Labor Market Effects of Obesity, Smoking, and Alcohol Use *†

Ilker Dastan ‡

Department of Economics

Rutgers University

November 15, 2009

Abstract

This paper analyzes the joint effects of obesity, smoking, and binge drinking on wages and on unemployment by using the National Longitudinal Survey of Youth data. The main objective of this study is to show that the effects of these behaviors on wages and unemployment are not measured accurately in analyses that consider only one or two since these behaviors are correlated or tend to cluster. My results illustrate that failing to include one or more of the risky behaviors in wage or unemployment regression would lead to an underestimation of the impact of being obese and an overestimation of the effect of binge drinking for both genders. However, when endogeneity is addressed by employing the Hausman-Taylor instrumental variable (HTIV) method in wage analyses and the multivariate probit method in unemployment analyses, I find that the estimated parameters of obesity or binge drinking are not statistically significantly different whether these behaviors are considered individually or simultaneously.

This study also conducts several sensitivity analyses. Firstly, the results reveal that the effects of these risky behaviors are not interactive. Secondly, the paper illustrates that the wage penalties for daily smoking are fairly constant over the wage distribution for both genders, but obesity affects the wages of males and females relatively more at lower quantiles of wages, and there is no wage penalty for being a binge drinker for either gender. Further, it is found that smokers are a heterogeneous group of people. In particular, the wage and unemployment effects of persistent smokers are different than beginning smokers and quitters. Moreover, obesity affects the wages and the likelihood of being unemployed of males only at the extremes of obesity. Lastly, I find evidence of wage penalties for being obese or a smoker in private sector jobs, but in the public sector only male smokers face lower wages.

JEL Classification: I1, J2, J3

Key words: obesity, smoking, alcohol, wages, unemployment, endogeneity, Hausman-Taylor instrumental variable, multivariate probit, quantile regression

---

* Please visit http://econweb.rutgers.edu/dastan for the most updated version of this paper.

† **Acknowledgements:** I am grateful to my dissertation advisor Louise B. Russell for her invaluable help and support, and many insightful suggestions and discussions. This study also benefitted enormously from the guidance and feedbacks of my dissertation committee members Mark Killingsworth, Jeffrey Rubin, Irina Grafova, and the participants of empirical microeconomics seminar at Rutgers. All mistakes are mine.

‡ **Contact Information:** Phd candidate, Department of Economics, Rutgers University. Telephone: 1-732-322-7298. Web: http://econ.rutgers.edu/dastan. Email: dastan@econ.rutgers.edu. Address: 75 Hamilton Street, New Brunswick, NJ 08901.

1. Introduction:

Current studies show that about 35% of the US population is obese, approximately 20% of Americans are smokers, and around 15% are excessive alcohol drinkers (CDC, 2007)[1]. Although there has been a decline in smoking rates in recent decades, the prevalence of obesity rates has risen rapidly while drinking rates have remained constant. Obesity is an important risk factor for numerous health problems including coronary heart disease, diabetes, osteoarthritis, hypertension, and stroke. Excessive alcohol consumption can lead to liver problems, infertility, cancer, and unintentional injuries due to accidents. Smoking has deadly consequences, including lung, larynx, esophageal, and oral cancers (WHO, 2002). Additionally, these health risk behaviors are widely documented to cause major problems with potentially important social and economic consequences, including increases in medical expenditures, lost productivity, work absence, unemployment, social penalties and discrimination (Harwood, 2000; Sturm, 2002).

Economists have given considerable attention to the effects of smoking, obesity and excessive alcohol use on labor market outcomes. However, their results on the effects of these risk behaviors on workers' wages differ widely. Despite the potentially large negative effects of health risk behaviors on unemployment, surprisingly little research has been done on this issue. Further, none of the earlier studies considered the effects of all three risk factors on labor market outcomes in the same analysis. Cigarette smoking, excessive alcohol use, and poor eating habits tend to reinforce each other (Betts, 2000), and it is known that smokers are less likely to be obese. Hence, when these risk factors are considered individually, estimates of their effects on wages or unemployment may be biased since the behaviors may be correlated. Therefore, the

[1] Obesity is defined as BMI (kg/m^2) of 30 or more, and identification for those under age 21 is through age- and gender-specific. Smoking is defined as smoking cigarettes daily or smoking an average of 1 or more cigarettes daily. Binge drinking is defined as drinking 6 or more alcoholic drinks in 4 or more occasions past month.

main focus of this paper is to estimate the joint effects of obesity, smoking and excessive alcohol use on wages and on unemployment using the National Longitudinal Survey of Youth (NLSY) data.

The NLSY consists of a nationally representative sample of 12,686 young men and women who were 14-21 years old when they were first surveyed in 1979. Data were collected yearly from 1979 to 1994, and biennially from 1996 to the present. The NLSY gathers information about this large group of young adults as they make the transition from school to work, and it provides a comprehensive overview of respondents' labor force experiences, career changes, labor market attachment, marriage, education investments, etc. (NLSY79 User's Guide, 2002). The use of a longitudinal data set enables the estimation of panel data methods, e.g. fixed effects model, which help avoid the potential bias caused by unobserved individual factors not captured in cross-sectional models. The use of panel data methods is an important extension to the literature because the data sets used in previous studies are mostly cross-sectional.

This study offers myriad additional contributions to literature: Firstly, the existing literature suggests that these health risk behaviors may be endogenous or, more specifically, missing or unobservable determinants of both risk behaviors and labor market outcomes may be correlated. Because the present data set contains extensive information on demographics, health risk behaviors, and related lifestyle behaviors, the analysis is able to examine risk behaviors/labor market outcomes relationships while also addressing the potential endogeneity of risk behaviors via Hausman-Taylor Instrumental Variable model in wage analyses and multivariate probit model in unemployment analyses, which appear to be the most appropriate methods for the purpose and the sample.

Another goal of this study is to examine whether the effects of these health risk behaviors on wages and unemployment are interactive or additive. For example, the effects of smoking and excessive drinking are interactive if the effects of excessive drinker smokers differ from the two effects added together, e.g. if they are due to the same unobserved factors. However, the effects would be additive if the health of the individual worsens more when the individual is attached to another behavior in addition to the current behavior. For this purpose, a set of interaction terms are included in the analyses and are tested for statistical significance.

Thirdly, the present study separates males and females, and blacks, Hispanics and whites into subsamples to conduct a richer examination of gender and racial/ethnic differences. Further, one may argue that obese, smokers or binge drinkers may tend to be in public sector jobs since workers are more exposed to wage discrimination in the private sector. The effects of obesity, smoking and binge drinking on wages are examined for private and public sector workers separately.

Lastly, higher wage workers may likely be less affected by these health risk behaviors, possibly because of the nature of their jobs, than lower wage workers or workers with less education, or vice verse. Hence, it is also examined to observe whether the effects of the three risk factors on wages might vary across the wage structure.

The OLS results of wage analyses reveal that the wage effect of obesity is underestimated if one fails to control for smoking and drinking behaviors of males. Moreover, the wage effects of binge drinking is underestimated for males (estimate is positive) and overestimated for females if alcohol use of the respondents is considered individually. Similarly, the probit results of unemployment analyses illustrate that failing to include one or more of the health behaviors in unemployment regressions would lead to an underestimation of the impacts of being obese and overestimation of the effect of binge drinking for both genders.

The results also show that once endogeneity is controlled for, the estimated parameters of obesity or binge drinking are not statistically significantly different whether these behaviors are considered individually or simultaneously in both wage and unemployment analyses. Therefore, the biases in the OLS and probit estimates are due to failure to account for endogeneity. Moreover, obesity has negative effects only on the wages of females (a penalty of 4.3%), smoking wage penalties are 4.8% and 2.5% for males and females, respectively. Being obese increases the probability of being unemployed by 1.8% for females only, smoking increases the likelihood of being unemployed by 2.5% and 1.7% for males and females, respectively. Binge drinking has no effect on wages or on unemployment for either gender.

Sensitivity analyses illustrate that the effects of risk behaviors are not interactive. But, once interaction terms are included into the models, binge drinking estimates for males become statistically significant both in wage and unemployment analyses, which shows the importance of including interaction terms into regressions. Thus, the effect of binge drinking could be correlated with the obesity and/or smoking behavior of the individual, for instance, if an obese person or smoker already perceives discrimination, becoming also a binge drinker may not result in any additional discrimination.

Furthermore, the results of ethnic/racial subsamples illustrate that although smoking has wage penalties for all subsamples, obesity affects the wages of only white females, and binge drinking affects wages of only white males. Also, I find evidence of wage penalties for obesity or smoking in private sector jobs. However, in the public sector, only males face lower wages only due to smoking.

Moreover, the results reveal that the wage penalties for daily smoking are fairly constant over the wage distribution for both genders. However, obesity affects the wages of males and females relatively more at lower quantiles, and there is no wage penalty for being a binge drinker for males and females at any quantile.

I also estimated wage and unemployment models using more detailed measures. My results when endogeneity is addressed demonstrate that smokers are a heterogeneous group of people: the wage and unemployment effects of persistent smokers are different than starters, quitters or young experimenters. Furthermore, obesity appears to affect the wages and the likelihood of being unemployed of males only at the extremes[2]. Moreover, contrary to the previous literature, drinking is found to have no positive effect on wages or on unemployment for either gender when endogeneity is accounted for.

The structure of the paper is as follows: The conceptual rationales for why these behaviors may affect labor market outcomes, the literature review and the empirical methodology are discussed in sections two, three and four, respectively. The data is described in section five together with the descriptive statistics. The results are presented in section six and the final section of the paper presents conclusions.

[2] Morbid obesity is defined as having a BMI or 35 or more.

2. Conceptual Rationales:

Existing studies explain the correlation between health risk behaviors and labor market outcomes in four ways (Morris, 2007):

(1) Health risk behaviors may cause lower wages or unemployment. This might arise for three reasons:

a) First, obesity, smoking or binge drinking can be weakening health conditions. Thus, the obese, smokers or binge drinkers are likely to be less productive than their counterparts and consequently less likely to be employed or more likely to receive lower wages.

Obesity is a major health problem related to a number of serious diseases. Furthermore, obesity can lead to some psychological problems occurring from discrimination, humiliation, social rejection or unhappiness with oneself. Existing research also shows that excessive alcohol use and smoking are associated with a long list of physical, psychological, and cognitive impairments. These problems at the personal level may also affect the individual's relationship to the labor market. Hence, obesity, smoking, and binge drinking may reduce work ability and increase absence from work, thereby reducing the individual's productivity, which in turn leads to lower wages and a lower likelihood of employment. (Mullahy and Sindelar 1993, 1996; Kenkel and Ribar 1994; Burkhauser and Cawley, 2004).

Furthermore, most health publications argue that there is a J-shaped association between alcohol consumption and the risk of cardiovascular disease, which suggests that alcohol consumption at moderate levels may be beneficial for health by relieving stress and reducing the incidence of heart disease (Sesso, 2001; Baum and Ford, 2006). Therefore, light or moderate drinking may be favorable to an individual's labor market productivity and might have positive effects on wages and on employment.

b) Second, there may be discrimination against the obese, smokers or binge drinkers which lessens their likelihood of being employed or negatively impacts their wages.

There may be prejudice by employers, employees, or customers reflecting their distaste or negative preferences for workers with health risk behaviors, particularly against obese women and smokers (Becker, 1971; Moon and McLean, 1980). It has been also documented in numerous experimental studies that there are weight-based discrimination at every stage of employment, from the hiring decision through wage-setting and promotion (Puhl and Brownell, 2001). Moreover, the studies also state that individuals who have these risk behaviors delay marriage, have reduced labor market experience, lower educational attainment, and greater probability of divorce (Yamaguchi and Kandel, 1985; Kenkel and Ribar, 1994), which indirectly affect their labor market success and reduce their productivities.

c) Third, these workers could be more costly for employers who provide health care. It is apparent that obese persons, smokers or drinkers will need more medical care than their counterparts. Economic theory implies that wages are decreasing in fringe benefits. Therefore, higher health care costs would lead to lower wages if employers provide health insurance benefits.

(2) Unemployment or lower wages may cause obesity, smoking or binge drinking:

Unemployment, low income level, social isolation or economic constraints may restrict access to healthy food and safe exercise, and may increase the likelihood of weight changes (Cawley, 2004). Unemployment may also affect weight if there is a relationship between long-term unemployment and mental health, and mental health is correlated with obesity (Bove and Olson, 2005). Additionally, sociology and alcohol research literatures state that unemployment causes emotional and financial stress, and stress is known to be related to weight gain, increased alcohol consumption, and smoking (Forcier, 1985; Harris et al., 1998).

(3) There may be unobserved variables that are correlated with both employment/wages and health risk behaviors:

A third factor may explain the relationship between health risk behaviors and employment/unemployment, e.g. time preference, self-control, sociability, reduced aspirations and lower commitment to values underlying to desire for success, etc. For instance, this factor could reflect preferences for current time as opposed to preferences for the future. Hence, people with health risk behaviors who are less likely to be concerned about the future and possible negative effects of these behaviors may invest less in human capital now and may be less likely to engage in job training. As a result, they would have low likelihood of being employed or lower earning profiles (Fuchs 1974; Becker et al. 1994). Alternatively, alcohol may play a networking role during the time spent with colleagues and could help the person to get additional information about promotion opportunities or may serve as a signal of commitment to the firm at company meetings (Montgomery, 1991; MacDonald and Shields, 2001).

(4) Health risk behaviors may be measured systematically with error due to unobserved factors correlated with employment or labor market success. For example, this may occur if individuals in lower socioeconomic groups are more likely to underreport their weights, smoking or alcohol use, or over report their heights, etc. (Greeve, 2008).

The goal of the paper is to identify the first effect and to produce unbiased estimates that are not contaminated by the other effects.

3. Literature Review:

3.1. Effects of Obesity:

The effects of obesity on labor market success, i.e. earnings, labor supply and occupation selection, have been analyzed in numerous US studies. The studies that do not use NLSY data find obesity wage penalties only for females (Sargent and Blanchflower, 1994; Behrman and Rosenzweig, 2001). The results of the studies that employ NLSY data and cross-sectional analyses are differing: some studies find wage penalties for both genders (Averett and Korenman, 1996; Maranto and Stenoien, 2000), some argue there is no penalty for either genders (Loh, 1993), and the remaining find obesity reduces female wages, but not those of males (Register and Williams, 1990; Pagan and Davila, 1997).

Cawley (2004) uses NLSY data and accounts for endogeneity employing fixed effects and instrumental variable (IV) models using BMI of sibling as an instrument. He finds obesity wage penalty only for white females. Employing NLSY and individual and sibling fixed effects models, Baum and Ford (2004) find obesity wage penalty for both genders but more for females.

The effect of obesity on employment has taken less attention and generally been less examined by European studies. Some of European studies that examine the effect of obesity on employment find insignificant effects for both genders (Harper, 2000). Some find positive effects on the long-term unemployment only for females (Sarlio-Lahteenkorva and Lahelma, 1999), and some find weak evidence that obese individuals are more likely to be unemployed (Garcia and Quintana-Domeque, 2007). Morris (2007) addresses the endogeneity by employing the bivariate probit (IV) model, uses prevalence of obesity in the area in which the respondent lives as the instrument, and finds that obesity has a negative effect on employment for both males and females. Greeve (2008) also uses IV models, employs whether the interviewee's father or mother had been prescribed medication for obesity as instrument and finds negative effects of BMI on employment.

Cawley (2000a) is the only US study that examines employment probabilities. He uses a conventional IV regression approach, the BMI of a biological child as an instrument, and finds that BMI has a positive effect on employment for mothers.

3.2. Effect of Alcohol Use:

Previous literature on the labor market effects of alcohol consumption concentrates on US data and differs in the alcohol consumption measures used. The choice of the alcohol measures are usually determined by data availability. The commonly used measures are binary indicators of alcohol consumption at different levels, frequency of consumption over some periods and clinical measures of alcohol consumption. Most of the studies use only cross-sectional samples and a few of them control for endogeneity or unobserved heterogeneity.

The results of the wage effects of alcohol use are contradictory. Some of the results state that light to moderate drinkers have higher wages than abstainers and heavy or binge drinkers (Berger, 1988; Hamilton and Hamilton, 1997; McDonald and Shields, 2001). While some of the studies do not find any penalties for heavy or binge drinking (Zarkin et al., 1998a), others find no evidence for benefits in terms of higher wages from moderate drinking over abstinence (Bryant et al., 1992).

Employing NLSY data and addressing endogeneity by using sibling and individual fixed effects, Kenkel and Ribar (1994) find that binge and heavy drinkers earn less than abstainers and much less than moderate drinkers, but Peters (2004) finds positive significant effects of drinking on wages, and Renna (2008) uses first difference regressions and fails to find any association.

Employing US data and accounting for endogeneity by IV models, Mullahy and Sindelar (1993, 1996) and Bryant et al. (1996) find insignificant effects of problem drinking on the probability of employment, but Terza (2002) finds negative significant effects for males. Kenkel and Ribar (1994) analyze NLSY data and account for endogeneity, but they find a positive and statistically significant association for females.

Employing non-US data and clinical measures of alcohol consumptions, some studies find lower employment probabilities for alcohol dependent men (MacDonald and Shields, 2004), but others find no significant effect for either gender when endogeneity is addressed via IV and fixed-effects models (Tekin, 2004; Feng et al., 2001).

3.3. Effect of Smoking:

The labor market effects of smoking have gained less attention but the results are quite similar. Studies that use international data (Lee, 1999) and that use US data (Leigh and Berger, 1989; Viscusi and Hersh; 2001) find that there is a wage penalty for smokers in all subgroups. Grafova (2009) finds that persistent smokers have lower wages than non-smokers, but smokers who will quit smoking in the future do not have a wage penalty even when they smoke. Levine et al. (1997) use NLSY data and accounts for endogeneity by using sibling fixed-effects. Their results support the earlier findings that smoking affects wages of both genders.

Although Zarkin et al. (1998) fails to find any impact of smoking on hours of labor supply of males, accounting for endogeneity via IV using the religiosity indicator as an instrument, French et al. (2001) finds a negative effect of smoking on employment. Further, employing NLSY and controlling for endogeneity, Kandel and Davies (1990) find negative effects of smoking on job mobility, employment gaps, and duration of unemployment.

3.4. Combined Effects of Alcohol Use and Smoking:

Studies that use international data (Lee, 1999; Auld, 2005) find that moderate drinking is associated with a higher income than drinking abstinence, while smoking is associated with lower income than non-smoking after correcting for endogeneity. Yet the effects of heavy drinking are still unclear. Moreover, it is found that failure to control for other health risk behaviors leads to a moderate underestimate of the effect of smoking or drinking.

3.5. Combined Effects of Obesity and Smoking:

Berger and Leigh (1989) find no wage penalty for either behavior, but Baum and Ford (2006) find an obesity wage penalty only for female workers when both health behaviors are included in the same analysis and when accounting for endogeneity. Controlling for both risk factors, obesity is significantly related to unemployment for women and heavy smoking is related to unemployment for men by Jusot et al. (2008).

3.6. Shortcomings of the Literature:

Most of the studies in the literature have some shortcomings and these shortcomings may be the reasons why they provide contradictory results. Firstly, to my knowledge, none of the studies have used all three risk behaviors in the same regression analysis. Secondly, some articles use small and unrepresentative samples. For instance, several studies use the very early years of longitudinal data in which the respondents were only very young and not representative of the general population. Additionally, although some of the studies account for endogeneity of health risk behaviors by employing instrumental variables (IV) methods, most of their identification strategies are somewhat poor, because they rely on very strong assumptions and the instruments they use may be weak. Furthermore, with the exception of a few studies, almost all articles rely on cross-sectional data and do not use panel data techniques to control for unobserved heterogeneity. Lastly, the results differ in the measures of employment and health risk behaviors used, since the choice of the measures is usually dictated by data availability.

4. Empirical Methodology:

The empirical work in this study is based on a theoretical framework developed by Mullahy and Sindelar (1996) and subsequently adopted by many economists. They offer the most commonly used and general model of labor supply that depends on wages, desired hours of work and the decision to participate. An individual maximizes his/her utility by allocating times for labor and leisure consumption activities, which yields the labor supply of the individual. But the individual's labor supply decision is turned into employment only if the individual is selected for employment, which is partly based on reservation wage, market wage and the employer's decision of offering employment. Hence, in short, the employment status of an individual is a function of human capital of the individual, economic conditions, demographic factors, and other factors that affect the individual's choice between labor and leisure consumption.

In light of earlier studies, in the simplest forms the employment equation and equations of health risk behaviors can be modeled as:

$$E_{it} = E(P_{it}, W_{it}, X_{it}, u_{it}) \quad (1)$$

$$O_{it} = O(P_{it}, W_{it}, X_{it}, o_{it}) \quad (2)$$

$$S_{it} = S(P_{it}, W_{it}, X_{it}, s_{it}) \quad (3)$$

$$BD_{it} = BD(P_{it}, W_{it}, X_{it}, bd_{it}) \quad (4)$$

where E, O, S, and BD are binary indicators of employment, obesity, smoking and binge drinking, respectively. P is all prices, W is wage rate, X is a vector of all observable factors, including non-labor income, and u, o, s, and bd are unobservable characteristics, i refers to the ith individual and t represents the time. To relate obesity, smoking and binge drinking to employment, equation 1 can be rewritten as a reduced-form relationship that can be expressed as:

$$E_{it} = E(O_{it}, S_{it}, BD_{it}, X_{it}, u_{it}) \tag{5}$$

This type of reduced-form equations eliminates the need to specify the underlying structure of the health risk behaviors, and they are also valid in the presence of corner solutions. Assuming linearity, the econometric equivalent of equation 5 can be expressed as:

$$E_{it} = \beta_1 O_{it} + \beta_2 S_{it} + \beta_3 BD_{it} + \beta_4 X_{it} + u_{it} \tag{6}$$

where β's are the parameters to be estimated. The employment equation can be estimated by a standard probit or logit model:

$$E_{it}^{*} = \beta_1 O_{it} + \beta_2 S_{it} + \beta_3 BD_{it} + \beta_4 X_{it} + e_{it} \tag{7}$$

where E_i^* is unobservable but $E_i = 1$ if $E_i^* > 0$. e is assumed to be a mean-zero, constant-variance random variable and it is uncorrelated with the explanatory variables. The coefficients can be obtained via maximum likelihood estimations.

In order to investigate the wage effects of obesity, smoking and binge drinking, a standard human capital earnings regression is estimated including some standard covariates. It is assumed that the effects of the three risk factors on wages can be estimated from the following 'wage level' model:

$$\ln(W_{\mathrm{it}}) = \beta_1 O_{it} + \beta_2 S_{it} + \beta_3 BD_{it} + \beta_4 X_{it} + \varepsilon_{it} \tag{8}$$

where W is the wage rate, X is a vector of personal and geographical characteristics that affect labor market outcomes, O, S, and BD are binary indicators of obesity, smoking and binge drinking and ε is the error term for observation i in time t. I assume that ε is a mean-zero, constant-variance random variable and it is uncorrelated with the explanatory variables in X. The coefficients can be estimated via ordinary least squares (OLS) regression.

Providing unbiased correlations between obesity, smoking, binge drinking and unemployment/wages is not straightforward because the health risk behaviors and unemployment/wages may simultaneously affect each other. In addition, an unobserved factor may be correlated with both health risk behaviors and labor market outcomes, such as time preference, self-control, sociability, etc. Technically these mean that the error term in unemployment/wages equation is correlated with obesity, smoking and binge drinking binary variables. Therefore, firstly a supplemental set of covariates are included into the model to control more widely for individual background characteristics that may affect labor market outcomes.

In a cross-sectional framework, previous studies have generally dealt with the problem of these kinds of endogeneity with conventional two-stage IV models. In such an approach one needs to instrument obesity, smoking and binge drinking with an instrument (or instruments) that is (are) correlated with health risk behaviors, but uncorrelated with labor market outcome.[3] However, the achievement of unbiased estimates via the IV method depends essentially on the predictive power and validity of the instruments. If there is weak correlation between instruments and health risk behaviors, or if the instruments are correlated with labor market outcome, then the IV estimates could still be biased. Moreover, mean square errors with an IV method may be

[3] The same approach has also been employed in this study. See appendix for the results.

large which implies a trade-off between bias and variance (Bollen et al., 1995; Norton et al., 1998). Furthermore, IV estimation is less efficient, because it does not fully exploit the specification and non-linearity of the unemployment model (Feng, 2001).[4]

A longitudinal data offers alternative solutions to the endogeneity problem: fixed-effects, random-effects, between-effects models.[5] The most commonly used model that deals with unobserved individual heterogeneity is fixed-effects, which lets one control for omitted or unobserved variables that differ between cases but constant over time by estimating the model as deviations from the means. However, the fixed-effects model does not eliminate bias in case of time-variant individual heterogeneity and takes out all the covariates that are time-invariant. Further, the model may worsen the bias caused by measurement error (Griliches and Hausman, 1986). Also, the model cannot identify the effects of individuals who do not change their behavior over time: it assumes individuals who change their behavior over time are similar to the ones who do not change their behavior over time.

To account for endogeneity, the Hausman-Taylor Instrumental Variable (HTIV) method is employed in wage analyses. This method is the most appropriate for the purpose and sample of this study. In the presence of correlation between some covariates and the unobserved individual characteristics, this method produces consistent estimates contrary to the OLS model. Furthermore, in contrast to the conventional IV methods, there is no need for finding external

[4] There is a large literature claiming that failing to control for selection into working could lead to sample selection bias in wage effects analyses. To account for this bias, a two-step Heckman correction method is employed for males and females separately. Non-wage family income is served as an instrument for the propensity of the individual to work or receive positive wage. The Mills Ratio correction terms which are employed to control for selection into working are not statistically significant in the wage equation either for males or females. Furthermore, the correction does not change the sizes and significance levels of the three health risk behaviors, hence the results with the Heckman correction are ignored and not presented.

[5] Some panel data techniques are also employed in this study. See appendix for the results.

instruments. In addition, in contrast to the fixed effects model, the Hausman-Taylor IV model allows producing the estimates for time-invariant covariates. Let the wage model be:

$$\ln(\boldsymbol{W}_{\mathrm{it}}) = \beta_1 \boldsymbol{X}_{1it} + \beta_2 \boldsymbol{X}_{2it} + \gamma_1 \boldsymbol{Z}_{1i} + \gamma_2 \boldsymbol{Z}_{2i} + \mu_i + \varepsilon_{it}, \tag{9}$$

where ε is a random error uncorrelated with explanatory variables, μ is unobserved individual characteristics, $\boldsymbol{X}_{1it}$ ($\boldsymbol{X}_{2it}$) are variables that are time-varying and uncorrelated (correlated) with μ_i, and $\boldsymbol{Z}_{1i}$ ($\boldsymbol{Z}_{2i}$) are variables that are time-invariant and uncorrelated (correlated) with μ_i. This model predicts potentially endogenous variables with a set of instruments obtained from within the model. Hausman and Taylor propose an instrumental variable approach using the following variables as instruments in the final GLS estimator: $\boldsymbol{X}_{1it}, \boldsymbol{Z}_{1i}, \boldsymbol{X}_{2it} - \overline{\boldsymbol{X}}_{2i}, \overline{\boldsymbol{X}}_{1i}$ which are individual-specific means of exogenous variables, individual-specific mean differences of exogenous variables, and individual-specific mean differences of endogenous variables (Hausman and Taylor, 1981). I considered obesity, smoking, and drinking variables as endogenous variables and all others as exogenous variables.[6]

In order to account for the endogeneity of health risk behaviors in the unemployment analyses, the study estimates a multivariate probit model, which appears to be more appropriate and more efficient than conventional two-stage IV methods:

$$\begin{gathered} U_{it}=1 \text{ if } U_{it}^{*}\geq 0 \text{ and } U_{it}=0 \text{ if } U_{it}^{*}\leq 0 \\ O_{it}=1 \text{ if } O_{it}^{*}\geq 0 \text{ and } O_{it}=0 \text{ if } O_{it}^{*}\leq 0 \\ S_{it}=1 \text{ if } S_{it}^{*}\geq 0 \text{ and } S_{it}=0 \text{ if } S_{it}^{*}\leq 0 \\ BD_{it}=1 \text{ if } BD_{it}^{*}\geq 0 \text{ and } BD_{it}=0 \text{ if } BD_{it}^{*}\leq 0 \end{gathered} \tag{10}$$

where $U_{it} = 1$ means the individual is unemployed and $U_{it} = 0$ means the person is employed. Similarly, $O_{it} = 1$, $S_{it} = 1$, and $BD_{it} = 1$ indicate the individual is obese, smoker, and binge

[6] I examined a couple other variables as endogenous, such as experience and marital status, but the results are unchanged.

drinker, respectively. U_{it}^*, O_{it}^*, S_{it}^*, and BD_{it}^* are unobserved latent variables that determine being unemployed, obese, smoker, and binge drinker, respectively. The empirical specification for the multivariate model can be written as:

$$\begin{bmatrix} U_{it}^* \\ O_{it}^* \\ S_{it}^* \\ BD_{it}^* \end{bmatrix} = \begin{bmatrix} \beta_1 O_{it} + \beta_2 S_{it} + \beta_3 BD_{it} + \beta_4 X_{it} \\ \alpha_1 X_{it} + \alpha_2 Z_{it} \\ \gamma_1 X_{it} + \gamma_2 Z_{it} \\ \delta_1 X_{it} + \delta_2 Z_{it} \end{bmatrix} + \begin{bmatrix} u_{it} \\ o_{it} \\ s_{it} \\ bd_{it} \end{bmatrix}$$

$$E[u] = E[o] = E[s] = E[bd] = 0$$
$$Var[u] = Var[o] = Var[s] = Var[bd] = 1$$
$$Cov[u, o] = \rho_{12},\ Cov[u, s] = \rho_{13},\ Cov[u, bd] = \rho_{14} \quad (11)$$

where Z is a vector of instruments, β's, δ's, γ's, and α's are parameters to be estimated, β_1, β_2, β_3 are the coefficients of interest, u, o, s, and bd are error terms distributed multivariate normally with mean equal to 0 for each variable and a variance-covariance matrix that has the value of 1 on the principal diagonal and the correlation terms $\rho_{mn} = \rho_{nm}$ as the off-diagonal terms. ρ_{12}, ρ_{13} and ρ_{14} are the correlations between the error terms in employment and obesity, smoking, and binge drinking equation, respectively.

A combination of frequently used instruments are also employed in the multivariate probit model[7]: several religion and religious attendance dummies, Rotter test score, youngest age started drinking dummies, alcoholic father in childhood (only satisfied for males), alcoholic mother in childhood (only satisfied for females) and cumulative years lived with alcoholic

[7] Earlier studies have widely used state cigarette tax, state alcoholic drinks tax (Kenkel and Ribar, 1994), and the BMI of siblings, parents, children (Cawley, 2004), prevalence of obesity, smoking, and drinking across individuals living in the same health authority area (Morris, 2007) as instruments for the individual's smoking behavior, drinking behavior, and obesity, respectively. However, it has been argued that most of these instruments are weak since they may be correlated with employment prospects as well, and adopting the health risk behavior of a sibling, child or parent would decrease the sample size largely. Moreover, the state taxes could not be used in this study, because information about the individuals' geographic location is not part of the public release version of NLSY79 data and these measures can only be obtained via Geocode data of NLSY79 data, which I could not yet acquire because of the strict policies of the BLS.

relative variables (Berger and Leigh, 1988; Bryant et al., 1996; Heien, 1996; Pagan and Davila, 1997; Hamilton and Hamilton, 1997).

This multinomial probit model is estimated according to the method of maximum likelihood using 'mvprobit' in Stata. Wald tests of significance of ρ's are used to test the endogeneity of unemployment and health risk behaviors (Wooldridge, 2002). If ρ's are significantly different from zero, then they are endogenous and probit/logit estimates are biased, hence the multinomial probit model should be used. If ρ's are zero, then the probit model of equations 7 would generate consistent estimates and there is no need for a multivariate model.

5. Data:

5.1. Data and Sample:

I use the National Longitudinal Survey of Youth (NLSY) data to study the effects of obesity, smoking and binge drinking on labor market outcomes. The NLSY is a comprehensive survey sponsored and directed by the Bureau of Labor Statistics of the U.S. Department of Labor. The main focus of the survey is labor force behavior. NLSY began in 1979 with 12,686 respondents (6,403 are male) ages between 14 and 21. These individuals were chosen from both civilian and military populations, and were interviewed annually from 1979 until 1994. After the 1994 survey, the NLSY began interviewing biennially. In 2004, these individuals ranged in age from 39 to 46. In each survey, the NLSY gathers information on each respondent's employment status, salaries, age, personal characteristics, family characteristics, etc.

A combination of commonly used data specifications are employed to construct the sample from the data. All available waves of the survey are used from 1979 to 2004 for wage analyses and all waves from 1979 to 1998 are used for unemployment analyses since unemployment statuses of the individuals cannot be determined after 1998. Most of the existing studies exclude respondents in years during which their education has yet to be completed and person-year observations in which the respondent is less than 18 years old. The same exclusions are employed for my sample. I also exclude respondents who were in the armed forces and who did not have any employment or wage information. Employing Cawley's (2004) specification, the females who were pregnant in the survey year are also excluded from the sample since weight (and obesity) may be affected by pregnancy. Additionally, for the wage analyses, the respondents working as part-time or received any wage in the survey year are included in my sample, but I exclude respondents who worked as self-employed, without pay or in a family business as most other authors did. For the unemployment analyses, respondents who are out of the labor force are also dropped from the sample.

The sample for the wage analyses consists of 12,106 individuals (6,147 male, 5,959 female) after dropping 580 respondents, and the sample for unemployment analyses consists of 12,009 individuals (6,076 male, 5,933 female) after dropping 677 respondents. Person-year observations are used in the regression analyses and each person-year is included as a separate observation.[8] There are 125,264 person-year observations used in wage regressions, 65,364 of whom are male and 59,899 of whom are female (141,142 person-year observations are dropped). After dropping 109,315 person-year observations, the sample for unemployment regressions contains 119,033 person-year observations, 64,712 of whom are male and 54,321 of whom are female.

[8] Respondents provide multiple observations since person-year observations are employed and I use 'cluster' command in Stata to control for correlation among observations that come from the same individual. This command defines an error structure where only errors between observations from different people are independent.

5.2. Descriptive Statistics:

5.2.1. Dependent Variables:

Hourly wages are calculated by dividing 'total income from wages and salary' to 'hours worked', and all wages are adjusted using the Consumer Price Index from the US Census Bureau with a 2003 base period. The average real wage, in 2003 dollars, for all person-year observations is \$14.79 per hour. Nominal wages are increasing over time for both men and women, except the first five years when the respondents were still young or not employed yet as shown in Figure 1. Results indicate that there is a considerable wage differential between men and women in almost all years; for example the real wage for men in 2004 is \$25.90 and it is \$19.49 for women.

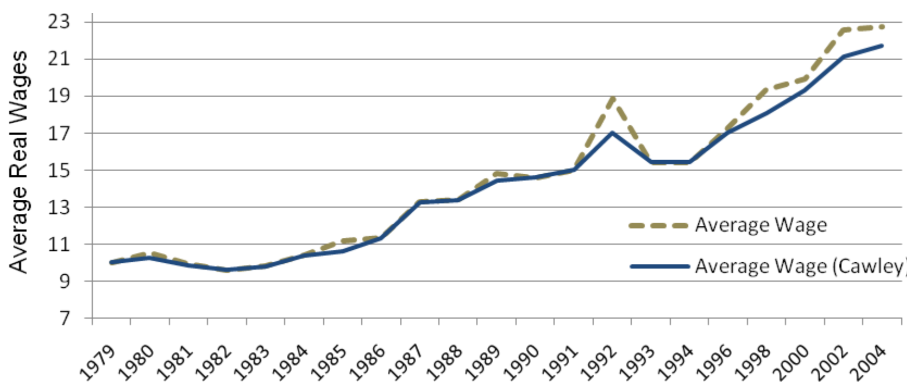

Figure 1: Average real wages by survey years

Because of the different question specifications in 1992, respondents' total net income from wages and salaries (and therefore average real wages) are not truncated, this is why there is a peak in this year. This problem is solved in Cawley's (2004) specifications. To get rid of the outliers, as Cawley (2004) did in his article, I recoded all the hourly wage information less than \$1 as \$1 and greater than \$500 as \$500. As a result, Cawley's specifications produce smoother wage distributions, and peaks in average wages, such as in 1992, disappeared.

NLSY provides detailed information about the labor force statuses of respondents by the questions asked in the Current Labor Force Status sections of the surveys. NLSY creates various employment statuses using the answers given to the question asked in all years of NLSY until 1998: 'What were you doing most of last week?'

Employed: (1) All civilians who did any work as paid employees, or in their own business or farm, or who worked 15 hours or more as unpaid workers in an enterprise operated by a member of the family; and (2) persons who were temporarily absent from work because of various personal reasons, whether they were paid for the time off or were seeking other jobs. Persons whose only activity consisted of work around the house or volunteer work are excluded.

Unemployed: All civilians who had no employment during the survey week, were available for work, and (1) had made specific efforts to find employment some time during the prior four weeks, (2) were waiting to be recalled to a job from which they were laid off, or (3) were waiting to report to a new wage and salary job scheduled to start within 30 days.

5.2.2. Obesity Variables:

Questions about weight were asked in every year except the 1979, 1980, 1983, 1984, 1988, and 1991 surveys and questions about height were asked only in the 1981, 1982, and 1985 survey years. I assumed that the heights of the respondents after the 1985 survey year are the same as the 1985 height observations since the youngest respondent was 21 years old in 1985. A linear increase of weight is assumed in the years, in which the respondent did not answer any weight question, and a linear increase of height is assumed in the 1983 and 1984 years since there is no height question asked in these years. Height information for 1981 is imputed for the first two years of the survey, 1979 and 1980. In 1981, respondents (between the ages of 16 and

23) weighed an average of 147.68 pounds. They weighed 185.44 pounds in 2004, when they were between the ages of 39 and 46.

Each respondent's body mass index (BMI) is used to measure obesity. BMI is defined as weight in kilograms divided by height in meters squared. A BMI less than 18.5 is considered underweight, a BMI of 18.5-24.9 is normal and a BMI greater than 25 is overweight according to the Centers for Disease Control and Prevention (CDC). Obesity is defined as a BMI of 30 or more. The CDC's current method for identifying obesity in those under age 21 is age- and gender-specific (CDC, 2006b) and only a few recent papers used this methodology. I use CDC's BMI growth charts to create obesity variables for the individuals under 21 years old.[9]

Health economics and medical studies demonstrate that there is a tendency for people to under-report their true weight but over-report their height (Cawley, 2004). Employing the third NHANES of 1988-1994, Cawley finds that on average men between the ages of 17 and 40 are inclined to over-report their true BMI by 0.02% and woman under-report their BMIs by around 1.5%. Therefore, to predict the true BMIs of the respondents from their reported BMIs and to correct the measurement error, Cawley's information and specifications are used in my study.

The average BMI over all survey years is 25.34 which is accepted as an overweight measure. In 2004, 1,634 respondents are obese and 2,149 are overweight out of 5,013 respondents. Table 1 shows that 19,125 person-year observations out of 125,264 have obese characteristics.

Figure 2 shows the average yearly BMI and obesity prevalence by year. In 1979, the average BMI was about 23.31, then average BMI increases monotonically and after 1989 the

---

[9] This method leads 150 person-year observations for men under 21 years old to change from overweight to obese, but leads around 100 person-year observations for women under 21 years old to change from obese to overweight categories.

average BMI reaches 25 meaning that respondents have an average BMI which is considered above the range currently considered best for health. The average wage for the full sample with a BMI of 30 or more is $14.65 and average wage with BMIs between 25 and 30 is $15.24.

Table 1. *Person-year observations*

Underweight	7,263 (6%)
Normal	57,254 (46%)
Overweight	41,622 (33%)
Obese	19,125 (15%)
Mild Obese	13,152 (10%)
Morbid Obese	6,013 (5%)
	125,264

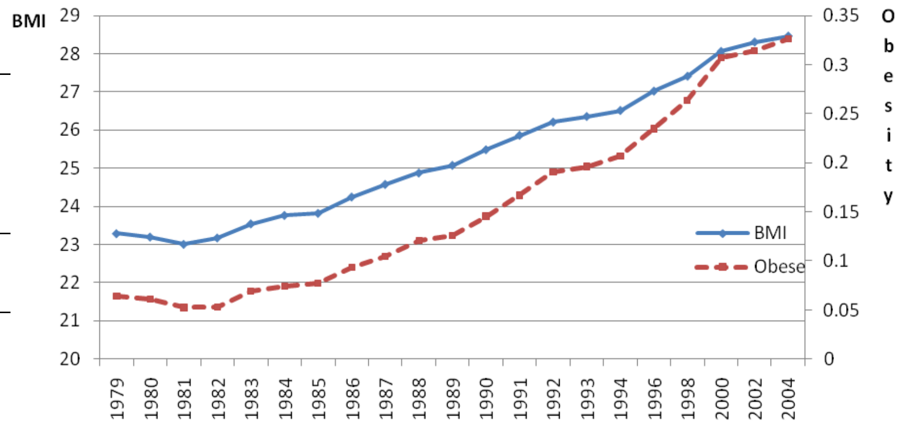

Figure 2: Body Mass Index (BMI) by year

5.2.3. Alcohol Consumption Variables:

Respondents are asked questions related to their alcohol consumption in the 1982, 1983, 1984, 1985, 1988, 1989, 1992, and 1994 survey years. However, the questions were not identical across the years. I employ Peters' (2004) and Keng and Huffman's (2007) methods to create alcohol consumption variables.

The 'current drinker' dummy variable is equal to one if the respondent answers positively to the question: "*did you drink any alcoholic beverages in last month*" or if the respondent gives an answer of at least one to the question in 1992: "*number of days drank alcohol in last month*". For the missing years, the earlier and later years' information about the drinking is imputed. For example, the 1986 and 1987 years do not have any questions about the drinking behavior of respondents. It is assumed that the 1986 drinking observations are the same with 1985, and the 1987 observations are the same with 1988. I also assume that the respondent is a current drinker in 1979, 1980 and 1981 if he is a current drinker in 1982.

Binge drinking is recently defined as drinking 5 or more drinks during a single occasion for men and 4 or more drinks during a single occasion for women by CDC. However, there is too little information in NLSY surveys to create binge drinking variables in compliance with this definition. Instead, I use Peters' and Keng and Huffman's specifications to create dummy variables indicating excessive alcohol use of the individuals. It is supposed that if the respondent gives an answer of 4 or more to the question "*number of days had 6 or more drinks in last month*", then the 'binge drinker' variable is equal to one. If the respondent drinks 6 or more drinks on one to three occasions, the respondent is defined as a 'heavy drinker'.

Table 2 illustrates that 65% of person-year observations are current drinkers and around 12% are binge drinkers. Figure 3 displays that the prevalence of current drinkers and binge drinkers are slightly decreasing over time.

Table 2. *Person-year observations*

Drinker	81,979 (65%)
Heavy Drinker	23,556 (19%)
Binge Drinker	14,820 (12%)
Permanent Abstainer	9,982 (8%)
Permanent Drinker	33,677 (27%)
	125,264

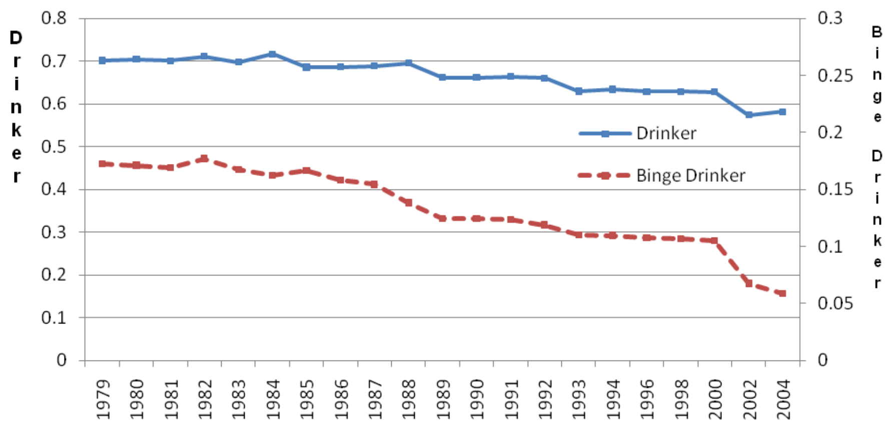

Figure 3: Percentage of Drinker by year

5.2.4. Smoking Variables:

Smoking questions are asked of respondents in the 1984, 1992, 1994 and 1998 survey years. The surveys are different across the years in the types of the questions asked. Therefore, I employ the similar methodologies that Levine et al. (1997) and Baum and Ford (2006) used in their papers.

'Daily smoker', 'heavy smoker' and 'light smoker' are created for all 18 years. The 'daily smoker' variable is equal to one if the respondent answers positively to the question asked in 1992, 1994 and 1998: "*does respondent currently smoke daily*". In 1984, respondents were asked the average number of cigarettes smoked per day, and I assume that the variable is also equal to one if the respondent averaged one or more cigarettes per day. If the answer to the question "*number of cigarettes smoked per day*" asked in all years is 20 or more, then the respondent is assumed to be a 'heavy smoker'. If the answer is between one and 20, then the respondent is a 'light smoker'.

Respondents are also asked when they quit smoking. If they smoke in the year the questions are asked, the 'daily smoker' variables for the earlier (missing) years are assumed to be one. If they respond that they never quit and they do not smoke in the year the questions are asked, the individuals are presumed to be nonsmokers in previous years. Respondents who reported quitting within the previous year are recorded as smokers in the previous year and respondents are assigned as daily smokers from 1979 to the year they quit.

Not all the respondents answer the quitting smoking question. For the missing years, earlier and later years' information, the average assumption, is imputed. For example, the smoking information for the 1985, 1986 and 1987 years are presumed to be the same in 1984, and smoking information for 1988-1991 is the same as 1992's smoking information.[10]

Levine et al. state that smokers often attempt to quit unsuccessfully. For example, using the 1991 NHIS survey, Levine estimates that over 40% of current smokers attempted to quit

[10] If some of the respondents smoke heavily on some days and do not smoke on the other days, my definition of daily smoking may not match that implied by the 1992, 1994 and 1998 survey years and it may lead to some false transitions in smoking status between the survey years. However, using data from the 1991 National Health Interview Survey, Levine estimates that frequency of heavy smoking on some days and not smoking on other days is only 4% for individuals between the ages of 26 and 33, the same age as NLSY respondents in that year.

smoking in the previous year. Therefore, to observe the effects of smoking in longer periods and to minimize the measurement error of imputations, some additional smoking variables are created by considering the smoking behavior of the respondents for the years in which smoking questions are asked; smoker (in all four years), non-smoker (in all four years), quitter, starter, young-experimenter (only smoked during the year 1984) and unsuccessful quitter (quit before but started again).

Table 3 shows that 35% of the sample is daily smokers and 13% of the respondents smoked in all four years in which the smoking questions are asked. Figure 4 shows that the percentages of daily smokers, heavy smokers, and light smokers are decreasing over time. Percentages are high in first couple of years due to young experimenters. After 1990, although the number of respondents who answer the smoking questions in each survey year varies little, there are slight increases or decreases in the percentages because most of the current smokers attempt to quit smoking in these years. However, as Levine asserts, their attempts are often unsuccessful.

Table 3. *Person-year observations*

Daily smoker	39,699 (32%)
Light smoker	18,756 (15%)
Heavy smoker	20,666 (17%)
Ever smoked daily	58,823 (47%)
Smoker	15,709 (13%)
Non-smoker	53,848 (43%)
Quitter	4,890 (4%)
Starter	4,301 (3%)
Young Experimenter	7,260 (6%)
Unsuccessful Quitter	10,154 (8%)
	125,264

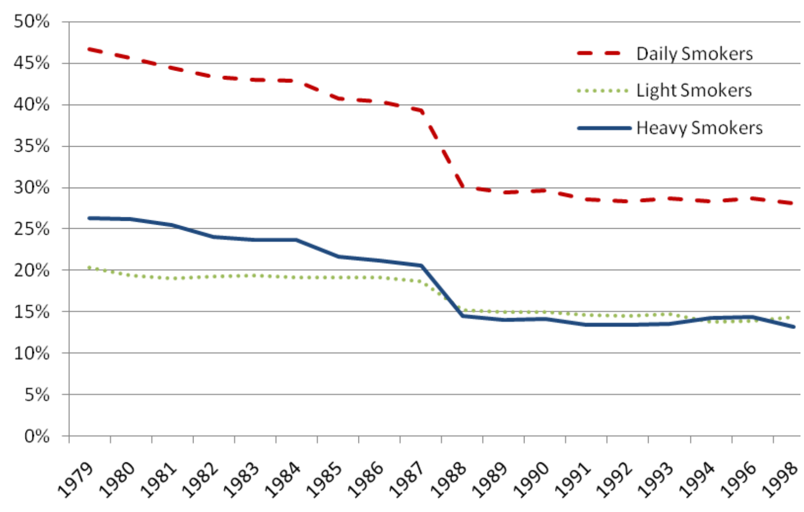

Figure 4: Percentage of Smokers by year

5.3. Average Wages and Prevalence of being Employed by Risk Factor Characteristics:

Table 4 shows the average wages and prevalence of employed/unemployed individuals of my sample by single characteristics and demographic groups. According to Table 4, blacks and Hispanics are more likely to be unemployed than whites. Only 67% of females are employed which is around 15% less than employed males. Further, the average wage of black workers is $2.26 less than white workers and $0.90 less than Hispanic workers. Females receive an average of $12.53, which is $3.10 less than for males.

Table 4 also demonstrates the wage penalties for individuals having health risk behaviors. For example, daily smokers in all years receive $14.40, which is $4.35 less than nonsmokers. Binge drinkers receive an average of $15.63, which is lower than the average of nondrinkers and the average of light to moderate drinkers. My sample consists of individuals from the early age of 18 to the oldest age of 46, and the obesity rates in early survey years are low as the real wage rates. Hence, the descriptive analyses are done for the survey years after 1990 in order not to conclude wrongly that obesity has a positive wage effect. Table 4 shows that the average wage of obese workers is $16.19 which is almost $2.00 less than nonobese workers. Moreover, the table also reveals that daily smokers, binge drinkers, and obese individuals are less likely to be employed than non-smokers, light to moderate drinkers, and the non-obese.

The descriptive statistics for the combined risk factor groups support the importance of controlling for one health behavior when exploring the wage effect of another behavior since attaching to another health risk behavior increases the wage penalty or the prevalence of being unemployed. For instance, the average wage for obese workers is $16.77 when they are also nonsmokers, but the average wage decreases to $13.96 for obese workers as they become

Table 4. *Average Real Wages by Single Characteristics and Demographic Groups*

	Average Wages (After 1990)			Prevalence of Employment Variables		
	%	Wage	S.D.	*%*	Employed, %	Unempl., %
Full sample	*100*	17.65	23.331	*100*	74.16	9.22
By Single Characteristics						
Whites	*52.96*	19.47	25.524	*38.77*	90.73	3.97
Blacks	*28.41*	14.92	20.755	*27.19*	67.04	13.85
Hispanics	*18.63*	16.80	20.020	*17.53*	71.54	8.78
Males	*52.43*	19.59	25.393	*49.32*	81.99	9.92
Females	*47.57*	15.53	20.628	*50.68*	66.54	8.55
Smokers-all years	*13.54*	14.44	19.943	*12.84*	69.21	11.55
Nonsmokers-all years	*50.54*	19.33	25.277	*40.15*	79.26	6.67
Daily smokers (Smokers)	*26.01*	14.60	19.170	*33.22*	68.28	12.64
Non-daily smokers (Nonsmokers)	*70.07*	18.75	24.388	*61.38*	77.31	7.52
Obese	*22.69*	16.19	20.724	*13.21*	72.65	8.14
Nonobese	*75.32*	18.17	24.158	*84.16*	74.47	9.38
Drinker	*61.65*	18.81	24.595	*63.25*	78.68	9.19
Nondrinker	*36.14*	15.76	20.876	*36.75*	66.37	9.27
Binge drinker (B.Drinker)	*9.74*	15.63	18.772	*12.18*	76.80	12.22
By Combined Combined Characteristics						
Nonobese & Nonsmoker	*51.75*	19.49	25.307	*50.63*	77.86	7.57
Obese & Nonsmoker	*16.96*	16.77	21.438	*9.28*	77.73	7.40
Nonobese & Smoker	*20.54*	14.78	20.167	*29.04*	68.42	12.89
Obese & Smoker	*5.06*	13.96	15.337	*3.39*	66.92	10.36
Nondrinker & Nonsmoker	*27.28*	16.14	19.985	*25.03*	69.54	8.34
Nondrinker & Smoker	*7.60*	14.16	22.156	*9.41*	57.94	12.19
B.Drinker & Nonsmoker	*4.994*	17.13	19.851	*5.34*	81.18	9.85
B.Drinker & Smoker	*4.368*	13.95	17.514	*6.21*	72.87	14.46
Nondrinker & Nonobese	*25.22*	16.07	20.961	*29.48*	66.26	9.51
Nondrinker & Obese	*9.97*	15.01	20.024	*6.05*	66.59	8.38
B.Drinker & Nonobese	*7.48*	15.81	18.898	*10.43*	76.93	12.40
B.Drinker & Obese	*2.16*	15.15	18.725	*1.46*	76.75	10.05
Nonobese & Nonsmoker & Nondrinker	*18.53*	16.61	20.136	*19.67*	69.74	8.55
Nonobese & Nonsmoker & B.Drinker	*3.65*	17.54	18.639	*4.45*	81.45	10.02
Nonobese & Smoker & Nondrinker	*5.75*	14.29	23.247	*8.01*	57.62	12.42
Nonobese & Smoker & B.Drinker	*3.55*	14.11	19.101	*5.46*	73.03	14.59
Obese & Nonsmoker & Nondrinker	*7.98*	15.05	18.472	*4.60*	68.66	7.67
Obese & Nonsmoker & B.Drinker	*1.30*	16.23	22.918	*0.79*	79.80	8.71
Obese & Smoker & Nondrinker	*1.73*	13.83	19.054	*1.18*	58.76	11.22
Obese & Smoker & H.B.Drinker	*0.78*	13.29	8.282	*0.59*	71.96	11.90

The sample for wage analyses contains 125,264 person-year observations and the sample for unemployment analyses contains 119,033 person-year observations. Wages are in year-2003 dollars.

smokers. Further, the prevalence of binge drinkers being unemployed is around 12.22%. The percentage of unemployment for binge drinkers rises to about 14.46% when they are also smokers, but the percentage decreases to 9.85% for binge drinkers as they become nonsmokers.

Although the correlations between the health risk behaviors are low in my sample (highest is 0.25 between binge drinking and daily smoking, and the lowest is -0.09 between being obese and binge drinking) due to the panel data properties of the data, descriptive statistics still assert that these health risk behaviors tend to cluster. For example, drinkers are less likely than nondrinkers to be obese (Z test=8.55). Additionally, obese people are less likely to be drinkers than non-obese people (Z test=7.78). Furthermore, smokers are less likely than nonsmokers to be obese, and the obese persons are less likely to than the nonobese persons to be smokers (Z test=10.02). Moreover, smokers are more likely to be drinkers, and drinkers are more likely to be smokers (Z test=15.45). To control for all the possible factors that may be correlated with health risk behaviors and wages/unemployment, multivariate regressions are employed. A set of standard covariates (Table A1) which are used in most wage/unemployment regressions in the literature, and several supplemental covariates (Table A2) are included into the multivariate regression analyses.[11]

6. Results:

6.1. Results when All Three Behaviors are Considered in the Same Analysis:

The main aim of this study is to demonstrate that obesity, smoking, and binge drinking behaviors are correlated or tend to cluster, so their effects on wages or on unemployment are not measured accurately in analyses that consider only one or two. Cigarette smoking, excessive

[11] See appendix for the full set of standard and supplementary variables used in regression analyses.

alcohol use and poor eating habits tend to reinforce each other (Betts, 2000), and it is known that smokers are less likely to be obese. Although previously the combined effects of obesity and smoking, and the combined effects of smoking and drinking have been analyzed in the economics literature, no author has considered the effects of all three risk factors on labor market outcomes at the same time and in the same analysis.

With the intention of observing the changes in the effects of these behaviors on wages or on the likelihood of being unemployed, four different regressions are run by using the OLS for wages and the probit regression model for unemployment analyses. In the first three regression analyses, each risk factor is included on the right-hand side of the wage or the unemployment model in addition to standard and most commonly used explanatory variables, but without the other two risk factors. All three health risk factors are included and controlled for each other in the last model. Table 5 shows the effects of these three health risk behaviors on log wages and Table 6 shows the probit estimates of the risk behaviors on the probability of being unemployed for males and females separately.

Table 5: *The effects of obesity, daily smoking and binge drinking on log wages when different specifications are used*

	Males				Females			
OLS results:	Model 1	Model 2	Model 3	Model 4	Model 1	Model 2	Model 3	Model 4
Obese	-0.023 (0.017)			-0.038 ** (0.016)	-0.119 *** (0.018)			-0.128 *** (0.018)
Daily smoker		-0.079 *** (0.012)		-0.085 *** (0.012)		-0.026 ** (0.012)		-0.032 *** (0.011)
Binge Drinker			0.009 (0.012)	0.019 * (0.011)			-0.032 * (0.018)	-0.022 (0.019)
Adj R^2	*0.266*	*0.267*	*0.266*	*0.273*	*0.258*	*0.257*	*0.257*	*0.262*

Regression models contain 65,365 person-year observations for males and 59,899 person-year observations for females. Clustered robust standard errors are in parentheses. Sampling weights are controlled. *** Significant at the 0.01 level. ** Significant at the 0.05 level. * Significant at the 0.10 level. See Table A1 for other explanatory variables.

Table 5 illustrates that the estimates noticeably change when these health risk factors are controlled for each other. Specifically, the results reveal that the wage effect of obesity is underestimated if one fails to control for the smoking and drinking behaviors of males. The coefficient of obesity decreases from -0.023 to -0.038 and becomes statistically significantly different from zero at the 0.05% level (P value=0.032). Moreover, the wage effects of binge drinking is underestimated for males (estimate is positive) and overestimated for females if alcohol use of the respondents is considered individually. The positive estimate of binge drinking increases and becomes statistically significant at the 0.10 level (P value=0.024) for males, and the binge drinking wage penalty decreases and becomes insignificant for females (P value=0.029).[12]

Table 6: *The effects of obesity, daily smoking and binge drinking on the likelihood of being unemployed when different specifications are used*

	Males				Females			
Probit results:	Model 1	Model 2	Model 3	Model 4	Model 1	Model 2	Model 3	Model 4
Obese	0.037 (0.044) [0.004]			0.078 * (0.041) [0.009]	0.052 (0.047) [0.006]			0.073 * (0.042) [0.009]
Daily smoker		0.212 *** (0.027) [0.024]		0.217 *** (0.027) [0.025]		0.185 *** (0.028) [0.023]		0.187 *** (0.028) [0.024]
Binge Drinker			0.056 ** (0.027) [0.007]	0.024 (0.026) [0.002]			0.126 *** (0.046) [0.018]	0.085 * (0.048) [0.011]

Regression models contain 64,712 person-year observations for males and 54,321 person-year observations for females. People who reported their employment status as out of labor force are excluded from the analyses. Clustered robust standard errors are in parentheses. Marginal effects are in brackets. Sampling weights are controlled. *** Significant at the 0.01 level. ** Significant at the 0.05 level. * Significant at the 0.10 level. See Table A1 for other explanatory variables.

[12] To support my ideas, the first three models are compared to the fourth model by using F-tests. All F-test values are greater than 20, thus, the data gives evidence of a statistically significant departure from the reduced model to the full model, and all three health risk variables should be included in the analysis. Moreover, the argument is also supported since adjusted R^2 increases each time when an additional risk factor is controlled for, for both males and females.

Similar to the results of wage analyses, Table 6 illustrates that failing to include one or more of health behaviors in unemployment regressions would lead to an underestimation of the impacts of being obese and an overestimation of the effect of heavy binge drinking for both males and females. Specifically, the estimates of the obesity coefficients for males and females increase and become statistically significantly different from zero at the 10% level when smoking and binge drinking behavior of individuals are controlled for (P values are 0.019 and 0.027 for males and females, respectively). Further, the estimates of binge drinking and their significance levels decrease when obesity and smoking behaviors are included in the same regression for both genders (P values are 0.013 and 0.017 for males and females, respectively).[13]

6.2. Endogeneity of the Health Risk Behaviors:

The OLS estimates in wage analyses and the probit estimates in unemployment analyses do not correctly characterize the effects of obesity, smoking or binge drinking since health risk behaviors may be endogenous: wages may simultaneously affect risk behaviors of the individuals, or missing or unobservable determinants of both risk behaviors and wages/unemployment may be correlated. In order to limit the potential for heterogeneity bias or the issue of omitted variables and to control more widely for individual background characteristics that might affect labor market outcomes, firstly the supplemental background variables are included into the employment analyses.[14]

[13] To support my ideas, above models are compared by using likelihood ratio tests which test the null hypothesis that removing some variables (namely obesity, smoking or binge drinking) has no effect; it does not lead to a poorer fitting model. The chi square statistics for all comparisons are above or very close to 10 and the corresponding P values are less than 0.05 which are statistically significant. This means that the health risk factors that were removed to produce the reduced models resulted in a model that has a significantly poorer fit, and therefore the variables should be included in the models.

[14] See Table A3 in Appendix for the results. Furthermore, to correct the possible specification error that might be created by employing numerous supplemental covariates, the Stata command 'linktest' is used which detects if the model is not properly specified. Further, to correct the possible multicollinearity problem, another Stata command 'vif' is employed. There seems to be no multicollinearity with health risk behaviors even when early marijuana-cocaine usage is included in the wage model. To reduce the multicollinearity, age and years of tenure are centered by subtracting the means from their predictor values before generating the square terms.

To account for endogeneity, the Hausman-Taylor Instrumental Variable (HTIV) method is employed in wage analyses. The method appears the most appropriate for the purpose and sample in this study. In the presence of correlation between some covariates and the unobserved individual characteristics, this method produces consistent estimates contrary to the OLS model. Furthermore, in contrast to the conventional IV methods, there is no need for finding external instruments. Additionally, in contrast to the fixed effects model, the HTIV model allows producing the estimates for time-invariant covariates.[15] Table 7 shows the results of the HTIV method when three health risk behaviors are considered individually and simultaneously for males and females separately.

Table 7: *The estimates of HTIV models of the effects of obesity, daily smoking and binge drinking on log wages when different specifications are used*

	Males				Females			
HTIV results:	Model 1	Model 2	Model 3	Model 4	Model 1	Model 2	Model 3	Model 4
Obese	-0.012 (0.015)			-0.017 (0.012)	-0.031 ** (0.014)			-0.043 *** (0.013)
Daily smoker		-0.045 *** (0.009)		-0.048 *** (0.009)		-0.024 ** (0.011)		-0.025 ** (0.011)
Binge Drinker			-0.010 (0.007)	-0.006 (0.007)			-0.013 (0.012)	-0.007 (0.013)

Regression models contain 65,365 person-year observations for males and 59,899 person-year observations for females. Clustered robust standard errors are in parentheses. Sampling weights are controlled. *** Significant at the 0.01 level. ** Significant at the 0.05 level. * Significant at the 0.10 level. See Table A2 for other explanatory variables.

Table 7 illustrates that the wage effects of obesity and binge drinking are associated with endogeneity biases. In the HTIV models, the estimates of obesity for males and the estimates of binge drinking for both genders become statistically insignificant. Obesity has negative effects only on the wages of females (a penalty of 4.3%), smoking wage penalties are 4.8% and 2.5% for males and females, respectively, and binge drinking has no effect on wages for either gender.

[15] See Appendix for the results of conventional IV methods, panel-data techniques and comparison of the OLS and HTIV methods.

Moreover, the results reveal that, once endogeneity is controlled for, the estimated parameters of obesity or binge drinking are not statistically significantly different whether these three risk behaviors are considered individually or simultaneously. Therefore, the biases in the OLS estimates are due to OLS's failure to account for endogeneity.

The probit estimates of unemployment analyses are also likely to be biased since the health risk behaviors may be endogenous. I addressed the potential endogeneity by using the multivariate probit method. This method appears to be the most appropriate for this study because conventional IV methods do not exploit fully the nonlinearity of unemployment model.[16] Table 8 shows the results of the multivariate probit method when three health risk behaviors are considered individually and simultaneously for males and females separately.

Table 8: *The estimates of the multivariate probit models of the effects of obesity, daily smoking and binge drinking on unemployment when different specifications are used*

	Males				Females			
	Model 1	Model 2	Model 3	Model 4	Model 1	Model 2	Model 3	Model 4
Obese	0.086 (0.092) [0.008]			0.155 (0.096) [0.016]	0.093 (0.091) [0.009]			0.191 * (0.099) [0.018]
Daily smoker		0.252 *** (0.051) [0.024]		0.271 *** (0.049) [0.025]		0.118 *** (0.043) [0.015]		0.207 *** (0.041) [0.018]
Binge Drinker			0.056 * (0.033) [0.007]	0.039 (0.032) [0.002]			0.146 * (0.075) [0.012]	0.102 (0.070) [0.008]

Regression models contain 64,712 person-year observations for males and 54,321 person-year observations for females. People who reported their employment status as out of labor force are excluded from the analyses. Clustered robust standard errors are in parentheses. Marginal effects are in brackets. Sampling weights are controlled. *** Significant at the 0.01 level. ** Significant at the 0.05 level. * Significant at the 0.10 level. See Table A2 for other explanatory variables.

Table 8 shows that the effects of the health risk behaviors on unemployment are associated with endogeneity bias associated with probit models. Being obese increases the

[16] See Appendix for the results of conventional IV methods, panel-data techniques and comparison of probit and multivariate probit models.

probability of being unemployed by 1.8% for females only, smoking increases the likelihood of being unemployed by 2.5% and 1.8% for males and females, respectively, and binge drinking appears to have no effect for either gender.[17] Similar to wage analyses, when endogeneity is accounted for, the estimated coefficients of obesity or binge drinking are the same whether these three behaviors are considered individually or simultaneously for both genders. Therefore, the biases in the probit estimates are due to failure of probit to account for endogeneity.

6.3. Sensitivity Analyses:

6.3.1. Interactive/Additive Characteristics of Health Risk Behaviors:

Previously, it has been assumed that the impact of one health risk behavior is the same regardless of another health risk behavior. The risk factors could be additive, for instance if health worsens when a binge drinker smokes or a smoker becomes obese; hence the wages may be decreased proportionally by poorer health. However, the health risk behaviors would not be additive, for example if they are due to the same unobserved factors (Baum et al., 2006). In order to investigate whether the effects of these risk behaviors are additive or interactive, four interaction terms are added into the wage and the unemployment regressions.[18] Table 9 shows the effects of the health risk factors and their interactions on wages and on unemployment accounting for endogeneity by using the HTIV and multivariate probit models.

The table shows that none of the interaction terms are statistically significantly different from zero; hence the effects of risk behaviors are not interactive. But, when interaction terms are

[17] The correlations between the error terms in employment and obesity, smoking, and binge drinking equation (ρ_{12}, ρ_{13} and ρ_{14}) are statistically significantly different from zero, hence risk behaviors are endogenous, and the multivariate probit model should be used.

[18] Although no formal presence of multicollinearity is found with the health risk factors and their interactions, to reduce the unobserved potential multicollinearity, the health risk factor variables are centered by subtracting the means from their predictor values before generating the interaction terms.

included into the models, binge drinking estimates for males become statistically significant both in wage and unemployment analyses, which shows the significance of including interaction terms. If an individual is neither obese nor a smoker, being a binge drinker results in a 1.7% wage penalty and increases the probability of being unemployed by 0.5%. Hence, the effect of binge drinking could be correlated with the obesity and/or smoking behavior of the individual, for instance, if an obese or smoker already perceives discrimination, also becoming a binge drinker may not result in any additional discrimination.

Table 9: *The estimates of HTIV and Multivariate Probit models of the effects of obesity, smoking and binge drinking and their interactions on log wages and on unemployment, for males and females separately*

	HTIV		Multivariate Probit	
	Males	Females	Males	Females
Obese	-0.010 (0.014)	-0.051 *** (0.015)	0.158 (0.097) [0.016]	0.195 * (0.100) [0.018]
Daily smoker	-0.051 *** (0.010)	-0.031 *** (0.011)	0.266 *** (0.050) [0.025]	0.199 *** (0.041) [0.018]
Binge Drinker	-0.017 * (0.010)	-0.034 (0.021)	0.051 * (0.030) [0.005]	0.114 (0.074) [0.010]
Obese * Smoker	0.039 (0.025)	0.005 (0.057)	-0.029 (0.097) [-0.002]	-0.075 (0.075) [-0.006]
Obese * Binge Drinker	0.014 (0.015)	0.037 (0.027)	0.059 (0.114) [0.005]	-0.016 (0.108) [-0.001]
Smoker * Binge Drinker	-0.001 (0.022)	0.023 (0.023)	-0.028 (0.053) [-0.002]	0.042 (0.085) [0.003]
Obese * Smoker * Binge Drinker	0.001 (0.041)	0.070 (0.077)	0.063 (0.011) [-0.005]	0.088 (0.092) [0.008]

Regression models of wage analyses for males contain 65,365 person-year observations, for females contain 59,899 person-year observations. Employment regression models contain 64,712 male and 54,321 female person-year observations. People who reported their employment status as out of labor force are excluded from employment analyses. Marginal effects are in brackets.

6.3.2. Wage Comparisons:

The impacts of three health risk factors on wages might vary across the wage structure with higher wage workers less likely to be affected, possibly because of the nature of their jobs, than lower wage workers or workers with less education, or vice verse. Table 10 displays the differences in the effects of three health risk factors on log wages at different points of the conditional wage distribution using simultaneous-quantile regression for males and females separately.

Table 10. *The estimates of Simultaneous Quantiles Regression model of the impact of obesity, smoking and binge drinking on log wages, for different quantiles separately, for males and females*

Males	10th Quantile	25th Quantile	50th Quantile	75th Quantile	90th Quantile
Obese	-0.004 (0.016)	-0.026 ** (0.011)	-0.020 *** (0.007)	-0.008 (0.009)	-0.015 (0.011)
Daily smoker	-0.055 *** (0.010)	-0.052 *** (0.005)	-0.050 *** (0.004)	-0.050 *** (0.004)	-0.052 *** (0.011)
Binge Drinker	0.006 (0.013)	0.009 (0.007)	0.009 (0.006)	0.004 (0.007)	0.001 (0.008)
Females	**10th Quantile**	**25th Quantile**	**50th Quantile**	**75th Quantile**	**90th Quantile**
Obese	-0.091 *** (0.018)	-0.076 *** (0.010)	-0.071 *** (0.008)	-0.059 *** (0.006)	-0.055 *** (0.012)
Daily smoker	-0.010 (0.010)	-0.031 *** (0.006)	-0.027 *** (0.005)	-0.025 *** (0.006)	-0.024 *** (0.005)
Binge Drinker	-0.003 (0.031)	0.013 (0.017)	-0.001 (0.006)	-0.020 ** (0.009)	-0.010 (0.011)

Regression models contain 65,365 person-year observations for males and 59,899 person-year observations for females. Standard errors with bootstrapping with 500 replications are in parentheses.

The table reveals that the wage penalties for daily smoking are fairly constant over the wage distribution for both genders. However, obesity affects the wages of males and females relatively more at lower quantiles, and there is no wage penalty for being a binge drinker for males and females at any quantile.[19]

6.3.3. More Detailed Measures of Health Risk Behaviors:

I also estimated wage and unemployment models using more detailed measures. The obesity dummy variable is divided into two dummies: 'morbid obesity' is defined as having a BMI of 35 or more and 'mild obesity' is defined as having a BMI of between 30 and 35. 'Smoker' is defined as being a daily smoker in all years. A 'young experimenter' is an individual who smokes only in the 1984 survey year, when smoking questions are asked for the first time, and then quits smoking in later survey years. A 'heavy smoker' is defined as smoking more than 20 cigarettes per day and a 'light smoker' is someone who smokes at least one but less than 20 cigarettes per day. The 'Current drinker' dummy variable is equal to one if the respondent drinks any alcoholic beverage in the previous month. 'Heavy drinking' is defined as drinking 6 or more drinks in one to three occasions in the previous month. The results for both genders are displayed in Table 11.

The results demonstrate that smokers are a heterogeneous group of people: the wage and unemployment effects of persistent smokers are different than the effect of starters, and quitters and young experimenters seem to have no effect on wages or unemployment. Furthermore, even though mild or morbid obesity leads to wage penalties and a higher likelihood of unemployment for females, only morbid obesity appears to affect the wages and the likelihood of being unemployed of males. Moreover, contrary to the previous literature, drinking is found to have no

[19] A Hausman test fails to reject the hypothesis that the full set of daily smoking coefficients are equal at the five quantiles for males and females. Another Hausman test that tests the null hypothesis that all obesity coefficients are equal at the five quantiles rejects the null hypothesis for both genders. The insignificant wage effects of obesity for males and daily smoking for females at the 10th quantile could possibly be due to the unstable working characteristics of the individuals that they might be working in different jobs as part-time or it could be because they may be working in jobs that offer low wages but less or no discrimination against obese or smoker in hiring or firing conditions.

positive effect on wages or any effect on unemployment for either gender when endogeneity is accounted for.

Table 11: *The estimates of HTIV and Multivariate Probit models of the effects of more detailed measures of health risk behaviors on log wages and unemployment, for males and females separately*

	HTIV		Multivariate Probit	
	Males	Females	Males	Females
Smoker	-0.059 *** (0.008)	-0.031 *** (0.009)	0.289 *** (0.045) [0.027]	0.212 *** (0.039) [0.019]
Quitter	0.031 (0.030)	-0.007 (0.031)	0.032 (0.045) [0.003]	0.015 (0.036) [0.001]
Starter	-0.063 * (0.032)	-0.024 (0.030)	0.106 ** (0.046) [0.009]	0.085 ** (0.035) [0.007]
Young Experimenter	-0.019 (0.026)	-0.022 (0.023)	0.075 (0.055) [0.006]	0.009 (0.047) [0.001]
Unsuccessful Quitter	-0.051 ** (0.021)	-0.014 (0.021)	0.178 ** (0.071) [0.015]	0.089 (0.068) [0.007]
Mild Obese	0.007 (0.013)	-0.029 ** (0.014)	0.077 (0.096) [0.007]	0.126 * (0.078) [0.011]
Morbid Obese	-0.029 ** (0.015)	-0.075 *** (0.020)	0.196 * (0.102) [0.018]	0.261 *** (0.109) [0.024]
Light Smoker	-0.061 *** (0.010)	-0.018 * (0.011)	0.168 *** (0.038) [0.016]	0.164 *** (0.044) [0.015]
Heavy Smoker	-0.031 *** (0.011)	-0.043 *** (0.014)	0.302 *** (0.057) [0.028]	0.233 *** (0.041) [0.021]
Current Drinker	0.007 (0.006)	0.013 (0.009)	-0.073 (0.053) [-0.006]	-0.085 (0.055) [-0.07]
Heavy Drinker	0.005 (0.006)	0.004 (0.007)	-0.016 (0.026) [-0.001]	-0.009 (0.037) [-0.001]

Regression models of wage analyses for males contain 65,365 person-year observations, for females contain 59,899 person-year observations. Employment regression models contain 64,712 male and 54,321 female person-year observations. People who reported their employment status as out of labor force are excluded from employment analyses.

6.3.4. Ethnic/Racial Differences:

To conduct a richer examination of ethnic/racial differences of health risk behavior wage penalties, I separated whites, blacks and Hispanics into subsamples and examined the wage effects separately. The results are displayed in Table 12.

Table 12: *The estimates of Hausman-Taylor IV models of the impact of obesity, smoking and heavy binge drinking on log wages, for whites , blacks and Hispanics separately*

	Males			Females		
	Whites	Blacks	Hispanics	Whites	Blacks	Hispanics
Obese	0.005 (0.016)	0.004 (0.023)	-0.001 (0.027)	-0.062 *** (0.018)	-0.024 (0.025)	-0.017 (0.031)
Daily smoker	-0.041 *** (0.012)	-0.070 *** (0.018)	-0.037 * (0.020)	-0.031 ** (0.013)	-0.006 (0.022)	-0.033 (0.027)
Binge Drinker	-0.018 ** (0.008)	0.002 (0.016)	0.025 (0.015)	-0.018 (0.015)	0.022 (0.032)	-0.008 (0.035)

Regression models contain 65,365 person-year observations for males and 59,899 person-year observations for females.

The table reveals that smoking has wage penalties for all subsamples of males but affects the wages of only white females. Further, there appears to be an obesity wage penalty only for white females, and a binge drinking wage penalty only for white males. However, the results must be assessed with caution since the person-year observations of blacks and Hispanics are relatively smaller than whites in my sample.

6.3.5. Public-Private Sector Differences:

The effects of obesity, smoking, and binge drinking on wages are also examined for private and public sector workers separately. In public sector, wages are generally determined by a priority, thus the sector is a relatively compact and fixed wage sector. On the contrary, wages of the private sector workers are determined by the employer, so workers are more exposed to wage discrimination in a private sector. Table 13 displays the estimates of the three health risk behaviors on wages of males and females working in a private or public sector, separately.

Table 7: *The estimates HTIV models of the effects of obesity, smoking and binge drinking on log wages, for public and private sector workers separately*

	Males		Females	
	Private	Public	Private	Public
Obese	0.021 (0.015)	0.016 (0.025)	-0.036 ** (0.017)	-0.006 (0.026)
Daily smoker	-0.045 *** (0.012)	-0.039 * (0.021)	-0.036 *** (0.013)	-0.031 (0.026)
Binge Drinker	-0.008 (0.008)	0.037 (0.023)	-0.020 (0.015)	0.005 (0.040)

Regression models contain 65,365 person-year observations for males and 59,899 person-year observations for females.

The table shows evidence of wage penalties for being obese or a smoker for females and a smoking wage penalty for males in private sector jobs. However, in the public sector, only males face lower wages only due to smoking, which proves the argument that private sector jobs are more prone to wage penalties.

7. Conclusions and Discussion:

This paper aims to find the joint effects of obesity, smoking and binge drinking on labor market outcomes, using the National Longitudinal Survey of Youth data. The main objective of this study is to show that the effects of obesity, smoking and binge drinking on wages or on unemployment are not measured accurately in analyses that consider only one or two health behaviors since these behaviors are correlated or tend to cluster. The OLS results of wage analyses reveal that the wage effect of obesity is underestimated for males, and the wage effects of binge drinking is underestimated for males (estimate is positive) and overestimated for females if one fails to control for other health risk behaviors in the analyses. Similarly, the probit results of unemployment analyses illustrate that failing to include one or more of the health behaviors in unemployment regressions would lead to an underestimation of the impact of being obese and an overestimation of the effect of binge drinking for both genders.

However, the OLS and probit estimates do not correctly characterize the effects of obesity, smoking, or binge drinking since these behaviors may be endogenous. I address the potential endogeneity by employing the Hausman-Taylor instrumental variable (HTIV) method in wage analyses and the multivariate probit method in unemployment analyses. The results illustrate that the wage and unemployment effects of risk behaviors are associated with endogeneity biases. Further, when endogeneity is controlled for, the estimated parameters of obesity or binge drinking are not statistically significantly different whether these behaviors are considered individually or simultaneously. Moreover, the results show that obesity has negative effects only on the wages of females (a penalty of 4.3%), and smoking wage penalties are 4.8% and 2.5% for males and females, respectively. Being obese increases the probability of being unemployed by 1.8% for females only, and smoking increases the likelihood of being unemployed by 2.5% and 1.8% for males and females, respectively. However, binge drinking appears to have no effect on wages or unemployment for either gender.

I have also conducted several sensitivity and robustness analyses on the labor market effects of obesity, smoking, and binge drinking. Firstly, this study examines whether the effects of the three health risk behaviors on wages or unemployment are interactive. None of the interaction terms are statistically significantly different from zero; hence the effects of risk behaviors are not interactive. However, when interaction terms are included into the models, binge drinking estimates for males become statistically significant both in wage and unemployment analyses which shows the significance of including interaction terms in regressions. Hence, binge drinking affects the wage of a male or probability of being unemployed if he is neither obese nor a smoker.

I examined whether the impacts of these risk factors on wages might vary across the wage structure using quantile regressions. The results reveal that the wage penalties for daily smoking are fairly constant over the wage distribution for both genders. However, obesity affects the wages of males and females relatively more at lower quantiles, and there is no wage penalty for being a binge drinker for males and females at any quantile.

I also estimated wage and unemployment models using more detailed measures. My results when endogeneity is addressed demonstrate that smokers are a heterogeneous group of people: the wage and unemployment effects of persistent smokers are different than starters, quitters or young experimenters. Furthermore, obesity appears to affect the wages and the likelihood of being unemployed of males only at the extremes. Moreover, contrary to the previous literature, drinking is found to have no positive effect on wages or on unemployment for either gender when endogeneity is accounted for.

I separated whites, blacks, and Hispanics into subsamples to conduct a richer examination of ethnic/racial differences. I found that, although smoking has wage penalties for all subsamples, obesity affects the wages of only white females, and binge drinking affects wages of only white males.

Lastly, health risk behaviors may affect the wages of individuals working in private sector more than those working in the public sector. I find evidence of wage penalties for being obese or a smoker in private sector jobs. However, in the public sector, only males face lower wages only due to smoking.

My results draw attention to methodological issues involved in estimating relationships between health risk behaviors and labor market outcomes, especially the need to account for unobserved heterogeneity or endogeneity. Failure to consider the endogeneity or unobserved factors that affect wages/unemployment and risk behaviors could lead to incorrect inferences. However, the study has some limitations. For instance, currently I am extending the analyses by using semi-parametric estimators to avoid making strong functional form and/or distributional assumptions. Furthermore, relying on self-reported alcohol or cigarette consumption may lead to biases since respondents may give inaccurate or biased information. Also, defining obesity as having BMI of 30 or more could result in a bias since BMI is a poor predictor of body fat. I plan to re-define these health risk behaviors and re-examine the labor market effects using these new measures.

The findings have direct implications and policy proposals for the labor market behavior of obese individuals, smokers and excessive alcohol users. The current study can also be expanded to give more insight for policy makers. For instance, I plan to combine the results of wage and unemployment effects of these three health risk behaviors to examine their total effects on life-time earnings of the individuals. Also, I want to focus on the impacts of health behaviors on spells of unemployment using duration regression specifications. In one line of work, I plan to analyze whether changes in these behaviors over time may lead to different results. For example, it can be examined to see if people who lose/gain weight, or who currently smoke, or who currently drink experience different outcomes from those who have or don't have those behaviors consistently throughout the time period. In another line of work, I plan to expand this work by considering additional labor market outcomes such as educational attainment or total number of jobs of the individuals.

References:

Auld MC. Smoking, drinking, and income. *Journal of Human Resources* 2000; 40:505-518.

Averett S, Korenman S. The economic reality of the beauty myth. *Journal of Human Resources* 1996; 31: 304-330.

Baum CL, Ford WF. The wage effects of obesity: a longitudinal study. *Health Economics* 2004; 13: 885-889.

Baum CL, Ford WF, Hopper JD. The Obese Smoker's Wage Penalty. *Social Science Quarterly* 2006; 87:863–881.

Becker G. *The Economics of Discrimination*, vol I, second ed. 1971. Chicago: The University of Chicago Press.

Becker GS, Grossman M, Murphy KM. An Empirical Analysis of Cigarette Addiction. *American Economic Review* June 1994; 84: 396-418.

Behrman J, Rosenzweig M. The Returns to Increasing Body Weight. Penn Institute for Economic Research Working Paper No.01-052. 2001.

Berger MC, Leigh JP. The effect of alcohol use on wages. *Applied Economics* 1988; 20: 1343-1351.

Betts N. Smoking, Heavy Drinking and Poor Nutrition Tend To Cluster. *The American Journal of Health Promotion* 2000.

Bollen K, Guilkey D, Mroz TA. Binary Outcomes and Endogenous Explanatory Variables: Tests and Solutions with an Application to the Demand for Contraceptive Use in Tunisia. *Demography* 1995; 32-1: 111-131.

Bove C, Olson C. Factors Contributing to Obesity in Rural, Low-Income New York Women. Working paper from the conference Rural Women's Health, 2005.

Bryant R, Samaranayake V, Wilhite A. Alcohol use and the wages of young men: whites vs. nonwhites. *International Review of Applied Economics* 1992; 6: 184–202.

Bryant R, Jayawardhana A, Samaranayake V, Wilhite A. The Impact of Alcohol and Drug Use on Employment: A Labor Market Study Using the National Longitudinal Survey of Youth. Institute for Research on Poverty. Discussion Paper no. 1092-96. 1996.

Burkhauser R, Cawley J. Obesity, Disability, and Movement Onto the Disability Insurance Roll. Working Paper 2004-089 University of Michigan, Retirement Research Center.

Cawley J. Body weight and women's labor market outcomes. Working paper no. 7841, National Bureau of Economic Research, Cambridge, MA. 2000a.

Cawley, John. The Impact of Obesity on Wages. *Journal of Human Resources* 2004; 39-2: 451-74.

Centers for Disease Control and Prevention (CDC). 2007. Behavioral Risk Factor Surveillance System (BRFSS). Washington, DC: Centers for Disease Control and Prevention, Department of Health and Human Services. Available at http://www.cdc.gov/brfss/index.htm

Centers for Disease Control and Prevention (CDC). 2007. Obesity Among Adults in the United States--No Statistically Significant Change Since 2003-2004. Washington, DC: Centers for Disease Control and Prevention, Department of Health and Human Services. Available at http://www.cdc.gov/nchs/data/databriefs/db01.pdf

Feng W, Zhou W, Butler JS, Booth BM, French MT. The impact of problem drinking on employment. *Health Econ* 2001; 10: 509–521.

Forcier MW. Labor force behavior of alcoholics: A review. *The International Journal of Addictions* 1985; 20: 253-268.

French M, Roebuck C, Alexandre P. Illicit drug use, employment, and labor force participation. *Southern Econ J* 2001; 68: 349–368.

Fuchs, VR. *Who Shall Live? Health, Economics, and Social Choice*. New York: Basic Books. Gottschalk 1974; 168.

Garcia J, Quintana-Domeque C. Obesity, Employment and Wages in Europe, In K. Bolin and J. Cawley (Eds.), Advances in Health and Health Services Research, Vol. 17: The Economics of obesity. New York: Elsevier. 2006.

Grafova IB, Stafford FP. The Wage Effects of Personal Smoking History. *Industrial and Labor Relations Review* 2009; 62: 381-393.

Greeve J. Obesity and Labour Market Outcomes: New Danish Evidence. wp 07-13, 2007, Department of Economics, Aarhus School of Business, University of Aarhus. Statistical Discrimination of Disabled Workers? Working paper SOLE 2005 Available: http://gsbwww.uchicago.edu/labor/E.08.4.%20Greve%20&%20Tranaes%20&%20Jensen.pdf

Griliches Z, Hausman J. Errors in variables in panel data. *Journal of Econometrics* 1986; 31: 93-118.

Hamilton V, Hamilton BH. Alcohol and earnings: does drinking yield a wage premium? *Canadian Journal of Economics* 1997; 97: 135-151.

Harper B. Beauty, stature and the labour market: a British cohort study. *Oxford Bulletin of Economics and Statistics* 2000; 62, 771–801.

Harris R, Zhou J, Youngblood BD, Rybkin I, Smagin G, Ryan D. Effect of repeated stress on body weight and body composition of rats fed low- and high-fat diets. *Am J Physiol Regul Integr Comp Physiol* 1998; 275: R1928-R1938.

Harwood H. 2000. *Updating Estimates of the Economic Costs of Alcohol Abuse in the United States: Estimates Update Methods and Data.* Report prepared by the Lewin Group for the National Institute on Alcohol Abuse and Alcoholism.

Hausman JA, Taylor WE. Panel data and unobservable individual effects. *Econometrica* 1981; 49: 1377–1398.

Heien DM. Do drinkers earn less? *Southern Economic Journal* 1996; 63(1): 60–8.

Jusot F, Khlat M, Rochereau T, Serme C. Job loss from poor health, smoking and obesity: a national prospective survey in France. *J Epidemiol Community Health* 2008; 62; 332-337.

Keng S, Huffman WE. Binge drinking and labor market success: a longitudinal study on young people. *J Popul Econ* 2007; 20: 35-54.

Kenkel DS, Ribar D. Alcohol consumption and young adults' socioeconomic status. *Brookings Papers on Economic Activity-Micro* 1994; 119-161.

Leigh JP, Berger MC. Effects of smoking and being overweight on current earnings. *American Journal of Preventive Medicine* 1989; 5: 8-14.

Levine PB, Gustafson TA, Velenchik AD. More bad news for smokers? The effects of cigarette smoking on wages. *Industrial and Labor Relations Review* 1997; 50: 493-509.

Loh ES. The economic effects of physical appearance. *Soc Sci Q* 1993; 74: 420-438.

MacDonald Z, Shields MA. The impact of alcohol consumption on occupational attainment in England. *Economica* 2001; 68:427-53.

Maranto CL, Stenoien AF. Weight discrimination: a multidisciplinary analysis. *Emp Respons Rights J.* 2000; 12:9–24.

Montgomery JD. Social networks and labor-market outcomes: Toward an economic analysis. *American Economic Review* 1991; 81:1408-18.

Moon M, McLean R. Health, obesity and earnings. *American Journal of Public Health* 1980; 70, 1006–1009.

Morris S. The impact of obesity on employment. *Labour Economics*. 2007; 14: 413-33.

Mullahy J, Sindelar JL. Alcoholism, work, and income. *Journal of Labor Economics*. 1993; 11: 494-520.

Mullahy J, Sindelar JL. Employment, unemployment, and problem drinking. *Journal of Health Economics*. 1996; 15: 409-34.

National Longitudinal Survey of Youth Data, 2002 [Internet]. [Cited 2002]. Available from: http://www.bls.gov/nls/79guide/2002/nls79g0.pdf

Norton EC, Lindrooth RC, Ennett ST. Controlling for the Endogeneity of Peer Substance Use on Adolescent Alcohol and Tobacco Use. *Health Economics* 1998; 7: 439-53.

Pagan JA, Davila A. Obesity, occupational attainment, and earnings. *Soc Sci Q* 1997; 78: 756-770.

Peters BL. Is there a wage bonus from drinking? Unobserved heterogeneity examined. *Applied Economics* 2004; 36: 2299-2315.

Puhl R, Brownell KD. Bias, Discrimination and Obesity. *Obesity Research* 2001; 12: 788-805.

Register CA, Williams DR. Wage effects of obesity among young workers. *Soc Sci Q* 1990; 71: 130-141.

Renna F. Alcohol Abuse, Alcoholism and Labor Market Outcomes: Looking for the Missing Link. *Industrial and Labor Relations Review* 2008; 62: 92-103.

Sargent, J. D., Blanchflower, D. G. Obesity and stature in adolescence and earnings in young adulthood. *Archives of Pediatric Adolescent Medicine* 1994; 148: 681–687.

Sarlio-Lahteenkorva S, Lahelma E. The Association of Body Mass Index with Social and Economic Disadvantage in Women and Men. *International Journal of Epidemiology* 1999; 28: 445-449.

Sesso HD. Alcohol and cardiovascular health: recent findings. *American Journal of Cardiovascular Drugs* 2001; 1: 167–172.

Sturm R. The Effects of Obesity, Smoking, And Drinking on Medical Problems. *Health Affairs* 2002; 245-253.

Tekin E. Employment, wages, and alcohol consumption in Russia. *Southern Economic Journal* 2004; 71: 397-417.

Terza J. Alcohol abuse and employment: A second look. *Journal of Applied Econometrics* 2002; 393-404.

Viscusi WK, Hersch J. Cigarette Smokers as Job Risk Takers. *Review of Economics and Statistics* 2001; 83: 269-80.

WHO. 2002. The world health report 2002 - Reducing Risks, Promoting Healthy Life.

Wooldridge, J.M., 2002. Econometric Analysis of Cross-Section and Panel Data. MIT Press, Cambridge, MA.

Yamaguchi K, Kandel DB. On the resolution of role incompatibility: Life event history analysis of family roles and marijuana use. *American Journal of Sociology* 1985; 90: 1284-1325.

Zarkin GA, French MT, Mroz T, Bray JW. Alcohol use and wages: new results from the National Household Survey on Drug Abuse. *Journal of Health Economics* 1998; 17: 53-68.

Appendix:

Table A1. *Descriptive statistics of standard variables*

	Variables	Mean	Std. Dev.
Dependent Variables			
Wage	Wage, *in 2003 $*	14.79	50.081
	Wage, *Averett, in 2003 $*	13.67	11.278
	Wage, *Cawley, in 2003 $*	14.15	18.444
	Log wage, *Cawley, in 2003 $*	2.37	0.737
Employment	Employed, *%*	74.16	
	Unemployed, *%*	9.22	
	Out of labor force, *%*	16.62	
	Part-time employed, *%*	7.14	
	Full-time employed, *%*	67.02	
	Self-employed, *%*	5.07	
	Employed-not-self, *%*	69.09	
Explanatory Variables			
Obesity Variables	Obese, *%*	15.27	
	Mild obese, *%*	10.51	
	Morbid obese, *%*	4.76	
	Body mass index, *kg/m²*	25.34	4.965
	Overweight, *%*	33.23	
	Normal, *%*	45.71	
	Underweight, *%*	5.80	
Drinking Variables	Current drinker, *%*	66.47	
	Binge drinker, *%*	12.01	
	Heavy drinker, *%*	31.17	
Smoking Variables	Smoker, *%*	12.54	
	Nonsmoker, *%*	42.99	
	Quitter, *%*	3.90	
	Starter, *%*	3.43	
	Young experimenter, *%*	5.80	
	Unsuccessful quitter, *%*	8.11	
	Daily smoker, *%*	33.35	
	Light smoker, *%*	15.77	
	Heavy smoker, *%*	17.37	
Personal Characteristics			
Age	Age, *years*	28.66	6.648
Race-Ethnicity	Black, *%*	25.86	
	Hispanic, *%*	17.28	
	Non Black non Hispanic, %	56.86	
Gender	Male, *%*	52.18	
Marital Status	Married, *%*	44.82	
	Separated-divorced-widowed, *%*	12.80	
	Never married, *%*	42.38	
Children	Children in household	0.86	1.111
Family Size	Family size	3.17	1.764
Health Status	Health limitation for work, *%*	6.62	
Acquired Human Capital			
Education Variables	Education Level, *years*	12.71	2.291
	Less than high school, *%*	3.28	
	Some high school, *%*	12.14	
	High school, *%*	48.13	
	More than high school, %	36.42	

	Variables	Mean	Std. Dev.
	Some college, *%*	20.22	
	College or more, *%*	16.20	
Experience	Experience, *years*	7.41	4.948
Tenure	Tenure, *years*	3.39	4.084
Nonwage Income	Nonwage income, *in 2003 $*	17,308.98	46,257.853
Environmental Conditions			
Residence	South, *%*	39.49	
	Northeast, *%*	17.61	
	West, *%*	19.47	
	North central, %	23.43	
Urban-Rural	Urban, *%*	78.64	
	Rural, %	21.36	
Labor Market Conditions			
Unemployment Rate	Unemployment rate, *0-6%*	41.42	
	Unemployment rate, *6-9%*	36.16	
	Unemployment rate, *9%*	22.42	

Wage and log wage are hourly measures. Body mass index (BMI) is weight/(height^2). Separated-divorced-widowed dummy equals one if the respondent is separated or divorced or widowed. Children dummy is the number of children in household. Health limitation for work equals to one if the current health of respondent limits the kind or amount of work respondent can do. Education level is the highest grade in school respondent completed. South, northeast, west and north central dummies equal to one if the respondent's current residence is in south, northeast, west and north central, respectively. Urban and south are equal to one if the respondent's current residence is urban or south. Experience is cumulative hours worked since respondent completed schooling. Tenure is total years of tenure with employer. Nonwage income is total net family income minus total income from wages and salary.

Table A2. *Descriptive statistics of supplemental background variables*

	Variables	Mean	Std. Dev.
Supplemental Background Variables			
Personal Characteristics			
Native Born	Native born, *%*	93.21	
Foreign Language	Foreign language, *%*	1.57	
AFQT Score	AFQT score	37.60	27.562
Rotter External Control Score	Rotter score	8.75	2.409
Rosenberg Self-Esteem Score	Rosenberg score	33.17	3.977
Attitudes Towards Family Roles Score	Attitudes towards family roles score	16.52	3.316
Shyness	Shyness, *%*	29.09	
Anyone Present During Survey	Anyone present, *%*	20.47	
Substance Use			
Marijuana or Cocaine Use	Marijuana, 0, *%*	70.78	
	Marijuana, 1, *%*	16.07	
	Marijuana, 2-3, *%*	10.47	
	Marijuana, 4-5, *%*	2.69	
Childhood			
Foreign Language in Childhood	Foreign language in childhood, *%*	22.41	
Lived with Both Parents	Lived with parents *(at age 14)*, *%*	67.44	
Received Magazine	Magazine *(at age 14)*, *%*	53.71	

	Variables	Mean	Std. Dev.
Received Newspaper	Newspaper *(at age 14)*, *%*	74.12	
Received Library Card	Library card *(at age 14)*, *%*	68.60	
Lived in South	South *(at age 14)*, *%*	36.35	
Lived in Urban	Urban *(at age 14)*, *%*	78.56	
Illegal Activity in 1980	Illegal activity, 1980, *%*	35.74	
Family			
Age of Youngest Child	No children, *%*	52.37	
	Youngest child, 0-1, *%*	8.67	
	Youngest child, 1-5, *%*	22.89	
	Youngest child, 5-, *%*	16.07	
Number of Siblings	No sibling, *%*	2.86	
	1 sibling, *%*	12.03	
	2-3 siblings, *%*	37.76	
	4 or more siblings, *%*	47.35	
Older Sibling	Older sibling, %	76.93	
Family in Poverty	Family poverty status, %	12.14	
Family Information			
Mother's Work Information	Mother worked part-time *(at age 14)*, *%*	16.41	
	Mother worked full-time *(at age 14)*, *%*	40.88	
Father's Work Information	Father worked part-time *(at age 14)*, *%*	3.70	
	Father worked full-time *(at age 14)*, *%*	71.11	
Mother's Education	Mother, less than high school, *%*	19.48	
	Mother, some high school, *%*	23.04	
	Mother, high school, *%*	37.04	
	Mother, more than high school, *%*	13.70	
	Mother's education missing, *%*	6.74	
Father's Education	Father, less than high school, *%*	22.74	
	Father, some high school, *%*	15.66	
	Father, high school, *%*	28.91	
	Father, more than high school, *%*	17.59	
	Father's education missing, *%*	15.10	
Environmental Conditions			
SMSA Variables	Not in SMSA, *%*	23.12	
	SMSA, in central city, *%*	17.89	
	SMSA, not in central city, *%*	30.61	
	SMSA, central city unknown, *%*	28.37	
Work Information			
Experience	Experience, *years*	5.95	4.362
Tenure	Tenure, *years*	2.46	3.270
Nonwage Income	Nonwage income, *in 2003 $*	17,308.98	46,257.853

Native born is equal to one if respondent is born in the US. AFQT is the armed forces qualifications test given in 1980 (between 1-100). Rotter score (between 4 and 16): higher score means higher external control, lower score means higher internal control. Rosenberg self-esteem score (between 10 and 40): lower score denotes lower self-esteem. Attitudes towards family roles score (between 8 and 32): lower score indicates non-traditional views. Anyone present is if anyone was present during the survey. Marijuana questions are asked in 5 times, Marijuana, 0 is never used, Marijuana, 1 is answered positively only once, Marijuana, 2-3 means twice or three times, Marijuana, 4-5 is 4 or 5 times. Magazine and newspaper indicate if someone in respondent's household receives at least one magazine or newspaper subscription. Family poverty status is 1 if family is in poverty. White collar is 1 if respondent's occupation is professional, technical, manager, official, proprietor, sales worker or clerical. Experience is cumulative hours worked since respondent completed schooling. Tenure is total years of tenure with employer. Nonwage income is total net family income minus total income from wages and salary.

Table A3: *The OLS estimates of the effects of obesity, smoking and binge drinking on log wages, with standard covariates only, and with standard and supplemental covariates*

	Males		Females	
OLS results:	Standard	Standard + Supplemental	Standard	Standard + Supplemental
Obese	-0.038 **	-0.022	-0.128 ***	-0.096 ***
	(0.016)	(0.014)	(0.018)	(0.016)
Daily smoker	-0.085 ***	-0.074 ***	-0.032 ***	-0.023 **
	(0.012)	(0.011)	(0.011)	(0.011)
Binge Drinker	0.019 *	0.017 *	-0.022	-0.007
	(0.011)	(0.010)	(0.020)	(0.016)

Regression models contain 65,365 person-year observations for males and 59,899 person-year observations for females. Clustered robust standard errors are in parentheses. Sampling weights are controlled. *** Significant at the 0.01 level. ** Significant at the 0.05 level. * Significant at the 0.10 level. See Table A1 and Table A2 for standard and supplementary covariates.

Table A4: *Probit estimates of the effects of obesity, smoking and binge drinking on unemployment, with standard covariates only, and with standard and supplemental covariates*

	Males		Females	
Probit results:	Standard	Standard + Supplemental	Standard	Standard + Supplemental
Obese	0.078 *	0.118 ***	0.073 *	0.026
	(0.041)	(0.039)	(0.042)	(0.038)
	[0.009]	[0.010]	[0.009]	[0.002]
Daily smoker	0.217 ***	0.115 ***	0.187 ***	0.102 ***
	(0.027)	(0.026)	(0.028)	(0.027)
	[0.025]	[0.009]	[0.024]	[0.010]
Binge Drinker	0.024	0.013	0.085 *	0.009
	(0.026)	(0.026)	(0.048)	(0.042)
	[0.002]	[0.001]	[0.011]	[0.001]

Regression models contain 64,712 person-year observations for males and 54,321 person-year observations for females. People who reported their employment status as out of labor force are excluded from the analyses. Clustered robust standard errors are in parentheses. Marginal effects are in brackets. Sampling weights are controlled. *** Significant at the 0.01 level. ** Significant at the 0.05 level. * Significant at the 0.10 level. See Table A1 and Table A2 for standard and supplementary covariates.

Table A5: *The estimates of OLS and HTIV models of the effects of obesity, smoking and binge drinking on log wages, for males and females separately*

	Males		Females	
	OLS	HTIV	OLS	HTIV
Obese	-0.022	-0.017	-0.096 ***	-0.043 ***
	(0.014)	(0.012)	(0.016)	(0.013)
Daily smoker	-0.074 ***	-0.048 ***	-0.023 **	-0.025 **
	(0.011)	(0.009)	(0.011)	(0.011)
Binge Drinker	0.017 *	-0.006	-0.007	-0.007
	(0.010)	(0.007)	(0.016)	(0.013)

Regression models contain 65,365 person-year observations for males and 59,899 person-year observations for females. Clustered robust standard errors are in parentheses. Sampling weights are controlled. *** Significant at the 0.01 level. ** Significant at the 0.05 level. * Significant at the 0.10 level. See Table A2 for other explanatory variables.

Table A6: *Probit and Multivariate Probit results of being unemployed for obesity, smoking and binge drinking, for males and females separately*

	Males		Females	
	Probit	Multivariate Probit	Probit	Multivariate Probit
Obese	0.118 ***	0.155	0.026	0.191 *
	(0.039)	(0.096)	(0.038)	(0.099)
	[0.010]	[0.016]	[0.002]	[0.018]
Daily smoker	0.115 ***	0.271 ***	0.102 ***	0.207 ***
	(0.026)	(0.049)	(0.027)	(0.041)
	[0.009]	[0.025]	[0.010]	[0.017]
Binge Drinker	0.013	0.039	0.009	0.102
	(0.026)	(0.032)	(0.042)	(0.070)
	[0.001]	[0.002]	[0.001]	[0.008]

Regression models contain 64,712 person-year observations for males and 54,321 person-year observations for females. People who reported their employment status as out of labor force are excluded from the analyses. Clustered robust standard errors are in parentheses. Marginal effects are in brackets. Sampling weights are controlled. *** Significant at the 0.01 level. ** Significant at the 0.05 level. * Significant at the 0.10 level. See Table A2 for other explanatory variables.

Table A7. *The estimates of IV and IV for fixed effects models of the effects of obesity, smoking and binge drinking on log wages, for males and females separately*

	Males			Females		
	OLS	IV	IV-Fixed Effects	OLS	IV	IV-Fixed Effects
Obese	-0.022	-0.045 **	-0.010	-0.096 ***	-0.123 **	-0.026 *
	(0.014)	(0.019)	(0.012)	(0.016)	(0.071)	(0.013)
Daily smoker	-0.074 ***	-0.046 *	-0.037 ***	-0.023 **	-0.035 **	-0.013
	(0.011)	(0.024)	(0.010)	(0.011)	(0.015)	(0.010)
Binge Drinker	0.017 *	0.022	0.005	-0.007	0.011	-0.005
	(0.010)	(0.020)	(0.008)	(0.016)	(0.019)	(0.013)

Regression models for males contain 65,365 person-year observations. Regression models for females contain 59,899 person-year observations. Standard errors are in parentheses. Sampling weights are controlled. *** Significant at the 0.01 level. ** Significant at the 0.05 level. * Significant at the 0.10 level. See Table A2 for the list of explanatory variables.

Table A8. *The estimates of IV and IV for fixed effects models of the effects of obesity, smoking and binge drinking on unemployment, for males and females separately*

	Males			Females		
	OLS	IV	IV-Fixed Effects	OLS	IV	IV-Fixed Effects
Obese	0.118 ***	0.129	0.066	0.026	0.138 *	0.076
	(0.039)	(0.097)	(0.057)	(0.038)	(0.072)	(0.058)
	[0.010]	[0.011]	[0.005]	[0.002]	[0.011]	[0.006]
Daily smoker	0.115 ***	0.211 **	0.178 *	0.102 ***	0.198	0.155 *
	(0.026)	(0.101)	(0.095)	(0.027)	(0.117)	(0.089)
	[0.009]	[0.016]	[0.014]	[0.010]	[0.016]	[0.11]
Binge Drinker	0.013	0.022	0.017	0.009	0.077	0.042
	(0.026)	(0.033)	(0.029)	(0.042)	(0.068)	(0.048)
	[0.001]	[0.002]	[0.001]	[0.001]	[0.006]	[0.003]

Regression models contain 64,712 person-year observations for males and 54,321 person-year observations for females. People who reported their employment status as out of labor force are excluded from the analyses. Clustered robust standard errors are in parentheses. Marginal effects are in brackets. Sampling weights are controlled. *** Significant at the 0.01 level. ** Significant at the 0.05 level. * Significant at the 0.10 level. See Table A2 for other explanatory variables.

Table A9. *Fixed effects, between effects and random effects estimates of the impact of obesity, smoking and drinking measures on log wages, for "males" only*

Males:	OLS	Fixed Effects	Between Effects	Random Effects
Obese	-0.022	-0.013	-0.004	0.010
	(0.014)	(0.015)	(0.019)	(0.012)
Daily smoker	-0.074 ***	-0.037 ***	-0.058 ***	-0.050 ***
	(0.011)	(0.012)	(0.013)	(0.009)
Binge Drinker	0.017 *	0.006	0.042 **	-0.001
	(0.010)	(0.009)	(0.018)	(0.008)

Regression models for males contain 65,365 person-year observations. Standard errors are in parentheses for between effects. Clustered robust standard errors are in parentheses for OLS, fixed effects and random effects. Sampling weights are controlled in OLS. *** Significant at the 0.01 level. ** Significant at the 0.05 level. * Significant at the 0.10 level. See Table A2 for the list of explanatory variables.

Table A10. *Fixed effects, between effects and random effects estimates of the impact of obesity, smoking and drinking measures on log wages, for "females" only*

Females:	OLS	Fixed Effects	Between Effects	Random Effects
Obese	-0.096 ***	-0.031 *	-0.076 ***	-0.057 ***
	(0.016)	(0.017)	(0.020)	(0.013)
Daily smoker	-0.023 **	-0.013	-0.005	-0.017 *
	(0.011)	(0.013)	(0.013)	(0.009)
Binge Drinker	-0.007	-0.005	-0.036	-0.009
	(0.016)	(0.015)	(0.032)	(0.013)

Regression models for males contain 59,899 person-year observations. Standard errors are in parentheses for between effects. Clustered robust standard errors are in parentheses for OLS, fixed effects and random effects. Sampling weights are controlled in OLS. *** Significant at the 0.01 level. ** Significant at the 0.05 level. * Significant at the 0.10 level. See Table A2 for the list of explanatory variables.

Table A11. *Fixed effects and random effects estimates of the impact of obesity, smoking and drinking measures on unemployment*

	Male			Female		
Logit results:	Logit	Fixed	Random	Logit	Fixed	Random
Obese	1.246 ***	1.052	1.135 **	1.045	0.877	1.018
	(0.094)	(0.098)	(0.067)	(0.073)	(0.084)	(0.060)
Daily smoker	1.238 ***	1.061	1.242 ***	1.199 ***	1.163 **	1.206 ***
	(0.058)	(0.068)	(0.046)	(0.059)	(0.090)	(0.049)
Binge Drinker	1.031	1.095 **	1.073*	1.021	1.138	1.030
	(0.049)	(0.052)	(0.041)	(0.078)	(0.099)	(0.070)

Regression models contain 64,712 male and 54,321 female person-year observations. People who reported their employment status as out of labor force are excluded from the analyses. Clustered robust standard errors are in parentheses. Sampling weights are controlled. *** Significant at the 0.01 level. **Significant at the 0.05 level. *Significant at the 0.10 level. See Table A2 for the list of all explanatory covariates.