	0.6106	0.9899
School Enrollment	-0.1692	0.7226	0.8153
Seatbelt 1	-0.2282	0.4239	0.5913
Seatbelt 2	0.2921	0.3401	0.3921
Seatbelt 3	-0.2888	0.1878	0.1267
Seatbelt 4	-0.0728	0.1873	0.6982
Seatbelt 5	Reference	Level	
Shy 1	0.0375	0.5053	0.9410
Shy 2	-0.0375	0.2504	0.8813
Shy 3	0.0367	0.2533	0.8850
Shy 4	0.1263	0.2182	0.5639

Effect	Estimate	Standard Error	p-value
Shy 5	Reference	Level	
Tenth Grade	0.0823	0.4248	0.8466
Twelfth Grade	-0.1503	0.6717	0.8233
U.S. Born	0.5207	0.3507	0.1402
Virgin	-0.2384	0.1401	0.0913
Weight Change: Gain	0.1004	0.1495	0.5032
Weight Change: Lose	0.1490	0.2063	0.4716
Weight Change: Not trying to change weight	0.0285	0.2339	0.9034
Weight Change: Stay the same weight	Reference	Level	
White	0.0799	0.5129	0.8765
Smokeless User	0.1149	0.2181	0.5993
Weekly Smokeless User	-0.1689	0.2768	0.5429
Pseudo-R^2	**0.7320**		

Table 9.

Initiation: Weighted Probit Regression Matched on Propensity Scores

Effect	Estimate	Standard Error	p-value
Intercept	2.8739	0.6031	<.0001
Baseball	-0.4454	0.0910	<.0001
Drive High	-0.1725	0.0712	0.0167
Emotional	-0.0563	0.0330	0.0910
Expectations 21	0.1671	0.0530	0.0020
Expectations 35	-0.1487	0.0361	<.0001
Expelled	-0.5847	0.1243	<.0001
Fixed Drinking	0.8387	0.1442	<.0001
Fixed Area Rural	0.4127	0.0813	<.0001
Fixed Area Suburban	-0.0945	0.0772	0.2228
Football	0.3688	0.0668	<.0001
Hang Out	0.0248	0.0392	0.5282
Height Feet	-0.3947	0.0867	<.0001
Household Size	0.0172	0.0325	0.5990
Income	0.0046	0.0020	0.0204
Loving Dad	-0.0549	0.0557	0.3264

Effect	Estimate	Standard Error	p-value
No Extra Curricular	-0.1586	0.1080	0.1444
Non-summer Employment	0.0196	0.0045	<.0001
Seatbelt	-0.1220	0.0202	<.0001
Seventh Grade	-0.3582	0.0697	<.0001
Shy	-0.0416	0.0296	0.1614
Smokeless Weekly Use	0.0155	0.0074	0.0395
Smokeless User	0.3430	0.1059	0.0015
Tenth Grade	0.4761	0.0981	<.0001
U.S. Born	0.2043	0.1845	0.2701
Virgin	0.1461	0.0877	0.0980
White	0.1709	0.1121	0.1300
Pseudo-R^2	**0.7261**		

Table 10

Cessation: Comparison of OLS and Mixed Model Propensity Score Regressions

OLS Effect	OLS Estimate	OLS p-value	Mixed Model Effect	Mixed Model Estimate	Mixed Model p-value
Intercept	-0.5745	0.8431	Intercept	-0.0224	0.8145
Treatment	-0.1314	< .0001	Treatment	-0.2314	0.0031
Propensity Score	19.2145	< .0001	Propensity Score	19.2145	0.0043
R^2	0.3425		-2 Res Log Likelihood	-423.4	

Note. Treatment indicates smokeless tobacco usage (= 1) if an individual was in the smokeless tobacco exposure group, and (= 0) otherwise.

Table 11

Initiation: Comparison of OLS and Mixed Model Propensity Score Regressions

OLS Effect	OLS Estimate	OLS p-value	Mixed Model Effect	Mixed Model Estimate	Mixed Model p-value
Intercept	0.2320	0.8462	Intercept	-0.0224	0.8328
Treatment	0.3516	< .0001	Treatment	0.4514	< .0001
Propensity Score	11.6244	0.0072	Propensity Score	11.6244	0.0028
R^2	0.2440		-2 Res Log Likelihood	-313.3	

Note. Treatment indicates smokeless tobacco usage (= 1) if an individual was in the smokeless tobacco exposure group, and (= 0) otherwise.

CHAPTER 5

DISCUSSION

This study set out to provide the best possible estimates of the effect of using smokeless tobacco as a tool to fight youth smoking. In general, the findings show that smokeless tobacco does not represent a viable mechanism for human application in cessation programs or as a tool for youth smoking prevention. It is important to note, however, that while this study informs us that smokeless tobacco would be a poor tool among the general American adolescent population, findings varied in a special population of medical and business interest – dual users (see Appendices A and B).

There have been two distinct and conflicting theoretical pathways whereby use of smokeless tobacco may impact youth smoking cessation and initiation. On one hand, when a youth begins to use smokeless tobacco one would expect that this may signal underlying motivations, such as a lack of foresight or increased utility derived from psychoactive substances. Alternatively, it suggests that the use of smokeless tobacco increases the overall need for tobacco or nicotine. Both of these phenomena pertain to an increase in the probability of smoking (initiation) or a decrease in the probability of quitting (cessation). In this study, an array of regression parameters and a propensity score-matching algorithm were used to separate out the endogenous effects from the causal result of smokeless tobacco use to the best degree attainable. This is never possible, particularly when using survey data – and represents a major limitation of this study. However such endogeneity is commonly unavoidable in such non-

clinical studies and this research goes beyond all past research in designing an endogeneity-minimizing identification strategy.

On the other hand, there also are consequences that have "beneficial" impacts on smoking – that is, there are effects that increase the probability of cessation and decrease the probability of initiation. Theoretically, we think of these in terms of a reduced short-term need for tobacco (substitution to MST, which is generally cheaper and provides a much higher level of nicotine). Statistically we have witnessed these two conflicting forces revealed in the regressions. The binary use term, which indicated that someone is a smokeless user, had a consistent statistically significant effect that was positively correlated with smoking. Simultaneously, the beneficial (negatively correlated) effects were observed in variables that measure the degree of smoking usage.

In order to most accurately interpret the results in this study, and to internalize all implications, it is essential to review them in the context of this thesis' purpose. The aim of this thesis is to inform the study of youth smoking prevention on the potential value of smokeless tobacco as a tool for policy and programs. Specifically, I set out to determine if smokeless tobacco may be beneficial to the health of youth consumers. This investigation was designed to use observational longitudinal data for analysis of American youth tobacco use.

In the most credible models of this study – those minimizing unobserved heterogeneity via matching and an array of regression controls, as in Tables 10 and 11 – the statistical analyses show that the net impact of smokeless tobacco use on smoking cessation is an increase in the probability of successful smoking cessation. This is in contrast to what simpler methods would have led researchers to believe; for instance the opposite effect may be observed in Table 2, by comparing Models 2 and 3 with Model 1 – when the variance is left to be explained only by one

term the effect which is positively correlated with smoking dominates that which is negatively correlated with smoking.

Due to the significantly higher current prevalence of smokeless tobacco use in males, another limitation of the analyses are that they may not be robust to the female gender. The scope of analysis included both cessation and smoking prevention. This paper describes the "average" American male student, as well as for a special subgroup of increasing interest – dual users (those who currently use both categories of tobacco – flammable and smokeless).

Past studies have produced little and conflicting evidence of the projected efficacy in smokeless tobacco as a cessation treatment for adults. Fewer have attempted closely related research with an American sample – which has many cultural and legal rationale to be observed directly, rather than to rely solely on the international studies which began the scientific discussion of this topic. Furthermore, past studies of smokeless tobacco demand among American youth have not yet investigated any cessation treatment effects. It had also not yet examined smokeless tobacco as a smoking prevention mechanism or applied the most appropriate identification strategies for the study of this notoriously controversial subject (i.e., Rubin's propensity score approach). The findings presented in Chapter 4 provide some initial data that can be used to promote better information among consumers, policymakers, health professionals and anti-smoking program administrators in the search for effective harm-free and harm-reduced smoking cessation aids.

Overview of the Role of Smokeless Tobacco in Youth Smoking Cessation and Initiation

The estimates in this study provide consistent evidence that the "average" American male youth would be exposed to an *increased* likelihood of smoking initiation but also in successful cessation when using smokeless tobacco, *ceteris paribus.* The reason for this appears to be driven by the correlates of increased nicotine tolerance and cravings – although, as discussed when interpreting the key regression variables' coefficients, there was some evidence of a strong opposing force (conflicting consumer reactions, which may exist within individual consumers or possibly distinguish them as comprised of two categories of users, slightly disproportionate in size). These findings present several interesting implications.

Overall, the interpretation of statistical analyses performed in this study provided evidence that reformation of government intervention may be warranted. There are several ways that one might act upon this. At the extreme, one might construe this as reason to allow the legal sale of tobacco products to youth, if cigarettes are taxed much more heavily than smokeless tobacco given the response of youth to price (DeCicca et al., 2008). However any number of alternative methods of action may be possible. For instance, if further research confirms these findings then perhaps policymakers should allow medical doctors to allow the purchase of smokeless tobacco among the appropriate young patients (a plan perhaps similar in structure to California's Medical Marijuana Program). An honorable socioeconomic goal is to minimize the harm to youth in their consumption behavior. Some individuals and special groups of interest may benefit much more than others by increased access to MST and other smokeless tobacco products – "dual users," for example.

Finally, the effect of smokeless tobacco among dual users was informed by this study (although a more thorough investigation of dual usage is reserved for future research). A major contribution of this study was the identification of the unique and distinguishable effects of (a) being a smokeless tobacco user in general with (b) the level of smokeless tobacco use among users. It would seem reasonable that, given time and budget constraints – if not natural constraints on the immediate desire for tobacco or nicotine – dual users consume a finite amount of tobacco, and the proportions of the most dangerous to less dangerous forms of tobacco could be altered by policy (as the discussion of government intervention illustrates). The findings of this study seem to support this reasoning. The weekly (and pseudo-continuous quantity) variables indicated that within the realm of smokeless tobacco users, increases in smokeless tobacco is associated with *decreased* probabilities of initiation and of failed cessation.

Discussion of Findings with a Subgroup of Interest - Dual Users

As always, it is important for readers to be very cognizant of the study's limitations – particularly in that the key statistical findings only truly describe a single overall "average" American male adolescent, whereas there is a great deal of variation in consumer behavior on the individual level and by user (respondent) category. Application of these broad findings would be expected to be inaccurate for many smokers. As in most pharmaceutical studies, this leaves the medical/smoking professionals the task of helping consumers understand their individual characteristics relative to treatments. This is broadly true in NRT and prescription-drug based cessation and initiation programs.

What this research has provided to these interested parties (such as policymakers and voting consumers) is information to help better tweak the youth tobacco ban policies and

treatment-option considerations for cessation. This was accomplished by providing some surprising consumer insights related to a small but growing sub-group of business and social importance – those who regularly consume both smokeless and lit-end tobacco products (dual users). There have been major new alternative tobacco products unveiled by companies such as U.S. Smokeless Tobacco, Philip Morris USA, R.J. Reynolds, and Swedish Match to market new "hip" and more flavorful products such as *snus* into America; thus, there is major industry investment in reacting to (and presumably perpetuating) a growth in the acceptability of occasional smokeless use by smokers. While smokers and smokeless tobacco users are traditionally loyal to their style of consumption, this study unveiled an interesting reason why tobacco companies may be leading the search to the "smoking cure" for some. The ability to switch between products exists in these users, and an increase in smokeless tobacco consumption in this group *does* increase the probability of successful smoking cessation. In this light, the companies' new products may serve to benefit some individuals wishing to quit smoking.

Synthesis of Implications for Consumers, Policymakers and Future Researchers

To evaluate the meaningful information that the consumers, professionals and policy makers should take away from this study, a brief review of the extent of government intervention in youth tobacco is necessary. An all-encompassing recitation of the current regulation of tobacco would be very complex, involving tax policies, legal settlement payments, F.D.A. authority forthcoming as well as many levels of governments (and private) use restrictions. With respect to adolescents, however, the primary legal policy is simply an all-out ban for those under the legal age.

Political "heat" in the debate on tobacco policy may stem from overwhelming agreement among adults regarding the seriousness of youth smoking. A side effect of this national opinion has been an enormous amount of direct campaigning against smoking with focus on children. Ironically, the most notable examples – such as the "Truth" television commercial series – are funded by the post-MSA tobacco industry. Going even beyond national advertisement, both public and private groups have made an extraordinary number of cessation programs, philosophies, and "tools" (NRT devices or medications used to help induce a change in consumption).

This study provides policymakers and others evidence to research use of smokeless tobacco in therapy, yet still consider it as a potential form of "medicine" to be studied for benefits among appropriate candidates. Although this could be construed as contradictory to some aforementioned studies' conclusions, this study may differ in its findings due to true underlying differences in its scope – all findings, interpretations and conclusions are limited and specific to young American males. Within the specified groups of interest however, findings show consistency across all models and key statistics; thus these findings represent credible estimates on the actionable issue of youth smoking prevention.

The thesis was also written to provide some initial empirical investigation of potential harm reduction strategies for the dual-user consumer group; and to raise awareness of the importance of these consumers to all readers, particularly those who are associated with the consumer health and tobacco industries. This provides initial evidence of some significantly important youth groups – male dual-users in particular – being negatively affected by inaccessibility of smokeless tobacco. Although companies have entered the market and continue to test products – such as American snus – which could be the most harm-reducing (and

sensorially pleasurable) consumer health products for some individuals at the moment, no medical research company has produced solutions catered to the unusual degree and common responses of dual users. Furthermore, those firms that are producing new forms of harm-reduced tobacco do not conduct research on youth, smokeless tobacco in cessation, or promote their own products' potential health improvement characteristics due to government (and self) restrictions, which culminated in the 1998 Master Settlement Agreement.

Yet this special group of respondents provokes a need for further research on sub-groups of users of special interest (e.g., patients medically advised to continue tobacco use, such as those whose doctors may be afraid of cessation's impact on depression, anxiety, Parkinson's Disease, etc.) as well as motivation for more user-appropriate cessation and prevention product/program developments.

If the pharmaceutical and medical service industries do not detect sufficient demand to inspire innovation, then this study may have policy implications for the Food and Drug Administration (FDA). The FDA has recently received what will inevitably become historically notable power via forthcoming regulation of the tobacco industry – they may find the combination of knowledge and power to selectively reform limitations on the study of youth by regulated companies.

The findings in this study on dual-users present incentives for future research to be conducted on this group, although it may be difficult to identify respondents for original data collection or experimentation. However, much more credibility and treatment/prevention information could be yielded with macro-economic data, or – ideally – with restricted-use large nationally representative datasets of several thousand-tobacco users. Simultaneously, more research conducted to determine if the main findings in this study of American males – which are

in stark contrast to those of the many papers published Bask and Melkersson – would be a useful means of determining if the conflicting findings are a result of statistical design or of the age of the population in this study; in other words, if youth are reacting differently to smokeless tobacco than adults.

REFERENCES

Ault, R. W., Ekelund, R. B., Jackson, J. D., & Saba, R. P. (2004). Smokeless tobacco, smoking cessation and harm reduction: An economic analysis. *Applied Economics, 36*(1), 17- 29.

Ault, R. W., Beard, R. T., Jackson, J. D. & Saba, R. P. (2005). On the (mis)use of cross- price effects to gauge the effectiveness of smokeless tobacco in smoking cessation. *The European Journal of Health Economics, 6*(1), 83 - 86.

Bask, M., & Melkersson, M. (2003). Should one use smokeless tobacco in smoking cessation programs? *The European Journal of Health Economics, 4*(4), 263 - 270.

Bask, M., & Melkersson, M. (2000). *Rational addiction when there are two addictive goods: Cigarettes and smokeless tobacco*. Umeå University, Department of Economics.

Becker, G. S. & Murphy, K. M. (1988). A Theory of Rational Addiction. *The Journal of Political Economy, 96*(4), 675-700.

Becker, G. S., Grossman, M., & Murphy, K. M. (1994). An Empirical Analysis of Cigarette Addiction. *American Economic Review, 84*(3), 396-418.

Becker, S.O. & Ichino, A. (2002). Estimation of Average Treatment Effects Based on Propensity Scores. *The Stata Journal. 2*(4), 358 – 377.

Benowitz, N. L. (1988). Pharmacological Aspects of Cigarette Smoking and Nicotine Addiction. *The New England Journal of Medicine. 319*(20), 1318.

Bolinder, G., Alfredsson, L., Englund, A., & de Faire, U. (1994). Smokeless tobacco use and increased cardiovascular mortality among Swedish construction workers. *American Journal of Public Health, 84*(3), 399-404.

Borio, G. (2008). *The tobacco timeline.* Retrieved 2/14/2008, from http://www.tobacco.org/History/Tobacco_History.html

Chan, J. S. K., & Kuk, A. Y. C. (1997). Maximum likelihood estimation for probit-linear mixed models with correlated random effects. *Biometrics, 53*(1), 86-97.

Comprehensive Smokeless Tobacco Health Education Act (CSTHEA), 15 U.S.C. §4403(a).

CDC, 2007. *Smokeless tobacco (fact sheet).* Retrieved 2/21/2008, from http://www.cdc.gov/tobacco/data_statistics/Factsheets/index.htm

CDC, 2008. *Health effects of cigarette smoking (fact sheet)*. Retrieved 2/21/2008, from http://www.cdc.gov/tobacco/data_statistics/Factsheets/health_effects.htm

CDC Data Table. (2008). *Economics: Total and per capita yearly consumption of manufactured cigarettes and percentage changes in per capita consumption, 1900-–2004* No. 2008.

Chaloupka, F. J., & Tauras, J. A. (2004). The impact of nicotine replacement therapies on cigarette demand. *Journal of Economics and Finance, 28*(3), 395-403.

Chaloupka, F. J., Tauras, J. A., & Grossman, M. (1997). Public policy and youth smokeless tobacco use. *Southern Economic Journal, 64*(2), 503-516.

DeCicca, P., Kenkel, D., & Mathios, A. (2008). Cigarette taxes and the transition from youth to adult smoking: Smoking initiation, cessation, and participation. *Journal of Health Economics, 27*(4), 904-917.

Foulds, J., Ramstrom, L., Burke, M., & Fagerstrom, K. (2003). Effect of smokeless tobacco (snus) on smoking and public health in Sweden. *Tobacco Control, 12*, 349 - 359.

Gilljam, H., & Galanti, M. R. (2003). Role of snus (oral moist snuff) in smoking cessation and smoking reduction in Sweden. *Addiction, 98*(9), 1183-1189.

Goel, R. K., & Nelson, M. A. (2005). Tobacco policy and tobacco use: Differences across tobacco types, gender and age. *Applied Economics, 37*(7), 765-771.

Greene, W. H. (2000). *Econometric analysis* (4th ed.). Upper Saddle River, N.J.: Prentice Hall.

Gruber, J., & Kőszegi, B. (2004). Tax incidence when individuals are time-inconsistent: The case of cigarette excise taxes. *Journal of Public Economics, 88*(9-10), 1959-1987.

Hoffmann, D., Brunnemann, K. D., Prokopczyk, B., & Djordjevic, M. V. (1994). Tobacco-specific N-nitrosamines and areca-derived N-nitrosamines: Chemistry, biochemistry, carcinogenicity, and relevance to humans. *Journal of Toxicology and Environmental Health, 41*(1), 1-52.

Jackman, S. (2000). Estimation and Inference Are Missing Data Problems: Unifying Social Science Statistics via Bayesian Simulation. *Political Analysis, 8*(4), 307-330.

Kleinbaum, D.G. (1994). *Logistic regression: A self-learning text.* New York, N.Y.: Springer-Verlag.

Keeler, T. E. (2002). The benefits of switching smoking cessation drugs to over-the-counter status. *Health Economics, 11*(5), 389-402.

Keeler, T. E., Marciniak, M., & Hu, T. (1999). Rational addiction and smoking cessation: An empirical study. *Journal of Socio-Economics, 28*(5), 633-643.

Maddala, G. S. (1983). *Limited-dependent and qualitative variables in econometrics.* New York, N.Y.: Cambridge University Press.

Melkersson, M. (2000). Smoking habits and quitting costs in the presence of a close substitute. Manuscript, University of Stockholm, 21 March 2000.

National Institutes of Health. (1986). The Health Consequences of Using Smokeless Tobacco: A Report of the Advisory Committee to the Surgeon General. Bureau of Maternal and Child Health and Resources Development. Office of Maternal and Child Health.

Ohsfeldt, R. L., & Boyle, R. G. (1994). Tobacco excise taxes and rates of smokeless tobacco use in the US: An exploratory ecological analysis. *Tobacco Control, 3*(4), 316-323.

Ohsfeldt, R. L., Boyle, R. G., & Capilouto, E. (1997). Effects of tobacco excise taxes on the use of smokeless tobacco products in the USA. *Health Economics, 6*(5), 525-531.

Ohsfeldt, R. L., Boyle, R. G., & Capilouto, E. I. (1999). Tobacco taxes, smoking restrictions, and tobacco use. In F. J. A. Chaloupka (Ed.), *The economic analysis of substance use and abuse: An integration of econometric and behavioral economic research* (pp. 15-29). NBER Conference Report series; Chicago and London; University of Chicago Press.

Piko, B. F., Luszczynska, A., Gibbons, F. X., & Tekozel, M. (2005). A culture-based study of personal and social influences of adolescent smoking. *European Journal of Public Health, 15*(4), 393-398.

Rock, V. J., Malarcher, A., Kahende, J. W., Asman, K., Husten, C., & Carabollo, R. (2006). *Cigarette smoking among adults - United States, 2006.* Retrieved 3/11/2008, from http://www.cdc.gov/mmwr/preview/mmwrhtml/mm5644a2.htm

Rogozinski, J. (1990). *Smokeless tobacco in the western world, 1550-1950.* Westport, Conn. and London; Greenwood Press, Praeger.

Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika, 70*(1), 41-55.

Rubin, D. B. (1997). Estimating causal effects from large data sets using propensity scores. *Annals of Internal Medicine, 127*(8), 757-763.

Rubin, D.B. (2001). Using propensity scores to help design observational studies: Application to the tobacco litigation. *Health Services and Outcomes Research Methodology 3*(2), 169-188.

Sabia, J. J., & Rees, D. I. (2008). Youth smoking and addiction: Evaluating the wisdom and efficacy of government intervention. *Expert Review of Pharmacoeconomics & Outcomes Research, 8*(3), 213-217.

Schafer, J.L. & Graham, J.W. (2002). Missing Data: Our View of the State of the Art. *Psychological Methods, 7*(2), 147-177.

Singer, J. D. & Willett, J. B. (2003). *Applied Longitudinal Data Analysis: Modeling Change and Event Occurrence.* Oxford, N.Y.: Oxford University Press.

Sloan, F.A., Ostermann, J., Picone, G. Conover, C. & Taylor Jr., D.H. (2004). *The price of smoking.* Cambridge, M.A.: The M.I.T. Press.

Stratton, K. R. (2001). *Clearing the smoke: Assessing the science base for tobacco harm reduction*. Washington, D.C.: National Academy Press.

Suranovic, S. M., Goldfarb, R. S., & Leonard, T. C. (1999). An economic theory of cigarette addiction. *Journal of Health Economics, 18*(1), 1-29.

Tauras, J., Powell, L., Chaloupka, F., & Ross, H. (2007). The demand for smokeless tobacco among male high school students in the United States: The impact of taxes, prices and policies. *Applied Economics, 39*(1-3), 31-41.

Udry, J.R. (1998). *A User's Guide to the Machine-Readable Files and Documentation*. Los Altos: Sociometrics Corporation.

Udry, J.R. (2003). *A User's Guide to the Machine-Readable Files and Documentation*. Los Altos: Sociometrics Corporation.

U.S. Public Health Service. (1964). *Smoking and health: Report of the advisory committee to the surgeon general of the public health service.* Retrieved 3/11/2008, from http://profiles.nlm.nih.gov/NN/B/C/X/B/

Verbeke, G., & Molenberghs, G. (2000). *Linear mixed models for longitudinal data*. New York: Springer-Verlag.

Viscusi, W.K. (1992). *Smoking: Making the risky decision.* Oxford, N.Y.: Oxford University Press.

APPENDIX A:

American Snus Tobacco Products

American Snus has recently emerged into the marketplace, ostensibly targeting cigarette and smokeless tobacco users with a "hassle free" alternative. The products are commonly sold side-by-side with traditional smokeless products in cigar shops and side-by-side with cigarettes in convenience stores. These products are dually targeted at smokers and smokeless tobacco users, not only in terms of product design and placement, but also by promotional coupon distribution via cigarette packages, smokeless tobacco containers and both smokeless and cigarettes mailing lists.

Figure 2. Marlboro Snus convenience store advertisement produced by Philip Morris USA, 2008.

Figure 3. Advertisement for Camel Snus produced by R.J. Reynolds (2009).

APPENDIX A3

Figure 4. Example of American snus marketing tactics in the form of product give-away (coupons) to current tobacco users. This photo was taken from a commonplace Camel cigarette package (2008).

APPENDIX B:

Non-Snus New Smokeless Tobacco Products with Youth and Dual-User Appeal

Figure 5. The new "Camel Orbs" test market products avoid FDA regulation of flavored tobacco and are compared in design to youth-friendly products (source: http://www.meltdownva.com, retrieved online July 11, 2009).

APPENDIX B2

Figure 6. Camel Sticks and Camel Orbs instructional inserts, graphics produced by R.J.R., 2009.

APPENDIX B3

Figure 7. Camel Strips emulate the design of breath strip products. Image produced by R.J. Reynolds, 2009.