WHO ARE THE CIGARETTE SMOKERS IN ARIZONA?

by

Mei-kuang Chen

_____

A Thesis Submitted to the Faculty of the

DEPARTMENT OF PSYCHOLOGY

In Partial Fulfillment of the Requirements
For the Degree of

MASTER OF ARTS

In the Graduate College

THE UNIVERSITY OF ARIZONA

2007

STATEMENT BY AUTHOR

SIGNED: Mei-kuang Chen

APPROVAL BY THESIS DIRECTOR

This thesis has been approved on the date shown below:

Lee Sechrest .
Lee Sechrest
Professor of Psychology

July 12, 2007
Date

ACKNOWLEDGEMENTS

This has been a long process. Without all the help provided by God, Lee, AJ, Michele, EGAD, my family, and friends, I know that I cannot get here. Thank you all for your patience, love and support. Academic is a long journey, but full of fun. I am encouraged by every step in this exciting process. This is the start of my expertise in English here in the U.S. and hopefully will lead to a broad future career in the world.

TABLE OF CONTENTS

TABLE OF CONTENTS - Continued

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

The purpose of this study was to investigate the relationship between cigarette smoking and socio-demographic variables based on the empirical literature and the primitive theories in the field. Two regression approaches, logistic regression and linear multiple regression, were conducted on the two most recent Arizona Adult Tobacco Surveys to test the hypothesized models. The results showed that cigarette smokers in Arizona are mainly residents who have not completed a four-year college degree, who are unemployed, White, non-Hispanic, or young to middle-aged adults. Among the socio-demographic predictors of interest, education is the most important variable in identifying cigarette smokers, even though the predictive power of these socio-demographic variables is small. Practical and methodological implications of these findings are discussed.

INTRODUCTION AND LITERATURE REVIEW

Background and Research Questions

Who are the cigarette smokers in Arizona? As public awareness of the harmful effects of smoking has increased, growing public health effort has been directed toward controlling tobacco, especially in the form of cigarette smoking. Even with all these efforts, however, researchers have found that disparities in cigarette smoking still exist by education, income, sex, race/ethnicity, occupation/work status/employment, age, disability, geographic location, and sexual orientation (Barbeau et al., 2004; Barbeau, Krieger, & Soobader, 2004; Centers for Disease Control and Prevention (CDC), 2006; Fiore et al., 1989; Rogers, Nam, & Hummer, 1995; Ryan, Wortley, Easton, Pederson, & Greenwood, 2001). We know that in the socio-economically disadvantaged populations, such as those with low income or low education, smoking behavior has become a more recalcitrant problem for them than for members of other better-off populations. Empirical findings have shown that education is related to current smoking prevalence and smoking cessation (Barbeau et al., 2004; Centers for Disease Control and Prevention (CDC), 2005; Escobedo & Peddicord, 1996; J. P. Pierce, Fiore, Novotny, Hatziandreu, & Davis, 1989; Sorensen, Barbeau, Hunt, & Emmons, 2004; Wetter, Cofta-Gunn, & Fouladi et al., 2005; Winkleby, Jatulis, Frank, & Fortmann, 1992). Lower-educated people tend to smoke more, and fewer of them have quit smoking. Similar phenomena exist for the subpopulations with low income or lower work status (Bang & Kim, 2001; Coleman, 2004; Wetter, Cofta-Gunn, & Fouladi et al., 2005). Nationally, American Indian and African American male adults, between age 18 and 44, have a higher smoking prevalence

than other groups (Centers for Disease Control and Prevention (CDC), 2006; U.S. Department of Health and Human Services, 1998). Our question is whether the accumulated findings regarding the relationship between cigarette smoking and demographic variables can help us to target the tobacco control effort in Arizona more effectively.

The main purpose of the present study is to investigate the relationship between demographic variables and cigarette smoking based on existing literature, identify the high risk populations, and, hopefully, shed some light on cigarette smoking of the high risk groups. There are two focuses of this study. First, from the standpoint of tobacco control programs, the aim is to identify the important demographic predictors of cigarette smoking in the order of their predictive power. The demographic variables of interest are education, income, employment, sex, race and age. We would like to know in Arizona, which demographic factors can better predict current smoking status and it can direct the tobacco control resources, including smoking prevention and cessation program, to specific target populations. Second, from the perspective of a research methodologist, my own, that is, the goal is to investigate how we can answer the question of interest appropriately. So far, most of the epidemiological studies have defined "current cigarette smoking" in a fairly consistent, if not fully informative, way (for example, Centers for Disease Control and Prevention (CDC), 2006), and have used rather similar statistical methods, mainly univariate or multivariate logistic regression, to investigate the relationship between cigarette smoking and relevant demographic variables, (for example, Barbeau et al., 2004). However, results of logistic regression may not give us sufficient

and clear information about the relative importance of the demographic predictors of cigarette smoking. Therefore, in addition to the traditional analysis, an alternative approach is explored here for supplementing information to answer the research question of interest in this study.

Review of the Literature

The purpose of this section is to review the literature related to the smoking behaviors of various demographic strata in the general population. Included are the empirical findings of the relationship between demographic characteristics and cigarette smoking, and the relevant theories that may explain the relationships among smoking and social-economic positions. Important demographic characteristics related to adult smoking behavior include education, income, occupation/work status/employment, sex, age, racial or ethnic identity, and some of the interactions among these demographic variables (Bergen & Caporaso, 1999; Brummett et al., 2002; Centers for Disease Control and Prevention (CDC), 2005; Fiore et al., 1989; Husten et al., 1997; Rogers et al., 1995; Wetter, Cofta-Gunn, & Fouladi et al., 2005; Winkleby, Fortmann, & Rockhill, 1992). The existing literature provides the foundations for the models tested in this study.

Education

Consistently, education has been shown to be perhaps the most important demographic predictor of current smoking status among adults (Bang & Kim, 2001; Bergen & Caporaso, 1999; Nelson et al., 1994; J. Pierce, Giovino, Hatziandreu, & Shopland, 1989; Wetter, Cofta-Gunn, & Irvin et al., 2005; Winkleby, Fortmann, & Barrett, 1990; Winkleby & Jatulis et al., 1992; Zhu, Giovino, Mowery, & Eriksen, 1996).

One of the earlier studies, Pierce et al. (1989) found that from 1974 to 1985, across all educational levels, smoking prevalence declined steadily, but this declining trend was more obvious for higher-educated people. Each year these better-educated persons smoked at a markedly lower rate. Pierce's data showed that smoking rate among college graduates decreased almost 5 times faster than among those with less than a high school education. Also, the quit ratio, proportion of former-smokers among ever-smokers, increased during that period of time, with the most highly-educated individuals having a higher quitting rate than other less-educated persons. The same relationship existed for older smokers (Husten et al., 1997), and work sites populations (Gritz et al., 1998). Across racial groups, the General surgeon's report on tobacco and racial/ethnic minority groups showed that among the high school graduates, the declines in smoking prevalence were greater than for the lower-educated, especially for men (U.S. Department of Health and Human Services, 1998). However, Zhu et al. (1996) showed somewhat different smoking patterns among people with different education levels. They found that after controlling all other demographic variables, people with 9 to 11 years of education had the highest rate of being current, ever, or heavy smokers, and also the lowest rate of having quit smoking. Their cigarette smoking status was different from that of people with 0 to 8 years of education. This group of 0 to 8 years of education was much less likely to have smoked and more likely to have quit smoking. In contemporary American society, not progressing beyond eight grades of schooling is fairly remarkable, and it is likely that this greatly undereducated group is quite different from the mainstream population. The 8-grades-or-less group probably includes a fair number of disabled

persons, recent immigrants, and others for whom smoking cigarettes is simply not part of the way they do or must live their lives.

Recent research also showed that smoking status continues to show a strong educational gradient (Centers for Disease Control and Prevention (CDC), 2005; Wetter, Cofta-Gunn, & Fouladi et al., 2005; Wetter, Cofta-Gunn, & Irvin et al., 2005). The higher the educational level a person attains, the less likely the person is to smoke. The current smoking rate generally decreased as school years increased (Centers for Disease Control and Prevention (CDC), 2005). As something of an anomaly to be addressed later, adults with a General Educational Development (GED) diploma had a higher current smoking prevalence than did adults who never finished high school. Wetter, Cofta-Gunn, & Fouladi et al., (2005) conducted research in a community-based adult sample, and they also found that higher education is strongly and significantly related to a lower current-smoking rate and a higher never-smoking rate. In their study, the current smoking prevalence was almost three times higher among individuals with less than high school education than among those with at least a college degree, while the percentage of "never smoked" among the lower-educated individuals was less than half that of the never-smoked rate of higher-educated individuals. Moreover, as in the earlier finding, the rate of quit smoking increased linearly with educational levels. Upon testing the different multivariate models, these researchers concluded that the relationship between education and smoking status was not attributable to demographic, environmental, or job-related variables. The evidence supported the conclusion that education is a specific predictor of smoking status.

Researchers have proposed several different mechanisms to explain the relationship between education and smoking. Education can be viewed as a marker for intelligence; and to cognitive and non-cognitive traits (such as emotion, motivation, self-esteem and etc.) (Heckman, Stixrud, & Urzua, 2006), manifesting the difference in willingness or ability to invest in human capital (Fuchs, 1979). The relationship between schooling and smoking would be the same whether schooling results in cognitive and non-cognitive skills or those skills result in more schooling. It may also be that schooling influences risk-taking behavior, such as smoking, through acquisition of positive social, psychological, and economic skills and assets, including positive attitudes about health, access to preventive health services, higher self-control, self-esteem and self-efficacy, and that schooling may provide protection from adverse influences (Winkleby et al., 1992). Also, educational attainment may be related to response to delayed rewards (Jaroni, Wright, Lerman, & Epstein, 2004). Ability to delay gratification may link to success at smoking cessation, since individuals have to give up the immediate reward of smoking, and then focus on the future health benefits of smoking abstinence. Jaroni's research showed that smokers with no college education discounted delayed rewards more than those who had at least some college education. Even within the smoker sample, education still had an effect on discounting delayed rewards. These results supported the plausible link between education and delay gratification as a mechanism of how education is related to smoking cessation. Still other writers have suggested that environmental factors may be related to education attainment, e.g., that school dropouts end up in environments that are conducive to and supportive of smoking.

Overall, the literature shows a rather strong relationship between education and smoking, and the possible mechanisms involve cognitive, non-cognitive abilities or delay gratification. Exactly how these mechanisms actually work is still a puzzle.

Income

Income, moderately correlated with educational levels (N. E. Adler & Snibbe, 2003; Winkleby et al., 1992), is also related to current smoking status. Adults living below the poverty level tend to have a higher current smoking prevalence than those at or above the poverty level (Bergen & Caporaso, 1999; Centers for Disease Control and Prevention (CDC), 1993; Centers for Disease Control and Prevention (CDC), 2005; Flint & Novotny, 1997). Flint & Novotny (1997) analyzed year to year trends of cigarette smoking prevalence and cessation from 1983 to 1993. They found a consistent pattern of smoking prevalence such that the subpopulation below the poverty threshold had a higher smoking rate across all the years in this study period. In their logistic regression models for different years, both the unadjusted and the adjusted odds ratio of current smoking given poverty status showed that people below the poverty level were more likely to smoke (Although, when sex, age, education, race, employment status, marital status, and geographical region were included in the models, the odds ratio based on poverty status became smaller.). Also, when other demographical variables were included in the model, the results showed that people at or above poverty level were more likely to have quit smoking compared to the poor people. The authors concluded that from 1983 to 1993, persons below the poverty level had been, and continued to be, at higher risk for smoking in terms of higher current smoking prevalence and lower. Obviously, these findings

indicated that lower income was associated with higher risk for smoking, but income may be a less consistent predictor of smoking than education (Winkleby & Jatulis et al., 1992). In summary, income probably is one of the important predictors of current smoking status, but the limited available research investigating the relationship between smoking and income in general population makes it difficult to derive any firm conclusion.

The reason why low income (poverty) should be associated with smoking remains unclear. Some plausible explanations reviewed by Bobak et al. (Bobak, Jha, Nguyen, & Jarvis, 2000) are as follows. First, poor and less educated people may not be aware of the health hazards of smoking compared to economically better off and educated people, and are thus more likely to take up and continue smoking. That explanation, of course, ties the income effect closely to education. Second, people in poverty or with lower social economic status are exposed to more stress than people not in poverty or with higher SES, and stress may have some effect on the development of the habit of smoking as a self-medication used for mood regulation, stress management, or as some other coping mechanism for material deprivation. Third, smoking may be a replacement reward, one of a few things that poor people can do for themselves. Fourth, a hypothesis from economic perspective is that a person whose income is lower would have less to lose from future health problems than a person with a higher income, given the same benefits of smoking perceived. Fifth, some evidence shows that poor smokers may have higher nicotine dependence than other non-poor smokers, and it may be more difficult for them to quit smoking. Also, some researchers think that poverty may be a marker of under-participation in the social norm changes related to smoking behaviors (Flint & Novotny,

1997) or a marker of neighborhood disadvantage (Bergen & Caporaso, 1999). Poverty status may represent something beyond the household or individual income level, such as available resources or spending power, but just what income level represents in the smoking equation remains uncertain.

Work Status/Employment/Occupation

Work status, employment type, or occupation, has been found to be related to smoking status as well. Empirical research showed that smoking prevalence of the currently unemployed people was significantly higher than that of people who were currently working or not in labor force, even after controlling for age, education, race, and marital status (Bang & Kim, 2001; Brackbill, Frazier, & Shilling, 1988; Nelson et al., 1994). Also, some occupational groups, such as blue-collar workers (including transportation equipment operatives, craftspeople, and laborers and etc.), have had higher smoking rates than other occupational groups, such as teachers or white-collar workers (Bang & Kim, 2001; Barbeau et al., 2004; Brackbill et al., 1988; Covey & Wynder, 1981; Nelson et al., 1994; Sterling & Weinkam, 1978; U.S. Department of Health and Human Services, 1985; Winkleby & Jatulis et al., 1992). In a longitudinal study, early unemployment, defined by being unemployed for at least 13 weeks during the age of 16 to 21, was one of the important predictors for later smoking at age of 30 (Hammarstrom & Janlert, 2003). The relative risk for future daily smoking was 1.97 and 1.86 for early unemployment for men and women respectively. Early unemployment was the highest predictive factor for both current smoking status and alcohol consumption. Young people who became unemployed at an early age may be a high-risk group for later destructive

health behavior. However, overall, the predictive power of the variables of interest in this study was not high, and hence the findings need to be interpreted with caution.

Occupation, as one of the indicators for social-economic status, in addition to education and income, represents social class level, prestige, responsibility, physical activity, and work exposures (N. E. Adler et al., 1994; Winkleby & Jatulis et al., 1992). Barbeau, Krieger and Soobader (2004) investigated the relationship between smoking and occupation, taking education, income, sex, and race into consideration, and found that how social class is conceptualized and measured is important to understanding smoking patterns. As other studies, the results of Barbeau, et al. (2004) supported the notion that people in working-class occupations are more likely to smoke. Marmot, Smith, Stansfeld et al. (1991) found clear employment-grade differences in smoking behavior. These investigators used the participants' civil service title and salary to categorize participants into six groups. Category one represents the highest job status and category six the lowest. Smoking prevalence consistently decreased for both male and female from category six to category one, with the higher smoking prevalence of category one females the only exception.

Also, smoking prevalence among white-collar workers continues to decline faster than does the smoking prevalence among blue-collar workers and hence is widening the difference in smoking prevalence by occupational groups (Nelson et al., 1994; U.S. Department of Health and Human Services, 1985). Similarly, unemployed people are less likely to have quit smoking than are people who are at work (Nelson et al., 1994). The higher smoking prevalence of both blue-collar workers and unemployed people reflects

the lower quit rates for both groups. In the report of the Surgeon General 1985 (U.S. Department of Health and Human Services, 1985), based on 1978-1980 National Health Interview Surveys (NHIS), it was shown that for both sexes, blue-collar workers started to smoke at an earlier age than white-collar workers (on average, approximately 14 months earlier, from 16.5-17.4 year-old for male and female laborers, respectively, to 17.7-19.4 year-old for male and female professionals, respectively), and a considerable portion of smokers report having started smoking at about the time they entered the workforce.

Nelson et al. (1994) discussed the possible reasons for the higher smoking prevalence among blue-collar workers than among white-collar workers. They believe that a birth cohort effect may account part of the decline of smoking prevalence observed across most occupations, but it was not clear how the birth cohort effect may be related to the higher smoking prevalence of some sections of occupational groups, such us unemployed people or blue-collar workers. One plausible explanation of the higher smoking prevalence among blue-collar workers is workplace environment. Nelson, et al. (1994) proposed that job strain, a combination of high psychological job demands and a low level of work control, and occupational stress may be more common among blue-collar workers; this higher job strain and perceived high levels of occupational stress may account for the increased intensity of smoking and inhibition of smoking cessation among this occupational group (Conway, Vickers, Ward, & Rahe, 1981; Green & Johnson, 1990; Nelson et al., 1994). Many smokers report using cigarette smoking to cope with their stress. Also, social influences at the work site may be another factor influencing workers'

smoking habits. Smoking is more likely part of the social norm among the blue-collar worker groups, and it adds to the difficulty for group members to quit smoking.

Few studies specifically explored the reasons unemployed people tend to smoke more than people who are working. Some researchers suspected that some personal characteristics or early experiences may predispose some people both to initiate smoking and find it difficult to obtain and keep jobs (Nelson et al., 1994), but the sample these investigators used as an example, people with some sort of mental illness, was only a weak explanatory factor at best. We think that high levels of stress associated with unemployment and extra time available for smoking may be two of the possible reasons.

Sex

In a historical trend, cigarette smoking became commonplace among women well after it caught on among men. In the early 20th century, smoking was not common among women. In the 1950s, smoking prevalence among men increased rapidly. But by the time of the release of the first Surgeon General's report on smoking, smoking rates of men had already started to decline, whereas smoking rates for women did not begin to decline substantially until the 1970s or later (Fiore et al., 1989; U.S. Department of Health and Human Services, 2001). Smoking prevalence among men is usually distinctly higher than the prevalence in women (Barbeau et al., 2004; Bergen & Caporaso, 1999; Brackbill et al., 1988; U.S. Department of Health and Human Services, 2001). However, in the 1990s, it was reported that there was some change among women's smoking. Smoking initiation had declined steadily in men, but it seemed to increase in women in the early to mid 1990s, especially in teenaged girls, and it declined again in late 1990s (Bergen &

Caporaso, 1999; Fiore et al., 1989; U.S. Department of Health and Human Services, 2001). The gap in smoking rates between men and women narrowed from 1965 to the 1980s, and since the mid-1980s it somewhat stabilized at about a five percent higher prevalence among men than among women (U.S. Department of Health and Human Services, 2001). The most recent Morbidity and Mortality Weekly Report (MMWR) showed that current smoking prevalence among men was 23.9%, higher than the rate among women, 18.1% (Centers for Disease Control and Prevention (CDC), 2006).

Some studies also indicated that women tend to smoke fewer cigarettes per day than do men (Benowitz & Hatsukami, 1998; Rogers, 1991; Rogers et al., 1995; Sterling & Weinkam, 1976). The surgeon general's report on smoking and women concluded that women smokers appear to be equally nicotine-dependent as men in terms of number of cigarettes smoked per day, and there is not much evidence of sex-specific differences in nicotine dependence (U.S. Department of Health and Human Services, 2001). Also, the evidence regarding sex difference in smoking quit rate are not always consistent. Some research showed that male smokers have had a higher probability of quitting smoking, especially at higher levels of cigarette consumption (Rogers et al., 1995). The lung health study evaluating the efficacy of a smoking cessation program found that men had higher cessation rates at 12 and 36 months follow-up. Women seemed to have more difficulty quitting smoking than men (Bjornson et al., 1995). In Gritz's study (1998), for short-term smoking cessation, there was significant sex difference between male and female quitters. In this worksite sample, male smokers with high-school level education or less were more likely to quit smoking than were female smokers with the same educational level. Also,

women were not as ready to quit as men in this worksite sample. Some researchers found that there might be sex differences in withdrawal symptoms while quitting tobacco use. More women than men reported withdrawal symptoms in a group of young smokers after receiving cessation services in a smoking cessation health center in China during the one-and-half year of study period (Abdullah, Lam, Chan, & Hedley, 2006). For the smokers who reported any withdrawal symptoms, it is more likely for them to have failed in their quitting attempt than those who did not report any withdrawal symptom.

U.S. national survey data showed, however, that quitting rates are not different for female and male smokers. The effectiveness of prevention and cessation interventions are also generally similar for women and men. Few sex differences in terms of smoking initiation and successful quitting have been identified so far (U.S. Department of Health and Human Services, 2001). In Benowitz & Hatsukami's (1998) review paper on sex difference in the pharmacology of nicotine addiction, the studies they cited showed that women seemed to be more nicotine-dependent based on DSM-IV criteria, and smoking cessation was more difficult for women than for men. They suspected that there may be sex difference in the nature of nicotine addiction. However, after reviewing the relevant empirical evidences in literature, they concluded that in general, not much in the way of sex differences in pharmacological and behavioral processes in nicotine addiction had been found. The basic processes determining nicotine addiction are highly similar in men and women.

Overall, smoking patterns among men and women are somewhat different, but the differences are not systematic or consistent across studies. Historical trends or cohort

effects were often used to explain these sex differences in smoking prevalence and quit rate. Husten et al. (1997) surmised that measures of cessation are cumulative over time, and since men began quitting cigarette smoking earlier in the century than did women, the cessation prevalence is higher for men. These researchers suspected that the differences in smoking cessation across sexes was due to these historical patterns and may not be related to any substantial sex difference in smoking. Empirical data seemed to suggest, however, that sex differences in smoking initiation may be responsible for the converging rates of smoking prevalence among different sexes (Fiore et al., 1989). The Surgeon General's report on smoking and women in 2001 (U.S. Department of Health and Human Services, 2001) reviewed relevant studies of sex differences in critical social, cultural, and personal influences on smoking initiation. The report showed that even though not many consistent findings have been obtained so far, girls seem to be more affected than boys by parental influences, desire to smoke for weight and mood control, and rebelliousness or rejection of conventional values. Girls and boys seem to start and continue to smoke for different reasons, and these differences may be reflected in differential rates of smoking and cessation.

Race/Ethnicity

According to the Surgeon General's report on tobacco use among U.S. racial/ethnic minority groups (U.S. Department of Health and Human Services, 1998), among the four major racial/ethnic minority groups, African Americans, American Indians and Alaska Natives, Asian Americans and Pacific Islanders, and Hispanics, African Americans bore the greatest health burden related to smoking at that time. The

national tobacco surveys have consistently shown that American Indians and Alaska Natives have the highest smoking prevalence among these major racial groups (Bolen, Rhodes, Powell-Griner, Bland, & Holtzman, 2000; Centers for Disease Control and Prevention (CDC), 2006; U.S. Department of Health and Human Services, 1998). African Americans and Whites also had a high smoking prevalence, but Asian Americans and Hispanic women have the lowest smoking prevalence (Centers for Disease Control and Prevention (CDC), 2006). The pattern of smoking prevalence among different racial groups seems to have remained rather stable during this decade.

Compared with White smokers, smokers among the racial/ethnic minority groups smoke fewer cigarettes every day, and most of them are more likely to report smoking occasionally, not daily (U.S. Department of Health and Human Services, 1998). A higher proportion of White smokers are heavy smokers than smokers among other racial/ethnic groups. Even though the national tobacco surveys have consistently shown that American Indians and Alaska Natives have the highest smoking rates among the major racial/ethnic groups, the more detailed breakdown of American Indians data in the Strong Heart Study indicated that American Indians in Arizona had a lower smoking prevalence among Indian Americans, especially for American Indian women in Arizona (Welty et al., 1995), consistent with the 1988-1992 aggregate data from the Behavioral Risk Factor Surveillance System (BRFSS) (U.S. Department of Health and Human Services, 1998). In BRFSS data, much lower overall cigarette smoking prevalence was found among Indian Americans in California and the Southwest. The majority of American Indian smokers responding to the BRFSS smoked 15 or fewer cigarettes per day. State and

regional surveys showed that the smoking cessation rate among American Indian smokers remained relative low compared to smokers in other racial/ethnic smokers, but the American Indian smokers in the Southwest and the Pacific Northwest had the highest quit rate among American Indian smokers.

For Asian Americans and Pacific Islanders, few data are available for assessing cigarette smoking, and the aggregate data usually cannot provide precise estimates. In addition, this racial category is heterogeneous in terms of both culture and health behaviors (U.S. Department of Health and Human Services, 1998), thus, the smoking behavior of this specific racial group will not be further discussed here.

For Hispanics, smoking prevalence had continued to decrease during 1978 to 1995 (U.S. Department of Health and Human Services, 1998), and the most recent MMWR showed that this declining trend of smoking prevalence seems to continue for the Hispanics (Centers for Disease Control and Prevention (CDC), 2006; Maher et al., 2005). The effect of acculturation on cigarette smoking has been investigated more in Hispanic population (U.S. Department of Health and Human Services, 1998). Generally, research has shown that level of acculturation, defined by self-reported acculturation level or language spoken at home, is positively related to prevalence of cigarette smoking, especially for women (Abraido-Lanza, Chao, & Florez, 2005; Bethel & Schenker, 2005; U.S. Department of Health and Human Services, 1998). Hispanics with a higher acculturation level were more likely to smoke. Overall, cigarette consumption remained stable or decreased somewhat, and the quit rate increased moderately among Hispanic smokers between 1978 and 1995. Hispanic smokers were slightly more likely to have quit

smoking for at least one month than smokers among White (U.S. Department of Health and Human Services, 1998). The level of smoking severity seems to be lower for the Hispanic compared to the White smokers.

Among the major racial groups, African Americans have a moderate cigarette smoking prevalence, but a disproportionally high incidence of lung cancer and cerebrovascular disease (Gadgeel & Kalemkerian, 2003; U.S. Department of Health and Human Services, 1998). Since 1965, African Americans had a higher smoking prevalence than Whites, but smoking prevalence decreased for both groups along the years. A recent MMWR report showed that similar smoking prevalence has been observed among the non-Hispanic White and African Americans (Centers for Disease Control and Prevention (CDC), 2006), but the African American smokers had a higher proportion than Whites of "light" smokers (Novotny, Warner, Kendrick, & Remington, 1988; U.S. Department of Health and Human Services, 1998). This report indicated that African Americans had a stronger preference for mentholated cigarettes than did smokers among other racial/ethnic groups. Studies also showed that the African American smokers had a lower quit rate compared to the White smokers (Husten et al., 1997; Novotny et al., 1988).

The patterns of cigarette smoking among racial/ethnic minority groups reflected the complex interactions of multiple factors, including socioeconomic status, cultural characteristics, acculturation, stress, biological elements, targeted advertising, price of tobacco products, and other community factors (U.S. Department of Health and Human Services, 1998). The investigators attributed the racial difference in smoking cessation to some genetic differences, including higher level of nicotine dependence, and higher

nicotine and tar levels in the body, and other factors, such as limited access to preventive tobacco control services, and lack of appropriate cessation services for smokers among different racial/ethnic minority groups (Husten et al., 1997). For example, researchers found that the metabolism of cotinine (nicotine is mainly metabolized to cotinine) is different for different racial/ethnic groups. Cotinine levels were found to be higher in African-American than in White or Mexican Americans. It suggested that the metabolism of cotinine may be related to the differences in the numbers of cigarettes smoked among African-American and Caucasian-American smokers (Ahijevych, Tyndale, Dhatt, Weed, & Browning, 2002; Caraballo et al., 1998; Perez-Stable, Herrera, Jacob, & Benowitz, 1998).

Age

The relationship between age and smoking is nonlinear in populations with at least a moderately wide range of ages (Rogers, 1991; Rogers et al., 1995). Adolescence is the most common period when smoking is initiated. Most people start to smoke when they are teenagers, before graduating from high school (U.S. Department of Health and Human Services, 1994). An earlier US national survey showed that smoking prevalence was low among young adults (ages of 18 to 24), and highest among the group of ages 25 to 44 (Centers for Disease Control and Prevention (CDC), 1993). There were more young people who have never smoked, but a smaller proportion of never smokers in the middle years. Among the young smokers, most of them smoked few cigarettes (Rogers et al., 1995). However, the most recent MMWR showed that young adults, aged 18 to 24, and adults aged 25 to 44, had the highest cigarette smoking rate in the US (Centers for

Disease Control and Prevention (CDC), 2006). Apparently more of them are smoking occasionally but not necessarily in large quantities.

Over the period of 1965 to 1994, current smoking prevalence was lower among the elderly people (age of 65 and above) than those aged 18 to 64 (Husten et al., 1997). One of the recent MMWR (Centers for Disease Control and Prevention (CDC), 2005) showed that this oldest group of adults had the lowest current cigarette smoking prevalence, and among all age groups, this group was the one most likely to have quit (Rogers et al., 1995). Usually, cigarette consumption increases with age from the teenage years to the ages of 40s or 50s, but tends to be lower for the older group (Rogers, 1991).

Researchers explained the relationship between age and smoking with cumulative quitting over time (Husten et al., 1997), and this cumulative quit rate will result in a lower smoking prevalence and a higher cessation rate among the elderly. Researchers also pointed out that since it would be more difficult for continuing smokers to survive to old age, this differential mortality across ages may result in an even more apparent decline in smoking rate and more apparent increase in smoking quit rate among older age groups. In addition, older smokers may be more likely to cut down their cigarette consumption due to economic or health reasons.

Interactions between Demographic Variables

The empirical data indicated that many interaction terms between demographic factors may be related to cigarette smoking, but how strong the connections were is often unclear in the studies. The patterns of smoking prevalence among subgroups in different demographic strata have been highly variable, and the unsystematic results do not

enhance by much our understanding of the prevalence of cigarette smoking among the different subgroups in the population. Only a few studies specifically investigated the relationships between cigarette smoking and the demographic interactions, even though the findings of tobacco studies usually presented complex smoking patterns among different groups of people (Novotny et al., 1988; Rogers, 1991; Rogers et al., 1995).

In Novotny et al.'s study (1988), race-sex and race-education interaction were included in their original multivariate logistic regression model, but these two interaction terms were dropped in the final model because of statistical non-significance. In the study of Rogers et al (1995), they investigated the interactions among age, sex, ethnicity and smoking habit, which included different levels of cigarette smoking, from never smokers to moderate-to-heavy smokers. They found that age-ethnicity, age-sex, and ethnicity-sex interactions were statistically significant in their polychotomous logistic regression model. This study provided great detail about the smoking levels in different subgroups. For example, regarding age and sex, they found that there was a sex difference between peaks of ages for smoking statuses. Women's smoking peaked at earlier ages than men's. Also, several age cross-over effects of the two sexes in the smoking statuses were displayed in the study. Similarly, the authors created the same graphs for age-ethnicity and ethnicity-sex interactions, and they provided great detail about smoking habits in the various subgroups related to different demographic factors. However, these detailed results could be sample-dependent and the majority of them seemed not to be interpretable. Because of lack of converging evidences and better theories for understanding other interaction terms, only the sex-ethnicity interaction is investigated in this study.

Among all the demographic interactions, studies have shown that the relationship between cigarette smoking and sex varies by different racial/ethnic groups. Consistently, studies have shown that Hispanic women were much less likely to smoke cigarettes than other racial/ethnic subgroups (Maher et al., 2005; Markides, Coreil, & Ray, 1987; Rogers, 1991; Rogers et al., 1995). Relevant studies suggested that acculturation is related to cigarette smoking among Latina women as mentioned in the race/ethnicity section. The results of the recent study on Latino population (Maher et al., 2005) indicated that the extremely low smoking prevalence among Latina women taking the survey in Spanish had driven the low prevalence of smoking rate among Latinos. It is probable that cigarette smoking is not common among traditional Hispanic women and that the social norm is strongly against cigarette smoking for this group (Rogers, 1991).

Relevant Theories

Summarizing and integrating related findings, researchers proposed different theories to explain the relationships between cigarette smoking and important demographic factors. Social-contextual model (Barbeau et al., 2004; Sorensen et al., 2004), cigarette diffusion models (Pampel, 2005), psycho-social stress model (Elstad, 1998), and life history theory are among the few efforts to triangulate the complex smoking disparity among socio-economically disadvantaged populations. These models provide systematic ways to investigate the plausible mechanisms underlying the relationship between smoking and related demographic characteristics. All of these models will be reviewed in this section briefly to provide a systematic understanding of the relationship between smoking and related demographic characteristics.

Sorensen et al. (2004) proposed a social contextual model for cessation intervention in tobacco use to reduce class-based smoking disparities. These researchers thought that smoking disparities reflect different social and material factors within the larger societal structure where individuals are located. Social contexts of people's lives, related to the socioeconomic position of an individual, are the fabric and texture of day-to-day experiences and realities, and it may be influenced or limited by the structural forces in the community or society. Ultimately, these social, material, or structural factors may have effects on health and health behaviors, such as cigarette smoking. Socioeconomic position can be defined as a social relationship based on individuals' structural location within the economy or society, and it determines one's prospects in life, exposures to stressors in life, and access to social, educational, and economic resources. As mentioned before, smoking is associated with low education, blue-collar class, or low income, and some social-context factors tend to cluster with these things, such as unemployment, lack of social or material support, living in unsafe neighborhoods, and no medical insurance. In this specific study (Sorensen et al., 2004), the researchers used this social context model to develop worksite-based smoking cessation approaches for blue-collar workers, the lowest occupational ranks. They combined the social-ecological perspective, relevant mediating mechanisms, and the modifying conditions based on the social context model in their intervention design. The modifying conditions, representing the influence of socioeconomic status on smoking patterns, include material circumstances, daily stressors, and functional meaning of smoking in the individual level, social ties, social norms, and family roles in the interpersonal level, job conditions,

organizational support for behavior change, and hazardous workplace exposures in the organizational level, access to neighborhood resources, housing, and tobacco advertising in the neighborhood/community level, and lastly, regulations and policies at the societal level. These researchers presented the empirical data to support the potential effect of the cessation intervention program for blue-collar workers based on the social-context model, and it seems to be a promising and useful framework to enhance our understanding of the relationship between social-economic position and tobacco use.

The cigarette diffusion approach was proposed by Pampel (Pampel, 2001; Pampel, 2005) to explain not only the relationship between smoking and relevant socioeconomic status or social characteristics, such as sex, race, or age, but also the fact that these social determinants of smoking tend to vary over time. This model is used to explain why people in socio-economically disadvantaged groups tend to smoke more and why the smoking trend first emerged among high status groups but has become increasingly concentrated among low status groups. Two closely related mechanisms based on social status were suggested: first, diffusion of innovations in a population via status-based communication networks and, second, a status-based cycle of adoption and replacement of fashions. The high status groups are most open to innovations and have more resources to adopt things like smoking, which costs money. Also, in a cycle of fashions, the upper classes, who use symbolic means to distinguish themselves from other classes, are often the first to adopt fashions. Later on, the middle and lower classes successively imitate and follow the higher class to adopt the new fashion. However, as the fashion becomes widespread across all classes and loses its distinctiveness, new fashions emerge among

the upper classes, and the cycle repeats. When smoking first emerged, it was viewed as a new fashion by high status groups who first adopted the habit, but later smoking came to lose its distinctiveness among this elite group as it diffused to the rest of the population. In the later stages, smoking declined first among high status people, because the allure of cigarette smoking has faded by diffusion, and the upper class have learned about the results of tobacco research, become concerned with health hazards of smoking. The high status groups first separate themselves from other groups via quitting smoking or giving up other unhealthy lifestyles. Again the innovation of adopting healthy lifestyles first emerges among the high socioeconomic groups. Pampel showed that the historical literature supported the idea that high status groups began smoking earlier than the general public, and later the patterns of imitation lead to diffusion of smoking practice, and the lower status group took up smoking after the high status groups. Based on this cigarette diffusion approach and recent studies, Pampel hypothesized that education, parents' education, and parents' income should lower smoking, while urban residence during adolescence, being male, and being White should raise smoking. More importantly, he hypothesized that much variation in smoking across cohorts exists and that the decline in smoking across cohorts should be larger among high status groups than low status groups. Pampel defined cohort in terms of the stage of cigarette diffusion when a cohort reaches adolescence and used multilevel modeling to test the cohort differences in the social patterns of smoking. He expected that education, parents' education and income, adolescent urban residence, sex, and race would influence smoking at the individual level, and that cohort characteristics would affect the determinants of smoking

at the aggregate level. Cohort was the unit of the second level in the multilevel analysis, and the effect of status variables on smoking was allowed to vary across cohorts. The results of the multinomial logistic model for the combined multilevel model basically supported the hypotheses based on the cigarette diffusion approach. In older cohorts and early stages of cigarette diffusion, smoking increased with socioeconomic status for young adults, but in younger cohorts and later stages of cigarette diffusion, smoking decreased with pre-adult socioeconomic status. The only exception in this study was the effect of race. Smoking changed little across races for different cohorts and stages of diffusion. The match between theory and this empirical study showed some explanatory value of this cigarette diffusion approach.

Psycho-social perspective has been used to explain current social inequalities in health (Elstad, 1998). Based on psycho-social perspective, an indirect pathway was proposed from stress through health-damaging behavior to poor health. People may react to adverse circumstances, such as deprivation in a socio-economic situation, by excessive alcohol use, smoking or other self-destructive behaviors as coping functions or mood-regulating tools (Kassel, Stroud, & Paronis, 2003; Spielberger & Reheiser, 2006). Some researchers specifically investigated the relationship between stress and tobacco use and suggested that cigarette smoking can be a habit in response to stressful conditions (Colby, Linsky, & Straus, 1994). Studies have showed that smokers tended to smoke more heavily under stressful situations (Conway et al., 1981; Schachter, 1977; Schachter, Silverstein, & Perlick, 1977). It is possible that socio-economically disadvantaged

population experiences more life stressors and lack of sufficient material or social support, and hence they adopt smoking more to cope with the stressful situations.

From the evolutionary perspective, life history theory provides a broad framework to understand the relationship between cigarette smoking and demographic variables, assuming that cigarette smoking reflects risk-taking behavior or social problem behaviors in general (Figueredo et al., 2006; Hill & Chow, 2002). Life history theory proposes an r/K continuum to represent reproductive behavior patterns, which are related to social problem behaviors. The theory predicts that people living in unstable and unpredictable environments tend to develop clusters of 'r-selected' or low-K traits, which include reproductive strategies such as high reproductive rates, low parental investment, and relatively short intergeneration time. This r/K continuum can be used to predict the cluster of 'social problem' behaviors. For example, young adults usually are higher in r or low in K, and tend to be more risk-taking than women or older men. Based on life history theory, the relationship between cigarette smoking and relevant demographic variables can be specified and further tested. We predict that people who are low-educated, low-income, blue-collar workers or unemployed, Black, or young male may be more likely to smoke.

Summary and Framework of Present Study

Based on the empirical evidence and pertinent theories, education, income, employment status, sex, age, race/ethnicity, and the interaction between race and sex are deemed relevant to predict current smoking status. However, the order of importance for these demographic factors and the different categorizations of the variables in predicting

cigarette smoking remain uncertain with all the relevant information at hand, since not much relevant information has been revealed in the literature. In order to target the right population for tobacco control effort, the main purpose of this study is to investigate the order of predictive power of these variables in relation to cigarette smoking and classify the smokers in a conceptually or statistically meaningful way. Alternative hypotheses would be tested in two approaches of data analysis in this study. Multiple working hypotheses would be specified to test the relative importance of the demographic variables of interest.

Overall Hypotheses

The two hypotheses regarding the order of predictive power of the relevant demographic variables are as follows:

A) Education, employment status, income, race, sex, age, and interaction between sex and race/ethnicity

B) Education, income, employment status, race, sex, age, and interaction between sex and race/ethnicity

Specific Hypotheses for Individual Predictor

Within each demographic predictor, hypotheses related to the important predictive categories were specified based on our understanding. The hypothesized ignorable categories are not listed here. Among these demographic factors, the best predictors in education, race and age are not always consistent; therefore, multiple working hypotheses regarding these variables are specified.

Education. Three working hypotheses regarding the best categorization of education in predicting cigarette addiction are as follows:

A. Considering high school graduates as the most important category in education, the following groups will be compared:
    1) At least high school graduation (including Master, Bachelor, AA Academic, AA Technical/vocational, some college, and high school graduate) versus GED and lower education (including GED, Grades 9 through 11, Grades 1 through 8, and never attended school)
    2) Among high school graduates and above, respondents with at least a Bachelor degree (including Master and Bachelor) versus participants without a Bachelor degree (including AA academic, AA technical/vocational, some college and high school graduates)
    3) Among the respondents with GED and lower education, GED and some high school (including GED and Grades 9 through 11) versus participants with no high school education (including Grades 1 through 8 and never attended school)
    4) Participants with GED versus Grades 9 through 11
B. Considering the Bachelor degree as the most important category in education, the following groups will be compared:
    1) Respondents with at least a Bachelor degree (including Master and Bachelor) versus no Bachelor degree (including AA academic, AA

Technical/vocational, some college, high school graduate, GED, Grades 9 through 11, Grades 1 through 8 and never attended school)

2) Among the respondents with no Bachelor degree, high school graduates and above (including Master, Bachelor, AA academic, AA technical/ vocational, some college, and high school graduate) versus GED and lower education (including GED, Grades 9 through 11, Grades 1 through 8, and never attended school)
3) Among the respondents without a high school degree, participants with some high school education (including GED, Grades 9 through 11) versus no high school education (including Grades 1 through 8 and never attended school)
4) Participants with GED versus Grades 9 through 11

C. Considering the GED as the most important category in education, the following groups will be compared:

1) GED and above (including Master, Bachelor, AA academic, AA technical/ vocational, some college, high school graduates and GED) versus some high school or less education (including Grades 9 through 11, Grades 1 through 8 and never attended school)
2) Among respondents with some high school or lower education, some high school (including only Grades 9 through 11) versus lower than high school (including Grades 1 through 8 and never attended school)

3) Among respondents with at least GED education, high school graduates and above (including Master, Bachelor, AA academic, AA Technical/vocational, some college and high school graduates) versus GED only
4) Among high school graduates, participants with at least Bachelor degree (including Master and Bachelor) versus without a Bachelor degree (including AA Academic, AA Technical/vocational, some college, and high school graduates)

Employment. The only distinction related to employment is of Unemployed/unable to work (including Out of work for > 1 year, out of work for < 1 year, and unable to work) and non-unemployed (including employed for wages, self-employed, homemaker, student, and retired) is important in predicting cigarette addiction.

Income. The cutting point of $20,000 in annual income (approximately the poverty level for a family of four (U.S. Department of Health and Human Services, 2007)) was used as the threshold for low-income in this study. We hypothesized that only the differentiation of low-income was important in predicting cigarette smoking. So the low-income group (including less than $10,000, between $10,000 and $14,999, and between $15,000 and $19,999) was compared with the non-low-income group (including between $20,000 and $24,999, between 25,000 and $34,999, between $35,000 and $49,999, between $50,000 and $74,999, and more than $75,000).

Race/Ethnicity. Two working hypotheses for categorizing race are set up as follows:

A. Considering White as the most important category of race/ethnicity in predicting cigarette addiction, the following groups will be compared:
    1) White, non-Hispanic versus non-White (including Black or African American, Hispanic, Asian, Native Hawaiian or Pacific Islanders, American Indians and other racial groups)
    2) Among participants who are non-White, Hispanic versus non-Hispanic (including Asian, Native Hawaiian or Pacific Islanders, American Indians and other)

B. Considering Hispanic as the most important racial predictor for cigarette addiction, the following groups will be compared:
    1) Hispanic versus non-Hispanic (including White, non-Hispanic, Black or African American, Asian, Native Hawaiian or Pacific Islander, American Indian, and other)
    2) Among participants who are White, non-Hispanic versus non-White (including Black or African American, Asian, Native Hawaiian or Pacific Islander, American Indian, and other)

Age. Three hypotheses are specified as follows:

A. Considering the age group 18-24 year-old as the most important category in predicting cigarette addiction, the following groups will be compared:
    1) Age of 18-24 versus age 25 and older
    2) Among age group 25 and older, age of 25-44 versus age of 45 and older)

3) Among age group 45 and older, age of 45-64 versus age of 65 or older

B. Considering the age group 65 or older as the most important category, the following groups will be compared:

1) Age of 65 or older versus other younger groups (including age of 18-64)

2) Among respondents younger than 65 , age of 18-24 versus age of 25-64

3) Among age 25-64, age of 25-44 versus age of 45-64

C. Considering the age group 25-44 as the most important category, the following groups will be compared:

1) Age 25-44 versus the other age groups (including age of 18-24, and age of 45 and above)

2) Among respondents whose age were not between 25-44, younger groups (including age of 18-24, and 45-64) versus age of 65 and older

3) Age of 18-24 versus age of 45-64

Interaction between sex and race/ethnicity. Two combinations were created because of the two hypotheses regarding racial categorization, but only Hispanic Women is hypothesized to be important in predicting cigarette smoking.

METHOD

Design and Procedure

This study is based on a secondary analysis of survey data, the Adult Tobacco Survey (ATS), a cross-sectional survey conducted in Arizona every 3 years by the Arizona Department of Health Services Office of Tobacco Education and Prevention Program (ADHS TEPP). Arizona ATS is a computer-assisted telephone interview, based on a random-digit dial (RDD) sample, conducted by the Northern Arizona University Social Research Laboratory (SRL). ATS 2002 and 2005 data sets were used for this study. The sampling schemes used in ATS 2002 and 2005 were different. The details of the sampling procedures were documented in the official reports of ATS (Arizona Department of Health Services Office of Tobacco Education and Prevention Program, 2006; Walsh et al., 2003), which are available on-line. In order to obtain a better representative sample of Arizona residents, the two data sets of ATS 2002 and 2005 were merged and analyzed for this study.

Participants

The respondents of ATS were at least 18 years of age or older, English speaking, and residing in households with telephones. In the ATS 2002 and 2005 sample, 6,025 and 6,071 persons were interviewed, respectively (The number of respondents of 2002 is different from the ATS 2002 executive report because of some error found in the dataset after the report was released. The number 6,025 was the corrected number for the final dataset of ATS 2002). The completion rates for the surveys were 67% for ATS 2002, and 63% for 2005. In total, 12,096 respondents were interviewed in Arizona across both years,

but not all of them answered all of the questions regarding smoking status and related demographic questions. There were 354 respondents who did not answer the education question, 1986 for income, 84 for work status, 451 for race, 87 for sex, and 490 for age. The regression analyses were restricted to the complete data set in terms of the current smoking status and the demographic variables of interest. The final, complete merged data set included 9,081 respondents from ATS 2002 and 2005.

Measures

The questionnaires used for ATS 2002 and ATS 2005 were slightly different, reflecting revisions of some items in ATS 2005. The variables of interest in this study include questions related to smoking behaviors and demographic characteristics. Most of these questions were asked in the same way in the ATS 2002 and ATS 2005, but a few of the questions were changed. The modified questions involved education, working status, race, and smoking policy at work. The definitions of measures and the reclassification rules for merging the variables from the modified questions are listed as follows.

Smoking Behaviors

The standardized measure of smoking status developed by Centers for Disease Control and Prevention (CDC) was assessed by two questions, asking “Have you smoked at least 100 cigarettes in your entire life?” and “Do you now smoke cigarettes everyday, some days, or not at all?” Current smokers were defined as having smoked at least 100 cigarettes in their lives and smoke “everyday” or “some days” at the time of the interview.

Other questions related to individual smoking behaviors include smoking frequency, duration, and smoking history. Respondents who were current smokers

answered the questions about how many days they smoked in the past 30 days (at the time of interview), on average how many cigarettes they smoked in the past 30 days, when they usually have the first cigarette after they wake up (within 5 minutes, 6-30 minutes, 31-60 minutes, or after 60 minutes), at what age they smoked their first cigarette, the age at which they became regular smokers, and the number of years they have been smoking or since they became regular smokers.

Demographic Characteristics

Education was measured by the highest level of education completed. Ten levels were listed as the responses for the question of education in ATS 2002, but eleven levels were listed in 2005. The highest two levels in ATS 2005 were collapsed into one level to be compatible to the highest level for the ATS 2002. The ten levels of education were, "Never attended school or only attended kindergarten', "Grades 1 through 8 (Elementary)", "Grades 9 through 11 (Some high school)", "Grade 12 (High school graduate)", "GED (General Education Degree)", "Some college, no degree", "AA (Associate of Arts, 2 years degree) technical/vocational", "AA (Associate of Arts, 2 years degree) academic", "BA, BS (Bachelor of Arts or Science; 4 years; college graduate), and "Master of Arts, Master of Science (Graduate School) and above".

Working status/unemployment was asked differently in ATS 2002 and 2005. The categorization of responses was finer in ATS 2005, so the responses of working status in ATS 2005 were collapsed into the categories in ATS 2002, including "employed for wages", "self-employed", "out of work for more than 1 year", "out of work for less than 1 year", "a homemaker", "a student", "retired", "unable to work".

Income was categorized in 8 levels in the questionnaires, “Less than $10,000”, “Between $10,000 and $14,999”, “Between $15,000 and $19,999”, “Between $15,000 and $19,999”, “Between $20,000 and $24,999”, “Between $25,000 and $34,999”, “Between $35,000 and $49,999”, “Between $50,000 and $74,999”, and “More than $75,000”. Based on the 2007 Health and Human Services (HHS) Poverty Guidelines (U.S. Department of Health and Human Services, 2007), for a family or household of 4 persons, the simplified poverty threshold is $20,650. Therefore, the cut point for low income was set at $20,000.

The race and ethnicity question was asked differently in ATS 2002 and 2005. In ATS 2002, two separate questions were asked regarding race and ethnicity, but only one question was asked in ATS 2005. The two questions in 2002 were about whether a respondent is Hispanic, Latino, or Spanish in origin, and what race the respondent considers himself/herself to be. In 2005, the question was “which of the following categories describes you best? Are you…”, and seven categories of responses were listed, including “White, non-Hispanic”, “Black or African American, Non-Hispanic”, “Hispanic”, “Asian”, “Native Hawaiian or other Pacific Islander”, “American Indian, Alaska Native”, and “Other”. In order to combine the race/ethnicity information into one variable in the merged dataset, two race/ethnicity variables in ATS 2002 were recoded based on the categories used in ATS 2005. The respondents who considered themselves to be White and ‘not’ Hispanic in origin were re-categorized as “White, non-Hispanic”; the respondents identified themselves as Hispanic in origin were re-categorized as “Hispanic”; and the respondents considering themselves to be Black or African American

and 'not' Hispanic in origin were re-categorized as "Black or African American" in the merged dataset.

Age was calculated according to the year the respondent was born and it was also categorized into different age groups, "18-24 years", "25-34 years", "35-44 years", "45-54 years", "55-64 years", and "65 years or older".

Statistical Analysis

Weighting was performed in the original ATS reports to adjust for region, age, sex, and race so as to reflect the proportion of various population groups in the state of Arizona. In this study, our main interest is in the relationship between cigarette smoking and demographic predictors, so the un-weighted, merged dataset of ATS 2002 and ATS 2005 was used for our analysis.

Datasets merging for ATS 2002 and 2005 was carried out in the statistical package R.2.2.1 (R Development Core Team, 2005). Logistic and liner multiple regression analyses were performed using SAS/STAT software, version 9.1 of the SAS System for Windows (SAS Institute Inc., 2003).

Data Merging

Data merging of ATS 2002 and 2005 was an important and challenging task for this study, because of the different numbers of questions across the years and slightly different wording for some questions in ATS 2002 and 2005 datasets. In the original datasets, the variable names were the number of the questions in the questionnaires of the two different ATSs, and, hence, the variables did not match with each other. New common variable names were assigned for the questions asking about the same

information. If the response categories were different across the surveys, the response categories were collapsed based on the common categories of the two surveys. The two recoded datasets of the same variables were merged into one combined dataset of ATS 2002 and 2005.

Logistic Regression Analyses

The binary outcome variable 'current smoker' for logistic regression was coded 1 for the respondents who were current smokers at the time of interview, and 0 for non-current smokers, which included former smokers and never smokers. The predictors were respondent's highest educational attainment, income level, work status/unemployment, race, sex, age, and the interaction between race and sex. In order to differentiate the relative importance of these predictors, sequential logistic regression models were performed to test the hypotheses specified in the section 'summary and framework of present study'.

Multiple Regression Analyses

For performing appropriate multiple regression analyses, a new continuous criterion variable was created to reflect the strength of the cigarette smoking habit. There were seven questions asking about cigarette smoking in the ATS 2002 and 2005 questionnaires, including 'Have you smoked at least 100 cigarettes in your entire life? (100 cigarettes)', 'Do you now smoke cigarettes everyday, some days, or not at all? (smoking frequency)', 'How soon after you wake up do you have your first cigarette? (first cigarette of the day)', 'How old were you the first time you smoked a cigarette, even one or two puffs? (age of smoking the first cigarette)', 'How old were you when

you first started smoking cigarettes regularly? (age of becoming a regular smoker)', 'On how many of the past 30 days did you smoke cigarettes? (days of smoking in last month)', and 'On the average, on days when you smoked during the past 30 days, about how many cigarettes did you smoke a day? (number of cigarettes per day)'. In order to make the best use of the information we obtained from both ATS 2002 and 2005 to create the new outcome variables, the relevant analyses were performed on the full merged dataset, in which missing data were included.

The skip pattern of the ATS was contingent on the question about 100 cigarettes for the entire life, so, this question was excluded when creating the new outcome variable. Two new variables regarding smoking legally or illegally were generated based on the age of first smoking a cigarette and the age of becoming a regular smoker. If the respondent had smoked the first cigarette (first cigarette under age of 18) or become a regular smoker under the age of 18 (becoming a regular smoker under age of 18), the new corresponding variables were coded 1 to signify the status of smoking cigarettes illegally at the beginning of the respondent's smoking habit. The two variables regarding age of the first cigarette and first becoming regular smokers were also recoded into years of smoking and years of being a regular smoker to reflect the strength of the cigarette smoking habit. Since years of smoking and years of being a regular smoker are highly correlated with the ages of the smokers, the two recoded variables are divided by the ages of the respondents to correct for the influences of age. The values for the variables 'smoking frequency' and 'first cigarette of the day' were recoded inversely to match the direction of the cigarette smoking habit, consistent with other smoking variables, with a

smaller number representing weaker cigarette smoking habit and a larger number for stronger cigarette smoking habit. Only smokers were asked to answer these smoking related questions, so there was no information in these variables for the non-smokers, and thus these variables for them became missing data. Logical imputation was carried out to deal with the missing data problem.

The details of the logical imputation are as follows. If the respondents responded that they have not smoked at least 100 cigarette for the entire life, the value of the variable 'smoking frequency' was recoded as '3', meaning 'not at all' for the non-smokers. If the respondents reported having not smoked at least 100 cigarettes for the entire life or 'not at all' for smoking frequency, the value of the variables 'first cigarette under age of 18', 'becoming a regular smoker under age of 18', 'days of smoking in last month', 'number of cigarettes per day', 'years of smoking', and 'years of being a regular smoker' were all recoded as '0', and the value of the variable 'first cigarette of the day' was recoded as '5', one level above the original highest response category 'after 60 minutes', which was recoded to '0' in the inverse scoring of the raw data, to indicate no cigarette smoking habit.

In total, there were eight variables related to strength of cigarette smoking habit in the merged dataset. The correlations between these variables are quite high, from .6, correlation between 'number of cigarettes per day' and 'becoming a regular smoker under age of 18', to .98, correlation between 'days of smoking in last month' and 'smoking frequency'. The high correlations between these smoking variables indicated that latent variable model can be an appropriate model to construct a scale for strength of cigarette

smoking habit. Exploratory Factor Analysis (using the SAS command PROC FACTOR, largest absolute correlation for a variable with any other variable for initial communality estimate, and primax for oblique rotation), was conducted on these eight observed variables to generate the smoking habit factors. Based on the content domain of the cigarette smoking items, one-factor and three-factor EFA models were performed on the raw data of cigarette addiction. The results of these factor analysis models are presented in the Table 1. Since most of the respondents were not smokers, the strength of cigarette smoking habit variable is determined mainly by the non-smokers. In case these variables only reflected the status of being a current smoker, we residualized these variables on the current smoking status. Current smoker status was used to predict all the cigarette smoking habit variables, and the residuals of these variables were factor analyzed to test the correspondence of smoking habit variables and current smoking status. One-factor and three-factor EFA models were performed on the residualized cigarette smoking habit variables and the results are listed in the Table 2.

The results of the Exploratory Factor Analysis showed that one-factor model fit the original eight variables of cigarette smoking habit better, and the three-factor model fit the residualized variables of smoking habit better. It represents that these eight variables are overlapped with the current smoking status, but they also carry unique information after being residualized on current smokers. Reliability analysis of internal consistency, coefficient alpha, was carried out for one-factor and three-factor scales. Table 3 displays the standardized coefficients alphas of different scale compositions. The coefficient alphas of these scales are fairly high, all around .90. Based on the content

Table 1. Factor Loadings for Cigarette Smoking

EFA models	One-factor model	Three-factor model		
Variables	Cigarette addiction	Current smoking	Smoking Duration	Illegal Smoking
Smoking frequency	0.98	0.44		
First cigarette of the day	0.90	0.79		
Days of smoking in last month	0.97	0.46		
Number of cigarettes per day	0.87	0.76		
Percent of years smoking	0.97		0.75	
Percent of years being a smoker	0.96		0.79	
First cigarette under age of 18	0.89			0.56
Being a regular smoker under 18	0.77			0.82

Table 2. Factor Loadings for Cigarette Smoking Residualized on Smoker Status

EFA models	One-factor model	Three-factor model		
Variables after being residualized on current smoking status	Cigarette addiction	Current smoking	Smoking Duration	Illegal Smoking
Smoking frequency R.	0.76	0.89		
First cigarette of the day R.	0.77	0.80		
Days of smoking in last month R.	0.70	0.83		
Number of cigarettes per day R.	0.62	0.58		
Percent of years smoking R.	0.62		0.91	
Percent of years being a smoker R.	0.66		0.91	
First cigarette under age of 18 R.	0.25			0.74
Being a regular smoker under 18 R.	0.32			0.72

domains and the results of factor analysis and coefficient alphas, we think that the three-factor model is more informative in helping us to understand the construct of cigarette smoking habit. Four of the eight variables, including smoking frequency, first cigarette of the day, days of smoking in the last month, and the average number of cigarettes per day, composed the subscale *current smoking*. Two variables, percent of years smoking (after the first cigarette) and percent of years being a regular smoker, composed the subscale *smoking duration*. The two other variables, first cigarette under age of 18 and being a

Table 3. Coefficient Alphas for the Overall Cigarette Smoking Habit and Subscales

Constructed Scales	Number of variables	Standardized Coefficient Alpha
Cigarette smoking habit (one-factor)	8	0.97
Current smoking	4	0.97
Smoking Duration	2	0.99
Illegal Smoking	2	0.89

regular smoker under age of 18, composed the subscale *illegal smoking*. Three composite scores were constructed based on these component variables, and one score for the higher-order factor was created accordingly. Since these eight variables were not in the same metric, the three subscale scores were calculated by taking the means of the standardized component variables. Then the mean of the three subscale scores was taken as the overall, higher-order score representing strength of cigarette smoking habit.

Based on understanding that most of the participants were non-smokers, we know that the distribution of the cigarette smoking habit variable will not be normal. The distributions of the three subscales and the overall cigarette smoking habit show that these scores are not normally distributed. Also, the three subscale scores were residualized on overall smoking habit, and overall cigarette smoking habit was residualized on all the demographic predictors, using the contrast coding described earlier. All of these residuals were not normal-distributed. However, the new composite cigarette smoking variables were not transformed to correct the problem of non-normal

distribution, since significance testing is not really informative in this study because of the large sample size. Therefore, the new non-transformed composite variables were used as the criterion variables in the multiple regression analysis.

In order to decide whether the overall cigarette smoking habit should be the only dependent variable, or whether the three subscales scores should be considered as well in the linear multiple regression models, we used an exploratory approach here to look at the relationship between the residuals of each subscale after regressing on the total cigarette smoking habit and the demographic predictors. The results are shown in Table 4. After taking out the overall cigarette smoking habit factor from the subscale of current smoking, almost no additional variance can be accounted for by the demographic predictors. Only 1% of the remaining variance is explained by the demographic predictors. Race/ethnicity, education, interaction and age turn out to be more important in predicting the unique component of current smoking, findings different from, although only trivially so, the specified overall hypotheses. For the unique component of subscales of smoking duration and illegal smoking, about 5 to 10% of the variance can be explained by the demographic predictors. The patterns of importance of the demographic predictors are very similar for these two subscales, and age turns out to be the most important predictor.

It is possible that the three subscales capture different information in addition to the overall cigarette addiction factor. However, not much variance in residualized current cigarette smoking is left to be accounted for by the demographic variables of interest, and the variance accounted for in the subscales for smoking duration and illegal smoking

Table 4. Type I SS for the Models of Residuals

DV / Contrast	Current smoking	Smoking duration	Illegal smoking
Education	1.8 / 21%	2.7 / 4%	5.4 / 8%
Employment	0.5 / 6%	11.7 / 18%	13.7 / 21%
Income	0.8 / 9%	0.4 / 1%	0.9 / 1%
Race/ethnicity	3.1 / 36%	1.7 / 3%	1.0 / 1%
Sex	0.1 / 1%	0.4 / 1%	0.1 / 0%
Age	1.1 / 13%	47.2 / 73%	43.0 / 66%
Interaction	1.2 / 14%	0.6 / 1%	1.5 / 2%
R^2	0.01	0.05	0.09

seems to be negligible. Therefore, only the new overall variable, cigarette smoking habit, is used as the main dependent variable in the linear multiple regression analyses to test the hypotheses of this study.

Contrast Coding

Because of the nominal nature of the demographic predictors, orthogonal contrast coding with nested design was used to make the best use of the information from the data in this study. A set of contrast coefficients was written to represent the categorization within each individual demographic predictor. For the variables education, race, and age, multiple sets of contrasts were specified for alternative hypotheses. Within each set of the contrasts, g-1 contrasts were used to exhaustively represent the membership of the g

groups in the specific predictor. The sum of the contrast coefficients within a contrast needs to be zero, and also, for each pair of the contrasts, the sum of the products of the coding coefficients needs to be zero (Gorsuch, 1983). The contrasts were set up according to the hypotheses specified in this study (See Appendix A).

RESULTS

Logistic Regression

The statistical 'model -2 Log Likelihood (-2LL)' is analogous to the error sum of squares, SSE, in linear regression analysis (Menard, 1995). The model -2LL for intercept and covariates for each added predictor would be used for comparison in this logistic regression analysis. The difference in reduced -2LL that a set of contrasts for a demographic predictor makes reflects the effect size of that demographic factor in predicting the status of current smoker, and hence is used as the main result in the analyses that follow.

Multiple Working Hypotheses for Education, Race/Ethnicity, and Age

Education. The -2LLs of the inclusive sequential logistic regression models with different sets of contrasts for the predictor 'education' are shown in table 5. Among the three contrast sets of education, the contrast set B (Ec) worked as hypothesized. The first contrast, bachelor degree or not, which was hypothesized to be the most important contrast in the set, explained the most variance in current smoking among the three contrast sets. Having a bachelor degree is the best predictor of current smoking status. Therefore, the contrast set Ec is used in the subsequent logistic regression analysis.

Race. Table 6 shows the -2LLs for the two sets of contrasts for race. Contrast A (R) performed slightly better than contrast B (Rh). It suggests that being a White, non-Hispanic or non-White can predict current smoking better than being a non-Hispanic or Hispanic, but the predictive power of these two categorizations is similar. Even though the resulting difference between these two contrast sets is small, the contrast set R is used

Table 5. Reduced -2LLs for the Contrast Sets of Education

Contrast set A	-2LL	Contrast set B	-2LL	Contrast set C	-2LL
Ed1	52.5	Ec1	212.5	Eg1	23.7
Ed2	181.4	Ec2	21.4	Eg2	3.5
Ed3	6.0	Ec3	6.0	Eg3	34.5
Ed4	3.1	Ec4	3.1	Eg4	181.4

Table 6. Reduced -2LLs for the Contrast Sets of Race/Ethnicity

Contrast set A	-2LL	Contrast set B	-2LL
R1	62.6	Rh1	60.9
R2	9.6	Rh2	11.2

Table 7. Reduced -2LLs for the Contrast Sets of Age

Contrast set A	-2LL	Contrast set B	-2LL	Contrast set C	-2LL
A1	6.3	Ag1	101.8	Am1	10.0
A2	18.9	Ag2	3.7	Am2	95.8
A3	88.9	Ag3	8.6	Am3	8.2

in the following logistic models.

Age. The reduced -2LLs for individual contrasts for different sets of age contrasts are presented in Table 7. Among the three contrast sets, contrast B (Ag) performed the

best. The first contrast of Ag explained the highest proportion of the variance in current smoking accountable by age. It shows that being an elderly (age of 65 or older) or not can predict current smoking status the best. Contrary to the hypotheses, within the younger groups (age below 65), the predictive power of being a young adult is not as good for being a middle-aged adult. The contrast set Ag will be used in the subsequent logistic regression models.

Model Comparison

Testing overall hypotheses- the inclusive model. The logistic regression analysis shows that the inclusive model of demographic predictors can only explain about 7% of the variance in current smoking status. Based on the two hypotheses regarding relative importance of the demographic predictors, two sequential logistic regression models were tested. The results are displayed in table 8. It shows that model A is a better model, that is, employment is a better predictor of current smoking than income. However, the relative importance of the predictors is not exactly the same as predicted. The results suggest that the order of importance of the demographic predictors for current smoking is education, employment, age, race, sex, income and interaction. Among these predictors, interaction, income, sex, and income are not good predictors, since all of them together explain less than 1% of the variance of current smoking status.

Testing specific hypotheses- the restricted model. Table 9 shows the results of the hypothesized restricted model. These hypothesized important contrasts can explain a total of 6% of the variance in current smoking, close to the amount (7%) accounted by the inclusive model. It suggests that most of the contrasts work as predicted in the hypotheses.

Table 8. Two Inclusive Models Regarding the Order of Employment and Income

Model A				Model B			
Predictor	df	-2LL	Added R^2	Predictor	df	-2LL	Added R^2
Education	9	273.9	0.028	Education	9	273.9	0.028
Employment	7	168.0	0.017	Income	7	29.9	0.0029
Income	7	40.7	0.0039	Employment	7	178.8	0.018
Race/ethnicity	6	74.1	0.0072	Race/ethnicity	6	74.1	0.0072
Sex	1	31.2	0.003	Sex	1	31.2	0.003
Age	5	124.3	0.012	Age	5	124.3	0.012
Interaction	6	6.3	0.0006	Interaction	6	6.3	0.0006

For example, the four deemed-to-be-important contrasts of education explain over 85% of the variance accountable for by education, and the three contrasts of age seemed to contribute to all of the variance of current smoking predicted by age. However, in order to make sure that no important contrasts were missed or no irrelevant contrasts were included in the restricted model, an exploratory approach was adopted for hypothesis testing and model comparison.

*An exploratory approach- the modified restricted model. A*n exploratory approach was adopted to select the contrasts for the modified restricted model. The statistic -2LL is used as an indicator for relative contribution of each contrast. A scree-type plot is generated for the relative contribution of each contrast in terms of Chi-square. Figure 1 shows the plot for the contrast set of education. Based on the plot, only the first two

Table 9. Reduced -2LLs for the Hypothesized Restricted Model

Predictor	df	-2LL	Added R^2	Sum of added R^2
Ec1	1	212.5	0.021	0.025
Ec2	1	21.4	0.0022	
Ec3	1	6.0	0.0006	
Ec4	1	3.1	0.0003	
Em1	1	77.5	0.0077	0.0077
I1	1	1.3	0.0001	0.0001
R1	1	30.2	0.0030	0.0039
R2	1	9.2	0.0009	
Sex	1	35.8	0.0035	0.0035
Ag1	1	183.9	0.0178	0.020
Ag2	1	10.2	0.0009	
Ag3	1	10.8	0.0011	
In2	1	2.4	0.0002	0.0002

contrasts are important in the whole set of education contrasts. Similarly, figures 2 to 6 are the plots for the rest of the contrast sets of the demographic predictors, and the cut points for employment, income, race, age, and interaction are 2, 2, 1, 1, and 3, respectively. For example, the most important contrast within income is the second contrast: among the respondents with income more than $20,000, the highest annual

income (above $75,000) versus below $75,000. The resulting sequential logistic regression model is presented in table 10.

The result shows that, sequentially and respectively, income, race, sex and interaction can explain less than 0.5% of the variance in current smoking status. Therefore, we include only education, employment, and age as our predictors in the final modified restricted model in logistic regression, excluding these four predictors and also the relatively not-so-important education contrast two. The results of the modified restricted model are shown in table 11.

The results suggest that for education, the accomplishment of a bachelor degree is the most important predictor for current smoking. Even though we did not include the second contrast in education in the final modified model, the exploratory results showed that among the respondents without a bachelor degree, GED/graduating from high school can predict current smoking to some extent (corresponding to the results of linear multiple regression models, and hence the result is presented here). The probability of being a current smoker was lower for the respondents with a bachelor degree than for those without a bachelor degree, and among the respondents without a bachelor degree, the GED group has a much higher likelihood of being current smokers in terms of odds ratio (results not shown, comparing GED with the group with a bachelor degree as the comparison group with dummy coding). Other categorizations nested within bachelor degree/not, however, cannot predict current smoking at any useful level. For employment, the first two contrasts are important in predicting current smoking. It showed that the respondents who were unemployed at the time of interview were more likely to be

Figure 1. Scree-type Plot for Education Based on -2LL

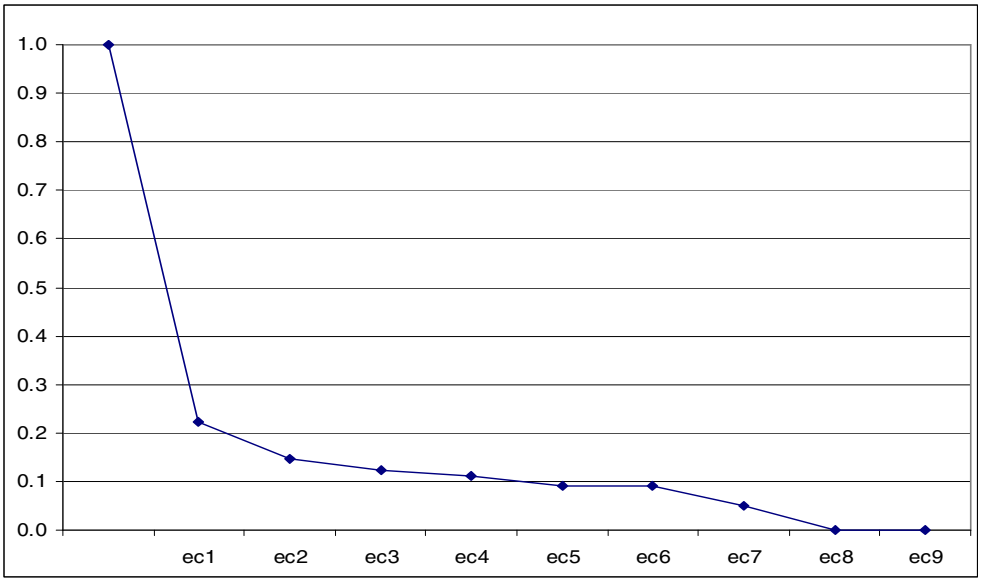

Figure 2. Scree-type Plot for Employment Based on -2LL

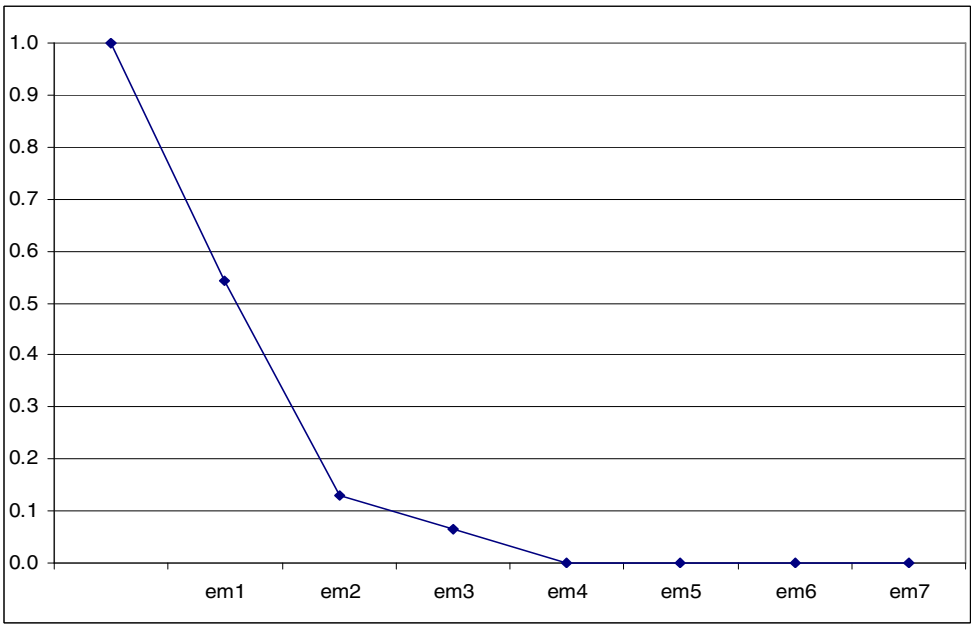

Figure 3. Scree-type Plot for Income Based on -2LL

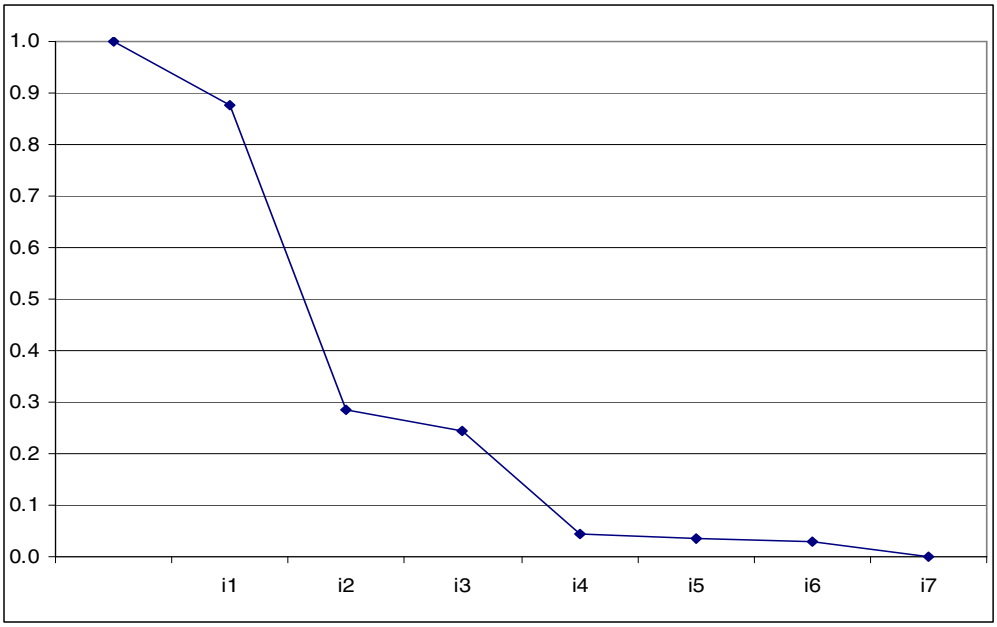

Figure 4. Scree-type Plot for Race/Ethnicity Based on -2LL

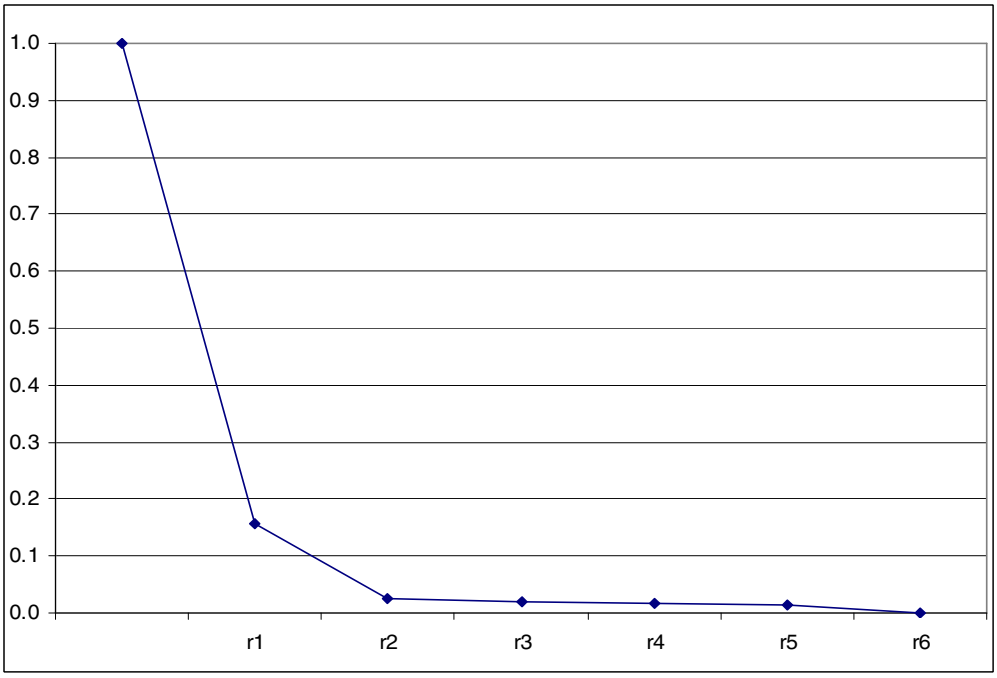

Figure 5. Scree-type Plot for Age Based on -2LL

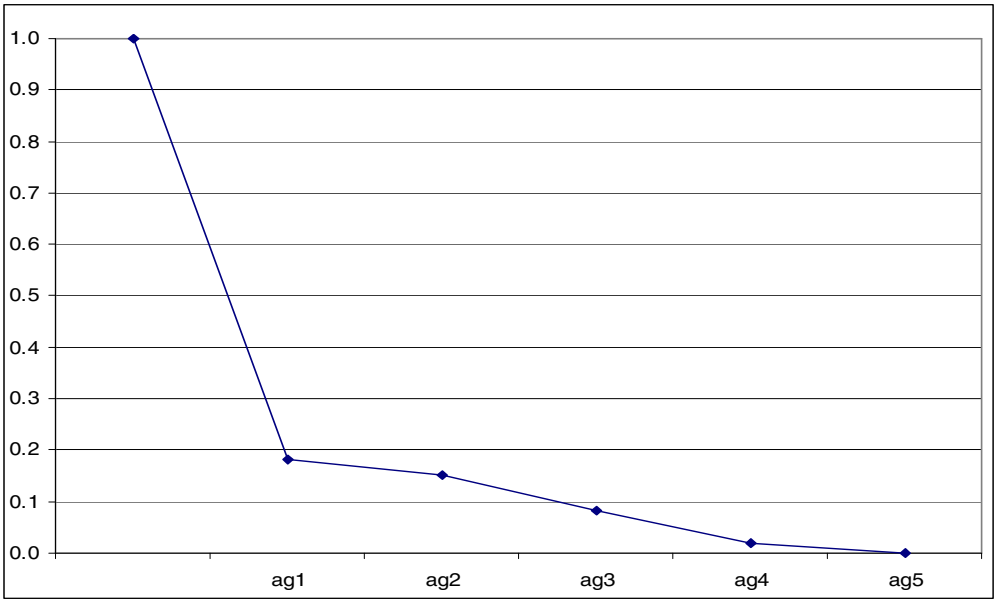

Figure 6. Scree-type Plot for Interaction Based on -2LL

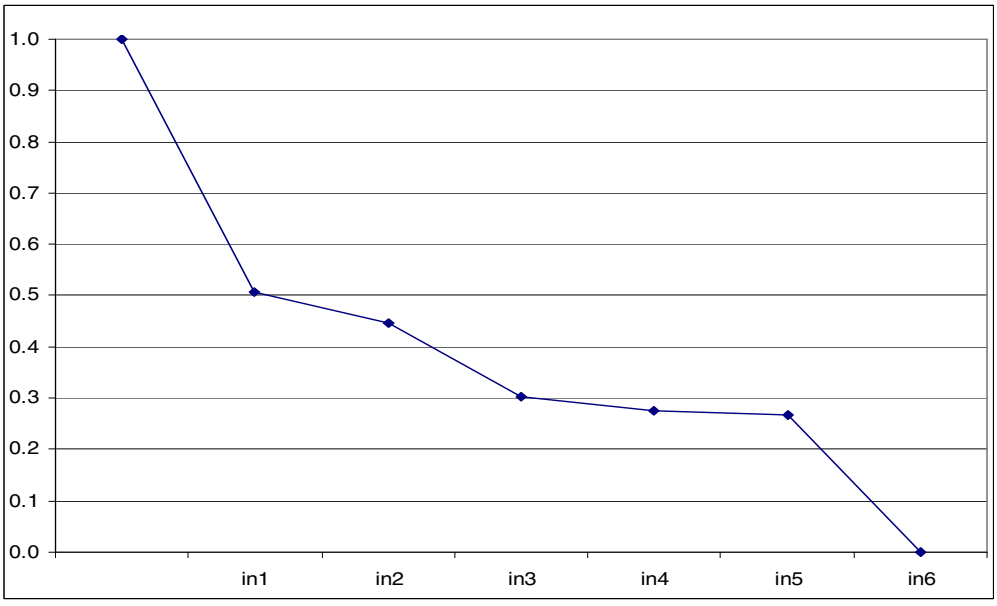

Table 10. Reduced -2LLs for the Exploratory Restricted Model

Predictor	Df	-2LL	Added R^2	Sum of added R^2
Ec1	1	212.5	0.021	0.023
Ec2	1	21.4	0.0022	
Em1	1	75.6	0.0075	0.014
Em2	1	68.1	0.0067	
I1	1	6.0	0.0006	0.0033
I2	1	27.7	0.0027	
R1	1	48.5	0.0047	0.0047
Sex	1	26.5	0.0026	0.0026
Ag1	1	128.7	0.0124	0.012
In1	1	4.0	0.0004	0.0014
In2	1	9.4	0.0009	
In3	1	0.8	0.0001	

Table 11. Reduced -2LLs for the Modified Restricted Model

Predictor	df	-2LL	Odds Ratio	Added R^2
Ec1	1	212.5	1.2	0.021
Em1	1	82.6	1.1	0.015
Em2	1	64.1	1.0	
Ag1	1	86.6	1.1	0.0085

current smokers than the respondents who were not unemployed, which included employees, employers, retirees, homemakers, and students. Also, within the non-unemployed group, the currently employed are slightly more likely to smoke than the respondents who were retirees, homemakers or students, but the odds ratio is close to 1, meaning that the difference in the likelihood of being a smoker is rather small.

Linear Multiple Regression

Multiple Working Hypotheses for Education, Race/Ethnicity and Age

In linear multiple regression models, the new overall variable, which is, we propose, related to strength of cigarette smoking habit, is used as the main dependent variable, and type I sum of squares (Type I SS) is used to compare the relative contribution of different demographic predictors. Tables 12 to 14 show the values of type I SS for different contrast sets for education, race, and age, respectively. The results of testing the multiple hypotheses for these three predictors are the same as those for the logistic regression models. Among the three contrast sets of education, the contrast set B (EC) performed the best. As predicted, the first contrast in the set explained the most variance of cigarette smoking habit. Regarding race, the first contrast in contrast set A has a slightly higher predictive power than the second contrast, which matched our hypothesis. For age, among the three contrast sets, the first contrast of the set B can explain the most variance in cigarette smoking habit, but the order of the importance of the second and third contrasts was not as predicted. As in the logistic regression model, the result shows that being elderly (age of 65 or older) or not is predictive (to any appreciable extent) of cigarette smoking habit, but within the younger group, the

Table 12. Type I SS for the Contrast Sets of Education

Contrast set A	Type I SS	Contrast set B	Type I SS	Contrast set C	Type I SS
Ed1	66.7	Ec1	180.3	Eg1	33.9
Ed2	146.6	Ec2	33.0	Eg2	4.9
Ed3	7.7	Ec3	7.7	Eg3	38.0
Ed4	2.4	Ec4	2.4	Eg4	146.6

Table 13. Type I SS for the Contrast Sets of Race/Ethnicity

Contrast set A	Type I SS	Contrast set B	Type I SS
R1	84.4	Rh1	68.0
R2	5.3	Rh2	21.7

Table 14. Type I SS for the Contrast Sets of Age

Contrast set A	Type I SS	Contrast set B	Type I SS	Contrast set C	Type I SS
A1	0.001	Ag1	75.9	Am1	5.5
A2	6.2	Ag2	0.2	Am2	72.0
A3	71.2	Ag3	1.4	Am3	0.002

predictive power of being young adults (age of 18-24) is not as good as for being older adults (age of 25-64). Therefore, the contrast set B of education, Ec, contrast set A of

race/ethnicity, R, and contrast B of age, Ag is used in the following linear multiple regression models.

Model Comparison

Testing overall hypotheses- the inclusive model. The linear multiple regression analysis shows that, as in the logistic regression model, the inclusive model of demographic predictors can only explain about 7% of the variance in level of cigarette smoking habit. Based on the two hypotheses regarding the relative importance of the demographic predictors, two multiple regression models were conducted. The results are displayed in table 15. They show that model A is a better model. That is, employment is a better predictor of level of cigarette addiction than income. However, the relative importance of the predictors is not exactly the same as predicted and as found in the logistic regression analysis. The results show that the order of importance of the demographic predictors for strength of cigarette smoking habit is education, employment, race, age, sex, income and interaction. Similarly to age, race can explain about 1% of the variance in smoking habit, but this result is different from the results of logistic regression. Among these predictors, interaction, income and sex seem not to be good predictors at all, since all of them together explain less than 0.5% of the variance in cigarette smoking habit.

Testing specific hypotheses- the restricted model. Table 16 shows the results of the hypothesized restricted model. These hypothesized important contrasts taken together can explain 6% of the variance of cigarette smoking habit, close to the amount accounted by the inclusive model. The same as for the logistic regression, these results suggest that

Table 15. Two Inclusive Models Regarding the Order of Employment and Income

Model A				Model B			
Predictor	df	Type I SS	Added R^2	Predictor	df	Type I SS	Added R^2
Education	9	246.4	0.028	Education	9	246.4	0.028
Employment	7	126.3	0.015	Income	7	18.3	0.0021
Income	7	21.7	0.0024	Employment	7	129.7	0.015
Race	6	92.5	0.011	Race	6	92.5	0.011
Sex	1	32.0	0.0037	Sex	1	32.0	0.0037
Age	5	85.2	0.010	Age	5	85.2	0.010
Interaction	6	2.4	0.0002	Interaction	6	2.4	0.0002

most of the contrasts work as we predicted. Similarly, in order to make sure that no important contrasts were missed in the restricted model, an exploratory approach was adopted for the purpose of comparison.

*An exploratory approach- the modified restricted model. A*n exploratory approach is adopted to select the contrasts for the modified restricted model. The type I SS is used to estimate the partial squared multiple correlation (SMC) between contrasts, and a cumulative partial SMC is calculated for each contrast. A scree-type plot is generated with 1-Cumulative SMC for each demographic predictor. Figures 7 to 12 show the plots for the contrast set of different predictors. Based on the plots, the cut points for the respective predictors are 2, 2, 4, 1, 1, and 6, respectively. Combining the results of the exploratory multiple regression model and the specified hypotheses, education,

Table 16. Type I SS for the Hypothesized Restricted Model

Predictor	df	Type I SS	Added R^2	Sum of Added R^2
Ec1	1	180.3	0.021	0.026
Ec2	1	33.0	0.0038	
Ec3	1	7.7	0.00088	
Ec4	1	2.4	0.00028	
Em1	1	86.4	0.010	0.010
I1	1	1.1	0.00013	0.00013
R1	1	56.8	0.0065	0.007
R2	1	5.1	0.00058	
Sex	1	33.3	0.0038	0.0035
Ag1	1	115.7	0.0133	0.014
Ag2	1	0.2	0.000018	
Ag3	1	1.7	0.000190	
In2	1	0.00021	0.00000	0.0000

employment, race and age are deemed important demographic variables in predicting cigarette smoking habit. The final modified restricted model of linear multiple regression is presented in table 17.

The results of the modified restricted model are slightly different from the results of modified logistic regression model. Race turns out to be more important than age in

Figure 7. Scree-type Plot for Education Based on Type I SS

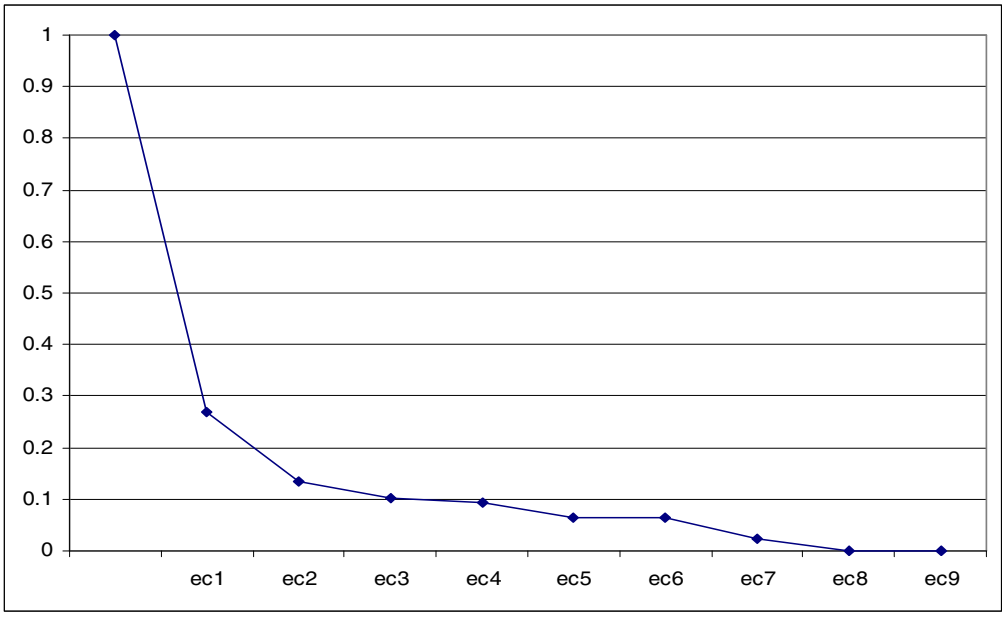

Figure 8. Scree-type Plot for Employment Based on Type I SS

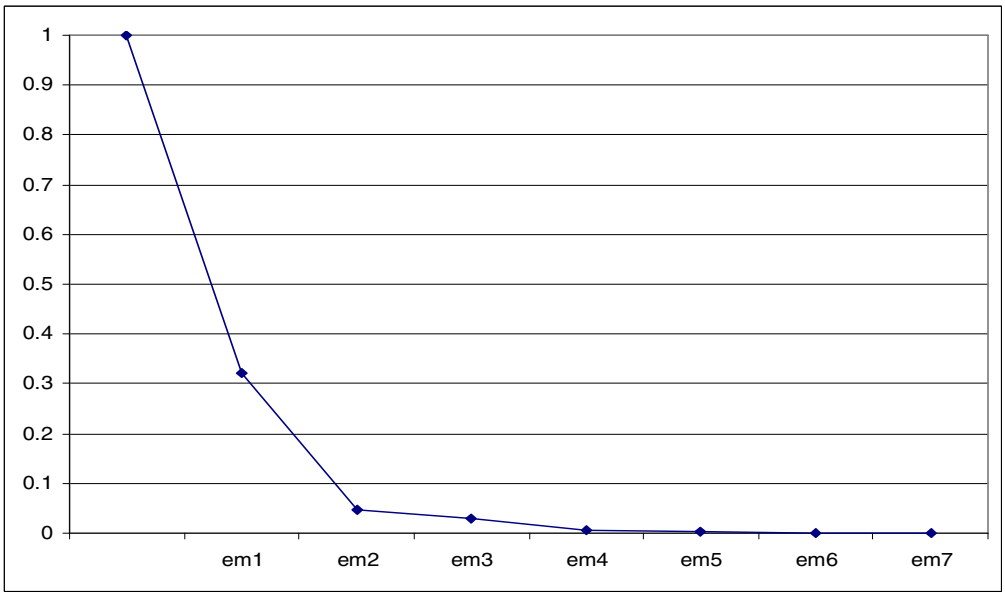

Figure 9. Scree-type Plot for Income Based on Type I SS

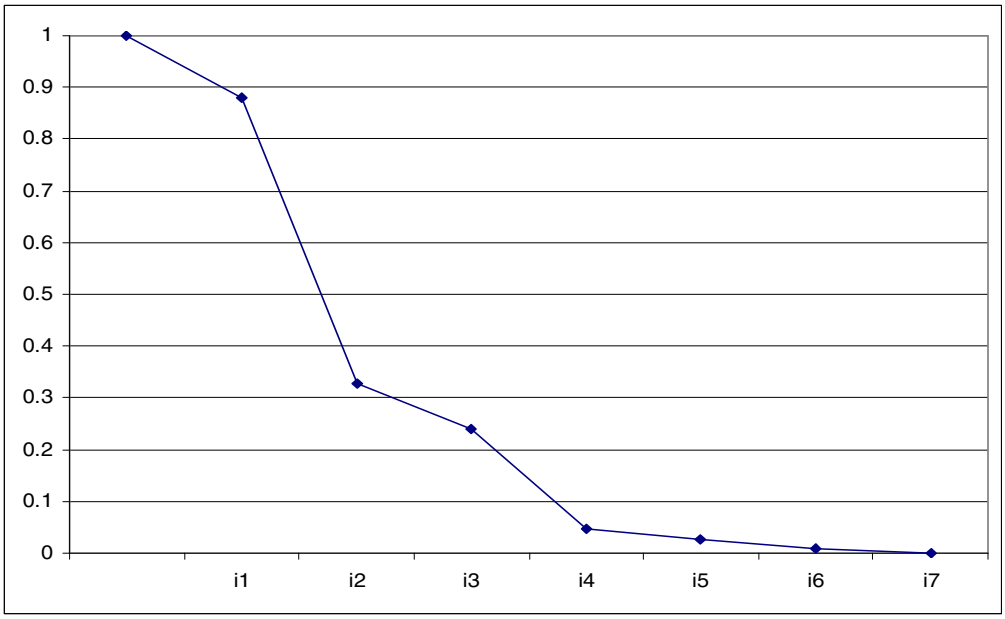

Figure 10. Scree-type Plot for Race/Ethnicity Based on Type I SS

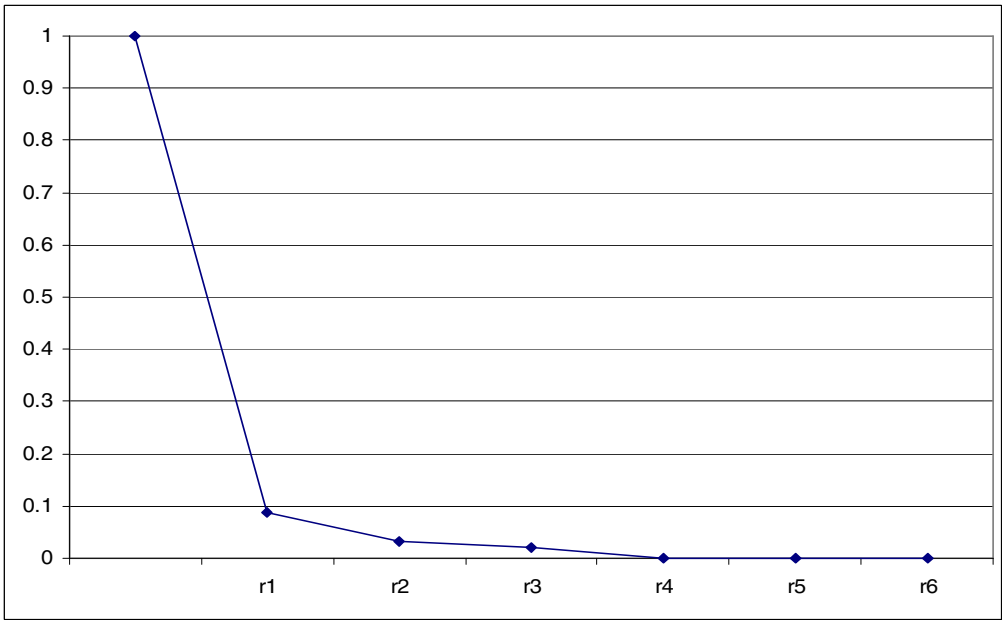

Figure 11. Scree-type Plot for Age Based on Type I SS

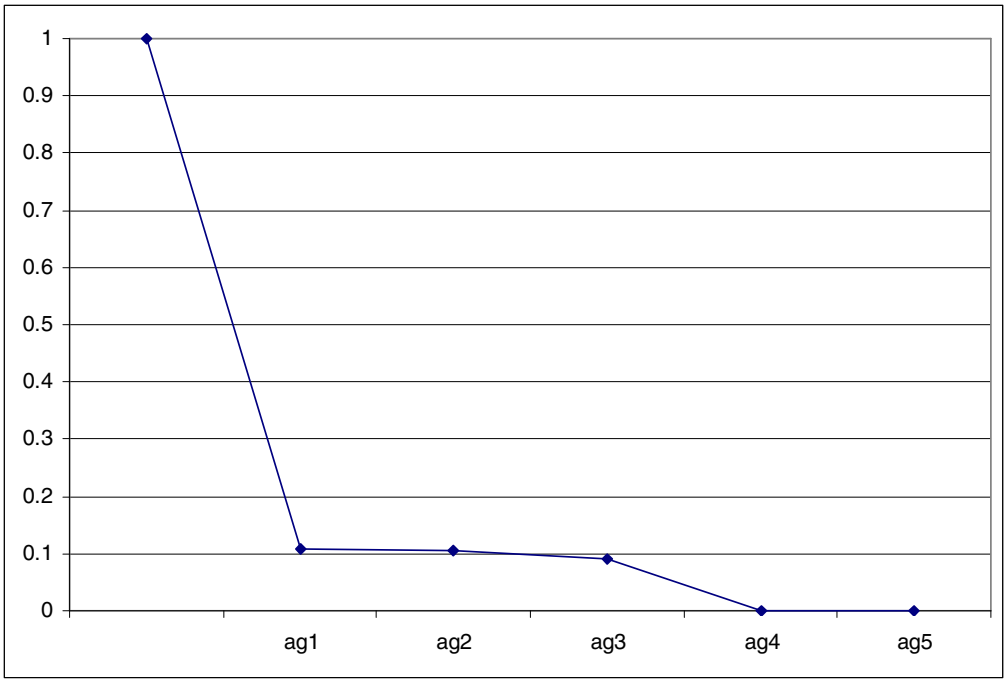

Figure 12. Scree-type Plot for Interaction Based on Type I SS

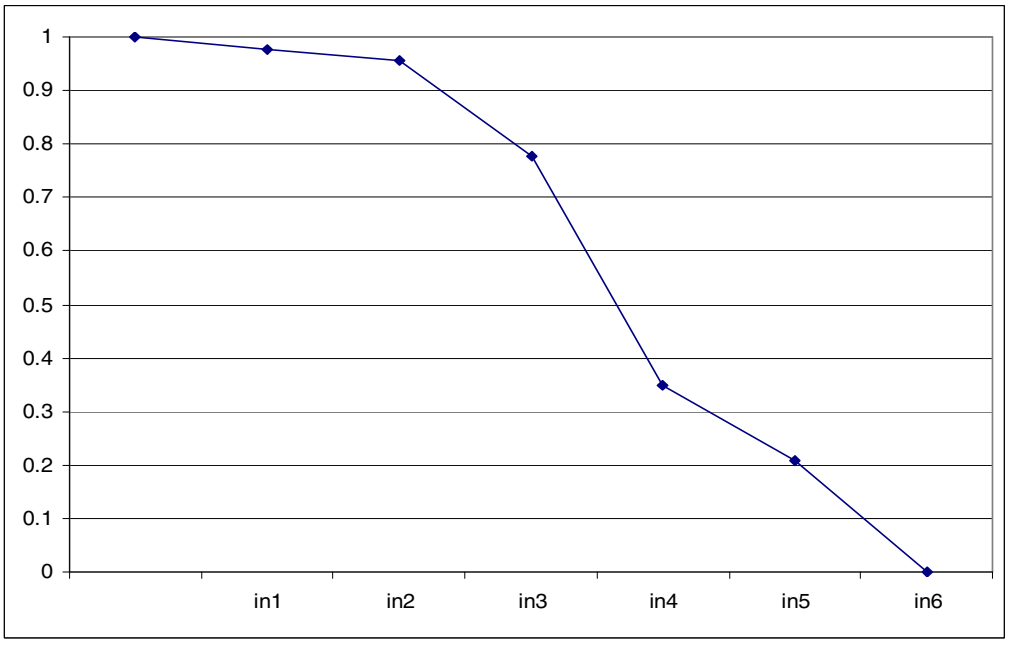

Table 17. Type I SS for the Modified Restricted Model

Predictor	Df	Type I SS	Added R^2
Ec1	1	180.3	0.025
Ec2	1	33.0	
Em1	1	85.0	0.014
Em2	1	34.0	
R1	1	70.0	0.009
R2	1	5.9	
Ag1	1	64.6	0.007

the linear multiple regression models. The White, non-Hispanic respondents were more likely than the non-White respondents to have a strong smoking habit. Also, among the non-White respondents, Hispanic respondents had a weaker cigarette smoking habit than other non-White groups. The other results correspond to the logistic regression model results.

DISCUSSION

Overview

The findings of this study are bound to be disappointing to those responsible for tobacco control efforts. Based on the 2002 and 2005 ATS data in the State of Arizona, the inclusive model of the demographic predictors can explain only about seven percent of the variance in cigarette smoking. Testing our specific hypotheses shows that education, employment, race, and age are in the best variables for predicting current cigarette smoking, but the variance accounted for in the most parsimonious model is even smaller, about five percent. The effect size for the predictive power of these demographic variables is in the range that, conventionally by Cohen's rule of thumb, would be regarded as "moderate", but the gain in prediction is too small to be of much practical use. These results apparently are, however, robust since the two different regression approaches give us similar big pictures of the current cigarette smokers in Arizona. The findings indicated that for adults in Arizona, four years of college education, not being unemployed or White, and being an elderly are protective factors for cigarette smoking. In Arizona, the adult cigarette smokers are mainly the people who have not completed a four-year college degree, or who are unemployed, White, non-Hispanic, or young to middle-aged adults.

Outcomes of Specified Hypotheses

Overall Hypotheses

The overall hypothesis is not fully supported. As predicted, education is the most important predictor for current cigarette smoking, accounting for almost 3 percent of the

variance in cigarette smoking. That is certainly a small effect, but it does suggest that interventions directed at the portion of the population with lower education will have a higher payoff in terms of the number of smokers reached. Beyond education, unemployment, age, and (probably) race/ethnicity are predictors of cigarette smoking, but the variance accounted by each of them is smaller than for education, about 1 percent at most. Employment turns out to be a more important explanatory variable than income; in fact, income may not be a predictor for cigarette smoking at all. In either logistic or linear multiple regression approach, income could not explain more than a trivial portion of the variance in cigarette smoking. The results regarding race/ethnicity are not consistent. Race/ethnicity accounts for only a small proportion of variance in logistic regression models, but its predictive value is similar to that of age in the multiple regression models. Since the results of two regression approaches did not differ substantially, we believe that multiple regression analysis would be the better analysis, because all the information regarding cigarette smoking is integrated into the outcome variable in the models.

Based on the final modified linear multiple regression model, the order of importance of these demographic predictors is education, employment, race, and age, not exactly as hypothesized. It is probably inherently difficult to understand and use demographic predictors when their relationships with the outcome variable are so limited. Tobacco researchers may select demographic variables for their models arbitrarily or simply for convenience, i.e., they happen to be available in the data sets. The theoretical ground for modeling relationships between demographic variables and smoking is

lacking. More research is needed to enhance the theoretical understanding about the underlying mechanisms for these proxy demographic variables.

Specific Hypotheses for Individual Predictor

Even though some of the proposed important demographic variables cannot explain much of the variance in cigarette smoking, the specific hypotheses regarding the individual predictor may still be informative about cigarette smoking of specific groups in the population in Arizona.

Education. Testing the three working hypotheses about education, the results show that college graduation (baccalaureate) does make a difference in terms of cigarette smoking behavior. Consistent with the literature (CDC, 2005; Wetter et al., 2005), respondents who completed college were less likely to be current smokers. One of the possible reasons is that smokers with a college degree are more likely to quit cigarette smoking than less well educated smokers (Brummett et al., 2002). However, the causal pathways by which education might influence cigarette smoking are poorly understood (Wetter et al., 2005; Yan, Liu, Daviglus, Kiefe, Sidney, Matthews, and Greenland, 2006). We believe that education is a proxy for some other important characteristics of individuals or environment, such as a smoking-unfriendly environment generally encountered in a four-year college, or a persevering personality, related to appreciation of delayed reward, which helps individuals succeed both in the hard work required to complete college and in smoking cessation. Probable underlying mechanisms need to be proposed and tested.

One specific category of educational level, GED, appears to be related to current smoking as well. When using GED/High school graduates as the cut point in both the logistic and linear multiple regression models, this contrast has a much lower predictive power than bachelor degree/not. However, the literature indicates (CDC, 2005) that compared to people with a college degree, people with GED are much more likely to smoke cigarette. In our analysis, comparing the respondents with different educational levels with those having bachelor degrees does show that the GED group has a highest probability of being current smokers. It is likely that the small number of cases in the GED group (N=153) decreases the predictive power of the GED/High school graduate contrast. The particularly high smoking rate of the GED group is possibly attributable to contextual factors, smoking social norms, and combination of higher impulsivity, reflected in discounting delayed rewards, and lower motivation toward conventional social rewards than other groups of people.

Employment. As hypothesized, the respondents who were unemployed and unable to work were more likely to smoke cigarette compared to others. Apparently, unemployment is a risk factor for cigarette smoking, based on both theoretical understanding and empirical evidence. Some studies focus on smoking prevalence for people with different occupations. It is possible that employment/unemployment distinction provides sufficient information in terms of cigarette smoking.

Among the respondents who were not unemployed, the distinction between employed and homemaker/student/retired is also related to cigarette smoking. Respondents who work -- and make money -- are slightly more likely to smoke than are

persons in the homemaker/student/retired category, although the odds ratio is close to 1. Arizona has a high proportion of retired people in the State, and the ATS samples usually include a high percentage of retired respondents. The relatively low smoking rate of the homemaker/student/retired group might be driven by the large number of retired respondents in ATS 2002 and 2005. It is likely that many retired people have given up cigarette smoking for health or social reasons. We need better understanding of the distinction between different groups with respect to variables for which employment is a proxy.

Income. The results of both the logistic and linear multiple regression models show that income is not much related to cigarette smoking, even when income was the only predictor in the model. That result is contradictory to extant literature and our hypotheses. The questionable quality of the income data, including the large amount of missing data, may be part of the reason. Respondents might not answer the income question precisely due to the complex structure of the survey questions or for personal reasons, and it might add errors into the observed data.

Even though the variance accounted by income is tiny, looking into the results of important contrasts within income itself, some interesting patterns may be informative about the connection between income and cigarette smoking. The best two predictive contrasts within income itself are, among the respondents with annual income above $20,000, the distinction between the highest income and the rest (more than $75,000/below $75,000 and above $20,000), and then among respondents between $25,000, and $74,999, the grouping between $25,000 and $34,999 (in one group), and

between $35,000 and $74,999 (the other group). It seems to us that higher income does to some extent protect Arizona residents from smoking cigarettes. The big picture of our findings is that higher education, better employment status, and higher income all inversely related to cigarette smoking.

Age. In both logistic and linear multiple regression models, the results show that age is more important than income as a predictor of smoking and also more important than race or the interaction between sex and race/ethnicity (even though we still put race earlier than age in the linear regression models due to our specified hypotheses). The consistent finding is that respondents older than 65 years old were more likely to be non-smokers at the time of interview. Being beyond the age usually considered as the beginning of old age is a protective factor for cigarette smoking.

Race/Ethnicity. Race/ethnicity might be an important predictor for cigarette smoking as well, but the results are not always consistent. A White, non-Hispanic respondent is more likely to be a current smoker compared to all other races. Being a Hispanic is somewhat a protective factor for cigarette smoking, but the relationship is much weaker. One possible reason for the unstable results here is that the statistical power may not be sufficient for other racial groups. When looking only at the binary outcome, current smoker or not, race/ethnicity is not a good predictor.

Interaction. None of the interactions turns out to be an important predictor in our models, whether because of the statistical nature of interaction terms or reality.

Practical Implication

There are two reasons why it may be interesting or useful to know the relationship between socio-demographic variables and smoking. One reason is that socio-demographic variables may be helpful in identifying and even isolating populations to be targeted by various smoking interventions. For example, if it is known that persons with lower levels of education are more likely to be smokers, then one might direct smoking cessation campaigns and services at geographic areas know to include higher proportions of less well educated persons. The second reason is that socio-demographic variables are almost certain to be proxies for variables of more theoretical and practical interest. Relationships between socio-demographic variables and smoking may help to direct our attention to questions of importance in relation to planning and carrying out of tobacco control efforts. Again, knowing that relatively limited education is related to smoking, we may be able to determine through what more interesting and proximal variable the effects of education are mediated, and we can then try to deal directly with that variable.

Current cigarette smokers are the very targets for smoking cessation programs. Based on the findings of this study, one important practical implication can be the targeting of cessation service. The state-funded cessation program should aim at serving people without a bachelor degree, especially those with a GED, or unemployed, young to middle-aged adults. High-educated (with at least a bachelor degree), or high income ($35,000 or above), or not unemployed, or old, non-White people are better protected from smoking. They are not the group with the greatest need for cessation service. We should put more effort into attracting and helping the disadvantaged smokers utilize

cessation services. More research should be directed at understanding the mechanism of connecting the disadvantaged groups and cigarette smoking, and the marketing strategies for attracting the disadvantaged groups.

Another implication is for the formulation of questions regarding socio-demographic information in the ATS questionnaire. Without a better theoretical understanding about the relationship between socio-demographic variables and cigarette smoking, based on the results of this study, we recommend that the socio-demographic questions in ATS be simplified. Only the important demographic predictors, such as education, employment status, age and race/ethnicity need to be included in the questionnaire. For each important demographic variable, only the categories shown to be related to cigarette smoking need to be included among the response options in the ATS questionnaire. For example, for education, questions about college completion and the GED should capture virtually all of the information needed to segment the population with respect to smoking habits. Similar response simplification could be achieved for employment and race/ethnicity questions. We include a list of recommended socio-demographic questions for future ATS questionnaire in appendix B.

For education, interviewers do not need to read or check the answer from the long list of the response categories (eleven substantive options for educational level plus two uncertain response categories in the ATS 2002 and 2005), but ask education question as an open-ended question instead. If the respondent's answer shows that he/she has not completed college, the interviewer checks the answer 'No bachelor degree'. Otherwise, the answer 'bachelor degree or above' is checked. If the respondent answers 'high

school' as the highest degree received, the interviewer further ask him/her whether it is a regular high school degree or GED. For employment status, if the respondent answers yes to the question 'are you working now?', a question about the number of people working for their employer is also asked for quality control purpose and also for more information about the working environment. If the respondent is not working at the time of interview, a more detailed question about the unemployment status will be asked to categorize the respondent into 'unemployed', 'student', 'homemaker', 'retired', or 'unable to work'. For race/ethnicity, only three response categories are listed in the question: 'White, non-Hispanic', 'Hispanic', and 'others'. For age, we ask which year the respondent was born and then recode the age into meaningful groups for data analysis.

If only the important demographic information is asked in ATS questionnaires, the spared item space can be dedicated to more important variables.

Contributions

Theoretical

It seems to us that most of epidemiologic or surveillance studies related to demographic information and cigarette smoking are based on convenience or arbitrary selection of samples. Not much theoretical foundation has been provided for this kind of research. For this study, related studies were reviewed and the hypotheses specified were based on the past research, even though the theoretical understanding was rather limited. Hopefully, the adopted approach in this study can stimulate more theoretical thinking about this kind of research and more theoretical advancement will be obtained in the field.

Methodological

This study has heavy methodological components in it. We hope that the new methodological procedures used here can be adopted by other researchers as well and different tools can be utilized for this type of research.

First, for the study with large sample size, significance testing or p values are not always informative, since most of the results would be significant. Also, a majority of similar studies in this area use logistic regression models for data analysis and often report an odds ratio as the criterion for importance of predictors. We have found, however, that odds ratios are highly unstable and can change dramatically as different predictors are present in the models. In order to overcome these problems, we focused on the measures of effect size in terms of the statistics -2LL or sum of squares and R square accounted for by the relevant predictors. It provides us with another tool to assess the importance of the predictors.

Second, the linear multiple regression approach seems to be a feasible approach for this type of research, too. The newly-created outcome variable makes the best use of the cigarette smoking information in the ATS questionnaire, and the linear regression works well in this study. The results of logistic and linear multiple regression models are basically consistent with each other and suggest that linear multiple regression analysis can be an alternative approach for this kind of research.

Third, the exploratory approach using the scree-type plots to decide the cut point of each contrast sets can be helpful for selecting the important contrasts. This exploratory

approach can complement the results of hypotheses testing and direct the possible model modification for a better understanding of the empirical data.

Limitations and future directions

Lacking a sound theoretical understanding of the relationship between the relevant socio-demographic variables and cigarette smoking limited the hypotheses specification in the first place. Most of the hypotheses specified in this study were only guesses at best and they might have completely missed the target. Especially when we come to defining the specific contrasts, we do not usually have adequate theoretical grounds for making all the detailed decisions. Most of the contrasts were decided based on primitive empirical findings and logical/subjective judgment, and it may limit the findings of this study.

The reached ATS sample in Arizona is mainly composed of residents with better living situations in terms of education, employment, and income, compared to the general population in Arizona. Lower-educated (without a high school degree), very poor, or minority adult residents were not represented in the survey. The small numbers of some specific groups in ATS samples may limit the validity of the findings in this study. Some grouping contrasts may appear unimportant just simply because there is no information available for the specific group in the contrast.

This study is a secondary analysis of existing surveillance data, and the original questionnaires decided what information was available for conducting this study. Especially given the fact that some items changed across the years, some merged responses may not convey exactly the same information. Also, this is a purely self-report

study. The information received through impersonal telephone interview may not reflect the real states of the respondents. The added error in the observed variables due to data merging and self-report nature may influence the findings of this study as well.

Based on the findings and limitations of this study, we propose some future directions to advance the relevant field. First, ATS should aim at collecting more data from the disadvantaged groups in the adult population of Arizona. Second, various research approaches should be adopted to have a better theoretical understanding of the connection of demographic predictors and cigarette smoking. Third, when we are more certain about the connection between cigarette smoking and socio-demographic predictors, we should investigate the underlying mechanisms for the protective effects of the relevant demographic categories to improve our tobacco control efforts for the disadvantaged groups in the population and make the intervention more effective. Last, but not least, since all the related demographic variables can explain only a rather small proportion of variance in cigarette smoking, we should focus on more important and relevant predictors in this kind of research, including state or even national level surveillance systems, and investigate what factors can be used to predict cigarette smoking effectively.

APPENDICES

APPENDIX A

ORTHOGONAL CONTRASTS WITH NESTED DESIGN

Education Contrast set A									
CATEGORY	Ed1	Ed2	Ed3	Ed4	Ed5	Ed6	Ed7	Ed8	Ed9
1 = never attended school	+3	0	+1	0	0	0	0	0	+1
2 = Grades 1 through 8	+3	0	+1	0	0	0	0	0	-1
3 = Grades 9 through 11	+3	0	-1	+1	0	0	0	0	0
4 = Grade 12 (high school graduate)	-2	+1	0	0	+3	0	0	0	0
5 = GED	+3	0	-1	-1	0	0	0	0	0
6 = Some college	-2	+1	0	0	-1	+2	0	0	0
7 = AA Technical/vocational	-2	+1	0	0	-1	-1	+1	0	0
8 = AA Academic	-2	+1	0	0	-1	-1	-1	0	0
9 = BA, Bachelor	-2	-2	0	0	0	0	0	+1	0
10 = Master or above	-2	-2	0	0	0	0	0	-1	0
Education Contrast set B									
CATEGORY	Ec1	Ec2	Ec3	Ec4	Ec5	Ec6	Ec7	Ec8	Ec9
1 = never attended school	+1	+1	+1	0	0	0	0	0	+1
2 = Grades 1 through 8	+1	+1	+1	0	0	0	0	0	-1
3 = Grades 9 through 11	+1	+1	-1	+1	0	0	0	0	0
4 = Grade 12 (high school graduate)	+1	-1	0	0	+3	0	0	0	0
5 = GED	+1	+1	-1	-1	0	0	0	0	0
6 = Some college	+1	-1	0	0	-1	+2	0	0	0
7 = AA Technical/vocational	+1	-1	0	0	-1	-1	+1	0	0
8 = AA Academic	+1	-1	0	0	-1	-1	-1	0	0
9 = BA, Bachelor	-4	0	0	0	0	0	0	+1	0
10 = Master or above	-4	0	0	0	0	0	0	-1	0
Education Contrast set C									
CATEGORY	Eg1	Eg2	Eg3	Eg4	Eg5	Eg6	Eg7	Eg8	Eg9
1 = never attended school	+7	+1	0	0	0	0	0	0	+1
2 = Grades 1 through 8	+7	+1	0	0	0	0	0	0	-1
3 = Grades 9 through 11	+7	-2	0	0	0	0	0	0	0
4 = Grade 12 (high school graduate)	-3	0	-1	+2	+3	0	0	0	0
5 = GED	-3	0	+6	0	0	0	0	0	0
6 = Some college	-3	0	-1	+2	-1	+2	0	0	0
7 = AA Technical/vocational	-3	0	-1	+2	-1	-1	+1	0	0
8 = AA Academic	-3	0	-1	+2	-1	-1	-1	0	0
9 = BA, Bachelor	-3	0	-1	-4	0	0	0	+1	0
10 = Master or above	-3	0	-1	-4	0	0	0	-1	0

*The gray highlights represent important contrasts for the categorization of predictors

Employment Contrast set							
CATEGORY	Em1	Em2	Em3	Em4	Em5	Em6	Em7
1 = Employed for wages	-3	+3	0	0	-1	0	0
2 = Self-employed	-3	+3	0	0	+1	0	0
3 = Out of work for >1 yr	+5	0	0	0	0	-1	+1
4 = Out of work for <1 yr	+5	0	0	0	0	-1	-1
5 = A homemaker	-3	-2	-1	+1	0	0	0
6 = A Student	-3	-2	-1	-1	0	0	0
7 = Retired"	-3	-2	+2	0	0	0	0
8 = Unable to work	+5	0	0	0	0	+2	0

Income Contrast set							
CATEGORY	I1	I2	I3	I4	I5	I6	I7
1 = Less than $10,000	+5	0	0	0	+2	0	0
2 = Between $10,000 and $14,999	+5	0	0	0	-1	+1	0
3 = Between $15,000 and $19,999	+5	0	0	0	-1	-1	0
4 = Between $20,000 and $24,999	-3	+1	+3	0	0	0	0
5 = Between $25,000 and $34,999	-3	+1	-1	+2	0	0	0
6 = Between $35,000 and $49,999	-3	+1	-1	-1	0	0	+1
7 = Between $50,000 and $74,999	-3	+1	-1	-1	0	0	-1
8 = More than $75,000	-3	-4	0	0	0	0	0

Race Contrast set A						
CATEGORY	R1	R2	R3	R4	R5	R6
1 = White, non-Hispanic	+6	0	0	0	0	0
2 = Black or African American	-1	+1	+4	0	0	0
3 = Hispanic	-1	-5	0	0	0	0
4 = Asian	-1	+1	-1	-1	-1	-1
5 = Native Hawaiian or Pacific	-1	+1	-1	-1	+2	0
6 = American Indian	-1	+1	-1	+3	0	0
7 = Other	-1	+1	-1	-1	-1	+1
Race Contrast set B						
CATEGORY	Rh1	Rh2	Rh3	Rh4	Rh5	Rh6
1 = White, non-Hispanic	+1	+5	0	0	0	0
2 = Black or African American	+1	-1	+4	0	0	0
3 = Hispanic	-6	0	0	0	0	0
4 = Asian	+1	-1	-1	-1	-2	0
5 = Native Hawaiian or Pacific	+1	-1	-1	-1	+1	+1
6 = American Indian	+1	-1	-1	+3	0	0
7 = Other	+1	-1	-1	-1	+1	-1

Age Contrast set A					
CATEGORY	A1	A2	A3	A4	A5
1 = 18-24 years	+5	0	0	0	0
2 = 25-34 years	-1	+3	0	0	+1
3 = 35-44 years	-1	+3	0	0	-1
4 = 45-54 years	-1	-2	+1	+1	0
5 = 55-64 years	-1	-2	+1	-1	0
6 = 65 or older	-1	-2	-2	0	0
Age Contrast set B					
CATEGORY	Ag1	Ag2	Ag3	Ag4	Ag5
1 = 18-24 years	+1	+4	0	0	0
2 = 25-34 years	+1	-1	+1	0	+1
3 = 35-44 years	+1	-1	+1	0	-1
4 = 45-54 years	+1	-1	-1	+1	0
5 = 55-64 years	+1	-1	-1	-1	0
6 = 65 or older	-5	0	0	0	0
Age Contrast set C					
CATEGORY	Am1	Am2	Am3	Am4	Am5
1 = 18-24 years	-1	+1	+2	0	0
2 = 25-34 years	+2	0	0	0	+1
3 = 35-44 years	+2	0	0	0	-1
4 = 45-54 years	-1	+1	-1	+1	0
5 = 55-64 years	-1	+1	-1	-1	0
6 = 65 or older	-1	-3	0	0	0

Contrast 1 for Interaction between sex and race	
Variables	Formula
IN1	Sex * R1
IN2	Sex * R2
IN3	Sex * R3
IN4	Sex * R4
IN5	Sex * R5
IN6	Sex * R6

Contrast 2 for Interaction between sex and race	
Variables	Formula
Inter1	Sex * Rh1
Inter2	Sex * Rh2
Inter3	Sex * Rh3
Inter4	Sex * Rh4
Inter5	Sex * Rh5
Inter6	Sex * Rh6

APPPENDIX B

RECOMMENDED DEMOGRAPHIC QUESTIONS

1. Education:
What is the highest level of school you completed or the highest degree you received?
_____ No bachelor degree
_____ GED (probe for GED when the respondent answers 'high school')
_____ Bachelor degree or above

2. Employment status:
Are you working now?
_____ Yes
Do more than 50 people work for your employer?
_____ Yes
_____ No
_____ Don't know/Not sure
(continue for a set of smoking policy questions related to work)

_____ No
Are you currently… READ LIST BELOW?
_____ A student (if Yes, skip the rest of the questions)
_____ A homemaker (if Yes, skip the rest of the questions)
_____ Retired (if Yes, skip the rest of the questions)
_____ Unable to work (if Yes, skip the rest of the questions)

3. Race/Ethnicity:
Which racial or ethnic group do you mostly identify yourself with?
_____ White, non-Hispanic
_____ Hispanic
_____ Others

4. Age:
What year were you born?
_____ (year)
_____ Don't know
_____ Refused

REFERENCES

Abdullah, A. S., Lam, T. H., Chan, S. S., & Hedley, A. J. (2006). Smoking cessation among chinese young smokers: Does gender and age difference matters and what are the predictors? *Addictive Behaviors, 31*(5), 913-921.

Abraido-Lanza, A. F., Chao, M. T., & Florez, K. R. (2005). Do healthy behaviors decline with greater acculturation? implications for the latino mortality paradox. *Social science & medicine (1982), 61*(6), 1243-1255.

Adler, N. E., Boyce, T., Chesney, M. A., Cohen, S., Folkman, S., & Kahn, R. L., et al. (1994). Socioeconomic status and health. the challenge of the gradient. *The American Psychologist, 49*(1), 15-24.

Adler, N. E., & Snibbe, A. C. (2003). The role of psychosocial processes in explaining the gradient between socioeconomic status and health. *Current Directions in Psychological Science, 12*(4), 119-123.

Ahijevych, K. L., Tyndale, R. F., Dhatt, R. K., Weed, H. G., & Browning, K. K. (2002). Factors influencing cotinine half-life during smoking abstinence in african american and caucasian women. *Nicotine & tobacco research : official journal of the Society for Research on Nicotine and Tobacco, 4*(4), 423-431.

Arizona Department of Health Services Office of Tobacco Education and Prevention Program. (2006). *Arizona adult tobacco survey 2005 report*Arizona Department of Health Services. from http://www.azdhs.gov/phs/tepp/pdf/adult_tobacco_survey.pdf

Bang, K. M., & Kim, J. H. (2001). Prevalence of cigarette smoking by occupation and industry in the united states. *American Journal of Industrial Medicine, 40*(3), 233-239.

Barbeau, E. M., Krieger, N., & Soobader, M. J. (2004). Working class matters: Socioeconomic disadvantage, race/ethnicity, gender, and smoking in NHIS 2000. *American Journal of Public Health, 94*(2), 269-278.

Barbeau, E. M., McLellan, D., Levenstein, C., DeLaurier, G. F., Kelder, G., & Sorensen, G. (2004). Reducing occupation-based disparities related to tobacco: Roles for occupational health and organized labor. *American Journal of Industrial Medicine, 46*(2), 170-179.

Benowitz, N. L., & Hatsukami, D. (1998). Gender differences in the pharmacology of nicotine addiction. *Addiction Biology, 3*, 383-404.

Bergen, A. W., & Caporaso, N. (1999). Cigarette smoking. *Journal of the National Cancer Institute, 91*(16), 1365-1375.

Bethel, J. W., & Schenker, M. B. (2005). Acculturation and smoking patterns among hispanics: A review. *American Journal of Preventive Medicine, 29*(2), 143-148.

Bjornson, W., Rand, C., Connett, J. E., Lindgren, P., Nides, M., & Pope, F., et al. (1995). Gender differences in smoking cessation after 3 years in the lung health study. *American Journal of Public Health, 85*(2), 223-230.

Bobak, M., Jha, P., Nguyen, S., & Jarvis, M. (2000). Poverty and smoking. In P. Jha, & F. Chaloupka (Eds.), *Tobacco control in developing countries* (pp. 41-61). USA: Oxford University Press.

Bolen, J. C., Rhodes, L., Powell-Griner, E. E., Bland, S. D., & Holtzman, D. (2000). State-specific prevalence of selected health behaviors, by race and ethnicity--behavioral risk factor surveillance system, 1997. *MMWR.CDC surveillance summaries : Morbidity and mortality weekly report.CDC surveillance summaries / Centers for Disease Control, 49*(2), 1-60.

Brackbill, R., Frazier, T., & Shilling, S. (1988). Smoking characteristics of US workers, 1978-1980. *American Journal of Industrial Medicine, 13*(1), 5-41.

Brummett, B. H., Babyak, M. A., Mark, D. C., Williams, R. B., Siegler, I. C., & Clapp-Channing, N., et al. (2002). Predictors of smoking cessation in patients with a diagnosis of coronary artery disease. *Journal of cardiopulmonary rehabilitation, 22*(3), 143-147.

Caraballo, R. S., Giovino, G. A., Pechacek, T. F., Mowery, P. D., Richter, P. A., & Strauss, W. J., et al. (1998). Racial and ethnic differences in serum cotinine levels of cigarette smokers: Third national health and nutrition examination survey, 1988-1991. *JAMA : the journal of the American Medical Association, 280*(2), 135-139.

Centers for Disease Control and Prevention (CDC). (2006). Tobacco use among adults--united states, 2005. *MMWR.Morbidity and mortality weekly report, 55*(42), 1145-1148.

Centers for Disease Control and Prevention (CDC). (2005). Cigarette smoking among adults--united states, 2004. *MMWR.Morbidity and mortality weekly report, 54*(44), 1121-1124.

Centers for Disease Control and Prevention (CDC). (1993). Cigarette smoking among adults--united states, 1991. *MMWR.Morbidity and mortality weekly report, 42*(12), 230-233.

Colby, J. P., Linsky, A. S., & Straus, M. A. (1994). Social stress and state-to-state differences in smoking and smoking related mortality in the united states. *Social science & medicine, 38*, 373-381.

Coleman, T. (2004). Special groups of smokers. *BMJ (Clinical research ed.), 328*(7439), 575-577.

Conway, T. L., Vickers, R. R.,Jr, Ward, H. W., & Rahe, R. H. (1981). Occupational stress and variation in cigarette, coffee, and alcohol consumption. *Journal of health and social behavior, 22*(2), 155-165.

Covey, L. S., & Wynder, E. L. (1981). Smoking habits and occupational status. *Journal of occupational medicine.: official publication of the Industrial Medical Association, 23*(8), 537-542.

Elstad, J. I. (1998). The psycho-social perspective on social inequalities in health. *Sociology of Health and Illness, 20*(5), 598-618.

Escobedo, L. G., & Peddicord, J. P. (1996). Smoking prevalence in US birth cohorts: The influence of gender and education. *American Journal of Public Health, 86*(2), 231-236.

Figueredo, A. J., Vásquez, G., Brumbach, B. H., Schneider, S. M. R., Sefcek, J. A., & Tal, I. R., et al. (2006). Consilience and life history theory: From genes to brain to reproductive strategy. *Developmental Review, 26*, 243-275.

Fiore, M. C., Novotny, T. E., Pierce, J. P., Hatziandreu, E. J., Patel, K. M., & Davis, R. M. (1989). Trends in cigarette smoking in the united states. the changing influence of gender and race. *JAMA : the journal of the American Medical Association, 261*(1), 49-55.

Flint, A. J., & Novotny, T. E. (1997). Poverty status and cigarette smoking prevalence and cessation in the united states, 1983-1993: The independent risk of being poor. *Tobacco control, 6*(1), 14-18.

Fuchs, V. R. (1979). Economics, health, and post-industrial society. *The Milbank Memorial Fund quarterly.Health and society, 57*(2), 153-182.

Gadgeel, S. M., & Kalemkerian, G. P. (2003). Racial differences in lung cancer. *Cancer metastasis reviews, 22*(1), 39-46.

Gorsuch, R. L. (1983). *Factor Analysis*. Hillsdale, NJ: Lawrence Erlbaum.

Gorsuch, R. L. (1991). *UniMult: For Uunivariate and Multivariate Data Analysis.* UniMult, Altadena, CA.

Green, K. L., & Johnson, J. V. (1990). The effects of psychosocial work organization on patterns of cigarette smoking among male chemical plant employees. *American Journal of Public Health, 80*(11), 1368-1371.

Gritz, E. R., Thompson, B., Emmons, K., Ockene, J. K., McLerran, D. F., & Nielsen, I. R. (1998). Gender differences among smokers and quitters in the working well trial. *Preventive medicine, 27*(4), 553-561.

Hammarstrom, A., & Janlert, U. (2003). Unemployment -- an important predictor for future smoking: A 14-year follow-up study of school leavers. *Scandinavian journal of public health, 31*(3), 229-232.

Heckman, J. J., Stixrud, J., & Urzua, S. (2006). *The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior*National Bureau of Economic Research, Inc.

Hill, E. M., & Chow, K. (2002). Life-history theory and risky drinking. *Addiction, 97*, 401-413.

Husten, C. G., Shelton, D. M., Chrismon, J. H., Lin, Y. C., Mowery, P., & Powell, F. A. (1997). Cigarette smoking and smoking cessation among older adults: United states, 1965-94. *Tobacco control, 6*(3), 175-180.

Jaroni, J. L., Wright, S. M., Lerman, C., & Epstein, L. H. (2004). Relationship between education and delay discounting in smokers. *Addictive Behaviors, 29*(6), 1171-1175.

Kassel, J. D., Stroud, L. R., & Paronis, C. A. (2003). Smoking, stress, and negative affect: Correlation, causation, and context across stages of smoking. *Psychological bulletin, 129*, 270-304.

Maher, J. E., Boysun, M. J., Rohde, K., Stark, M. J., Pizacani, B. A., & Dilley, J., et al. (2005). Are latinos really less likely to be smokers? lessons from oregon. *Nicotine & tobacco research : official journal of the Society for Research on Nicotine and Tobacco, 7*(2), 283-287.

Markides, K. S., Coreil, J., & Ray, L. A. (1987). Smoking among mexican americans: A three-generation study. *American Journal of Public Health, 77*(6), 708-711.

Menard, S. (1995). *Applied Logistic Regression Analysis.* Sage University Paper series on Quantitative Applications in the Social Sciences, 07-106. Thousand Oaks, CA: Sage.

Nelson, D. E., Emont, S. L., Brackbill, R. M., Cameron, L. L., Peddicord, J., & Fiore, M. C. (1994). Cigarette smoking prevalence by occupation in the united states. A comparison between 1978 to 1980 and 1987 to 1990. *Journal of occupational medicine.: official publication of the Industrial Medical Association, 36*(5), 516-525.

Novotny, T. E., Warner, K. E., Kendrick, J. S., & Remington, P. L. (1988). Smoking by blacks and whites: Socioeconomic and demographic differences. *American Journal of Public Health, 78*(9), 1187-1189.

Pampel, F. C. (2005). Patterns of tobacco use in the early epidemic stages: Malawi and zambia, 2000-2002. *American Journal of Public Health, 95*(6), 1009-1015.

Pampel, F. C. (2001). Cigarette diffusion and sex differences in smoking. *Journal of health and social behavior, 42*(4), 388-404.

Perez-Stable, E. J., Herrera, B., Jacob, P.,3rd, & Benowitz, N. L. (1998). Nicotine metabolism and intake in black and white smokers. *JAMA : the journal of the American Medical Association, 280*(2), 152-156.

Pierce, J., Giovino, G., Hatziandreu, E., & Shopland, D. (1989). National age and sex differences in quitting smoking. *Journal of psychoactive drugs, 21*(3), 293-298.

Pierce, J. P., Fiore, M. C., Novotny, T. E., Hatziandreu, E. J., & Davis, R. M. (1989). Trends in cigarette smoking in the united states. educational differences are increasing. *JAMA : the journal of the American Medical Association, 261*(1), 56-60.

R Development Core Team (2005). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL: http://www.R-project.org.

Rogers, R. G. (1991). Demographic characteristics of cigarette smokers in the united states. *Social biology, 38*(1-2), 1-12.

Rogers, R. G., Nam, C. B., & Hummer, R. A. (1995). Demographic and socioeconomic links to cigarette smoking. *Social biology, 42*(1-2), 1-21.

Ryan, H., Wortley, P. M., Easton, A., Pederson, L., & Greenwood, G. (2001). Smoking among lesbians, gays, and bisexuals: A review of the literature. *American Journal of Preventive Medicine, 21*(2), 142-149.

SAS Institute Inc. (2006). SAS 9.1.3 Help and Documentation, Cary, NC: SAS Institute Inc.

Schachter, S. (1977). Studies of the interaction of psychological and pharmacological determinants of smoking: IV. effects of stress on cigarette smoking and urinary pH. *Journal of Experimental Psychology: General, 106*, 24-30.

Schachter, S., Silverstein, B., & Perlick, D. (1977). Studies of the interaction of psychological and pharmacological determinants of smoking: V. psychological and pharmacological explanations of smoking under stress. *Journal of Experimental Psychology: General, 106*, 31-40.

Sorensen, G., Barbeau, E., Hunt, M. K., & Emmons, K. (2004). Reducing social disparities in tobacco use: A social-contextual model for reducing tobacco use among blue-collar workers. *American Journal of Public Health, 94*(2), 230-239.

Spielberger, C. D., & Reheiser, E. C. (2006). Psychological defense mechanisms, motivation and the use of tobacco. *Personality and Individual Differences, 41*, 1033-1043.

Sterling, T. D., & Weinkam, J. J. (1978). Smoking patterns by occupation, industry, sex, and race. *Archives of Environmental Health, 33*(6), 313-317.

Sterling, T. D., & Weinkam, J. J. (1976). Smoking characteristics by type of employment. *Journal of occupational medicine.: official publication of the Industrial Medical Association, 18*(11), 743-754.

U.S. Department of Health and Human Services. (2007). Department of Health and Human Services Office of the Secretary: Annual update of the HHS poverty guidelines. Federal Register, 72(15), 3147-3148.

U.S. Department of Health and Human Services. (2001). *Women and smoking: A report of the surgeon general-2001.* Rockville, MD.: Office on Smoking and Health, Public Health Service.

U.S. Department of Health and Human Services. (1998). *Tobacco use among U.S. Racial/Ethnic minority groups: A report of the surgeon general.* Rockville, MD.: Office on Smoking and Health, Public Health Service.

U.S. Department of Health and Human Services. (1994). *Preventing tobacco use among young people: A report of the surgeon general*. Rockville, MD.: Office on Smoking and Health, Public Health Service.

U.S. Department of Health and Human Services. (1985). *The health consequences of smoking--cancer and chronic lung disease in the workplace: A report of the surgeon general.* Rockville, MD.: Office on Smoking and Health, Public Health Service.

Walsh, M., Basta, J., Nodora, J., Solop, F., Hagen, K., & McCarrier, K. (2003). *Arizona adult tobacco survey 2002 executive report*Arizona Department of Health Services. from http://www.azteppdata.org/docs/presentations/2002_AZ_AdultTobaccoSurvey.pdf

Welty, T. K., Lee, E. T., Yeh, J., Cowan, L. D., Go, O., & Fabsitz, R. R., et al. (1995). Cardiovascular disease risk factors among american indians. the strong heart study. *American Journal of Epidemiology, 142*(3), 269-287.

Wetter, D. W., Cofta-Gunn, L., Fouladi, R. T., Irvin, J. E., Daza, P., & Mazas, C., et al. (2005). Understanding the associations among education, employment characteristics, and smoking. *Addictive Behaviors, 30*(5), 905-914.

Wetter, D. W., Cofta-Gunn, L., Irvin, J. E., Fouladi, R. T., Wright, K., & Daza, P., et al. (2005). What accounts for the association of education and smoking cessation? *Preventive medicine, 40*(4), 452-460.

Winkleby, M. A., Fortmann, S. P., & Barrett, D. C. (1990). Social class disparities in risk factors for disease: Eight-year prevalence patterns by level of education. *Preventive medicine, 19*(1), 1-12.

Winkleby, M. A., Fortmann, S. P., & Rockhill, B. (1992). Trends in cardiovascular disease risk factors by educational level: The stanford five-city project. *Preventive medicine, 21*(5), 592-601.

Winkleby, M. A., Jatulis, D. E., Frank, E., & Fortmann, S. P. (1992). Socioeconomic status and health: How education, income, and occupation contribute to risk factors for cardiovascular disease. *American Journal of Public Health, 82*(6), 816-820.

Yan, L. L, Liu, K., Daviglus, M. L., Colangelo, L. A., Kiefe, C. I., Sidney, S., Matthews, K. A., and Greenland, P. (2006). Education, 15-year risk factor progression, and coronary artery calcium in young adulthood and early middle age- the coronary artery risk development in young adults study. *JAMA,* 295(15), 1793-1800.

Zhu, B. P., Giovino, G. A., Mowery, P. D., & Eriksen, M. P. (1996). The relationship between cigarette smoking and education revisited: Implications for categorizing persons' educational status. *American Journal of Public Health, 86*(11), 1582-1589.