THE EFFECT OF OBESITY ON LABOR MARKET OUTCOMES

Euna Han

A dissertation submitted to the faculty of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Health Policy and Administration.

Chapel Hill
2006

Approved by:

Edward C. Norton, Ph.D. (Advisor)

Barry M. Popkin, Ph.D.

Kristin L. Reiter, Ph.D.

Sally C. Stearns, Ph.D.

Wilbert van der Klaauw, Ph.D.

ABSTRACT

Euna Han: Effect of Obesity on Labor Market Outcomes
(Under the direction of Edward C. Norton)

This dissertation investigates the effect of obesity on labor market outcomes. Obesity is important for labor market outcomes. Obese people may be discriminated against by consumers or employers due to their distaste for obese people. Employers also may not want to hire obese people due to the expected health cost if the employers provide health insurance to their employees. Because of those consumers' and employers' distaste for obese people or because of these different costs, being obese may result in poor labor market outcomes in terms of wages and/or the likelihood of being employed, as well as sorting of obese people into jobs where slimness is not rewarded. This study used the National Longitudinal Survey of Youth 1979 (NLSY79). The NLSY79 provides panel information for a nationally representative sample of 12,686 young men and women who were 14 to 22 years old when first surveyed in 1979. The sample was followed for 14 years. Labor market outcomes were measured by 1) the probability of employment, and 2) the probability of holding occupations where slimness potentially rewards hourly wages. Weight was measured by Body Mass Index (BMI). All results were assessed separately by gender as a function of BMI splines and other controls. The endogeneity of BMI was controlled in a two-stage instrumental variable estimation model with over-identifying exogenous individual and state-level instruments, controlling for individual fixed effects. The Heckman selection model was used to control for the selection into the labor force, with the state-level identifying instruments of the non-employment rate, the number of business establishments, and the number of Social Security Program beneficiaries. Results show that gaining weight adversely affects labor market outcomes for women, but the effect is mixed for men overall. The size and direction of the effects vary by gender, age groups, and type of occupations. Findings from this investigation could help our understanding of the economic cost of obesity to an individual beside its adverse effect on health. The spillover effect of obesity will increase the total cost of obesity to both individuals and society as a whole.

CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I: INTRODUCTION

Overview

The objective of this study is to understand the effect of obesity on labor market outcomes. The prevalence rate of obesity has increased by over 50% since the late 1970s. Chou, Grossman and Saffer (2002) showed a sharp upward trend in obesity between 1978 and 1991 using nationally representative data. During this thirteen-year period, the number of obese Americans grew by 55%. The previous literature has consistently reported health problems and a high health-related social cost caused by obesity (Sturm, 2002).

Obesity is important for labor market outcomes. Obese people may be discriminated against by consumers or employers due to their distaste for obese people in a job where slimness does not matter. Employers also may not want to hire obese people due to the expected health cost if the employers provide health insurance to their employees (Hamermesh and Biddle, 1994). These employer-side distastes may result in poor labor market outcomes in terms of wage earnings and the low likelihood of being employed, as well as sorting of obese people into jobs where slimness is not rewarded.

Accurate estimation of the effect of obesity on labor market outcomes supports the understanding of the economic cost of obesity to an individual beyond its adverse effect on health. The spillover effect of obesity to non-health sector, i.e. labor market outcomes, increases the total cost of obesity to individuals. The consequences of individuals' choice relevant to their body weight also to be borne by others who are not directly involved the choice. An example of such externality cost is an increase in the insurance premium due to large health services use by obese people (Bhattacharya and Sood, 2005). Individuals do not always make rational choices concerning body weight and the information they can use are not perfect, which provide a rationale for the public intervention (Cawley, 2004). Among several potential policy measures to help individuals' rational decision regarding the body weight control, disseminating information about the adverse effect of obesity will provide an incentive for an individual to change their behavioral choices relevant to healthy body weight.

This study addresses the following specific research questions: Ceteris paribus, does an increase in BMI: (1) decrease the likelihood of being employed; (2) decrease the likelihood of sorting into occupations where social interaction with customers or colleagues is required; (3) decrease wage earnings; (4) affect wage earnings differently at various stages of the life cycle; and (5) affect wage earnings differently in occupations where social interaction with customers or colleagues is required versus other occupations. These research questions are explored separately by gender.

This study uses the National Longitudinal Survey of Youth (NLSY79), which is suitable to address the relationship between obesity and labor market outcomes with sophisticated statistical techniques. The NLSY79 has detailed information on the labor market outcomes, height, and weight in a panel structure.

Labor market outcomes are only observed for the participants in the labor force. The Heckman selection model is used to control for selection into the labor force with the following identifying instruments at the state level: the non-employment rate, the number of business establishments, and the number of Social Security Program beneficiaries.

Obesity, which is the key explanatory variable of interest in the study, is endogenous. Individual fixed-effects control for the unobservable individual heterogeneity (like endowment). The endogeneity of obesity is controlled for in two-stage instrument variable estimation models with over-identifying instruments. The identifying instruments for obesity are three state-level variables (fast food prices, beer prices, and sales in restaurants), and two individual-level variables (siblings' BMI and a five-year lag of the respondents' BMI).

Trends and causes on obesity

This section contains two sub-sections. First, the growing epidemic of obesity is discussed to emphasize the importance of obesity for public policy. In the second subsection, the endogenous characteristics of obesity are discussed. Individuals make their behavioral choices that may affect their body weight, in particular, diet and exercise. The potential correlation of obesity with labor market outcomes is also discussed in the second sub-section because the correlation also makes obesity an endogenous explanatory variable.

The obesity epidemic

The dramatic growth in obesity (or the state of being overweight) has been an important concern for policymakers and the public over the last several decades. Previous literature has consistently reported health problems and a high social cost caused by obesity. For example, Stevens et al. (1998) estimated that excess body weight increases the risk of death for individuals between 30 and 74 years due to coronary artery disease, stroke, high blood pressure, cancers of colon, breast and prostate, and diabetes. This estimation implies that obesity is the second primary attribute to premature death, second only to smoking. In a study by Sturm (2002), social costs of obesity are reported to exceed those of cigarette smoking and alcoholism. Therefore, obesity also affects major public transfer programs such as Medicaid, Medicare, and Social Security (Lakdawalla and Philipson, 2002).

Although weight has been rising in the U.S. throughout the twentieth century, the rise in obesity since the 1980s is fundamentally different from past changes. That is, since the 1980s, weight has grown more than physicians recommend for healthy weight (Cutler, Glaeser, and Shapiro, 2003).

Figure 1.1 depicts annual trends in average Body Mass Index (BMI) (left vertical axis) and in the percentage (right vertical axis) who are obese for persons 18 years and older in the Behavioral Risk Factor Surveillance System during the period between 1984 and 1999. BMI is a measure of height-adjusted weight equal to weight in kilograms divided by squared height in meters. Persons with a BMI equal to or greater than 30 are classified as obese. A BMI between 25 and 30 is classified as overweight, and a BMI below 18.5 is underweight (National Heart, Lung, and Blood Institute, 1998). Annual trends show that the average BMI increased from 1984 to 1998 by 9%, and the number of obese adults more than doubled during the same period (Chou, Grossman, and Saffer, 2002).

More than half of Americans were overweight or obese in 1999, and the increase in the proportion of being overweight and obese affects all ages, racial and ethnic groups, and both genders (U.S. Department of Health and Human Services, 2001). The prevalence of obesity is higher than that of smoking, use of illegal drugs, or other risk factors for most of the highly prevalent chronic diseases, including heart disease, diabetes, and cancer (Philipson, 2001).

The endogeneity of obesity

Obesity is endogenous because individual choices partly affect the state of being obese besides endowed genetic factors. That is, obesity primarily is a choice variable, not a given variable. In technical terms, obesity as a regressor is not orthogonal to the unobserved characteristics in the error terms in models of labor market outcomes because unobserved individual heterogeneity will affect both labor market outcomes and obesity. Time preference is an example of unobserved individual heterogeneity. Individuals with a high discount rate for future will be less likely to invest in their own human capital such as education, which would be correlated with their own wage earnings. Those individuals with a high discount rate also are less likely to restrain from risky health behaviors including the consumption of fattening food. The discount rate affects both obesity and labor market outcomes. Thus, obesity is an endogenous explanatory variable in the econometric sense.

Several factors have been discussed as contributors to obesity as a choice variable in the previous literature. First, low income or poverty has been claimed to cause being overweight or obese particularly in women (Stunkard, 1996; Sobal and Stunkard 1989). The different distribution of obesity by income level may complicate policy measures for resolving any discrimination against obese people in the labor market (Averett and Koreman, 1996). It has been reported that fast food and convenience food are inexpensive and are high in calories compared to other healthier foods. If more fattening foods are generally cheaper than healthier, non-fattening food, the people with lower income will be more likely to consume fattening foods (Chou, Grossman, and Saffer, 2002).

The effect of wage income on the extent of obesity is indeterminate. A decline (or modest increases in some years) in monetary income appear to have stimulated the demand for inexpensive but fattening convenience and fast food. The real average hourly earnings in the private sector decreased from 1982 to 1995, and it was only 4.5% higher in 2002 than in 1982 (Chou, Grossman, and Saffer, 2002; U.S. Census Bureau, 2003; Bureau of Labor Statistics, 2005). However, previous studies have found that higher household income did not result in better weight outcomes (Chou, Grossman, and Saffer, 2002; Lakdawalla and Philipson, 2002). That result might imply that an increase in participation into the labor force or an increase in work hours for women contributes to weight gain through a decrease in leisure time. In fact, participation rate in the labor force for women increased 12.5% between 1982 and 2002 (Bureau of Labor Statistics, 2005). Those increasing trends of market work will reduce the time and energy available for home production including food preparation, which can also contribute in part to the increasing prevalence of convenience or fast food. Several studies have supported the effect of reduced leisure time for household production on weight gain. For example, a child is more likely to be overweight if her mother worked more hours per week over the child's life, and that adverse effect of work hours on child's excess weight is larger for those mothers in high socioeconomic status (Ruhm, 2004; Anderson, Butcher, and Levine, 2003).

Second, an increase in the number of fast-food restaurants in town will decrease the time cost for using those services, which will result in cheaper access to those places. According to the Census of Retail Trade, the per capita number of restaurants, lunchrooms, and cafeterias increased by 63.7% between 1987 and 1996 (Bureau of the Census, 1996). Also, full service restaurants and limited-service eating establishments increased 34.5% between 1992 and 2000 (Bureau of the Census, 2000). Previous literature has shown that the number of restaurants per capita had a positive and significant effect on the weighted sample means of the extent of obesity, while prices at fast-food restaurants, full-service restaurants had negative and significant effects on the extent of obesity (Chou, Grossman, and Saffer, 2004).

If the fast food market is competitive, then an increase in the number of fast food restaurants will decrease the average price of fast food. The previous literature has displayed systematic dispersion in the number of restaurants or grocery stores in a market by socioeconomics profiles of the market. For example, Stewart and Davis (2005) found that low population and low levels of income might be associated with limited access to restaurants, and thus, higher prices. If the demand for groceries or number of restaurants increases in a market, firms would supply more by opening new stores in the market, ceteris paribus. Kaufman and colleagues (2005) reported higher grocery prices in urban stores than in suburban markets. Spatial concentration of people with a common socioeconomic profile may complicate the dynamic relationship among obesity, socioeconomic status, and food prices. Individuals tend to choose to live near others like themselves, and thus, those with the best opportunities at economic success will cluster together. For example, a high proportion of low income, racial or ethnic minorities tend to live in urban centers, while people with high income tend to live in suburban areas (Toussaint-Comeau and Sherrie, 2002). This possible neighborhood selection in an area by socioeconomic profile implies people with low income may pay higher food prices due to their residential area characteristics, and therefore, more tend to demand cheap substitutes for expensive foods.

Third, increased calorie consumption also has contributed to the increase in obesity in U.S. over the last few decades. Technological change has dramatically shortened both fixed and variable costs for mass preparation of foods. The technological innovation that allows mass preparation of food decreases the price of food, where the price includes the time cost for preparing food. Furthermore, it reduces the time delay before actual consumption of food, which will particularly affect food consumption by people with self-control problems. For those people, a decrease in the delay of instant gratification from food consumption will make it more difficult to pass up current pleasure for future benefits (Cutler, Glaeser, and Shapiro, 2003).

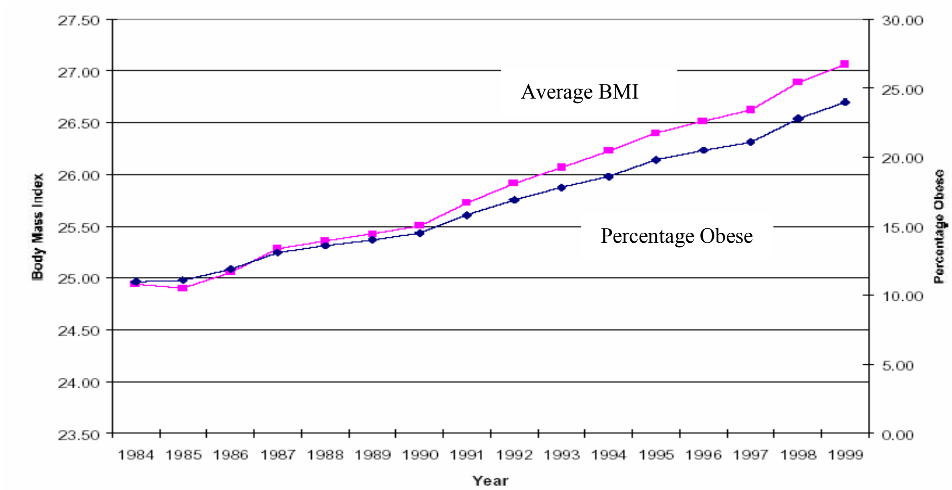
Trends in energy intake have also changed since the 1970s. For example, energy intake from sweetened beverage consumption increased by 135%, while energy intake from milk consumption fell by 38% from 1977 to 2001 for samples aged 2 to 60 years. This corresponds to a 278 total calorie increase during the same period (Nielsen and Popkin, 2004; Popkin, 1996). Other than the source of energy intake, the portions of food have increased as well. Between 1977 and 1996, food portion sizes increased both inside and outside the home for specific food items including salty snacks (from 1.0 to 1.6 oz), soft drinks (13.1 to 19.9 fl oz), french fries (3.1 to 3.6 oz), hamburgers (5.7 to 7.0 oz), and Mexican food (6.3 to 8.0 oz) (Nielsen and Popkin, 2003).

Fourth, other than the increased calorie consumption, a more sedentary lifestyle may generate a substantial growth in obesity. Lakdawalla and Philipson (2002) suggest that the income growth drives the sedentary life style by increasing the cost of physical activity in leisure time, as well as increasing the quantity of food intake. In their estimation, income growth has an inverted U-shape relationship with body weight, while the reduction in strenuousness of work raises weight in a linear fashion. They also separated job-related exercise from job strength when considering the effect of employment on obesity because job strength is predicted to have different effects on obesity, which is generally measured by

BMI. Workers in a job requiring physical strength may have strong muscle mass, and thus, greater BMI.

Fifth, smoking affects obesity, even though the direction and size of the effect remains undetermined (Gruber and Frakes, 2006). Individuals who quit smoking typically gain weight. Therefore, the anti-smoking campaign, which began to accelerate in the early 1970s may be an important trend affecting the increase in obesity. The increase in the real price of cigarettes contributed to the anti-smoking trend, which partly resulted from Federal excise tax hikes and a number of state tax hikes (Chou, Grossman, and Saffer, 2004).

Figure 1.1 Trends in Body Mass Index and percentage obese for persons 18 years of age and older in the Behavioral Risk Factor Surveillance System, 1984-1999



Source: Chou, Grossman, and Saffer (2002)

CHAPTER II: LITERATURE REVIEW

The first section of this chapter reviews the previous literature about the effect of obesity on labor market outcomes. Because education is a strong predictor of labor market outcomes, the second subsection discusses the effect of obesity on school-related outcomes. The previous literature studying the effect of physical appearance on labor market outcomes is reviewed in the third subsection, because obesity is one component of looks. In the last section, the significance of the current study including policy implications and how the current study can improve the previous literature is discussed.

How does obesity affect labor market outcomes?

Several studies have linked obesity to labor market outcomes, mostly wages. Even though all of those studies essentially used the same data, the NLSY79, their results differ markedly. These inconsistent trends in the previous literature may be attributed to the lack of valid control for the endogeneity of obesity (Register and Williams, 1992; Loh, 1993; Pagan and Davila, 1997; Gortmaker et al., 1993; Sargent and Blanchflower, 1994).

Recently, a few studies reported statistically significant negative effects of obesity on wages when they tried to control for the endogeneity of obesity. Averett and Korenman (1996) replaced current body weight with a lagged body weight when estimating the effect of body weight on wages using the 1988 survey of the NLSY79. Additionally, they produced within-sister estimates using a sister fixed-effects model. Both genders were estimated to suffer obesity penalties in the labor market in terms of low earnings, with a lesser extent for women than for men, even after self-esteem was controlled for. However, their way to control for endogeneity of obesity is not likely to produce valid parameter estimates for the effect of obesity on the labor market outcomes. Their estimation was based on the assumption of no serial inter-temporal correlation in the wage residuals, which is not likely. Although the sister fixed effects sweep out the unobserved permanent endowment factors at the family level belonging to the error term, individually heterogeneous endowment factors will remain unobserved in the error term.

Similar to Averett and Korenman's (1996) study, Conley and Glauber (2005) took the lag of 13 and 15 years of BMI as instruments for the current BMI, and used the sibling fixed-effects model to control for the endogeneity of BMI. Using the Panel Study of Income Dynamics (PSID) 1986, 1999, and 2001 data, they estimated the effect of obesity on three labor market outcomes: occupational prestige, labor earnings, and total family income. Their study results were consistent with Averett and Korenman (1996) in that obesity penalizes women in terms of not only their own earnings, family income, and occupational prestige, but also spouse's earnings, and spouse's occupational prestige. Their study is the only study that measured the effect of obesity on occupational prestige. They measured occupational prestige using Duncan's Socioeconomics Index (SEI) for 1970 U.S. census occupational classification codes. However, like Averett and Korenman's (1996) study, they took the lagged BMI as an instrument for obesity with sibling fixed effects. Thus, the limitations of Averett and Korenman's (1996) study also apply to Conley and Glauber's (2005) study.

A study by Behrman and Rosenzweig (2001) did not find a statistically significant effect of BMI on wages in the labor market with a survey on identical female twins from a sample from the Minnesota Twins Registry. In this study, the identical twin fixed-effects model was used, which eliminated the permanent but unobserved genetic endowments including earnings endowment from the error term. The unobserved earnings endowment in the error term would lead to biased estimators for the effect of BMI on wage, because earning endowment might be correlated with education and a genetic component of body weight. However, if the physical characteristics including BMI were also affected by contemporaneous wage shocks in the error term, then within-twin estimates would still be biased. Therefore, the authors used lagged consumption or the lagged physical characteristics as an instrument for current BMI and height after they swept out time-consistent endowment factors with twin fixed effects.

The results of this specification showed near zero effect of BMI on wages, while height has a statistically significant and strong positive effect on wages. These results may imply that the statistically significant inverse association between adult BMI and wages in other studies is due to a correlation between unmeasured earnings endowments and BMI as discussed by the authors. The statistically significant positive effect of height on wages may be explained by the positive correlation between height and weight at birth. As the authors discussed, weight at birth has a positive effect on adult height, while adult weight was mainly explained by genetic factors, but not by weight at birth. Weight at birth has been a proxy for good prenatal care. Although their estimation technique is more likely to produce an unbiased parameter estimate for the effect of BMI on wage, they used only 808 sample persons in their estimation (Behrman and Rosenzweig, 2001). This small sample size may be the main cause of statistically insignificant effect of BMI on earnings equation due to possibly overestimated standard errors (for example, the estimates for the twin fixed-effects on the 2SLS was .00197, and the corresponding absolute value of robust t statistics was 0.19).

Cawley (2000) estimated the effect of obesity on women's employment disability, which was measured by limitations on the amount of paid work and limitations on types of paid work using 12 years of data from the NLSY79. When the sample women's own child's body weight was used as an instrument for the sample women's body weight, obesity did not have a statistically significant effect on a limitation on the amount of paid work nor a limitation on the type of paid work. Assuming that the instrument in this study is valid, the results imply that loss of body weight among obese women might not reduce the employment disability. However, the validity of the instrument remains untested. Children's body weight will not be a valid instrument for mother's body weight if there is unobserved heterogeneity for mothers in the wage residual, which affects both children's body weight and the mother's employment disability. For example, smoking or alcohol consumption during pregnancy may reflect the pregnant women's inconsistent discount rate for the future, which will affect both their performance in their job and their children's health at birth.

In another study done by Cawley (2004), the effect of obesity on wage rate was estimated with 12 years of data from the NLSY79. Siblings' body weight was used as an instrument for sample persons' body weight, and the individual and sibling fixed-effects model were estimated. The results found obesity penalties in the labor market in terms of wages only for white women. However, the author discussed that siblings' body weight does not have a high power to identify the variation in the respondent's body weight. Moreover, the use of siblings' body weight as an instrument restricted the results to the sample who reported to have any sibling.

All of the previous studies focused on the total effect of obesity on wages without identifying indirect pathways linking obesity to different performance in the labor market. A study by Baum and Ford (2004) was one of the two studies trying to estimate the indirect effect of obesity on labor market outcomes.

In a study using 12 years of the NLSY79, they tested four potential pathways linking obesity to labor market outcomes: less productivity due to health problems from obesity, less investment on human capital by obese workers, employers' distaste for obese employees due to high health care cost for obese people, and consumers' distaste for obese workers. Their empirical evidence suggested that those pathways might mediate the effect obesity on labor market outcomes. However, the authors did not control for the endogeneity of obesity other than using the individual and family fixed-effects model. If any explanatory variables are not strictly exogenous, i.e., uncorrelated with current and earlier disturbance terms or shocks, then the fixed-effects estimators are inconsistent. Also, they estimated only the intensive marginal effects of those factors in the labor market outcomes in terms of wage, but did not consider the extensive marginal effect, i.e., the effect on labor market participation.

Bhattacharya and Bundorf (2004) focused on health care cost differentials as an explanation for wage differential between obese and non-obese people. In their estimation, they compared wage differentials between obese and non-obese individuals in employment with employer-provided health insurance to the wage differentials between obese and non-obese individuals in employment without employer-provided health insurance. Their difference-in-difference estimator using the NLSY79 data during 1989-1998 showed statistically different wage differential between obese and non-obese individuals when the difference-in-difference estimator (interaction between obesity and employer-provided health insurance) was included in the model. The authors argue that this indicates obese workers had lower wages compared to non-obese individuals because of their higher medical expenditure and not because of possible discrimination against obese individuals in the labor market. Their study is a welcome exception for pointing out one potential pathway under the negative effect of obesity on labor market outcomes.

However, most importantly, Bhattacharya and Bundorf's (2004) study did not control for the endogeneity of either employer-provided health insurance or obesity in their model. Obese individuals may choose a job with employer-provided health insurance over a job without employer-provided health insurance, recognizing that they have high need for health care services due to obesity. Those obese individuals being aware of their high medical need may have a lower productivity in their jobs than the non-obese individuals due to their health problems. In such a case, obese individuals' relatively lower wages compared to non-obese individuals may not be fully attributed to the incremental medical cost they incur to their employers. Jobs with employer-provided health insurance also may have different characteristics from the jobs without employer-provided health insurance. For example, those jobs with employer-provided health insurance may tend to belong to a larger firms, and tend to provide other non-health insurance fringe benefits, such as maternal leave or pension programs. There may be variation in job characteristics between the jobs with a high proportion of obese individuals and the jobs with a low proportion of obese individuals in a company with employer-provided health insurance. Thus, wage differential between obese and non-obese individuals in such a case would be explained by different job characteristics rather than workers' weight and subsequent medical costs. As stated in chapter 1, obesity is endogenous. Several factors can contribute to both obesity and poor labor market outcomes in terms of low wages. Individual heterogeneity time preference is an example of such factors. Therefore, endogeneity of obesity and employer-provided health insurance should be accounted for in the model. Otherwise, the estimation results would be biased.

How does obesity affect schooling-related outcomes?

One potential pathway for a causal effect of obesity on differentials in the labor market would be differentials in the level of human capital. The previous literature has consistently reported the significant effect of human capital, in particular education or schooling, over the other factors on the outcomes in labor market, although the validity of the estimated causal effect has not been agreed upon (for example, Card, 1994; Angrist and Krueger, 1991). Thus, identifying the effect of obesity on the development of human capital, particularly, schooling-related outcomes, will be highly relevant for policy measures to intervene in any social or economic adverse effect of obesity.

However, there has been little study on the effect of obesity on schooling. Moreover, the few existing studies have reported only simple differentials in the distribution of human capital between obese and non-obese people without controlling for other covariates nor the endogeneity of obesity. For example, Cawley (2004) showed unadjusted differentials in the education level and intelligence test scores using multiple years of the NLSY79. The magnitude of the effect of obesity on those achievements was different by race and gender in his samples. White lighter men had a higher value of human capital measures on average than white heavier men, while it was quite opposite for black men when BMI was used: heavier black men had higher education levels and test scores than lighter black men. For women, samples in the lighter group have more years of education and higher test scores on average than samples in the heavier group regardless of race.

Likewise, Sargent and Blanchflower (1994) found that obese girls at age 16 had poor performances in math and reading tests in later years than non-obese girls in their study using a British birth cohort. However, they did not find this differential in test scores between the obese and non-obese boys of the same age. This gender differential was also reported in the study by Gortmaker and colleagues (1993) using one year (1981) of the NLSY79. They found that the overweight women in their sample had less education than the non-overweight women, although a differential was not found in the sample of men.

How does beauty affect the labor market outcomes?

Considering that slimness is a component making an individual physically attractive, studies of premiums to overall beauty or stature in the labor market would help to identify any potential consumers' or employers' distaste for physically unattractive workers, including obese ones. However, obesity can be a good proxy for beauty in a study for estimating the effect of beauty on labor market outcomes. There is a potential error in the measurement of beauty because standards for beauty are various and rather subjective, while the measurement of obesity is more objective than beauty.

Four studies directly estimated the effect of beauty on labor market outcomes. Hamermesh and Biddle (1994) examined the effect of looks on earnings using interviewers' ratings of respondents' physical appearance with three different surveys in U.S. and Canada (the 1977 Quality of Employment Survey, the 1971 Quality of American Life Survey, and the 1981 Canadian Quality of Life Study). Their results found that the penalty of earnings for plainness is larger than the premium of earnings for beauty, and the effects for men are as large as for women. The effect of looks on earnings was found to be independent of the type of occupation. This result might imply no discrimination against plain-looking workers by consumers because consumer-based discrimination may lead to job sorting in the way obese workers choose a job with less interaction with consumers.

For marriage market outcomes, women's looks were not related to the likelihood of marriage. However, below-average-looking women were more likely to marry men with lower education level than their own attainment. If educational level has a positive effect on earnings in the labor market, this would imply that below-average-looking women face additional economic penalty for bad looks in the form of marrying a husband with potentially lower earnings.

Harper (2000) replicated the study by Hamermesh and Biddle (1994) using U.K. data, and found similar results. It is possible that people with higher earnings are more able to and willing to invest in beauty, and thus, beauty may be endogenous. However, the potential endogeneity problem of looks might be less crucial than obesity because an individual's looks hardly change during adulthood by any natural way.

Another study does not seem to support this assumption of exogeneity of looks. In a study using a Chinese survey data, Hamermesh, Meng, and Zhang (2002) estimated a system of structural equations of earnings, spending on beauty items, and the respondents' perceived beauty. The respondents' perceived beauty was scored by the interviewers' subjective criteria. The identification for earnings came from interviewer fixed effects, and total household expenditure. Spending on beauty items were identified by the interviewer fixed effects, measures of human capital, and the respondents' health and nutritional characteristics. The identification for perceived beauty was raised from the respondents' human capital, and occupational achievement. They found that additional spending on clothing and cosmetics has a positive marginal effect on a woman's perceived beauty. Moreover, such spending on beauty items was estimated to result in higher earnings for women. However, the sources of identification remained untested.

The effects of beauty on differentials in earnings and career choices were also found in a study by Biddle and Hamermesh (1998) with a longitudinal data on a large homogeneous sample of graduates from one law school. The beauty of each participant of the survey was rated using a book of photographs of matriculants in each entering class. In this study, they found that better-looking attorneys who graduated in the 1970s earned more than others after 5 years of practice. Attorneys in the private sector were better-looking than those in the public sector, and the monetary reward to beauty rose, in particular in the private sector. More attractive men obtained partnership early, and those who moved from public to private sector were more attractive ones while a switch from private to public sector was observed for less attractive ones. Both of the earnings differentials generated by beauty and sorting into the private sector grew as the respondents matured in their practices.

This study in part provides the sources of discrimination against plain looking employees, i.e., whether it comes from the consumer side or the employer side. Because this study did not find significant earning differentials among employed lawyers unlike the self-employed lawyers, employers' distaste for plain looking lawyers was not supported in this study. Rather, the empirical evidence from this study seems to point more toward customer distaste for plain looking employees. As the authors discussed, discovering of the underlying cause of the effect of beauty on labor market outcomes will be important to determine whether any public-policy intervention is required, and if so, how to implement it to alter any detrimental effect of beauty or other physical characteristics on labor market outcomes (Biddle and Hamermesh, 1998).

Significance of this study

Accurate estimation of the effect of obesity on labor market outcomes will support the understanding of the economic cost of obesity to an individual besides its adverse effect on health. Individuals' behavioral choices regarding body weight can impose costs not only to the individuals themselves, but also to the others that are not relevant to the choices. An example is the high health care costs for the obese individuals. The rising health expenditure for obese people tolls the other individuals in the same insurance pool via rising insurance premium. If the obese people are under a public insurance, their large consumption of health care resources cost the general public (Bhattacharya and Sood, 2005). Government regulation, such as a special tax for junk food may be needed (or able to) curb the obesity epidemic. However, given that individuals are the ultimate decision-makers for their body weight, raising awareness of the obesity costs to individuals may also be important to reduce the obesity epidemic. Individuals are known to change their behavioral choices more efficiently by a response to the incentives rather than their strong willingness or preferences for changes (French, Story, and Jeffery, 2001; Cawley, 2004). The spillover effect of obesity on labor market outcomes may be able to provide an additional incentive to the individuals to adjust their behavioral choices toward a healthier body weight.

This study will also improve the literature in the following ways:

1) None of the previous studies have used over-identifying instruments, which leaves the validity of the instruments untested in those studies. Over-identifying instruments allow

conducting the tests of the exogeneity of over-identifying instruments for obesity in the labor market outcome equations besides the tests of the quality of instruments. Thus, this study generates valid parameter estimates of the effect of obesity on labor market outcomes with supported instruments. The individual fixed-effects model is used in conjunction with the two-stage instrument variables estimation model to sweep out any unobserved permanent individual heterogeneity in the error term. Altogether, this study accounts for the feedback effect of labor market outcomes on wages. Within-variation estimators alone would be inconsistent if strict exogeneity fails by feedback effect, that is, if the current obesity level is affected by the prior error term in the equations on labor market outcomes.

2) This study distinguishes the effect of obesity on labor market outcomes at the extensive margin (i.e., employment and occupation choice), as well as at the intensive margin (i.e., wages for participating workers). Although one previous study investigated the effect of obesity on the probability of employment at a so-called white-collar job (Cawley, 2000), the parameter estimates in that study remained untested due to limited controls for the endogeneity of obesity, as discussed earlier. None of the previous literature has estimated the effect of obesity on the occupation choices where obesity may affect the job performances. Furthermore, this study investigates different marginal effect of obesity on hourly wages between occupations where obesity may penalize and other occupations.

3) This study identifies different effects of obesity on labor market outcomes at different points in the life cycle by estimating separate models by age groups. Considering labor market outcomes show diverse patterns over the life cycle, it would be reasonable to assume that the effect of obesity would be different at various stages of the life cycle. In addition,

age-specific models allow for examination of the cumulative effects of obesity on labor market outcomes at the intensive margin.

CHAPTER III: CONCEPTUAL FRAMEWORK

In this chapter, potential underlying factors linking obesity and labor market outcomes are discussed, using related economic theories. Four potential underlying factors for the total effect of obesity on labor market outcomes have been identified: health problems caused by obesity; myopia of obese individuals; consumer-based discrimination against obese workers in the labor market; and employer-based distaste for obese workers with regard to high health care cost for obese people or other factors associated with weight and job productivity. It should be noted, however, that this study focuses on the total effect of obesity on labor market outcomes, rather than estimating the direct effect of obesity on labor market outcomes through the suggested factors. How the potential instruments for obesity would work to identify its effect on labor market outcomes is discussed. Testable hypotheses suggested by the conceptual framework follow.

How does obesity affect labor market outcomes

Health problems for obese people

Obese people may have low productivity due to their health problems directly associated with obesity. For example, obese individuals are likely to have limitations in mobility, and thus, will be less productive and/or less likely to be hired if their jobs require a high level of mobility. Other than this visible mobility restriction, the previous literature has consistently reported health problems and a high health care cost caused by obesity. Therefore, obese people may have worse labor market outcomes than non-obese people, ceteris paribus, because their health problems caused by obesity limit the amount of work or types of work via higher absenteeism or sick leave.

However, the effect of health problems regardless of its causes is not the main interest of this study. An individual's health status is also endogenous because some health problems can cause obesity itself. Thus, this study excludes person-year observations for any respondent who answered that their health problems limit types of work or amount of work to exclude health problems as a potential mediating factor.

Myopia

Typically, it is assumed that an individual's time preference is persistent, and thus, her discount rate should be the same over time. However, individuals may discount the near future less heavily than the long-term future when they make decisions over time intervals. That is, those individuals discount future hyperbolically or quasi-hyperbolically (Becker and Murphy, 1988).

With hyperbolic discounting, the individual's future behavior may be inconsistent with the optimal plan of the present compared to an individual with a constant discount rate. If this inconsistency is recognized, the rational individual will either "precommit" to the future, or consistently plan. She may preclude future options so that it will conform to the present desire (precommitment). Alternatively, she may modify the chosen plan to take account of her future disobedience to the optimal plan of the present, realizing that the possibility of disobedience imposes a further constraint on the set of plans that are attainable at the present moment (consistent planning: thriftiness). However, if the individuals do not realize their inconsistent time-preference, they become “spendthrift” or myopic with inconsistent or imprudent planning (Strotz, 1955). Thus, myopic people discount the future at a higher rate than the pure time discount rate, while they trade off consumption in future states at the time discount rate (Cutler, Glaeser, and Shapiro, 2003).

If individuals are myopic, they ignore future effects when they make decisions about current consumption (Becker, Grossman, and Murphy, 1994). Food consumption brings immediate gratification, while costs of over-consumption of food occur in the future (Cutler, Glaeser, and Shapiro, 2003). Therefore, myopic workers are less likely to be concerned about long-term adverse health effects of consuming fattening foods at present than non-myopic workers, and accordingly, more likely to be obese (Cawley, 2000).

If the high discounting of future consequences of food consumption is also found in consumption of other goods associated with their human capital, those people will ignore future return to the investments on their human capital, such as on-the-job training, when they make decisions about current consumption of those investments. People with high future discount rates are also likely to participate in risky health behaviors, including smoking and heavy drinking, for the same reason they consume fattening foods. Those potentially less human capital and/or risky behaviors may cause poor labor market outcomes.

Discrimination in the labor market against obese people

Discrimination in the labor market has been widely studied by economists. It is difficult to separately identify discrimination against a group in the labor market from the effect of inter-group differences in unobserved productivity (Hamermesh and Biddle, 1994). This difficulty applies to the possible discrimination against people seeking unhealthy risk behaviors, including smoking or alcoholism. Consumption of unhealthy fattening foods (or failure to get sufficient exercise) is a risky health behavior that is more likely to be observable than consumption of other risky health behaviors. That is, consumption of fattening unhealthy foods mostly yields obesity, which is quite visible. An unhealthy risk behavior may be correlated with any other unhealthy behaviors if individuals with high discount rate are more likely to consume those behaviors. Therefore, any existing discrimination against obese people may come from only discrimination against consumption of fattening unhealthy foods (or failure to get sufficient exercise), overall discrimination against any type of risky health behaviors, or a high discount rate. Also, discrimination against obese people can result from consumer-based distastes for obese workers, or employer-based distaste regardless of their overall preference for people performing risky health behaviors. Regardless of the underlying reasons for discrimination against obese people, the discrimination in the labor market would result in poorer labor market outcomes for obese people than non-obese people.

Consumer based-discrimination

There may be some occupations where non-obese workers are more productive than obese workers due to consumer-based discrimination against obese workers. For example, consumers may prefer a slim sales representative in a beauty shop to an obese one. Results from an experiment demonstrated that employers perceived obese persons as unfit for public sales positions and as more appropriate for telephone sales involving little face-to-face contact. In another experiment for the same study, participants rated obese job applicants as lacking self-discipline, having low supervisory potential, and having poor personal hygiene and professional appearance (Puhl and Brownell, 2001; Martin, 1990).

Although Becker (1971) proposed the consumer-based discrimination theory based on the discrimination of white consumers against black sellers, that theory can be applied to explain possible consumer-based discrimination against obese workers. In order to focus on the aspect of consumer discrimination, it is assumed that some individuals have a propensity for discrimination against obese sellers, while obese sellers are indifferent about the sliminess of the buyers. Under those assumptions, if obese sellers charge monetary price P of an output, an individual with a distaste for obese sellers will perceive the price as being $P(1+d)$, where d is the discrimination coefficient. Discrimination coefficient d will measure the intensity of the propensity for discrimination against the obese seller (Becker, 1971).

This adjusted price is similar to the definition of hedonic price. Hedonic prices are defined as the implicit prices of attributes. The hedonic price is revealed to individuals from observed prices of different products and the specific amounts of characteristics associated with them. That is, consumers value goods based on the attributes or characteristics of those goods affecting their utility (Rosen, 1974). Following the hedonic price model, individual economic agents would consider a product as a whole package of observed price and other non-price characteristics entering their utility functions. Thus, those economic agents will choose a product with price and non-price benefit bundle maximizing their utilities. If those economic agents have a propensity for slim sellers, products sold by obese sellers may impose more constraint on their utility maximization due to high perceived price (Buffum and Whaples, 1995).

If consumers' propensity for discrimination against obese workers varies by type of occupation, obese workers will be systemically sorted into occupations where being non-obese is rewarded via consumer distaste for obese workers. Furthermore, if consumer-based discrimination against obese workers comes primarily from the appearance of the individuals, there may be gender differentials when obesity is measured with BMI. A large BMI for men may be capturing typical male traits, such as strength, because BMI does not measure actual body fat (Pagan and Davila, 1997). Also, there might be differentials based on employer size, assuming that large employers could carry out segregation between obese and non-obese workers in-house (Buffum and Whaples, 1995).

Job activities where body weight is likely to be important can be identified by observable job characteristics. Examples of those job observable job characteristics include the extent of strenuousness in a job or social interactions. If the empirical evidence shows that obese workers are more likely to work in a job where obesity is not penalized by consumers, consumer-based discrimination against obese workers will be supported. However, job sorting by obesity would not be complete. That is, those obese workers might take the penalizing jobs instead of taking non-penalizing jobs due to lack of skills. Non-obese workers could be found in a job where slimness is not the main feature for rewards in the job (Hamermesh and Biddle, 1994). If empirical evidence is found that obese workers earn a lower wages on their job than non-obese workers in a job where consumers may discriminate against obese workers, ceteris paribus, then the argument that the consumers' distaste for obese workers lowers the productivity of obese workers may be supported.

The actual composition of persons in the jobs penalizing obesity will also depend on the characteristics of the income distribution in the jobs that do not penalize obesity (Borjas and

Bronars, 1989). Skilled obese workers will have more incentives to take a job that does not penalize obesity than skilled non-obese workers if they can observe the different income distribution between the obese and non-obese group in a job penalizing obesity. Thus, the skill composition of workers in a job penalizing obesity and a job without penalizing obesity will differ between the obese and non-obese group.

Employer-based discrimination

If discrimination against obese people is illegal, then employer-based discrimination against obese workers will not explain the differences in labor market outcomes between obese workers and non-obese workers, assuming employers do not find ways around the law. This is because illegality of discrimination against obese people will result in little observable variation in employers' discrimination. Therefore, the identification of employer-based discrimination against obese workers will not be feasible. However, Michigan is the only state that prohibits employment discrimination on the basis of weight. In other states, the legality of discrimination against obese people depends on the content of each case (American Obesity Association, accessed in 2005). Discrimination against obese people due to their appearance may be legal as long as their obesity is not found to be a physical or mental disability that substantially limits one or more major life activities of the individual (Martin, 1994; Roehling, 1999).

Employers may have a propensity for distaste for obese employees for reasons including their own preferences for lean employees, consideration of their consumers' distaste for obese workers, belief of different ability to do jobs between obese and non-obese employees, or their concerns about rising employer-provided health insurance costs. An employer with a propensity for discrimination against obese employees will act as if a money wage rate is $\pi(1+d)$, where π is an actual money wage rate, and d is a discrimination coefficient (Becker, 1971). However, employers' distaste for obese workers would have a limited effect on the differential in labor market outcomes between the obese and non-obese group if the labor market is competitive unless firms maximize utility/welfare instead of profits and all employers have a disutility from hiring obese workers. Competition in the labor market requires that the price of an efficiency unit of each labor input be the same for all skill groups, assuming intangible aspects such as job satisfaction are reflected to the observed efficiency. Therefore, competition in labor market ensures that employers' distaste for obese workers would not affect the differential in outcomes at the average skill levels within each group of non-obese workers and obese workers. A competitive output market would also constrain employer-based discrimination against obese workers, as well as obese and non-obese employees have similar elasticities for labor supply (Borjas and Bronars, 1989; Buffum and Whaples, 1995).

Employers' concerns for rising health insurance costs due to obese employees may lead to a systematic sorting of non-obese workers into the jobs with employer-provided health insurance. Thus, if empirical evidence is found that obese workers are less likely to get a job providing health insurance via employers, ceteris paribus, employers' concerns about high health care costs for obese workers will be supported for a cause of the different labor market outcomes between obese and non-obese workers. Nonetheless, obese workers can be found in a job with employer-provided health insurance, if other characteristics of those obese workers (such as higher job skills) could generate positive net profits when subtracting the marginal increase in cost due to high health care costs for their obesity (or, it may be just because non-obese employees cannot choose not to subsidize obese employees' health care cost in the job with employer-provided health insurance). Obese employees would prefer a job providing a good health insurance plan than a job without it once they recognize their high risk of having health problems due to their obesity. Instead of not hiring those obese workers, employers may try not to give an increase in wage for obese people to compensate for the incremental health care costs for obese workers. However, this study does not empirically address the role of employer-provided health insurance in the causal effect of obesity and labor market outcomes.

Potential instruments for obesity

Potential instruments for obesity to identify causality on labor market outcomes are chosen from factors that have been discussed as contributing to obesity in the previous literature. Several exogenous factors have been discussed as contributors to obesity as a choice variable in the previous literature.

First, fast food and convenience foods are inexpensive and are high in calories compared to other healthier foods (Popkin, 2001). The increasing trends of labor market participation of women will reduce the time and energy available for home production including food preparation, which can also contribute in part to the increasing consumption of convenience or fast food.

Second, an increase in the number of fast-food restaurants in town will decrease the time cost for using those services, which will result in cheaper access to those places. The previous literature has shown that the number of restaurants per capita had a positive and significant effect on the weighted sample means of the extent of obesity, while price at fast-food restaurants, full-service restaurants, and the price of food at home had negative and significant effects on the weighted sample means of the extent of obesity (Chou, Grossman, and Saffer, 2004).

Third, smoking affects obesity although the effect or magnitude of the effect remains unsettled. Individuals who quit smoking typically gain weight. The anti-smoking campaign, which began to accelerate in the early 1970s, may be an important trend affecting increases in obesity (Chou, Grossman, and Saffer, 2004; Gruber and Frakes, 2006).

Fourth, several economic studies have pointed out that alcohol consumption is a contributor for weight gain. Alcohol is high in calories and addictive. A 12-ounce can of regular beer has more calories than other alcoholic beverages or regular soda of the same size. Thus, persistent consumption of alcohol would contribute weight, ceteris paribus. The relationship of alcohol and weight gain varies by age, gender, and weight level. The positive effect of alcohol consumption is clearer for women and higher weight categories (Maclean, Norton and French, 2006). Assuming that alcohol is a normal good, the high price of alcohol would lead to a decrease in consumption.

Based on suggested contributing factors to obesity, this study explored the following state-level variables as potential instruments: cigarette prices, per capita number of restaurants including fast food restaurants and full service restaurants, per capita number of food stores, per capita sales of food, per capita sales in all types of restaurants, cost of junk food, and cost of food.

The explanatory power (,as well as the potential endogeneity bias) of instruments for obesity increases as the increase as the unit of measure becomes close to individual level, which is the unit of analyses for this study. Therefore, the following two individual-level variables were also explored as potential instruments: siblings' BMI and five-year lags of respondents' BMI. Siblings with the same parents are likely to share parental genes affecting weight and height, which put siblings BMI as a potentially strong instrument for respondents BMI (Cawely, 2004). However, the exclusion restriction of siblings' BMI was not tested in the previous study, and it is possible that siblings' BMI is correlated with the error terms in the labor market outcomes equation. If genes affecting obesity are not exclusive from the genes affecting academic intelligence or time preference, siblings' BMI is not likely to be excluded from the labor market outcomes equations.

For this reason, this study explored one more individual-level instrument — five year lags of BMI — which allows the test of exclusion restriction for an over-identified variable. It is quite obvious that respondents' past BMI could be the most accurate predictor for the current BMI. Nevertheless, it could be a bit challenging to assume that the past BMI is excluded from the labor market outcomes model as discussed in the previous chapter. This study tries to overcome this hardship by canceling out time-invariant individual fixed effects from both the first- and second-stage equations. Also, by testing the exclusion restriction for both of the individual-level variables, this study was able to determine whether data support these two variables as valid instruments.

Testable Hypotheses

The conceptual framework leads to five testable hypotheses, which are investigated separately by gender:

Ceteris paribus, an increase in height-adjusted body weight would

1) Decrease the likelihood of being employed.

2) Decrease the likelihood of sorting into occupations where social interaction is required.

3) Decrease wage earnings.

4) Differently affect wage earnings at various stages of a life cycle.

5) Differently affect wage earnings at occupations where social interaction with customers or colleagues is required from other occupations.

CHAPTER IV: DATA

The National Longitudinal Survey of Youth

This study used the National Longitudinal Survey of Youth 1979 (NLSY79). The NLSY79 is a nationally representative sample of 12,686 young men and women who were 14 to 22 years of age when first surveyed in 1979. Blacks, Hispanics, and economically disadvantaged non-black and non-Hispanics were over-sampled. The cohort was interviewed annually through 1994, and after 1994, it has been surveyed biennially (U.S. Department of Labor, 2001).

The NLSY79 has excellent information about body weight, height, employment, marriage, investment on human capital, and other health behaviors in a panel structure. The NLSY79 is particularly useful to investigate the effect of obesity on labor market outcomes in the long term because the panel started to enter the survey when they were in the typical starting age for participating full-time in the labor market. This age distribution of the data would allow studying the effect of obesity on labor market outcomes at the extensive margin (i.e., labor market participation choice, and occupation choice), as well as at the intensive margin (i.e., a change in wage over time during their work).

Most of the information in the NLSY79 for this study is publicly available. The primary data files for this study were obtained from the public domain of the Bureau of Labor Statistics, U.S. Department of Labor. Data from 12 years (1981, 1982, 1985, 1986, 1988, 1989, 1990, 1992, 1993, 1994, 1996, and 1998) were pooled to create the samples for this study.

This study has obtained the following detailed confidential geographic information and county-level labor market condition variables in the NLSY79 by applying to the Bureau of Labor Statistics: 1) detailed geographic information: state, county and geographic region (including metropolitan area) of each respondent's location of residence at the age of the first interview at 1979; state, county and geographic region (including metropolitan area) of each respondent's location of current job; state, county and timing of up to five residential moves since January 1978 or since the last interview; and 2) labor market condition variables for county of residence from the Census *County and City Data Books* including labor force, business establishments, employment, and government programs.

The NLSY79 has maintained a high retention rate over the survey years. Around 90% of the NLSY respondents remaining eligible for interview participated in the survey during the survey years. All base-year respondents, including those reported to be deceased, are considered eligible for interview except those who have been dropped from the sample (see Table 4.1).

The final analysis included 91,435 person-years among the original eligible sample of 153,155 person-years covering the following 14 years: 1981, 1982, 1985, 1986, 1988, 1989, 1990, 1992, 1994, 1996, and 1998. The final sample of 91,435 person-years was obtained after applying the following exclusion criteria: 1) not interviewed; 2) at active military service at the time of interview; 3) pregnant within a year from the time of interview including pregnant at the time of interview; 4) younger than 18 years old; 5) interviewed less than 6 times in 12 interviews over 17 years; 6) upper or lower 1% of the overall distribution of BMI; and, 7) answered that health problems limited types or amount of works. Men and women are represented the final sample almost equally (47, 435 and 44,000 person-years, or 5,391 and 5,220 persons for men and women, respectively). Among those exclusion criteria, the number of person-year observations excluded due to their health problems limiting types or amount of works is 2,726 and 2,043 for men and women, respectively. Tables 4.2a and 4.2b show the overall distribution of the variables in the final sample by gender.

Dependent variables

This study used three measures of labor market outcomes as the dependent variables: employment, occupation, and wage. Table 4.3 summarizes the types of outcomes in the labor markets as the dependent variables, samples, models, and several of the statistical issues to be anticipated.

Employment

As displayed in Table 4.4, the probability of employment was estimated for total samples within gender, while the other two measurements of dependent variables (the probability of occupation where social interaction with customers or colleagues is required, and wages) was estimated for employed samples within gender.

Samples in the NLSY79 was coded as employed if: "(1) a sample individual did any work at all as paid employees in their own business or profession, or on their own farm, or who worked 15 hours or more as unpaid workers in an enterprise operated by a member of the family during the survey week; and (2) a sample individual was not working but had jobs or businesses from which she was temporarily absent because of illness, bad weather, vacation, labor-management disputes, or various personal reasons, whether they were paid for the time off or were seeking other jobs" (NLSY79 User's Guide, 2004). That is, any sample person out of the labor force except for temporary reasons was considered as unemployed in this study.

Men were employed slightly more than women in the final sample. Eighty two percent (39,021 person-years of 47,435 person-years) were employed, while 70% (30,871 person-years among 44,000 person-years) were employed.

Overall, men switched employment status either from employment to non-employment or vice versa less often than women (37.9% for men versus 41.2% for women) in the total sample. The proportion of the sample persons who ever changed the employment status from employment to non-employment was higher in women than men (26.4% for women versus 16.4% for men), while it was opposite in the cases for the sample persons who ever changed the employment status from non-employment to employment. This different pattern for the switch in the employment status may be related to child birth, following maternity leave, and child-rearing for women.

The proportion of the sample persons who ever switched their employment status in either direction was different by age groups: a higher proportion of men ever switched employment status than women for 18-24 age group, while the trend was opposite in the rest of the age groups (see table 4.4).

Wages

Wages were assumed to be missing for individuals who do not enter the labor market.

Having no wages may indicate that the wages those individuals could earn if they were to work were simply unobserved. Because non-working individuals with unobserved wages are likely to be systemically different than working individuals with observed positive wages, the Heckman selection model with controls for selection is appropriate (Puhani, 2000). The identifying instruments in the Heckman selection model included the following three state-level variables: unemployment rate, number of business establishments, and number of Social Security Program beneficiaries (Cawley, 2000; Puhani, 2000; MaCurdy, Green, and Paarsch, 1990).

The NLSY79 collected data on respondents' usual earnings including tips, overtime, and bonuses but before deductions during every survey year for each employer for whom the respondent worked since the last interview date (NLSY79 User's Guide, 2004). For this study, wages were measured by the hourly rate of pay at the most current or the most recent job (CPS job). Yearly inflation was adjusted in the hourly wages at the CPS job by GDP deflator. The hourly wages greater than $400 before adjusting for the GDP deflator were replaced as missing.

Hourly wages increased with age for both genders in the employed final sample. Hourly wages varied a lot for the underweight women sample (BMI<25) particularly for the samples in their mid and late 30s. Considering the small number of samples in the underweight group, and general weight gain over ages, underweight women sample may be more heterogeneous compared to other body weight groups in terms of labor market outcomes. On average, a man earned $13.19 per hour on average for GDP deflated hourly wages at the current job, while the average hourly wages was $10.74 for a woman (see Table 4.2a and 4.2b).

Occupations where social interaction with customers or colleagues is required

Like wages, occupations were observed only for people participating in the labor force. The difference in characteristics between working individuals and non-working individuals also applied to the estimation of the effect of obesity on the probability of having occupations where social interaction with customers or colleagues is required. Thus, the Heckman selection model was also be used to control for the selection in the measurement of occupations where social interaction with customers or colleagues is required. The identifying instruments in the Heckman selection model were the same as in the wages equation.

Information on occupations has been collected in the NLSY79 in a consistent coding scheme throughout the survey. The industry and occupation codes in the NLSY79 are a compilation of the 3-digit 1970 U.S. Census occupational classifications, the 3-digit 1980 U.S. Census occupational codes, and the 1977 military occupational specialty codes. The 1980 Census codes have been used in addition to the 1970 codes, beginning with the 1982 survey to classify the industry and occupation of respondents' CPS job, and the 1977 military occupational specialty codes used to classify responses to the 1979–85 questions on military jobs and military occupations (U.S. Department of Labor, 2001). This study used 1980 census codes. Each 1980 census occupations code was itemized as ten categories by industry, including: managerial and professional specialties; technical; sales; administrative support/clerical; service; farming, forestry and fishing; precision production, craft, repair; machine, assemblers, inspectors; transportation, material moving; and, handlers, helpers, laborers. Distribution of the ten categories of occupational codes in the 1980 census between men and women showed that women consisted of almost three fourth in the administrative support/clerical category. Women also composed more than half in the service category. For the sales, and managerial and professional specialties, both genders had similar proportions (see Table 4.5).

To identify occupations where slimness is rewarded, this study combined two different sources of information and generated a dummy variable representing occupations where social interaction with customers or colleagues is required. First, the information in the classification in the Dictionary of Occupational Titles (DOT) was used. The DOT was developed for standardizing occupational information by the U.S. Employment Service. Based on the data collected by occupational analysts, the first edition of the DOT was published in 1939, and the fourth revision was released in 1991. Blocks of jobs were assigned to the 9-digits occupational codes that are based on the nature of the work performed and the demands of such work activities upon the workers. These work requirements included eight separate classification components: training time, aptitudes, interests, temperaments, physical demands, working conditions, work performed, and industry. Among those 9-digits of each DOT codes, the 5th digit reflects relationship to people, which is categorized as a nine point scale: mentoring (scale 0), negotiating (scale 1), instructing (scale 2), supervising (scale 3), diverting (scale 4), persuading (scale 5), speaking-signaling (scale 6), serving (scale 7), and taking instructions-helping (scale 8) (Office of Administrative Law Judges Law Library, 1991). This study included all but the last (taking instructions-helping) as indictors that interpersonal interaction is an important aspect of the occupation (Hamermesh and Biddle, 1994). Therefore, a dummy variable representing occupations where social interaction with customers or colleagues is required was generated as a dependent variable for the occupations with scale 0 to 7 in the 5th digit of the DOT codes. Around half of the employed population had occupations where social interaction with customers or colleagues is required in the final sample (52.03% for men and 56.39% for women).

In the NLSY79, occupations were coded following the Census 1980 until 1998. Census 1980 codes cannot be linked to the DOT codes directly. Thus, for this study, several linking algorithms were adopted for assigning the 5th digit of the DOT codes to each of the occupation codes in the final sample by matching the occupation code in the Census 1980 system in the data and the DOT codes. The DOT codes are linked to the Occupational Information Network (O*NET) codes by a matching algorithm provided by the developer of the O*NET system. O*NET is a comprehensive database of worker attributes and job characteristics, which was developed as the replacement for the DOT. The first edition of O*NET was released in 1998. O*NET codes can be linked to the Census 2000 codes by its original design. Also, the Census Bureau provides a matching table for linking occupation codes in the Census 2000 to the old Census including the 1980 Census. Through these multiple matching algorithms, the DOT codes and the occupation codes in the O*NET system were linked to the 1980 census occupation codes.

Second, the O*NET was used. Job characteristics in O*NET include the followings: knowledge, skills, abilities, generalized work activities, work context, work styles, work interests, education and training levels, and occupation-specific tasks. Skills required in an occupation are further categorized as the following: basic skills, complex problem solving skills, resource management skills, social skills, system skills, and technical skills. Among those skills categories, social skills are defined as “developed capacities used to work with people to achieve occupational goals,” and further categorized as six subcategories including:

coordination, instructing, negotiation, persuasion, service orientation, and social perceptiveness (O*NET resource center, accessed in 2005). This study took the occupations that require those social skills as the occupations where slimness may reward performances in jobs. Therefore, a dummy variable representing occupations where social interaction with customers or colleagues is required was generated as a dependent variable for the occupations requiring social skills in the O*NET. The proportion of sample persons in the occupations where social interaction with customers or colleagues is required dropped to around one third when those occupations were identified using the O*NET system compared to the DOT codes (30.21% for women and 22.63% for men).

As a composite measure, a dummy variable was generated for characterizing occupations that were classified to require social interaction with customers or colleagues either in the DOT codes or O*NET system. As expected, the proportion of sample persons in occupations where social interaction with customers or colleagues is required increased in both genders compared to two different measures described above (62.23% women versus 70.20% for men). In the final analysis, only the composite measure of occupations was used.

In general, an almost equal proportion of men and women switched occupations status either from occupations where social interaction with customers or colleagues is required to the other occupations or vice versa (45.8% for men versus 47.5% for women) in the total sample. The proportion of the sample who ever changed from occupations where social interaction with customers or colleagues are required to other occupations without such requirements was lower for women than men in total sample (35.8% for women versus 41.1% for men). The proportion of the sample who ever switched in either direction was much smaller in the 35-41 years of age group than the other three younger age groups for both genders. A higher proportion of men ever switched occupations status than women for all age groups except the oldest 35-41 years of age group (see Table 4.6).

Explanatory Variables

The variable of primary interest is the extent of obesity, which was measured with body mass index (BMI). BMI is defined as weight in kilograms divided by height in meters squared. In the NLSY79, height (self-reported by the respondents) information was collected only three times, 1981, 1982, and 1985, although the respondent's current weight (self-reported by the respondents) has been collected in every round of the survey. However, given that respondents were between 20 and 27 in 1985, height in 1985 was used as the respondents' adult height on the assumption that height typically stops changing at those ages (Cawely, 2004).

Both height and body weight information in the NLSY79 are self-reported, which may contain measurement error. Several previous studies using the NLSY79 have used the third National Health and Nutrition Examination Survey (NHANES III) to correct the potential measurement error on self-reported height and weight, which has both measured and self-reported information on weight and height (for example, Cawley 2004). However, this study does not correct the potential measurement error for the following reasons: 1) the the NLSY79 is a representative sample for youths aged 16 to 22 at its starting year, while the NHANES III samples are not restricted to youths; 2) the respondents in the NHANES III were aware that their weight and height would be measured after their self-reports of weight and height (Evans et al, 2005). Thus, the size or magnitude of the errors in self-reported height and weight in the NHANES III may be different from other survey with only self-reported height and weight like the NLSY79.

Persons with BMI equal to or greater than 30 are classified as obese. A BMI between 25 and 30 is classified as overweight, and BMI below 18.5 is underweight (National Heart, Lung, and Blood Institute, 1998). Four BMI splines were generated with cut points 18.5, 25, and 30 to obtain the different marginal effect of BMI on each category of state of being obese on labor market outcomes.

BMI in the total sample increased over time as sample persons aged for both genders. This increase in BMI over ages also was clear when BMI were categorized into the four groups as described above. The proportion of being overweight and obese increased over age while the proportion of the normal weight declined over age for both genders (see figure 4.1 and 4.2).

In the final sample, BMI showed some extent of within-person variation over time for both men and women. Variations in BMI were checked for the continuous BMI measure, as well as for the following four categories: BMI < 18.5; $18.5 \leq$ BMI < 25; $25 \leq$ BMI < 30; and BMI ≥ 30. Table 4.7 shows within person variation in the four BMI categories in the total and employed samples by both genders. Overall, men switched BMI groups more than women in both the total and employed sample (47.5% for women versus 49.5% for men in the total sample, and 43.4% for women versus 46.7% for men in the employed sample). The proportion of the sample who ever switched group was different by the BMI group, and the difference was wider for women than men. When the final and employed sample was grouped into four age groups, the overall extent of within person variation in each of four BMI groups decreased to 20 to 30%. For each BMI group, the within-person variation diverged from 12% to 52% (see Table 4.8a and 4.8b). In another measure of within person and between persons variations of BMI, the continuous BMI per person was decomposed into the mean of BMI over time per person and the deviations of BMI of a person in each time from the average BMI over time per person. The standard deviation of those two components of BMI was displayed in Table 4.9 for the total and employed sample by men and women. Overall, average BMI varied between 17 and 45. Deviation of BMI at time t for each sample person ranged -18 and 15. Within person variation in BMI over time was more than half of the variation in BMI across sample persons in both the total and employed sample for both genders (see Table 4.9).

Other descriptive data distributions

The proportion of sample persons with a college or higher education was greater for women than men across ages over 25. For example, at age 30, around 30 to 34% of the sample had obtained college or higher education. Overall, up to 33 to 37% in the total sample received college or higher education (see Table 4.9).

In Table 4.10a and 4.10b, non-Hispanic Whites were 49% of the total sample, and non-Hispanic Blacks were 28%. Hispanic and Asian had a smaller proportion, which composed 17% and 6%, respectively, of the final sample. Among four regional areas inclusive of Northeast, North-central, South, and West, sample persons in South were 39% of the total sample, while the other three regional areas were composed almost evenly (18% in Northeast, 24% in North-central, and 19% in South).

Figure 4.3 shows hourly wages, which was deflated by average GDP, over ages among BMI groups by gender. Overall, adjusted hourly wages increase as sample persons get older in all BMI groups in both genders. Hourly wages were varied more in the underweight women group, in particular, from in their mid 30s. A slightly similar trend was displayed in the normal weight group for men. The sample in the heavier BMI group earned less than the lighter BMI group at a given age for both genders, which was as expected.

Table 4.1 Sample sizes, retention rates, and response rates in the NLSY79

Year	Eligible sample	Interviewed sample	# deceased	response rate	retention rate
1979	12,686	12686	0	-	-
1980	12,686	12141	9	95.8%	95.7%
1981	12,686	12195	29	96.3%	96.1%
1982	12,686	12123	44	95.9%	95.6%
1983	12,686	12221	57	96.8%	96.3%
1984	12,686	12069	67	95.6%	95.1%
1985[4]	11,607	10894	79	94.5%	93.9%
1986	11,607	10655	95	92.6%	91.8%
1987	11,607	10485	110	91.2%	90.3%
1988	11,607	10465	127	91.2%	90.2%
1989	11,607	10605	141	92.5%	91.4%
1990	11,607	10436	152	91.1%	89.9%
1991[5]	11,607	9018	144	91.8%	90.5%
1992	9,964	9016	156	91.9%	90.5%
1993	9,964	9011	177	92.1%	90.4%
1994	9,964	8891	204	91.1%	89.2%
1996	9,964	8636	243	88.8%	86.7%
1998	9,964	8399	275	86.7%	84.3%
2000	9,964	8033	313	83.2%	80.6%
2002	9,964	7724	346	80.3%	77.5%
Total	227,113	205,703	2,768	91.7%	90.6%

Notes:

1. Source: NLSY79 User's Guide: A Guide to the 1979–2002 National Longitudinal Survey of Youth Data.
2. Response rate is defined as "the percentage of base-year respondents remaining eligible and not known to be deceased who were interviewed in a given survey year".
3. Retention rate is calculated by "dividing the number of respondents interviewed by the number of respondents remaining eligible for interview." All 1979 (round 1) respondents including those reported as deceased are eligible for interviews, with the exception of those who have been permanently dropped from the sample.
4. After the 1984 surveys, interviewing ceased for 1,079 members of the military sub-sample; retained for continued interviewing were 201 respondents randomly selected from the original entire military sample of 1,280; 186 of the 201 participated in the 1985 interview. The total number of the NLSY79 civilian and military respondents eligible for interview (including deceased respondents) beginning in 1985 was 11,607.
5. The 1,643 economically disadvantaged non-Black/non-Hispanic men and women members of the supplemental sub-sample were not eligible for interview as of the 1991 survey year. The total number of the NLSY79 civilian and military respondents eligible for interview (including deceased respondents) beginning in 1991 was 9,964.
6. The year used for this study is shaded.

Table 4.2a Summary statistics for the final sample: women[1]

Variables	Mean	Std. Dev.	Min	Max
Dependent variable				
Employed	0.702	0.458	0	1
Ln(hourly wage) (N=30,871)	2.322	0.506	0	6.068
Occupations requiring social interaction (N=30,871)	0.572	0.495	0	1
Independent Variable of interest				
BMI	24.673	4.916	17.243	45.725
Individual instruments for BMI				
Five year lags of BMI (N=10231)	24.313	4.633	17.243	45.359
Siblings' BMI (N=17760)	25.304	4.587	17.270	45.725
State-level IV for BMI				
Cost of fast food[2]	3.709	0.689	1.129	7.423
Cost of beer[3]	3.488	0.647	2.129	5.815
Sales in restaurants[4]	0.728	0.147	0.326	1.840
Instruments for the Heckman model				
Unemployment rate: > 15%	0.023	0.149	0	1
Unemployment rate: 12 - 15%	0.059	0.236	0	1
Unemployment rate: 9 - 12%	0.145	0.352	0	1
Unemployment rate: 6 – 9%	0.372	0.483	0	1
Unemployment rate: < 3%	0.024	0.152	0	1
Number of beneficiaries receiving Social Security Benefits[5]	0.154	0.024	0.048	0.214
Total number of employment per 10,000 state populations	5451.937	685.040	3809.902	13818.140
Year				
1981	0.086	0.280	0	1
1982	0.082	0.275	0	1
1985	0.089	0.285	0	1
1986	0.085	0.279	0	1
1988	0.084	0.277	0	1
1989	0.089	0.285	0	1
1990	0.087	0.282	0	1
1992	0.076	0.264	0	1
1993	0.076	0.266	0	1
1994	0.084	0.278	0	1
1996	0.082	0.274	0	1
1998	0.080	0.271	0	1
Demographic variables				
Race				
Black	0.276	0.447	0	1
Hispanics	0.173	0.378	0	1
Asian	0.062	0.241	0	1
White	0.489	0.500	0	1
Age	28.164	5.421	18	41

Table 4.2a Summary statistics for the final sample: women – continued

Variables	Mean	Std. Dev.	Min	Max
Education				
< High school	0.684	0.465	0	1
College	0.152	0.359	0	1
> College	0.164	0.370	0	1
Regional Variables				
State per capita income[6]	20.134	3.316	11.864	31.726
Northeast	0.177	0.382	0	1
North-central	0.230	0.421	0	1
South	0.407	0.491	0	1
West	0.185	0.389	0	1
Urban	0.788	0.409	0	1
CPI	1.273	0.224	0.907	1.755
Cost of living	1.086	0.188	0.714	2.167
Variable for a sensitivity analysis				
Marital status (married)	0.453	0.498	0	1
Number of children	1.098	1.188	0	9
Pregnancy (never pregnant)	0.448	0.497	0	1

Notes:

1. Total observations are 44,000 unless otherwise noted.
2. Average cost of the following three items: a McDonald's Quarter-Pounder with cheese, a thin crusted cheese pizza at Pizza Hut or Pizza Inn, fried chicken at Kentucky Fried Chicken or Church's (Chou, Grossman, and Saffer, 2002).
3. Average price of a bottle of Budweiser Schlitz before the fourth quarter of 1989, and Budweiser and Miller Light as of the fourth quarter of 1989.
4. Sales in full-service and limited service restaurants in 1,000$ per 100 state populations.
5. Number of beneficiaries receiving Social Security Benefits per state population.
6. GDP deflated state per capita personal yearly income in $1,000.

Table 4.2b Summary statistics for the final sample: men[1]

Variables	Mean	Std. Dev.	Min	Max
Dependent variable				
Employed	0.823	0.382	0	1
Ln(hourly wage) (N=39,021)	2.493	0.534	0	6.145
Occupations requiring social interaction (N=39,021)	0.521	0.500	0	1
Independent Variable of interest				
BMI	25.839	3.924	18.794	40.687
Individual instruments for BMI				
Five year lags of BMI (N=14,293)	25.779	3.826	18.794	40.687
Siblings' BMI (N=21,578)	25.489	4.610	17.243	45.725
State-level IV for BMI				
Cost of fast food[2]	3.726	0.677	1.129	7.423
Cost of beer[3]	3.493	0.641	2.129	5.815
Sales in restaurants[4]	0.731	0.144	0.326	1.840
Instruments for the Heckman model				
Unemployment rate: > 15%	0.022	0.148	0	1
Unemployment rate: 12 - 15%	0.057	0.232	0	1
Unemployment rate: 9 - 12%	0.142	0.349	0	1
Unemployment rate: 6 – 9%	0.371	0.483	0	1
Unemployment rate: < 3%	0.023	0.149	0	1
Number of beneficiaries receiving Social Security Benefits[5]	0.153	0.023	0.048	0.214
Total number of employment per 10,000 state populations	5469.998	702.680	3809.902	13818.140
Year				
1981	0.070	0.256	0	1
1982	0.085	0.279	0	1
1985	0.089	0.285	0	1
1986	0.086	0.280	0	1
1988	0.091	0.287	0	1
1989	0.093	0.291	0	1
1990	0.092	0.289	0	1
1992	0.080	0.271	0	1
1993	0.080	0.271	0	1
1994	0.080	0.271	0	1
1996	0.078	0.269	0	1
1998	0.075	0.263	0	1
Demographic variables				
Race				
Black	0.279	0.448	0	1
Hispanics	0.175	0.380	0	1
Asian	0.050	0.218	0	1
White	0.497	0.500	0	1
Age	28.026	5.292	18	41

Table 4.2b Summary statistics for the final sample: men – continued

Variables	Mean	Std. Dev.	Min	Max
Education				
< High school	0.723	0.448	0	1
College	0.120	0.326	0	1
> College	0.157	0.363	0	1
Regional Variales				
State per capita income[6]	20.234	3.243	11.864	31.726
Northeast	0.181	0.385	0	1
North-central	0.242	0.428	0	1
South	0.379	0.485	0	1
West	0.199	0.399	0	1
Urban	0.792	0.406	0	1
CPI	1.274	0.219	0.907	1.755
Cost of living	1.093	0.197	0.714	2.167
Variable for a sensitivity analysis				
Marital status (married)	0.416	0.493	0	1
Number of children	0.706	1.041	0	8
Pregnancy (never pregnant)	N/A	N/A	N/A	N/A

Notes:

1. Total observations are 47,435 unless otherwise noted.
2. Average cost of the following three items: a McDonald's Quarter-Pounder with cheese, a thin crusted cheese pizza at Pizza Hut or Pizza Inn, fried chicken at Kentucky Fried Chicken or Church's (Chou, Grossman, and Saffer, 2002).
3. Average price of a bottle of Budweiser Schlitz before the fourth quarter of 1989, and Budweiser and Miller Light as of the fourth quarter of 1989.
4. Sales in full-service and limited service restaurants in 1,000$ per 100 state populations.
5. Number of beneficiaries receiving Social Security Benefits per state population.
6. GDP deflated state per capita personal yearly income in $1,000.

Table 4.3 Description of Dependent Variables

Types of outcomes	Sample	Model	Statistical Issues
1. Employment	Total Sample	Probit; Linear Probability Model; 2SRI	Endogeneity
2. Occupations requiring social interaction with customers or colleagues	Employed sample	Heckman Selection; Linear Probability Model; 2SRI	Endogeneity; Selection
3. Wages	Employed sample	Heckman Selection; Log linear; 2SRI	Endogeneity; Selection

Table 4.4 Variation in employment status over time by gender and age groups

Gender	Age group	Employed	Freq. (person-years)	%	Freq. (persons)	% of ever switchers[a]
Men	18-24	No	3876	28.84	2442	45.59
		Yes	9564	71.16	4361	21.34
		Total	13440	100.00	6803	30.04
	25-29	No	2101	14.10	1325	49.04
		Yes	12803	85.90	4548	9.83
		Total	14904	100.00	5873	18.68
	30-34	No	1748	13.25	1055	44.35
		Yes	11446	86.75	4145	8.39
		Total	13194	100.00	5200	15.68
	35-41	No	689	11.68	497	36.20
		Yes	5208	88.32	2674	5.52
		Total	5897	100.00	3171	10.33
	Total	No	8414	17.74	3220	71.66
		Yes	39021	82.26	5057	16.40
		Total	47435	100.00	8277	37.90
Women	18-24	No	4547	36.18	2874	37.67
		Yes	8021	63.82	3959	22.10
		Total	12568	100.00	6833	28.65
	25-29	No	3560	27.56	2075	34.31
		Yes	9358	72.44	3971	14.31
		Total	12918	100.00	6046	21.17
	30-34	No	3441	27.56	1894	33.30
		Yes	9044	72.44	3759	14.32
		Total	12485	100.00	5653	20.68
	35-41	No	1581	26.22	1033	30.54
		Yes	4448	73.78	2464	12.09
		Total	6029	100.00	3497	17.54
	Total	No	13129	29.84	3975	59.95
		Yes	30871	70.16	5023	26.38
		Total	44000	100.00	8998	41.21

Note:
a. The ever-switchers represent for the sample who ever switched the employment status.

Table 4.5 Distribution of industry categories of the 1980 Census codes in the final sample

Occupation	Women	%	Men	%	Total
Managerial and professional specialties	7,184	50.25	7,113	49.75	14,297
Technical	1,239	48.15	1,334	51.85	2,573
Sales	2,995	50.98	2,880	49.02	5,875
Administrative support/clerical	9,170	76.02	2,893	23.98	12,063
Service	6,245	55.79	4,949	44.21	11,194
Farming, forestry and fishing	218	13.44	1,404	86.56	1,622
Precision production, craft, repair	1,226	11.42	9,506	88.58	10,732
Operations – machine, assemblies, inspectors	1,733	34.03	3,360	65.97	5,093
Operations – transportation, material moving	223	7.64	2,697	92.36	2,920
Operations – handlers, helpers, laborers	578	16.69	2,885	83.31	3,463

Table 4.6 Variation in the status of holding occupations where social interactions is required over time by gender and age groups

Gender	Age group	Occupations requiring social interaction	Freq. (person-years)	%	Freq. (persons)	% of ever switchers[a]
Men	18-24	No	4111	42.98	2786	38.57
		Yes	5453	57.02	3320	30.09
		Total	9564	100.00	6106	33.96
	25-29	No	5579	43.58	3084	39.20
		Yes	7224	56.42	3595	31.84
		Total	12803	100.00	6679	35.24
	30-34	No	4874	42.58	2794	40.36
		Yes	6572	57.42	3318	31.41
		Total	11446	100.00	6112	35.50
	35-41	No	2206	42.36	1476	31.60
		Yes	3002	57.64	1839	23.20
		Total	5208	100.00	3315	26.94
	Total	No	16770	42.98	4522	54.07
		Yes	22251	57.02	4742	41.14
		Total	39021	100.00	9264	47.45
Women	18-24	No	3029	37.76	2219	39.57
		Yes	4992	62.24	3128	26.39
		Total	8021	100.00	5347	31.86
	25-29	No	3643	38.93	2313	38.63
		Yes	5715	61.07	3056	25.78
		Total	9358	100.00	5369	31.31
	30-34	No	3352	37.06	2139	40.59
		Yes	5692	62.94	3035	26.00
		Total	9044	100.00	5174	32.03
	35-41	No	1635	36.76	1179	31.53
		Yes	2813	63.24	1783	18.84
		Total	4448	100.00	2962	23.89
	Total	NO	11659	37.77	4112	57.23
		YES	19212	62.23	4688	35.78
		Total	30871	100.00	8800	45.80

Note:
a. The ever-switchers represent for the sample who ever switched the occupation group.

Table 4.7 Within-person variation in four BMI categories in the total and employed sample by gender

Gender	BMI group	Freq. (person-years)	%	Freq. (persons)	% of ever switchers[a]
Men	**Total sample**				
	BMI < 18.5	0	0.00	0	NA
	18.5 ≤ BMI < 25	21778	45.91	4068	41.73
	25 ≤ BMI < 30	18576	39.16	4050	51.74
	BMI ≥ 30	7081	14.93	1913	61.18
	Total	47435	100.00	10031	49.48
	Employed sample				
	BMI < 18.5	0	0.00	0	NA
	18.5 ≤ BMI < 25	17548	44.97	3705	40.01
	25 ≤ BMI < 30	15559	39.87	3661	48.49
	BMI ≥ 30	5914	15.16	1672	57.37
	Total	39021	100.00	9038	46.65
Women	**Total sample**				
	BMI < 18.5	1509	3.43	673	72.20
	18.5 ≤ BMI < 25	25684	58.37	4505	31.42
	25 ≤ BMI < 30	10150	23.07	3147	62.45
	BMI ≥ 30	6657	15.13	1674	52.81
	Total	44000	100.00	9999	47.51
	Employed sample				
	BMI < 18.5	1000	3.24	505	69.08
	18.5 ≤ BMI < 25	18524	60.00	3984	27.95
	25 ≤ BMI < 30	7014	22.72	2521	59.14
	BMI ≥ 30	4333	14.04	1337	49.88
	Total	30871	100.00	8347	43.37

Note:
a. The ever-switchers represent for the sample who ever switched the BMI group.

Table 4.8a Within-person variation in four BMI categories in the total and employed sample by age groups and gender: Total sample

Age group	BMI group	Women				Men			
		Freq. (person-years)	%	Freq. (persons)	% of ever switchers[a]	Freq. (person-years)	%	Freq. (persons)	% of ever switchers[a]
18-24	BMI < 18.5	598	4.76	442	51.62	0	0.00	0	NA
years old	18.5 ≤ BMI < 25	8454	67.27	4058	17.99	7667	57.05	3648	25.33
	25 ≤ BMI < 30	2439	19.41	1762	44.61	4190	31.18	2706	45.49
	BMI ≥ 30	1077	8.57	714	37.89	1583	11.78	1180	49.10
	Total	12568	100.00	6976	28.88	13440	100.00	7534	36.30
25-29	BMI < 18.5	502	3.89	335	43.28				
years old	18.5 ≤ BMI < 25	7891	61.09	3446	15.02	7347	49.30	2911	18.38
	25 ≤ BMI < 30	2791	21.61	1594	37.65	5745	38.55	2520	29.27
	BMI ≥ 30	1734	13.42	858	24.71	1812	12.16	834	31.78
	Total	12918	100.00	6233	23.66	14904	100.00	6265	24.54
30-34	BMI < 18.5	316	2.53	199	42.75	0	0.00	0	NA
years old	18.5 ≤ BMI < 25	6532	52.32	2770	16.01	4962	37.61	2075	21.08
	25 ≤ BMI < 30	3225	25.83	1729	34.65	5888	44.63	2530	24.48
	BMI ≥ 30	2412	19.32	1160	23.38	2344	17.77	1042	25.02
	Total	12485	100.00	5858	23.88	13194	100.00	5647	23.33
35-41	BMI < 18.5	93	1.54	71	35.86	0	0.00	0	NA
years old	18.5 ≤ BMI < 25	2807	46.56	1525	13.55	1802	30.56	1021	18.87
	25 ≤ BMI < 30	1695	28.11	1084	28.45	2753	46.68	1574	19.85
	BMI ≥ 30	1434	23.79	874	17.44	1342	22.76	816	20.07
	Total	6029	100.00	3554	19.50	5897	100.00	3411	19.61

Note:
a. The ever-switchers represent for the sample who ever switched the BMI group.

Table 4.8b Within-person variation in four BMI categories in the total and employed sample by age groups and gender: Employed sample

Age group	BMI group	Women				Men			
		Freq. (person-years)	%	Freq. (persons)	% of ever switchers[a]	Freq. (person-years)	%	Freq. (persons)	% of ever switchers[a]
18-24	BMI < 18.5	373	4.65	301	46.56	0	0.00	0	NA
years old	18.5 ≤ BMI < 25	5591	69.70	3092	14.52	5425	56.72	2989	21.78
	25 ≤ BMI < 30	1456	18.15	1096	39.66	2995	31.32	2055	38.99
	BMI ≥ 30	601	7.49	425	34.03	1144	11.96	892	43.65
	Total	8021	100.00	4914	23.78	9564	100.00	5936	31.03
25-29	BMI < 18.5	379	4.05	260	41.87	0	0.00	0	NA
year old	18.5 ≤ BMI < 25	5965	63.74	2774	12.56	6280	49.05	2680	17.12
	25 ≤ BMI < 30	1924	20.56	1176	34.78	4979	38.89	2289	27.41
	BMI ≥ 30	1090	11.65	593	23.08	1544	12.06	739	30.39
	Total	9358	100.00	4803	20.89	12803	100.00	5708	22.96
30-34	BMI < 18.5	198	2.19	135	40.54	0	0.00	0	NA
years old	18.5 ≤ BMI < 25	4860	53.74	2247	14.33	4277	37.37	1882	19.83
	25 ≤ BMI < 30	2375	26.26	1354	31.91	5127	44.79	2311	22.75
	BMI ≥ 30	1611	17.81	859	22.47	2042	17.84	942	23.23
	Total	9044	100.00	4595	21.80	11446	100.00	5135	21.77
35-41	BMI < 18.5	50	1.12	44	35.90	0	0.00	0	NA
years old	18.5 ≤ BMI < 25	2108	47.39	1226	11.58	1566	30.07	904	17.58
	25 ≤ BMI < 30	1259	28.30	858	25.59	2458	47.20	1420	18.28
	BMI ≥ 30	1031	23.18	664	15.70	1184	22.73	738	18.63
	Total	4448	100.00	2792	17.25	5208	100.00	3062	18.16

Note:
a. The ever-switchers represent for the sample who ever switched the BMI group.

Table 4.9 Variation in BMI in within and between persons by gender

Sample	Gender	Variation	N	Mean	Std. Dev.	Min	Max
Total samples	Men	Overall	44000	24.789	5.120	17.243	45.725
		Between	5391		4.668	17.360	44.236
		Within			2.384	−18.723	15.420
	Women	Overall	47435	25.896	3.979	18.794	40.687
		Between	5220		3.544	18.828	40.351
		Within			2.047	−12.949	12.147
Employed samples	Men	Overall	39021	25.971	3.956	18.794	40.687
		Between	5057		3.627	18.828	40.687
		Within			1.915	−12.949	11.571
	Women	Overall	30871	24.613	5.002	17.270	45.725
		Between	5023		4.734	17.324	44.353
		Within			2.167	−12.772	14.735

Table 4.10 Education level by BMI groups for both genders

	BMI groups				
Education groups	BMI < 18.5	18.5 ≤ BMI < 25	25 ≤ BMI < 30	BMI ≥ 30	Total
Women					
≤ High school or less	1,019	16,770	7,198	5,117	30,104
	(3.38)	(55.71)	(23.91)	(17.00)	(100.00)
College	207	4,127	1,467	891	6,692
	(3.09)	(61.67)	(21.92)	(13.31)	(100.00)
> College	283	4,787	1,485	649	7,204
	(3.93)	(66.45)	(20.61)	(9.01)	(100.00)
Total	1,509	25,684	10,150	6,657	44,000
	(3.43)	(58.37)	(23.07)	(15.13)	(100.00)
Men					
≤ High school or less	0	15,423	13,405	5,466	34,294
	(0.00)	(44.97)	(39.09)	(15.94)	(100.00)
College	0	2,631	2,252	832	5,715
	(0.00)	(46.04)	(39.41)	(14.56)	(100.00)
> College	0	3,724	2,919	783	7,426
	(0.00)	(50.15)	(39.31)	(10.54)	(100.00)
Total	0	21,778	18,576	7,081	47,435
	(0.00)	(45.91)	(39.16)	(14.93)	(100.00)

Note:
1. Proportions within rows are in the parentheses.

Figure 4.1 Distribution of BMI over age by gender

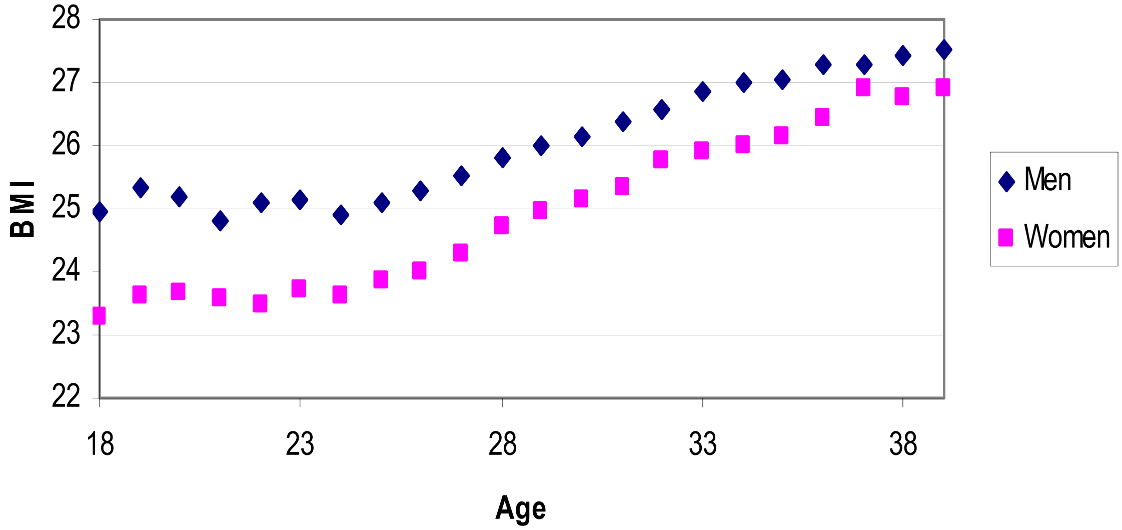
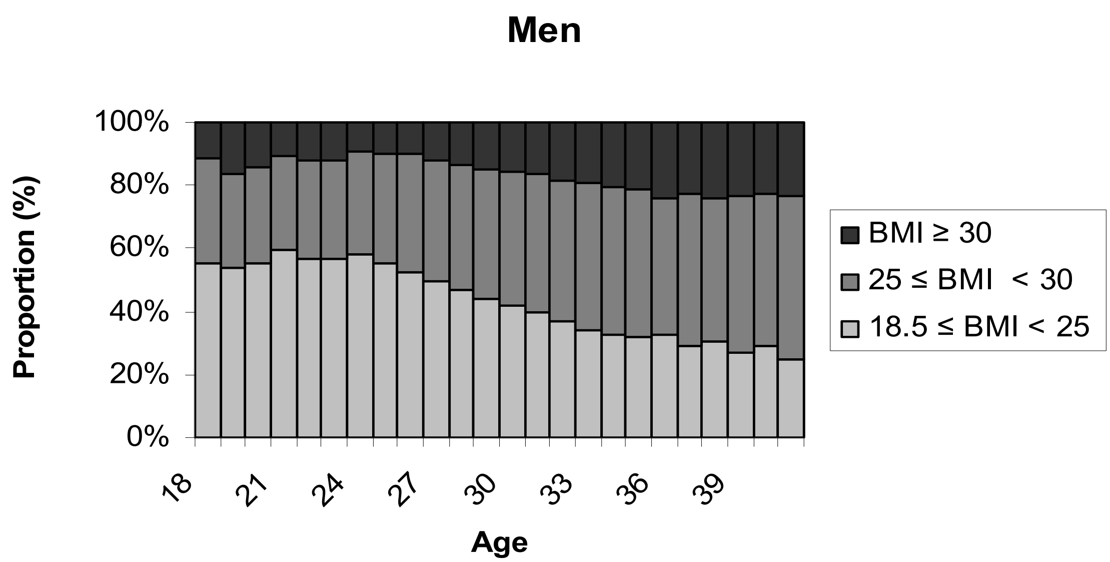


Figure 4.2 Distribution of BMI by BMI categories over age by gender



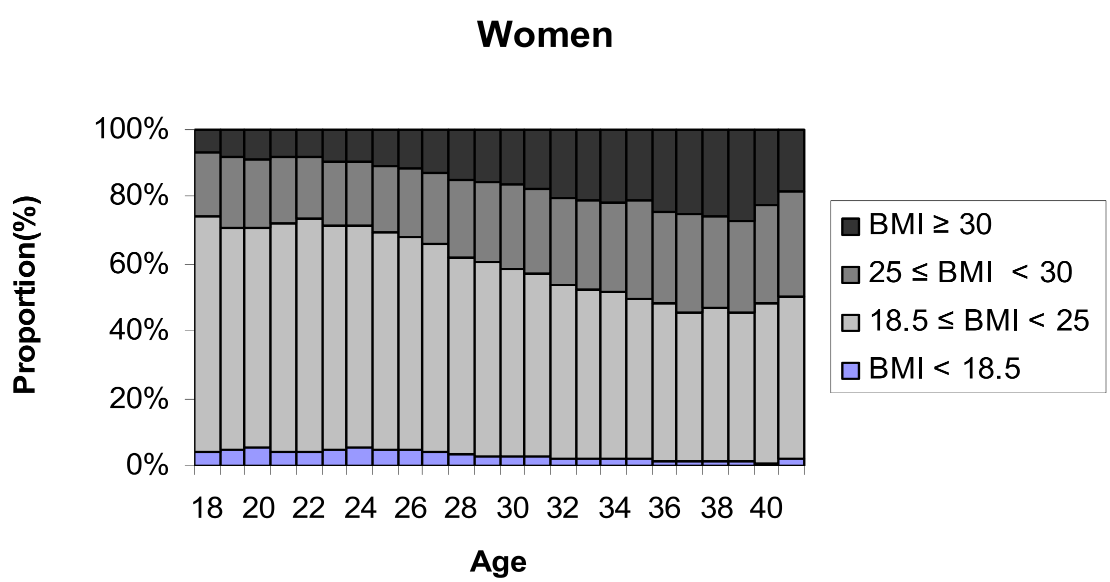
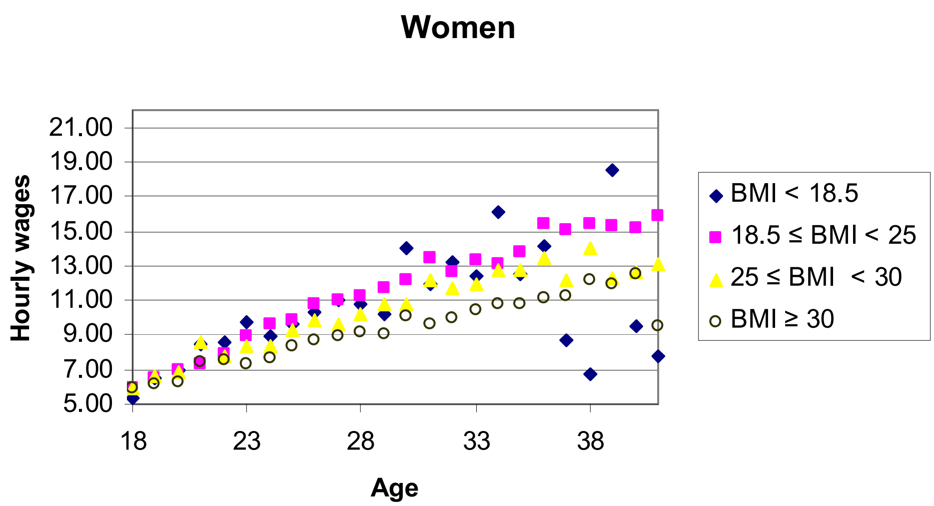
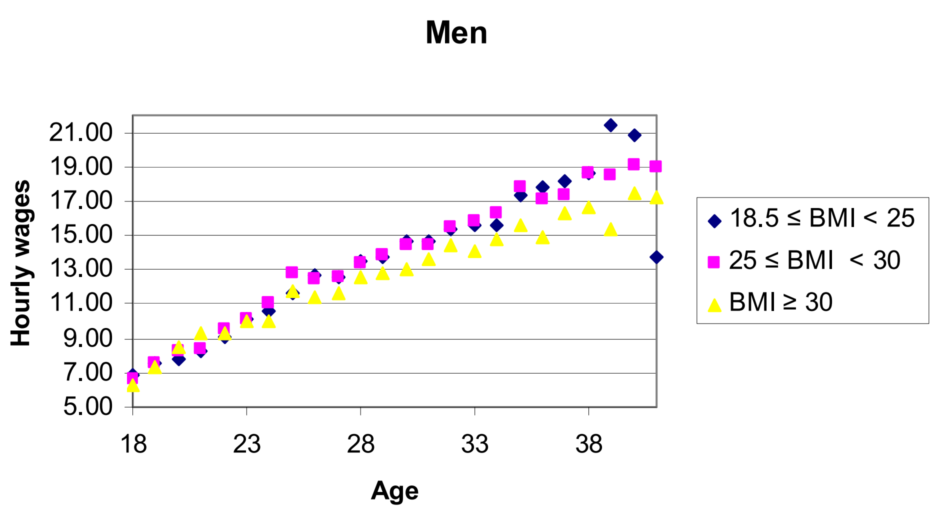


Figure 4.3 Distribution of hourly wages at the CPS jobs for men and women





CHAPTER V: RESEARCH DESIGN AND METHODS

Estimation models

Labor market outcomes are modeled as a function of three BMI splines, age, four age groups, and other exogenous variables including education levels and regional information.

The main equations are:

(1) $\Pr(Employed)_{it} = f(BMIspline_{it}, Age_{it}, Age\_group_{it}, X_{it}, \mu_i, \varepsilon_{it})$

(2) $\Pr(Occupation)_{it} = f(BMIspline_{it}, Age_{it}, Age\_group_{it}, X_{it}, \mu_i, \varepsilon_{it})$

(3) $\ln(Wage)_{it} = f(BMIspline_{it}, Age_{it}, Age\_group_{it}, X_{it}, \mu_i, \varepsilon_{it})$

(4) $\ln(Wage)_{it} = f(BMIspline_{it}, \ln(Wage)_{it-1} Age_{it}, Age\_group_{it}, X_{it}, \mu_i, \varepsilon_{it})$

where suffixes *i* and *t* stand for individual and time, respectively. In all equations, μ stands for time-invariant individual fixed effects, and ε stands for identically independently distributed error terms. Equations (2) through (4) were estimated only for employed people, while equation (1) was estimated for all sample persons. Every equation was estimated separately by gender to allow each group to have different coefficients for BMI splines, i.e., a continuous piecewise linear function with bend at 25 and 30, and other explanatory variables, as well as different intercepts for each measurement of dependent variables. In addition to separate estimation by gender, each equation was estimated separately by four age groups (18-24, 25-29, 30-34, and 35-41 years old).

This study excluded any woman respondents who were pregnant at the time of interview or had been pregnant within a year from the time of interview to control for the potential effect of pregnancy on women's BMI and labor market outcomes. Furthermore, a sensitivity analysis has been performed for the subgroup of women who have never been pregnant up until the time of interview.

All equations have two sources of variation: between-individual variation at a given time; and within-individual variation across time.

Statistical models in this study controlled for exogenously determined variables, which the previous literature has reported to affect labor market outcomes. These variables included: age, education level at the interview, and other demographic information such as region of residence (U.S. Department of Labor, 2001; Cawley, 2004). Marital status and child-bearing are known to affect the labor force participation, particularly, for women. For example, a woman may leave the job market after marriage or after a child-birth, but reenter the workforce after child-bearing or a divorce. However, this study primarily performed a reduced form model as far as the marital status and the child-bearing (or rearing) decisions by not controlling for those two variables due to the obvious endogeneity of those variables. Instead, models controlling for the martial status and the number of children were estimated as a sensitivity analysis. It should be noted that models for the sensitivity analysis treated the marital status and the number of children as exogenous, which prevent those models from being primary.

Statistical methods

<u>Statistical issue #1: Measurement error in height and weight</u>

Both height and body weight information in the NLSY79 are self-reported. It is well known that self-reported human body measurements contain measurement error. Obese people are likely to under-report their actual weight, while short people are likely to over-report their actual height (Chou, Grossman, and Saffer, 2002).

Several previous studies using the NLSY79 have used the third National Health and Nutrition Examination Survey (NHANES III) to correct the potential measurement error on self-reported height and weight, which contains information on weight and height both from physical examinations and self-reports. In those studies, actual height and weight was regressed on reported height and weight for predicting actual height and weight (Cawley, 2000; Cawley 2004; Chou, Grossman, and Saffer, 2002; Lakdawalla and Philipson, 2002).

However, the NLSY79 has a different sample composition from the NHANES III, i.e., the NLSY79 is a representative sample for youths aged 16 to 22 at its starting year, while the NHANES III samples are not restricted to youths. It is not known if the direction or size of the reporting errors would be different by the respondent's age. In addition, Evans and colleagues (2005) pointed out that respondents in the NHANES III were aware of the nature of the survey and that their weight and height would be measured after their self-reports of weight and height. Thus, the size or magnitude of the errors in self-reported height and weight in the NHANES III may be different from other survey with only self-reported height and weight like the NLSY79. Moreover, the reporting error in height and weight is not a classical measurement error that causes attenuation bias. Therefore, this study did not correct reporting errors in height and weight, unlike the previous literature.

Statistical issue #2: Selection bias in wage equations

If there is selection into the labor market, zero wages for non-workers will not be real zeros, but instead, wages for those people will be considered as missing (Greene, 2000). Therefore, if more entries into and exits from the labor market affect labor supply functions, taking into account those missing-wages into the labor supply function is more important (Heckman, 1993).

Studies about wage- and income-elasticity for men have shown that labor market participation was almost inelastic for individuals with higher wages and for greater hours worked. Instead, men who worked zero or near zero hours showed an elastic responsiveness to the wage and income in their choice for labor market participation. Those results imply that participants in the labor market might be systemically different from non-participants, and the probability of participating in the labor market contributed the most to the estimated wage- and income-elasticity of hours worked for marginal participants (Heckman, 1993).

The secular trends of participation into the labor force for men have also showed that men in the labor force are systemically different from men not in the labor force. The significant secular increase in unemployment among men since 1967 was found to be heavily concentrated among less skilled individuals (Juhn et al., 1991). The secular trend of participation in the labor force for black men was different from white men. First, secular declines in participation rate into the labor force are heavily concentrated in the youngest experience groups for black men. Second, within educations levels, blacks who are not working in a typical week are much more likely to have not worked the entire year. Those black men not in the labor market are much less likely to be married and live with spouses, and much more likely to live with relatives than black men in the labor force (Juhn, 1992).

Those observations may imply that the decision to participate is very relevant for estimating the men labor supply.

For married women, wage- and income- elasticity of hours worked are generally still larger in absolute value than the labor supply elasticity for married men (Heckman, 1993). Typically, men are wage earners in a married household, while women substitute hours worked in the market for worked hours in household production. The participation in the labor force for prime-working age (aged person between 25-54 years) women has been much lower than for men during the years, and more than 90% of prime-age men have participated in the labor force.

Selection into the labor market has been an important issue in most studies of women's labor supply (Killingsworth and Heckman, 1986). Marital patterns have major implications for women's participation in the labor force (Becker, 1973). For example, Devereux (2004) reported positive own-wage elasticity in hours worked in the labor market for women, and strong negative cross-wage elasticity, while both own-wage and cross-wage elasticity in hours worked in the labor market for men were very small. These results imply that married women's labor market participation strongly responds to their husbands' wage income.

This study used the Heckman selection model to correct for the selection bias in the labor market for evaluating the effect of obesity on labor market outcomes. The probability that an individual participates in the labor force is calculated as follows: $\Pr[y > 0 \mid W] = \Phi(W\gamma, \nu)$, where y stands for the actual wage, and W is a vector of observed explanatory variables including BMI and all other variables in outcome equations with coefficient γ and ν for errors. The Inverse Mills' Ratio, $\lambda(W\gamma) = \phi(W\gamma) / \Phi(W\gamma)$, is used to estimate the expected value of the error term conditional on the actual income y being positive as the following formula shows:

$$E[y \mid y > 0, X] = X\beta + E[\varepsilon \mid y > 0, X] = X\beta + \rho\sigma_{\varepsilon}\lambda(W\gamma),$$

where X is a vector of observed explanatory variables with coefficient β, and errors ε. The two error terms v and ε follow normal distributions with mean zero, and the variance-covariance matrix Σ is $\sum = \begin{bmatrix} 1 & \rho\sigma_{\varepsilon} \\ \rho\sigma_{\varepsilon} & \sigma_{\varepsilon}^{2} \end{bmatrix}$, where ρ is a correlation coefficient between v and ε. The variance of v is normalized to a unit.

Estimates from the probit model for the effect of obesity on the labor market participation were used to correct the selection bias in the wage equation, and the probability of having occupations where social interaction with customers or colleagues is required. Estimation models for those two outcomes were run by adding the Inverse Mills' Ratio as an additional explanatory variable.

In order to obtain a well-performing model, the Heckman selection model requires identifying instruments. The identifying instruments should be correlated with the propensity to participate in the labor force, but not be correlated with explanatory variables in the wage equations. If this exclusion restriction is not satisfied, the wage equation is only identified through the nonlinearity of the Inverse Mills' Ratio (Greene, 2000). However, the multicollinearity between other variables in the main model and the Inverse Mills' Ratio is likely to prevail if the Inverse Mills' Ratio is a linear function of the arguments of the main model over a wide range, or the arguments in the main model have a small range (Puhani, 2000; Dow and Norton, 2003).

The identifying instruments for the propensity to participate in the labor force included: the unemployment rate at the current residential area, number of business establishments, and number of Social Security Program beneficiaries (Cawley, 2000; Puhani, 2000; MaCurdy, Green, and Paarsch, 1990). The unemployment rate was provided as in the residential unit in the NLSY data. The other two instruments were obtained from *the Census of Retail Trade* at the state level.

Statistical issue #3: Endogeneity of weight

The endogeneity of obesity was controlled for with two methods: 1) two-stage instrumental variables estimation in conjunction with the individual fixed-effects model; and, 2) dynamic panel data models.

Individual fixed-effects model

The panel nature of the NLSY79 allows control for the time-constant unobserved individual heterogeneity via individual fixed effects. The individual fixed-effects model cancels out any time-invariant individual-level variable regardless of being observed or not. Therefore, any unobservable individual heterogeneity in the error term that causes simultaneity bias would be dropped out. Potential correlations of the unobservable individual heterogeneity with the observed explanatory variables in the model were examined using the Hausman test. For all estimation equations, the Hausman test rejected the null hypothesis of the random-effects model, and thus, the individual fixed-effects model was chosen over the random-effects model.

However, the individual fixed-effects model cannot control for all aspects of the endogeneity of obesity for the following reasons. First, time-varying unobserved individual heterogeneity would still remain in the error term, which can yield biased estimates. Examples of time-inconsistent individual heterogeneity include time-inconsistent hyperbolic discount rate, and an external shock affecting both body weight and labor market outcomes (e.g., stress).

Second, strict exogeneity is required for obtaining consistent fixed-effects estimators because the fixed-effects estimator transforms the original observations as deviations from the individual means over time, introducing all realizations of the error term series over times into the estimation. If there are any feedback effects of labor market outcomes on BMI, the fixed-effects estimators would be inconsistent with fixed time periods due to failure of the strict exogeneity condition (Bond, 2002).

BMI and other covariates at time $t+s$, where $s \geq 1$, are potentially correlated with the error term at time t in wage equation. For example, wages directly affect body weight via the individual's choice for food, and indirectly affect body weight via cost of physical activity. If the current consumption of weight-contributing goods such as high calorie food becomes realized at the near future, as well as instantly, current wage would be correlated with future weight status. Therefore, strict exogeneity potentially fails in the given estimation model on the effect of weight status on hourly wages.

Two-stage instrumental variables estimation method

Two-stage instrumental variables estimation was used for all estimation equations to control for the endogeneity of obesity. In the first stage, auxiliary (reduced-form regression) are estimated with the individual fixed-effects model. In the second-stage equation, the first-stage residuals are included as an additional regressor in the second-stage estimation. The previous literature has reported that the 2SRI yields a consistent estimator for the endogenous variables, and also identical to the two-stage least square (2SLS) model in the linear case (Terza, 2005).

In this study, the endogenous BMI variables are splined into three groups in the second stage, but linear BMI is estimated in the first stage. A Monte Carlo Simulation was performed to establish that the 2SRI model provides unbiased estimators when it is used for controlling for the endogeneity of BMI for the models on labor market outcomes when BMI was splined. The Monte Carlo Simulation results confirmed that inserting residuals from the linear estimation of the BMI in the first stage estimation into the second stage equation with splined BMI provides unbiased estimates (see Appendix 1).

To obtain asymptotically unbiased parameter estimates in either the 2SLS or the 2SRI, good instruments for BMI are required. Good instruments should be highly correlated with the BMI, but not be correlated with the error terms in the labor market outcomes equations (Greene, 2000).

This study explored three sets of instrumental variables in the estimation. First, state-wide variables were tested as the potential instruments including the following: cigarette prices, per capita number of restaurants including fast food restaurants and full service restaurants, per capita number of food stores, per capita sales of food, per capita sales in all types of restaurants, fast-food price, and food price. Among the suggested instruments, information outside of the NLSY79 was obtained following two previous studies (Chou, Grossman, and Saffer, 2002; Lakdawalla and Philipson, 2002). The number of fast food and full-service restaurants are taken from the 1982, 1987, 1992, and 1997 Census of Retail Trade. For other years not in the Census of Retail Trade, these variables are obtained by extrapolations of logarithmic time trends. Cigarette prices, fast-food prices, food price, and beer prices are constructed from quarterly food prices in the American Chamber of Commerce Researchers Association (ACCRA). Because the NLSY79 has collected data yearly until 1994, and biennially after 1994, quarterly data in the ACCRA was averaged by year. Both of the NLSY79 and ACCRA has random selection of county for each state, which restricts the unit of those variables from the ACCRA to the state level. In the final analysis, fast food price, beer price, and per capita sales in all types of restaurants were chosen among the state-level instruments after specification tests to check the quality of those potential instruments. (specifically, beer price and fast food price for men, and beer price and per capita sales in all types of restaurants for women were chosen following the first stage explanatory power with the assumption that beer price is the exactly identifying instrument).

Second, two individual-level variables were also explored as potential instruments including five-year lags of the respondents' BMI and siblings' BMI. Both of those variables are available in the NLSY data. One caveat for using those two individual-level instruments is that it decreases sample size significantly because siblings' BMI are available only for the respondents whose siblings also participated in the survey. Likewise, using five-year lags of the respondents' BMI restricted sample size. Even with that caveat, those two individual-level instruments passed the exclusion restriction test with the assumption of siblings' BMI as the exactly identify instrument, as well as have the reasonable power of the explanation for the respondents' BMI.

The state- and individual-level instruments define a different group of marginal sample. That is, the state-level instruments only identify a marginal sample who change their behavioral choices affecting body weight by the change in beer prices, fast food prices, or access to all types of restaurants. The individual-level instruments, particularly siblings' BMI identify a marginal sample with siblings. Siblings' BMI provides mixed source of nature and nurture effect on BMI, assuming that siblings share some genes predisposing obesity (in case of full-siblings) as well as they have common upbringings. Therefore, using siblings' BMI as an instrument identify a marginal sample who change their behavioral choices by the change in the genes predisposing obesity or their upbringings. This study also explores a combination of the three state-level instruments (beer prices, fast food prices, and per capital sales in all types of restaurants) with one of the individual-level instruments (siblings' BMI). The combined set of instruments (hereafter combined instruments) will allow for expanding the marginal sample for the individual-level instruments, as well as strengthening the state-level instruments in terms of first stage explanatory power.

The main equation in this study is over-identified. The over-identification is a strong feature of this study. The major restriction of the just-identifying instrument is that the validity of the instruments can not be tested. With over-identifying instruments, the exclusion restriction of the instruments from the main equation can be verified with specification tests with the important assumption of valid exactly identifying instrument. One of the caveats of the two-stage instrumental variables techniques is that it often produces much larger standard errors than OLS with weak instruments that are not strongly correlated with the endogenous explanatory variable of interest (Greene, 2000). Thus, adding more instruments in the first stage can also alleviate such problems by increasing the explanatory power.

Results from the first-stage individual fixed-effects model show that the individual-level instruments have much more explanatory power than the state-level instruments in terms of marginal R^2. An increase in R^2 by adding the individual-level instruments was 0.124 for men and 0.235 for women, while it was 0.003 and 0.001 for men and women, respectively, for the state-level instruments. For the combined instruments, the marginal R^2 was 0.085 for men, and 0.065 for women. The *F*-statistics for testing the null hypothesis that the first-stage coefficients on the instruments are jointly equal to zero exceeds the minimum *F* statistics of 10 suggested by Staiger and Stock (1997) for all sets of instruments. For the combined instruments, *F* statistics are larger than 20 for both genders (see Table 5.1).

The null hypothesis that the first-stage coefficients on the instruments are jointly equal to zero is rejected at the 1% level for all three sets of instruments (state-level, individual-level and combined instruments) for both genders. These results seem to confirm that both the individual and state-level instruments for this study are strong instruments for the BMI model. Full results of the first-stage individual fixed-effects model are described in Table 5.2 for three sets of instruments.

Across all measures of labor market outcomes, the OLS regression, as well as the individual fixed-effects model, was used to predict the effect of obesity on labor market outcomes. For the probability of employment, the probit model also was estimated because the Inverse Mills' Ratio from the probit model for the probability of employment was used for hourly wages and occupations models to control for the selection into the labor force.

First, the exogeneity of three BMI splines was tested by including both the actual BMI splines and the predicted error term in the second-stage estimation (Bollen, Guilkey, and Mroz, 1995). The statistical significance of the predicted error term tests the null hypothesis of the exogeneity of BMI. Results of the test of exogeneity of BMI splines are mixed for all three labor market outcomes, and the exogeneity of BMI splines are not always rejected.

However, failure to reject the null hypothesis of the exogeneity of BMI splines implies that the data does not have enough information to reject the null hypothesis, and it does not conclude that BMI is exogenous in some specifications. Therefore, the 2SRI model was explored for all labor outcome measures for both genders in the final analyses. The simple OLS regression and the individual fixed-effects model results were also shown together for the purpose of comparisons in the next chapter.

Second, the exclusion restriction of the over-identifying instruments was tested by the Lagrange Multiplier (LM) test, which examines whether all instruments are jointly excluded from the second-stage labor market outcomes equation. Because the null hypothesis to be tested is that the over-identifying instruments in each set are all valid, rejection of the null hypothesis implies that some of the over-identifying instruments are invalidly excluded from the second-stage equation. Nonetheless, the condition for exclusion restrictions of the other over-identifying instruments should be built on the assumption that the exactly identified instrument is validly excluded from the labor market outcomes equations (Greene, 2000). Beer price and siblings' BMI are assumed to be the exactly identified instrument in the individual fixed-effects model for the state (and combined) and individual instruments, respectively. Siblings' BMI is assumed to be the exactly identifying instrument for the combined set of instruments.

The test of the exclusion restriction did not reject the null hypothesis that the over-identifying individual-level instrument is validly excluded from the second-stage equations for all three outcomes measures for both genders. For state-level instruments, the null hypothesis of the over-identifying valid instruments was rejected for the OLS regression model on the probability of having occupations requiring social interaction with customers or colleagues for women, and barely failed to reject the same null hypothesis for the individual fixed-effects model on hourly wages for men (see Table 5.3). For the combined instruments, the validity of the over-identifying instruments was rejected for the OLS regression model for both genders for the probability of employment and log hourly wages. For the probability of having occupations requiring social interactions, the OLS regression for women rejected the validity of the over-identifying combined instruments. In the final analysis, the 2SRI model was applied to the specifications supporting the validity of the over-identifying instruments, and the results were compared to the other results from the specifications failing the validity of the over-identifying instruments, the simple OLS regression, or the simple individual fixed-effects model.

Table 5.1 Strength of instruments

Instruments	Gender	N	R^2			F^d
			With IV	Without IV	Marginal	
Individual level[a]	Men	7665	0.16	0.04	0.12	14.57***
	Women	5534	0.25	0.01	0.23	22.54***
State level[b]	Men	43017	0.083	0.080	0.003	13.81***
	Women	39175	0.06	0.06	0.00	12.52***
Combined[c]	Men	19732	0.112	0.085	0.028	20.99***
	Women	15996	0.0912	0.065	0.027	22.41***

Notes:

a. Individual level instruments include siblings' BMI and five year lags of the respondents' BMI.s
b. State-level instruments include fast food and beer price for men and beer price and sales in full-service and limited service restaurants in $1,000 per 100 state populations for women.
c. Combined instruments include siblings' BMI, fast food and beer price, and sales in full-service and limited service restaurants in $1,000 per 100 state populations for both genders.
d. Null hypothesis for this test is that the coefficients of the instruments in the first stage are jointly equal to zero.
e. P value < 0.01: ***

Table 5.2 Results from first-stage equation

First stage	Individual-level IV		State-level IV		Combined IV	
	Men	Women	Men	Women	Men	Women
Individual-level instruments						
Five year lags of BMI	0.055***	0.129***				
	-0.015	-0.022				
Siblings' BMI	0.043***	0.060***			0.051***	0.080***
	-0.011	-0.019			0.007	0.009
State-level instruments						
Fast food price			0.079**		0.188**	-0.114
			-0.036		0.078	0.109
Beer price			0.164***	-0.244***	0.272***	-0.181*
			-0.05	-0.062	0.07[3]	0.104
Sales in restaurants[2]				0.807***	0.267	0.910**
				-0.22	0.301	0.377
Demographic variables						
Age	0.225	0.256***	-0.026***	0.115***	0.192***	0.281***
	-0.007	-0.011	-0.008	-0.009	0.016	0.022
Education level						
College	-0.256	0.204	0.083	0.104	-0.049	-0.017
	-0.244	-0.338	-0.074	-0.072	0.102	0.129
> College	0.042	1.002**	-0.216**	-0.203**	-0.324***	-0.208
	-0.33	-0.405	-0.088	-0.089	0.121	0.154
Number of observations	7,665	5,534	43,017	39,175	19,732	15,996

Notes:
1. Other covariates include three dummies representing U.S. regions, and year dummies.
2. Unit of analysis is person-years.
3. Sales in full-service and limited service restaurants for women.
4. Standard errors are in parentheses.
5. P value < 0.1: *, P value < 0.05: **, P value < 0.01: ***

Table 5.3 Specification tests results (*P*-values) for instruments[1]

Models	Gender	Individual-level IV[2] Test of Exogeneity[5]	Individual-level IV[2] Test of Exclusion Restriction[6]	State-level IV[3] Test of Exogeneity[5]	State-level IV[3] Test of Exclusion Restriction[6]	Combined IV[4] Test of Exogeneity[5]	Combined IV[4] Test of Exclusion Restriction[6]
Pr (employment)							
Logit	Men	0.681	0.921	0.002	0.515	0.620	0.109
	Women	0.024	0.37	0.483	0.806	0.522	0.135
OLS	Men	0.754	0.765	0.003	0.279	0.731	0.001
	Women	0.025	0.243	0.521	0.64	0.537	0.004
FE	Men	0.053	0.42	0.009	0.866	0.630	0.971
	Women	0.022	0.214	0.002	0.46	0.216	0.166
Hourly wages							
OLS	Men	0.696	0.862	0.448	0.136	0.158	0.000
	Women	0.000	0.821	0.248	0.526	0.018	0.001
FE	Men	0.581	0.114	0.669	0.068	0.382	0.321
	Women	0.169	0.705	0.084	0.638	0.061	0.652
Pr (occupations)							
OLS	Men	0.983	0.254	0.13	0.467	0.982	0.122
	Women	0.654	0.688	0.275	0.014	0.326	0.044
FE	Men	0.987	0.541	0.653	0.704	0.652	0.716
	Women	0.749	0.383	0.147	0.168	0.058	0.432

Notes:

1. P-values are reported.
2. Individual-level instruments include siblings' BMI and the five years lag of the respondents' BMI.
3. State-level instruments include fast food and beer price for men and beer price and sales in full-service and limited service restaurants in \$1,000 per 100 state populations for women.
4. Combined instruments include siblings' BMI, fast food and beer price, and sales in full-service and limited service restaurants in \$1,000 per 100 state populations for both genders.
5. Null hypothesis is that the BMI splines are exogenous.
6. Null hypothesis is that the over-identifying instruments are excluded from the main second stage equations. Test statistics is χ^2 with degrees of freedom of one.

CHAPTER VI: RESULTS

Overview

This chapter presents the main results on how obesity (as measured with BMI) affects labor market outcomes, which are measured three ways: 1) the probability of being employed; 2) the probability of having occupations where slimness potentially reward; and, 3) hourly wages. For each outcome, the results are reported for two main models including: 1) the 2SRI; and 2) the 2SRI in conjunction with the individual fixed-effects model. I also display results for the simple OLS regression and the simple individual fixed-effects model at the same table for a comparison purpose.

All models were estimated for three sets of instruments (individual- and state level, and combined), and by gender. Separate estimation by age group was performed to find any different effect of obesity at a different point over the life cycle.

For hourly wages only, two more models are additionally estimated. First, the Arellano-Bond model was performed to recover any dynamic underlying relation of BMI with hourly wages. Second, separate models were run by occupations where social interactions with customers or colleagues are required versus other occupations.

In general, my econometric models support that BMI has a negative effect on labor market outcomes, particularly, hourly wages on the current or most recent job for women.

Employment

An increase in BMI has a significant effect on the likelihood of employment for both men and women. However, the direction of the effect confirms the hypothesis — that an increase in weight would decrease the probability of employment — only for men, but not for women.

Men

For men, the direction of the effect of BMI on the probability of employment is overall negative for the sample with the individual-level instruments (hereafter called the individual IV sample), the sample with the state-level instrument (hereafter called the whole sample), and the sample with the combined instruments (hereafter called the combined sample). Across BMI splines, the negative effect remains almost the same size. Thus, the results overall support the research hypothesis that an increase in BMI has a penalty for the probability of employment.

When the individual fixed effects were dropped from the 2SRI estimation, BMI in all three segments is estimated to significantly adversely affect the probability of employment at the 10% level in the individual IV and the whole sample. The magnitude of the effect is larger for the whole sample using the state-level instruments than the individual IV sample. For the individual IV sample, the 2SRI model in conjunction with the individual fixed-effects model shows that a one-unit increase in BMI decrease the likelihood of employment around 5 percentage points on average across three BMI splines. The magnitude of the negative effect is around 9 percentage points for the whole sample. For the combined sample, the size of the estimates for the BMI splines is reduced to extensively to around 0.7 percentage points, and statistically insignificant (see Table 6.1).

The variations in the size of the effect of BMI splines on the probability of employment confirm that the different groups of the marginal sample are identified by those three sets of instruments. As discussed in chapter V, the state-level instruments only identify the marginal sample who change their behavioral choices affecting their body weight by the change in beer prices, fast food prices, or access to all types of restaurants in the state-level. The individual-level instruments, particularly siblings' BMI identify a marginal sample with siblings. Siblings' BMI are likely to explain the respondents' BMI by either inherited genetic information predisposing obesity or shared upbringings. Therefore, using siblings' BMI as an instrument identify a marginal sample who will have different body weight by the change in their genes or their upbringings. The combined instruments identify a combined marginal sample for the individual IV sample and the whole sample.

Other than the difference in the response for body weight according to the change in a value of instruments, two differences among three samples may cause the variations in the size of the estimated effects. First, the sample size is different among three samples. The individual IV sample is only about one-sixth of the whole sample in terms of number of observations because only sample persons whose siblings were also the respondents in the survey were included in the final analysis. In addition, using five-year lags of the respondent's BMI as another individual-level instrument also restricted the sample size. Second, the age distribution largely differs among those samples. The youngest sample person in the individual IV sample is 23 years old, while it is 18 years old in the whole sample.

For the combined sample, the 2SRI model with controlling for the individual fixed effects were estimated by four age groups. As the results, the direction and size of the effect vary by age group. The directions of the effect of BMI splines on the likelihood of employment are negative in the older than 30 age group, while the effects are positive for men aged 18 to 29 years old. For the youngest group aged 18-24 years old, the absolute size of the effect is much larger than the size for other age groups and statistically significant (one unit increase of body weight within individual increases the probability of employment in the range of 13 to 17 percentage points). For the individual IV sample, the overall negative effect of BMI on the likelihood of employment appears in the older than 30 age group. For this sample, individual fixed effects are not controlled due to small sample size in each age group (see Table 6.2)

Women

On the contrary, for women, an increase in BMI generally raises the likelihood of employment, which contradicts the suggested hypothesis that an increase of BMI decreases the likelihood of employment.

The size of the effect for the whole sample is large compared to the individual IV sample and the combined sample like for men. The positive effect ranges between 3.0 to 5.0 percentage points for the individual IV sample across BMI splines, while it increases up to 11.9 percentage points for the whole sample using the state-level instruments. The size is the smallest for the combined sample, which ranges between 1.6 to 2.5 percentage points across BMI splines (see Table 6.1).

When the 2SRI with the individual-level instruments is separately run by age group for the individual IV sample, a strong and significant positive effect is found for the 30-34 years old group (4.7 percentage points increase in the likelihood of employment by a one-unit increase of BMI in the underweight or normal weight range). This positive effect becomes stronger and larger in both the overweight and obese range. For the combined sample, results for the 2SRI controlling for the individual fixed effects have the positive direction for the sample younger than 30 years old. The direction turns negative for the others (30 and older) (see Table 6.2).

As discussed in the chapter IV, the primary models for this study estimate a total effect of BMI on the probability of employment, and thus, represent a reduced form model for the omitted variables including the marital status and the number of children. Different directions of the effect of BMI on the probability of employment may be related to those life time events for women. A woman may leave the job market after marriage or after a child-birth, but reenter the workforce after child-bearing or divorce. Results for two sensitivity analyses for the combined sample do not support the effect of those life time events on the positive causality of BMI on the probability of employment. A sensitivity analysis for the sample women who have been never pregnant result in all positive direction of the effect in all age groups.

The second sensitivity analysis controls for the marital status and the number of children as additional covariates in the model but treated those two variables as exogenous. Results for the second sensitivity analysis are overall consistent to the primary results from the reduced model. For the combined sample, the direction of the effect is negative for the groups older than 25 years of age for overweight and obese spline, and for the underweight or normal weight spline, the direction is negative for all age groups except for the group aged 25-29 years old. For the individual IV sample, strong positive effect of BMI on the probability of employment is observed for the sample aged 30 and older (4.4 to 5.2

percentage points versus 3.0 to 3.7 percentage points increase in the probability of employment by a unit increase of BMI for 30-34 years of age group and 35-39 years of age group, respectively).

It is not clear why the results for women are contradictory to the suggested hypothesis. One probable reason may include the preciseness of the measurement of the dependent variable, which is the probability of employment. More specific categorization of the employment based on some detailed characteristics of the employment, e.g., fringe benefits may support the research hypothesis of penalty for labor market at the extensive margin. Second, the positive effect of BMI on the probability of employment is total effect. Various indirect factors are likely to affect the positive causal effect, including a marriage, child-birth or child-rearing. Even though the sensitivity analysis controlling for the marital status and the number of children does not support the effect of those life time events, the potential endogeneity of the marital status and the number of children are ignored in the sensitivity analysis. Therefore, results from the sensitivity analysis should be considered with a caution, and the potential indirect effect of a marriage or child-bearing (or rearing) should not be excluded.

Occupation requiring social interactions with customers or colleagues

In general, not only the size but also the direction of the effect changes by additionally controlling for the individual fixed effects in the 2SRI model. This seems to imply that some time-invariant individual characteristics, e.g., being extroverted, which is potentially correlated with the occupation choice, may still remain in the error term for the 2SRI model.

An increase in BMI negatively affects the probability of having occupations requiring social interactions for women for the 2SRI model with controlling for the individual fixed effects. For men, the direction differs by the marginal samples defined by each set of instruments for the same model.

Men

The direction of the effect varies by the marginal samples. For the individual IV and the whole sample, an increase in BMI across three splines positively affects the likelihood of having an occupation where social interactions are required. This does not confirm the proposed hypothesis that an increase in BMI would adversely affect the likelihood of having an occupation where slimness potentially rewards by requiring social interactions. However, for the combined sample, the direction is negative, which conforms to the hypothesized negative effect.

The estimated absolute size of the effect becomes larger for the whole sample with state-level instruments compared to the individual IV sample or the combined sample, which was also observed for the probability of employment (see Table 6.3). Potential reasons for this enlargement would be the same as for the probability of employment.

The direction and size of the effect vary by age group. For the individual IV sample, the penalizing effect only appears until the 20s, and after this age, an increase of BMI turns to have a rewarding effect for the occupation choice requiring social interactions. For the younger sample aged between 18 to 29 years old, an increase in BMI has a penalty for the probability of having occupations requiring social interactions in the range between 1.4 and 5.5 percentage points across BMI splined group. The magnitude of the penalty is larger for the group aged 25-29 years old than the group aged 18-24 years old. However, for the sample older than 29 years old, the effect is positive in the range of 1.4 to 13.5 percentage points. The positive effect gets larger and significant at the 10% level for the older group (aged 35-40 years old) compared to the other group (aged 30-34) (see Table 6.4).

The opposite direction of the effect for the combined sample across all ages compared to the other two marginal samples seems to be driven by the youngest and the oldest age group, which shows opposite direction between the combined sample and the individual IV sample. The absolute size of the effect increase as the age group gets older, even though the direction is not clear for an interpretation.

Women

The estimated directions for the coefficients of BMI splined group support the hypothesis that an increase in weight (as measured with BMI) would decrease the likelihood of sorting into occupations where social interactions are required (see Table 6.3).

The increase of the size of the effect for the whole sample is also observed compared to the individual IV sample or the combined sample. A unit increase of BMI decreases the probability of having occupations by 4.2 to 5.5 percentage points for the combined sample, while the magnitude of the effect increases around 1.5 times for the whole sample. Not only the size differs, but also the direction differs between the individual IV sample and the combined sample for the overweight and obese splines. This discre

Separate estimation of the model by age group confirms different marginal effect between the individual IV sample and the combined sample. The direction is all opposite in each age group between the two samples except the youngest group (aged 18-24 years old) where the effect is consistently negative for both samples. For the combined sample, the direction is negative in all age groups except the group aged 30-34 years old. No clear pattern is found in the absolute size of the effect for the combined sample (see Table 6.4).

Hourly wage

An increase in BMI negatively affects hourly wages for women, which supports the proposed hypothesis. The direction of the effect is unclear for men, even though it is generally positive.

The effect of BMI on hourly wages varies by age groups across three BMI splines. For women, the wage penalty for gaining weight remains in all age groups, and the size of the wage penalty increases as the age group becomes older until age of 34 years old. For men, the direction is negative only for age group younger than 35 years old. This confirms the hypothesis that age plays an important role in the causal association of BMI with wages.

The size or the direction of the effect varies by the types of occupations when the occupations were classified by the requirement of social interactions. Thus, the proposed hypothesis — that an increase in weight would differentially affect wages at occupations where social interactions with customers or colleagues are required from other occupations — is supported.

<u>Men</u>

Gaining weight does not seem to penalize hourly wages for men for the individual IV and the combined sample. A one-unit increase in BMI raises hourly wages by 1.3 to 5.3% for the individual IV sample, and 2.1-3.3% for the combined sample. However, the direction of the effect of BMI on hourly wages is opposite for the whole sample, for which the state-level instruments were used to identify the causal effect. An increase in BMI negatively affects hourly wages in the range between 4.8% (the normal weight range) and 5.8% (the obese range) for the whole sample (see Table 6.5).

Controlling for the individual fixed effects changes the direction of the effect for the overweight or the obese range of BMI. This may imply that the unobserved time-consistent individual heterogeneity still remains in the error term in the 2SRI model.

Different age groups have different effects of an increase in BMI on hourly wages, and even the direction as well as the size of the effect varies by age group. When the 2SRI is separately run by age group for the individual IV sample, the direction of the effect is positive only for the youngest (aged 18-24) and oldest (aged 35-40) group. The effect is negative for the mid age group (aged 25-34) (see Table 6.6). This implies that the causal association of an increase in body weight with wages is not linear over the life-cycle. The wage penalty for gaining weight seems to be adjusted by the conventional wisdom that people tend to gain weight as they get older. Results for the combined IV sample are generally consistent in that the wage penalty for gaining body weight disappears for the oldest group.

When the models are estimated separately by occupation group (by requirement of social interactions), the direction of the effects are opposite between two occupation groups for all three marginal samples. For the whole and the combined sample, the wage penalty for gaining body weight within individual appears only for occupations requiring social interactions. The size of the penalty effect ranges 8.3 to 9.0% for the whole sample, while the magnitude is much smaller for the combined sample (0.5 to 0.9%). For those two samples, an increase in body weight increases wages for occupations without requiring social interactions. For the individual IV sample, the directions of the effect in both occupation groups are opposite to other two samples (see Table 6.7a).

Women

The wage penalty for an increase in BMI is clear for women. An increase in BMI by one unit decreases hourly wages on the current or the most recent jobs for women by around 8% (for the whole sample) to 4% (for the combined sample). The magnitude of the negative effects is slightly larger for the overweight and obese BMI range than the underweight or normal weight range (see Table 6.5).

The adverse effect of an increase in BMI on hourly wages remains consistent in all age group. Overall, the magnitude of the negative effect starts to decrease pass the age of 30 (see Table 6.6).

Dividing the sample by types of occupations (by the requirement of social interactions) does not change the direction of the effect. For both types of occupations, an increase in BMI decreases hourly wages. Overall, the size of the effect is larger for the occupations where social interactions are not required. This supports the hypothesis that suggests a different effect of the weight gain on hourly wages by types of occupations. This also seems to imply that gaining weight harms wages via different pathways from social interactions with customers or colleagues. Given this study's finding that weight gain adversely affects the probability of having occupations requiring social interactions, overweight or obese women in those occupations may have higher job skills than the overweight or obese women who fail to enter those occupations. Thus, they may not be penalized at the internal margin once they enter those jobs. The difference in the size of the effect between the two occupation groups is much larger for the whole sample using the state-level instruments than the individual IV sample. The extent of negative effect of BMI on hourly wages was rather implausibly high (over 20%) on average for women with occupations where social interactions are not required for the whole sample (see Table 6.7b).

Table 6.1 Effect of BMI on the probability of employment using the linear probability model

Independent variable	Individual IV sample[4]				Whole sample[5]				Combined sample[6]			
	OLS	FE	IV	IV + FE	OLS	FE	IV	IV + FE	OLS	FE	IV[7]	IV + FE
Men												
BMI < 25	0.008**	-0.007	0.002	-0.064**	0.005**	0.000	-0.128***	-0.091***	0.008***	0.003	0.002	-0.006
	(0.004)	(0.006)	(0.020)	(0.030)	(0.002)	(0.002)	(0.045)	(0.035)	(0.003)	(0.003)	(0.018)	(0.020)
25 ≤ BMI < 30	-0.002	-0.001	-0.008	-0.057*	0.001	0.001	-0.131***	-0.089***	-0.001	0.000	-0.007	-0.009
	(0.003)	(0.005)	(0.020)	(0.030)	(0.002)	(0.002)	(0.045)	(0.035)	(0.003)	(0.003)	(0.018)	(0.020)
BMI ≥ 30	0.002	0.000	-0.004	-0.056*	-0.003	0.001	-0.137***	-0.089***	0.000	0.004	-0.007	-0.005
	(0.003)	(0.004)	(0.020)	(0.029)	(0.002)	(0.002)	(0.045)	(0.035)	(0.003)	(0.003)	(0.018)	(0.020)
N (person-years)	7,665	7,665	7,665	7,665	47,435	47,435	47,435	47,435	19,732	19,732	19,732	19,732
Women												
BMI < 25	-0.003	-0.011	0.030**	0.039*	0.002	-0.005***	0.034	0.115***	0.003	-0.001	0.014	0.021
	(0.004)	(0.007)	(0.015)	(0.023)	(0.002)	(0.002)	(0.049)	(0.040)	(0.003)	(0.003)	(0.018)	(0.018)
25 ≤ BMI < 30	-0.003	0.002	0.030*	0.050*	-0.007**	-0.005***	0.025	0.118***	-0.006	-0.005	0.006	0.016
	(0.006)	(0.007)	(0.015)	(0.022)	(0.003)	(0.002)	(0.049)	(0.040)	(0.005)	(0.003)	(0.019)	(0.018)
BMI ≥ 30	0.005	0.003	0.038**	0.049**	-0.001	-0.002	0.030	0.119***	0.003	0.003	0.014	0.025
	(0.004)	(0.006)	(0.015)	(0.021)	(0.002)	(0.002)	(0.049)	(0.040)	(0.003)	(0.003)	(0.019)	(0.018)
N (person-years)	5,534	5,534	5,534	5,534	44,000	44,000	44,000	44,000	15,996	15,996	15,996	15,996

Notes:

1. Unit of analysis is person-years.
2. Three dummy variables representing the following four age groups are included in every specification: 18-24 years old; 25-29 years old; 30-34 years old; 35-41 years old.
3. Standard errors are in parentheses.
4. Individual-level instruments include five years lag of BMI and siblings' BMI for both men and women sample.
5. State-level instruments include cost of beer and cost of fast food for men, and cost of beer and sales in restaurants (full-service and limited services) in 1,000 dollars per 100 state populations by states for women.
6. Combined (individual- and state-level) instruments include siblings' BMI, cost of beer, cost of fast food, and sales in restaurants (full-service and limited services) in 1,000 dollars per 100 state populations by states for both men and women sample.
7. For this model, over-identifying instruments did not pass the test of the exclusion restriction for both genders.

Table 6.2 Effect of BMI on the probability of employment by age group

Groups	Independent variable	Age 18-24	Age 25-29	Age 30-34	Age 35-40
Individual IV sample[3]: OLS					
Men	BMI < 25	0.057	0.012	-0.012	0.002
		(0.051)	(0.029)	(0.030)	(0.030)
	25 ≤ BMI < 30	0.052	0.011	-0.018	-0.020
		(0.050)	(0.030)	(0.030)	(0.029)
	BMI ≥ 30	0.069	0.017	-0.013	-0.016
		(0.048)	(0.028)	(0.030)	(0.031)
N (person-years)		503	2,509	3,103	1,550
Women	BMI < 25	0.061	0.021	0.047*	-0.009
		(0.092)	(0.024)	(0.024)	(0.035)
	25 ≤ BMI < 30	0.054	0.024	0.043*	-0.016
		(0.089)	(0.027)	(0.023)	(0.033)
	BMI ≥ 30	0.027	0.023	0.056**	-0.007
		(0.096)	(0.026)	(0.023)	(0.033)
N (person-years)		462	1,719	2,193	1,160
Combined sample[4]: FE					
Men	BMI < 25	0.166**	0.053	-0.022	-0.092
		(0.077)	(0.060)	(0.130)	(0.121)
	25 ≤ BMI < 30	0.161**	0.062	-0.023	-0.098
		(0.077)	(0.060)	(0.130)	(0.120)
	BMI ≥ 30	0.129*	0.052	-0.020	-0.092
		(0.077)	(0.060)	(0.130)	(0.119)
N (person-years)		4,790	6,793	5,775	2,374
Women	BMI < 25	0.011	0.024	-0.053	-0.074
		(0.120)	(0.086)	(0.100)	(0.130)
	25 ≤ BMI < 30	0.028	0.005	-0.060	-0.063
		(0.122)	(0.085)	(0.099)	(0.127)
	BMI ≥ 30	0.022	0.012	-0.059	-0.052
		(0.121)	(0.085)	(0.100)	(0.128)
N (person-years)		3,900	5,170	4,867	2,059

Notes:
1. Unit of analysis is person-years.
2. Standard errors are in parentheses.
3. Individual-level instruments include five years lag of BMI and siblings' BMI for both men and women sample.
4. Combined (individual- and state-level) instruments include siblings' BMI, cost of beer, cost of fast food, and sales in restaurants (full-service and limited services) in 1,000 dollars per 100 state populations by states for both men and women sample.

Table 6.3 Effect of BMI on the probability of having occupations where social interaction is required

Independent variable	Individual IV sample[4]				Whole sample[5]				Combined sample[6]			
	OLS	FE	IV	IV + FE	OLS	FE	IV[7]	IV + FE	OLS	FE	IV[7]	IV + FE
Men												
BMI < 25	0.002	0.009	0.010	0.002	0.001	0.007 **	-0.082	0.027	0.000	0.009*	-0.001	-0.006
	(0.007)	(0.011)	(0.032)	(0.052)	(0.003)	(0.003)	(0.053)	(0.052)	(0.006)	(0.005)	(0.028)	(0.031)
25 ≤ BMI < 30	0.006	-0.004	0.019	-0.004	0.001	-0.002	-0.081	0.022	0.001	-0.001	0.001	-0.015
	(0.006)	(0.008)	(0.032)	(0.051)	(0.003)	(0.003)	(0.054)	(0.052)	(0.004)	(0.004)	(0.028)	(0.031)
BMI ≥ 30	0.004	0.008	0.014	0.007	0.001	0.001	-0.079	0.031	0.003	0.013***	0.002	-0.001
	(0.006)	(0.009)	(0.031)	(0.050)	(0.003)	(0.003)	(0.054)	(0.052)	(0.005)	(0.005)	(0.028)	(0.031)
N (person-years)	6,632	6,632	6,632	6,632	39,021	39,021	39,021	39,021	16,294	16,294	16,294	16,294
Women												
BMI < 25	0.005	-0.009	0.002	-0.001	-0.002	-0.001	0.049	-0.080	-0.002	0.000	0.017	-0.049*
	(0.005)	(0.011)	(0.021)	(0.038)	(0.002)	(0.003)	(0.046)	(0.055)	(0.004)	(0.005)	(0.020)	(0.026)
25 ≤ BMI < 30	0.005	-0.010	0.002	0.002	-0.001	-0.004	0.046	-0.081	0.000	-0.005	0.018	-0.055**
	(0.007)	(0.012)	(0.021)	(0.036)	(0.003)	(0.005)	(0.043)	(0.052)	(0.006)	(0.005)	(0.020)	(0.026)
BMI ≥ 30	-0.004	-0.001	-0.005	0.009	-0.001	0.002	0.049	-0.075	0.003	0.006	0.022	-0.042
	(0.005)	(0.009)	(0.022)	(0.037)	(0.003)	(0.003)	(0.045)	(0.054)	(0.005)	(0.004)	(0.020)	(0.026)
N (person-years)	4,208	4,208	4,208	4,208	30,871	30,871	30,871	30,871	11,522	11,522	11,522	11,522

Notes:

1. Unit of analysis is person-years.
2. Three dummy variables representing the following four age groups are included in every specification: 18-24 years old; 25-29 years old; 30-34 years old; 35-41 years old.
3. Standard errors are in parentheses.
4. Individual-level instruments include five years lag of BMI and siblings' BMI for both men and women sample.
5. State-level instruments include cost of beer and cost of fast food for men, and cost of beer and sales in restaurants (full-service and limited services) in 1,000 dollars per 100 state populations by states for women.
6. Combined (individual- and state-level) instruments include siblings' BMI, cost of beer, cost of fast food, and sales in restaurants (full-service and limited services) in 1,000 dollars per 100 state populations by states for both men and women sample.
7. For this model, over-identifying instruments did not pass the test of the exclusion restriction for women.

Table 6.4 Effect of BMI on the probability of having occupations where social interaction is required by age group

Groups	Independent variable	Age 18-24	Age 25-29	Age 30-34	Age 35-40
The individual IV sample[3]: OLS					
Men	BMI < 25	-0.026	-0.039	0.015	0.083
		(0.067)	(0.052)	(0.048)	(0.063)
	25 ≤ BMI < 30	-0.014	-0.055	0.014	0.135 **
		(0.068)	(0.052)	(0.048)	(0.061)
	BMI ≥ 30	-0.024	-0.054	0.021	0.105 *
		(0.073)	(0.053)	(0.048)	(0.060)
N (person-years)		407	2,162	2,689	1374
Women	BMI < 25	-0.041	0.018	-0.026	0.015
		(0.113)	(0.039)	(0.033)	(0.044)
	25 ≤ BMI < 30	-0.038	0.014	-0.038	0.039
		(0.114)	(0.04)	(0.034)	(0.043)
	BMI ≥ 30	-0.094	0.011	-0.036	0.013
		(0.120)	(0.043)	(0.035)	(0.045)
N (person-years)					
Combined sample[4]: FE					
Men	BMI < 25	0.026	-0.075	0.199	-0.227
		(0.108)	(0.103)	(0.242)	(0.149)
	25 ≤ BMI < 30	0.000	-0.061	0.175	-0.224
		(0.108)	(0.102)	(0.244)	(0.148)
	BMI ≥ 30	-0.015	-0.074	0.185	-0.191
		(0.108)	(0.103)	(0.242)	(0.151)
N (person-years)		3,459	5,809	4,939	2,087
Women	BMI < 25	-0.069	-0.182	0.020	-0.155
		(0.150)	(0.132)	(0.163)	(0.136)
	25 ≤ BMI < 30	-0.037	-0.177	0.015	-0.127
		(0.149)	(0.131)	(0.162)	(0.134)
	BMI ≥ 30	-0.070	-0.167	0.014	-0.135
		(0.149)	(0.132)	(0.163)	(0.135)
N (person-years)		2,614	3,784	3,596	1,528

Notes:
1. Unit of analysis is person-years.
2. Standard errors are in parentheses.
3. Individual-level instruments include five years lag of BMI and siblings' BMI for both men and women sample.
4. Combined (individual- and state-level) instruments include siblings' BMI, cost of beer, cost of fast food, and sales in restaurants (full-service and limited services) in 1,000 dollars per 100 state populations by states for both men and women sample.

Table 6.5 Effect of BMI on the hourly wage

Independent variable	Individual IV sample[4]				Whole sample[5]				Combined sample[6]			
	OLS	FE	IV	IV + FE	OLS	FE	IV	IV + FE	OLS	FE	IV[7]	IV + FE
Men												
BMI < 25	0.026***	0.029***	0.012	0.053	0.024***	0.005**	-0.047	-0.048	0.030**	0.013***	-0.026	0.033
	(0.008)	(0.010)	(0.041)	(0.045)	(0.004)	(0.003)	(0.091)	(0.061)	(0.007)	(0.005)	(0.036)	(0.028)
25 ≤ BMI < 30	-0.006	-0.011	-0.021	0.013	-0.003	0.000	0.072	-0.052	-0.005	0.004	-0.052	0.028
	(0.007)	(0.008)	(0.040)	(0.044)	(0.004)	(0.002)	(0.091)	(0.061)	(0.006)	(0.004)	(0.035)	(0.028)
BMI ≥ 30	-0.011	-0.008	-0.027	0.016	-0.009**	-0.002	0.079	-0.058	-0.012*	-0.003	-0.060*	0.021
	(0.008)	(0.007)	(0.041)	(0.043)	(0.004)	(0.002)	(0.092)	(0.061)	(0.007)	(0.004)	(0.036)	(0.028)
N (person-years)	6,632	6,632	6,632	6,632	39,021	39,021	39,021	39,021	16,294	16,294	16,294	16,294
Women												
BMI < 25	0.000	0.006	-0.096***	-0.036	-0.006*	0.001	-0.086	-0.081	-0.006	0.002	-0.064**	-0.039*
	(0.006)	(0.010)	(0.025)	(0.031)	(0.003)	(0.002)	(0.070)	(0.050)	(0.005)	(0.004)	(0.025)	(0.022)
25 ≤ BMI < 30	-0.025***	-0.002	-0.117***	-0.041	-0.011***	-0.005*	-0.088	-0.092*	-0.018***	-0.012***	-0.073***	-0.049**
	(0.008)	(0.007)	(0.025)	(0.029)	(0.004)	(0.003)	(0.069)	(0.049)	(0.006)	(0.004)	(0.026)	(0.021)
BMI ≥ 30	0.008	0.001	-0.102**	-0.038	-0.006**	-0.003	-0.085	-0.089*	0.002	-0.001	-0.059**	-0.042*
	(0.005)	(0.005)	(0.027)	(0.029)	(0.003)	(0.002)	(0.069)	(0.049)	(0.004)	(0.004)	(0.026)	(0.022)
N (person-years)	4,208	4,208	4,208	4,208	30,871	30,871	30,871	30,871	11,522	11,522	11,522	11,522

Notes:

1. Unit of analysis is person-years.
2. Three dummy variables representing the following four age groups are included in every specification: 18-24 years old; 25-29 years old; 30-34 years old; 35-41 years old.
3. Standard errors are in parentheses.
4. Individual-level instruments include five years lag of BMI and siblings' BMI for both men and women sample.
5. State-level instruments include cost of beer and cost of fast food for men, and cost of beer and sales in restaurants (full-service and limited services) in 1,000 dollars per 100 state populations by states for women.
6. Combined (individual- and state-level) instruments include siblings' BMI, cost of beer, cost of fast food, and sales in restaurants (full-service and limited services) in 1,000 dollars per 100 state populations by states for both men and women sample.
7. For this model, over-identifying instruments did not pass the test of the exclusion restriction for both genders.

Table 6.6 Effect of BMI on the hourly wage by age group

Groups	Independent variable	Age 18-24	Age 25-29	Age 30-34	Age 35-40
The individual IV sample[3]: OLS					
Men	BMI < 25	0.052	-0.047	0.026	0.062
		(0.117)	(0.052)	(0.058)	(0.089)
	25 ≤ BMI < 30	0.068	-0.050	-0.044	0.070
		(0.129)	(0.051)	(0.056)	(0.083)
	BMI ≥ 30	0.032	-0.072	-0.018	0.037
		(0.116)	(0.053)	(0.057)	(0.084)
N (person-years)		407	2,162	2,689	1374
Women	BMI < 25	-0.073*	-0.040	-0.096**	-0.022
		(0.039)	(0.031)	(0.044)	(0.048)
	25 ≤ BMI < 30	-0.047	-0.054*	-0.124***	-0.058
		(0.049)	(0.031)	(0.041)	(0.047)
	BMI ≥ 30	-0.063	-0.035	-0.093**	-0.016
		(0.046)	(0.031)	(0.047)	(0.046)
N (person-years)		339	1,304	1,661	904
Combined sample[4]: FE					
Men	BMI < 25	-0.018	-0.043	-0.094	0.025
		(0.086)	(0.079)	(0.216)	(0.186)
	25 ≤ BMI < 30	-0.001	-0.043	-0.103	0.026
		(0.084)	(0.078)	(0.218)	(0.187)
	BMI ≥ 30	-0.018	-0.038	-0.095	0.027
		(0.085)	(0.077)	(0.217)	(0.190)
N (person-years)		3,459	5,809	4,939	2,087
Women	BMI < 25	-0.058	-0.059	-0.014	-0.031
		(0.128)	(0.083)	(0.168)	(0.157)
	25 ≤ BMI < 30	-0.063	-0.070	-0.018	-0.004
		(0.126)	(0.081)	(0.165)	(0.154)
	BMI ≥ 30	-0.052	-0.069	-0.026	-0.027
		(0.126)	(0.082)	(0.168)	(0.156)
N (person-years)		2,614	3,784	3,596	1,528

Notes:
1. Unit of analysis is person-years.
2. Standard errors are in parentheses.
3. Individual-level instruments include five years lag of BMI and siblings' BMI for both men and women sample.
4. Combined (individual- and state-level) instruments include siblings' BMI, cost of beer, cost of fast food, and sales in restaurants (full-service and limited services) in 1,000 dollars per 100 state populations by states for both men and women sample.

Table 6.7a Effect of BMI on hourly wage by occupation group: Men

Independent variable	Individual IV sample[4]				Whole sample[5]				Combined sample[6]			
	OLS	FE	IV	IV + FE	OLS	FE	IV	IV + FE	OLS	FE	IV[7]	IV + FE
Requiring social interactions												
BMI < 25	0.023**	0.035**	0.023	0.078	0.026***	0.011***	-0.016	-0.083	0.031***	0.017**	-0.029	0.010
	(0.010)	(0.016)	(0.050)	(0.076)	(0.005)	(0.004)	(0.145)	(0.120)	(0.008)	(0.007)	(0.038)	(0.037)
25 ≤ BMI < 30	0.006	0.005	0.008	0.038	-0.005	0.002	-0.044	-0.086	-0.002	0.002	-0.061	-0.005
	(0.009)	(0.012)	(0.050)	(0.074)	(0.004)	(0.003)	(0.145)	(0.120)	(0.007)	(0.005)	(0.037)	(0.037)
BMI ≥ 30	-0.012	-0.011	-0.010)	0.040	-0.006	-0.003	-0.045	-0.090	-0.010	-0.001	-0.069*	-0.009
	(0.009)	(0.014)	(0.050)	(0.071)	(0.005)	(0.004)	(0.146)	(0.120)	(0.008)	(0.006)	(0.038)	(0.036)
N (person-years)	3,214	3,214	3,214	3,214	20,345	20,345	20,345	20,345	9,164	9,164	9,164	9,164
Not requiring social interactions												
BMI < 25	0.028**	0.037 *	-0.000	-0.007	0.020***	0.007	-0.074	0.049	0.031***	0.002	-0.043	0.029
	(0.011)	(0.022)	(0.060)	(0.086)	(0.005)	(0.005)	(0.149)	(0.125)	(0.008)	(0.008)	(0.050)	(0.050)
25 ≤ BMI < 30	-0.017*	-0.018	-0.050)	-0.059	-0.001	0.000	-0.096	0.046	-0.010	0.003	-0.083*	0.030
	(0.009)	(0.016)	(0.060)	(0.085)	(0.005)	(0.004)	(0.149)	(0.125)	(0.008)	(0.006)	(0.049)	(0.050)
BMI ≥ 30	-0.011	-0.023 *	-0.050)	-0.040	-0.011**	-0.003	-0.108	0.041	-0.016**	-0.006	-0.089*	0.021
	(0.010)	(0.012)	(0.060)	(0.081)	(0.005)	(0.004)	(0.149)	(0.125)	(0.008)	(0.007)	(0.050)	(0.050)
N (person-years)	3,418	3,418	3,418	3,418	18,676	18,676	18,676	18,676	7,130	7,130	7,130	7,130

Notes:
1. Unit of analysis is person-years.
2. Three dummy variables representing the following four age groups are included in every specification: 18-24 years old; 25-29 years old; 30-34 years old; 35-41 years old.
3. Standard errors are in parentheses.
4. Individual-level instruments include five years lag of BMI and siblings' BMI for both men and women sample.
5. State-level instruments include cost of beer and cost of fast food for men, and cost of beer and sales in restaurants (full-service and limited services) in 1,000 dollars per 100 state populations by states for women.
6. Combined (individual- and state-level) instruments include siblings' BMI, cost of beer, cost of fast food, and sales in restaurants (full-service and limited services) in 1,000 dollars per 100 state populations by states for both men and women sample.
7. For this model, over-identifying instruments did not pass the test of the exclusion restriction for both genders.

Table 6.7b Effect of BMI on hourly wage by occupation group: Women

Independent variable	Individual IV sample[4]				Whole sample[5]				Combined sample[6]			
	OLS	FE	IV	IV + FE	OLS	FE	IV	IV + FE	OLS	FE	IV[7]	IV + FE
Requiring social interactions												
BMI < 25	0.000	0.005	-0.088***	-0.057	-0.005	0.008**	-0.067	0.003	-0.006	0.000	0.011	-0.007
	(0.007)	(0.017)	(0.033)	(0.051)	(0.003)	(0.003)	(0.076)	(0.066)	(0.005)	(0.006)	(0.027)	(0.029)
25 ≤ BMI < 30	-0.027**	0.026	-0.114***	-0.055	-0.010**	-0.026***	-0.071	-0.011	-0.011	-0.005	0.006	-0.012
	(0.011)	(0.022)	(0.032)	(0.048)	(0.005)	(0.006)	(0.075)	(0.065)	(0.008)	(0.006)	(0.027)	(0.029)
BMI ≥ 30	0.007	0.007	-0.091**	-0.061	-0.003	-0.002	-0.064	-0.004	0.002	-0.003	0.02	-0.010
	(0.007)	(0.009)	(0.036)	(0.051)	(0.004)	(0.003)	(0.076)	(0.066)	(0.005)	(0.005)	(0.027)	(0.029)
N (person-years)	2,194	2,194	2,194	2,194	17,658	17,658	17,658	17,658	7,063	7,063	7,063	7,063
Not requiring social interactions												
BMI < 25	-0.002	0.008	-0.105***	-0.015	-0.006	0.005	-0.07	-0.226***	-0.006	0.002	0.002	-0.032
	(0.009)	(0.018)	(0.032)	(0.064)	(0.004)	(0.004)	(0.080)	(0.078)	(0.006)	(0.008)	(0.032)	(0.039)
25 ≤ BMI < 30	-0.024**	-0.016	-0.122***	-0.041	-0.012**	-0.015*	-0.071	-0.241***	-0.029***	-0.014	-0.021	-0.047
	(0.010)	(0.029)	(0.032)	(0.058)	(0.005)	(0.008)	(0.080)	(0.078)	(0.009)	(0.009)	(0.032)	(0.037)
BMI ≥ 30	0.009	-0.005	-0.109***	-0.026	-0.010***	-0.005	-0.071	-0.234***	0.001	-0.007	0.010	-0.040
	(0.006)	(0.012)	(0.034)	(0.057)	(0.004)	(0.004)	(0.080)	(0.078)	(0.006)	(0.007)	(0.032)	(0.038)
N (person-years)	2,014	2,014	2,014	2,014	13,218	13,218	13,218	13,218	4,459	4,459	4,459	4,459

Notes:

1. Unit of analysis is person-years.
2. Three dummy variables representing the following four age groups are included in every specification: 18-24 years old; 25-29 years old; 30-34 years old; 35-41 years old.
3. Standard errors are in parentheses.
4. Individual-level instruments include five years lag of BMI and siblings' BMI for both men and women sample.
5. State-level instruments include cost of beer and cost of fast food for men, and cost of beer and sales in restaurants (full-service and limited services) in 1,000 dollars per 100 state populations by states for women.
6. Combined (individual- and state-level) instruments include siblings' BMI, cost of beer, cost of fast food, and sales in restaurants (full-service and limited services) in 1,000 dollars per 100 state populations by states for both men and women sample.
7. For this model, over-identifying instruments did not pass the test of the exclusion restriction for both genders.

CHAPTER VII: DISCUSSION

The first hypothesis in this study — that an increase in weight within-person would decrease the probability of employment — is supported for men, but not for women. The size, as well as the direction of the effect slightly varies across the BMI splined group for each gender, not to mention by gender overall. The magnitude of the effects is large and statistically significant for men, particularly for the 2SRI model controlling for individual fixed effects.

It is not clear why the effect is contradictory to the suggested hypothesis for women. One probable explanation is that strong life cycle effects, e.g., marriage, child-birth, and child-rearing. A probable scenario is that women leave the job market after a marriage or after a child-birth, but reenter the workforce after a child-bearing or a divorce. Another probable scenario is the penalty of being overweight or obese for outcomes in the marriage market. That is, overweight or obese women may be less likely to be married. Or, they may be more likely to get married with a spouse who have poor labor market outcomes, and thus, those women may need to participate in the labor market (regardless of the quality of employment). Even though two sensitivity analyses (one for restricting sample to the never pregnant women, and the other for controlling for the marital status and the number of children) do not support the role of those life cycle events, a structure model controlling for the potential endogeneity for all covariates will need for precise estimates. Third, the power of instruments may be not strong enough to recover the underlying unbiased relationship between BMI and the probability of employment for women, particularly for the state-level instruments. The partial *R*-squared contribution by the state-level instruments is much smaller compared to the individual-level or combined instruments. However, *F*-statistics for the state-level instruments are larger than 10, which has suggested as the minimum number to obtain the unbiased results by the two-stage instrument variable estimation model (for example, Stagier and Stock, 1997). Recovering the marginal effect of BMI on labor market outcomes at the extensive margin may require much stronger identifying variables than the marginal effect at the intensive margin. Third, even at the extensive margin, different measures of the dependent variable (the probability of employment) would help to find the underlying causality. For example, being overweight or obese may impede the entrance into a high quality job rather than the employment per se. Some examples of those quality factors in a job would be a full-time versus a part-time, paid vacation days, or fringe benefits. Lastly, the enlarged size of the estimated coefficients of BMI splines for the whole sample with the state-level instruments may imply a high correlation between the predicted residuals from the first-stage individual fixed-effects model and other covariates in the second-stage equation, even though simple correlation coefficients do not support that implication.

The second hypothesis — that an increase in weight would decrease the probability of having occupations where social interactions with customers or colleagues are required — is supported for women, but not for men. The results are overall statistically insignificant.

In general, the results are consistent with the different conventional public conception for the ideal body weight for men and women. For men, being slim may be sometimes interchangeable with being non-masculine (this is particularly probable when the body weight is measured with BMI), which may not be appreciated all the time. Non-significant results may be related to how to define the occupations where social interactions are required. As described earlier, this study used the information of a dichotomous measure whether an occupation requires social interactions with customers or colleagues, which might limit accurate identification of the underlying effect of BMI. It is reasonable to assume that the effect would be different at the intensive margin on the continuous spectrum of the extent of social interactions from the effect at the extensive margin.

The third hypothesis — that an increase in weight would decrease hourly wages — was confirmed for women. An increase in BMI results in a statistically significant decrease in hourly wages for women, which is found in all splined range of BMI. The results are rather ambiguous for men. Gaining weight within the underweight or normal weight range actually rewards hourly wages for men.

As discussed earlier, this different direction of the effect by gender seems to imply that the conventional public conception for an appropriate body weight works differentially by gender. For men, being petite does not seem to reward their labor market outcomes, while for women, being slim helps to earn higher wages. The results in this study are generally consistent with the reports from some key references including Cawley (2004) and Baum and Ford (2004). Although those two studies used the same data as this analysis, not only their methods for dealing with the endogeneity of BMI are different (and restricted), but also the model specifications are different. For the OLS regression and the individual fixed-effects model, which are common methods in all three studies including this, the direction and size of the effects are consistent among these studies, as predicted.

The fourth hypothesis — that an increase in weight would differentially affect hourly wage at various stages of the life cycle — is supported for both genders. Results from this study find the different direction or size of the effect of gaining weight on hourly wages by age group. For men, the effect is positive for the youngest (aged 18-24) sample, but becomes negative for the mid-range of the age group (aged 25-34), and then turns positive again for the oldest (aged 35-40) group. For women, the adverse effect of an increase in BMI on hourly wage consistently remains in all age group. However, the size of the negative effect becomes larger as the age group gets older until past the age of 34 years old.

It is probable that the strong negative effect for women may actually reflect the association of child-bearing and rearing with weight gain and a temporary leave from their jobs. Nevertheless, sensitivity analyses excluding ever pregnant women from the final analysis are not deviated from the results for all women. Another sensitivity results controlling for the marital status and the number of children (treating them exogenous) are also consistent from the primary results for the reduced model as far as the marital status and the number of children. This inquiry will be able to be answered precisely in a structure model treating those two variables as endogenous.

The results imply that the causal association of an increase in weight with wages is not linear over the life-cycle. The wage penalty for gaining weight seems to be lessened by the conventional public wisdom that people tend to gain weight as they get older.

The fifth hypothesis — that an increase in weight would differentially affect wage earnings at occupations where social interactions with customers or colleagues are required from other occupations without that requirement — was confirmed for both men and women. For women, weight gain penalizes hourly earnings in both types of occupations, but the size of the penalizing effect is surprisingly bigger for the occupations not requiring social interactions than the occupations requiring the social interactions. For men, the positive effect of an increase in BMI on hourly wages (for the normal weight range) is clear for occupations requiring social interactions.

The results for women may imply that non-slim individuals who enter an occupation requiring social interactions are likely to have higher job skills than non-slim individuals who fail to enter such an occupation. The results for men seem to imply that being underweight is equally penalized for men as being fat. Appropriately built men (not slim or fat) may be more appreciated in terms of higher hourly wages in the occupations where social interactions are required.

In summary, the evidence presented here suggests that BMI has an important effect on certain labor market outcomes. By and large, the results in this study are consistent with the previously reported effects of weight gain on labor market outcomes (wages) in the literature, but only stronger causality was found in this study. For men in the combined sample, an increase in BMI has a negative effect on labor market outcomes at the extensive margin (the probability of employment and the probability of having occupations requiring social interactions), while the direction is positive for log hourly wages. For women in the combined sample, an increase in BMI penalizes the probability of having occupations requiring social interactions and log hourly wages, while the direction is positive for the probability of employment. Some variations in the direction or the size of the effect of weight gain on hourly wages are found by age group for both genders. For the probability of employment, the direction is positive for the combined sample younger than 30 years of age, but turns negative for the sample aged 30 and older for both genders. The wage penalty is consistent across all age groups for women, while it disappears for the oldest group (age 35-40 years old) for men (see Table 7.1a, and 7.1b). The direction and size of the effect of BMI on hourly wages are found different between the occupations requiring social interactions with customers or colleagues and other occupations where those interactions are not required (see Table 7.2).

This study improves the previous literature investigating the effect of body weight on labor market outcomes in a number of ways. First, this study controls better for the endogeneity of weight, which allows estimating the unbiased causal effect of BMI on labor market outcomes. Above all, this study uses two sets of the over-identifying instruments, both of which pass the test of the exclusion restriction for most of the specifications. The magnitude of the estimated effect found in this study is larger than the results from some key references including Cawley (2004) and Baum and Ford (2004). Cawley's (2004) study also used the two-stage instrumental variable estimation, but the author used only exactly identifying instrument, i.e., siblings' BMI. The individual-level over-identified instruments for this study include siblings' BMI, too. However, it is important to note that this study advances the previous study by finding one more individual-level instrument as well as over-identifying state-level instruments so that the validity of the instruments is verified by the over-identification test and the test for the strength of instruments. In addition, this study controls for the individual fixed effects in the 2SRI model, so that any remaining time-consistent individual heterogeneity is controlled. Different results between the 2SRI and the 2SRI in combination with the individual fixed-effects model in this study shows the importance of combining both methods. Baum and Ford (2004) used only the individual fixed-effects model. However, the fixed-effects model can not be an independent measure to control for the endogeneity of BMI as previously discussed in the method section.

Second, this study explores three different over-identifying sets of instruments, and identifies different marginal effect for the different group of marginal sample. The standard two-stage instrumental variable estimation model recovers parameters for a group of marginal sample who change their behavioral choices affecting body weight according to the instruments (Basu, Heckman, and Navarro-Lozano, 2006). Therefore, different instruments will produce different marginal effects. The individual-level instruments (siblings' BMI and five year lags of the respondents' BMI) are likely to identify the causality for a group of marginal sample whose body weights are affected via genetic factors or shared upbringings among siblings. The state-level instruments (beer prices, fast food prices, and the sales in all types of restaurant) identify the causality for a marginal sample whose body weight are affected by those state-level variables representing access for food. Therefore, those state-level instruments only identify individuals who change their consumption by a change in the monetary or geographic access for food or beer.

Third, this study applies stricter exclusion restrictions to better control for the potential endogeneity bias even though this causes a smaller sample size for this study than the previous studies with the same data. An example of those additional exclusion criteria for the final sample is to exclude any sample person who reported subjectively-determined adverse effects of their health on the amount or types of works. Applying stricter exclusion restrictions helps to strengthen the causality in the association of BMI with labor market outcomes for non-health reasons, which is more relevant to the interest of this study. This study also excluded women who were pregnant within a year from the time of interview to remove the pregnancy effect. Controlling for the pregnancy information in the model as Cawley's (2004) study is likely to interfere to recover the causal relationship of BMI with hourly wages unless the endogeneity of pregnancy is controlled. This study also performs a sensitivity analysis for women who have been never pregnant to verify the control for the pregnancy effect.

Fourth, this study measures the marginal effect of BMI not only at the intensive margin but also at the extensive margin, including the probability of employment and the probability of having occupations where social interactions are required. A very limited number of previous studies have estimated the effect of BMI on the probability of employment. However, those studies used limited controls for the endogeneity of BMI. None of the previous literature has investigated the effect of BMI (or body weight) on the occupation choices. Estimating how the BMI affects labor market outcomes at the extensive margin by characteristics of employment, such as whether social interactions are required or not, would help to understand the underlying mechanism of the association of BMI with labor market outcomes.

Fifth, this study finds a different effect of BMI on hourly wages by the occupations where social interactions with colleagues or customers are required versus other occupations without requiring the social contact. This helps to understand one important pathway for the causality of BMI on hourly wages, i.e., the conventional adverse conception by public for being fat in the modern society. It should be noted that this study carefully excludes one obvious pathway of the wage penalty for being overweight or obese, i.e., the adverse health effect of being overweight or obese.

Sixth, none of the previous studies have measured the different linear effects of BMI by splined segments, which allows the different size and direction of the effect by BMI ranges on the continuous spectrum of BMI. In fact, this study finds a different direction of the effect of an increase in body weight on labor market outcomes, particularly for men.

Further research could address more clearly the link between body weight and labor market outcomes by dealing with some limitations of this study. First, addressing the role of other factors associated with both body weight and labor market outcomes would help to directly understand how BMI affects labor market outcomes. For example, marriage is known to affect BMI for both genders. Especially women tend to gain weight due to pregnancy, and the child-rearing following a child-birth. It also affects an individual's choice regarding labor market participation and performances in the market. The previous literature has dealt with this association on a very limited level. This study does not control for those factors (including the martial status and the number of children) in the primary models, which limits the estimation for this study to the reduced forms. A sensitivity analysis controlling for the marital status and the number of children provides similar results for the reduced form estimation as far as the marital status and the number of children. However, the sensitivity analysis assumes those two variables as exogenous. The endogeneity of the marital status and the number of children remains untested and uncorrected (if the endogeneity truly exists). Being overweight or obese may also penalize outcomes in the marriage market including the probability of marriage, the probability of divorce, or spouse's income. The potential interaction between marriage market outcomes and labor market outcomes, particularly for women, should be addressed in a structural model in a future study.

Second, continuous measures of the extent of social interactions required for an occupation could allow better identification of any causal effect of BMI on the likelihood of having occupations requiring social interactions with customers or colleagues. In addition, even a nonlinear measurement of types of occupations requiring social interactions may need to be refined.

Exploring differences in job characteristics, e.g., part-time versus full-time, fringe benefits, or paid vacations, over the life cycle would help to more accurately understand the causality of BMI on employment or hourly wages. Occupation codes had not been synchronized until 1998 when the O*NET system took over the DOT codes system. The new system has a great potential to increase the strength of causality because characteristics of occupations can be directly linked to the most recent 2000 census occupation codes. Clearly, types of occupations continue to evolve over time, as well as individuals' conception of the appropriate body weight. Therefore, it would be interesting to see how the association of BMI with types of occupations changes over time. Answering this question would be possible as the NLSY79 data has been updated. It should be noted that characteristics of occupations of the DOT codes were matched to the 1980 census occupation codes for this study via several steps of linking process.

Third, it would be important to see the effect of BMI on labor market outcomes for middle-aged or elderly sample as majority of the respondents for the NLSY79 approaches a middle age in the near future.

Accurate estimation of the effect of obesity on labor market outcomes will support the understanding of the economic cost of obesity to an individual besides its adverse effect on health. Individuals' behavioral choices regarding body weight can impose costs not only to the individuals themselves, but also to the others that are not relevant to the choices. Given that individuals are the ultimate decision-makers for their body weight, raising awareness of the obesity costs to individuals would be important to reduce the obesity epidemic in addition to other potential public intervention such as fast food tax. The spillover effect of obesity on labor market outcomes may be able to provide an additional incentive to the individuals to adjust their behavioral choices toward a healthier body weight.

Table 7.1a Summary of results by age group: 2SRI with the individual fixed-effects model for the combined sample[1]: Men

Dependent variable	Independent variable	Age group				
		All	Age 18-24	Age 25-29	Age 30-34	Age 35-40
Probability of employment	BMI < 25	-0.091***	0.166**	0.053	-0.022	-0.092
		(0.035)	(0.077)	(0.060)	(0.130)	(0.121)
	25 ≤ BMI < 30	-0.089***	0.161**	0.062	-0.023	-0.098
		(0.035)	(0.077)	(0.060)	(0.130)	(0.120)
	BMI ≥ 30	-0.089***	0.129*	0.052	-0.02	-0.092
		(0.035)	(0.077)	(0.060)	(0.130)	(0.119)
	N (person-years)	19,732	4,790	6,793	5,775	2,374
Probability of occupations requiring social interactions	BMI < 25	-0.006	0.026	-0.075	0.199	-0.227
		(0.031)	(0.108)	(0.103)	(0.242)	(0.149)
	25 ≤ BMI < 30	-0.015	0.000	-0.061	0.175	-0.224
		(0.031)	(0.108)	(0.102)	(0.244)	(0.148)
	BMI ≥ 30	-0.001	-0.015	-0.074	0.185	-0.191
		(0.031)	(0.108)	(0.103)	(0.242)	(0.151)
	N (person-years)	16,294	3,459	5,809	4,939	2,087
ln (hourly wages)	BMI < 25	0.033	-0.018	-0.043	-0.094	0.025
		(0.028)	(0.086)	(0.079)	(0.216)	(0.186)
	25 ≤ BMI < 30	0.028	-0.001	-0.043	-0.103	0.026
		(0.028)	(0.084)	(0.078)	(0.218)	(0.187)
	BMI ≥ 30	0.021	-0.018	-0.038	-0.095	0.027
		(0.028)	(0.085)	(0.077)	(0.217)	(0.190)
	N (person-years)	16,294	3,459	5,809	4,939	2,087

Note:
1. Combined (individual- and state-level) instruments include siblings' BMI, cost of beer, cost of fast food, and sales in restaurants (full-service and limited services) in 1,000 dollars per 100 state populations by states for both men and women sample.

Table 7.1b Summary of results by age group: 2SRI with the individual fixed-effects model for the combined sample[1]: Women

Dependent variable	Independent variable	Age group				
		All	Age 18-24	Age 25-29	Age 30-34	Age 35-40
Probability of employment	BMI < 25	0.021	0.011	0.024	-0.053	-0.074
		(0.018)	(0.120)	(0.086)	(0.100)	(0.130)
	25 ≤ BMI < 30	0.016	0.028	0.005	-0.06	-0.063
		(0.018)	(0.122)	(0.085)	(0.099)	(0.127)
	BMI ≥ 30	0.025	0.022	0.012	-0.059	-0.052
		(0.018)	(0.121)	(0.085)	(0.100)	(0.128)
	N (person-years)	15,996	3,900	5,170	4,867	2,059
Probability of occupations requiring social interactions	BMI < 25	-0.049*	-0.069	-0.182	0.020	-0.155
		(0.026)	(0.150)	(0.132)	(0.163)	(0.136)
	25 ≤ BMI < 30	-0.055**	-0.037	-0.177	0.015	-0.127
		(0.026)	(0.149)	(0.131)	(0.162)	(0.134)
	BMI ≥ 30	-0.042	-0.07	-0.167	0.014	-0.135
		(0.026)	(0.149)	(0.132)	(0.163)	(0.135)
	N (person-years)	11,522	2,614	3,784	3,596	1,528
ln (hourly wages)	BMI < 25	-0.039*	-0.058	-0.059	-0.014	-0.031
		(0.022)	(0.128)	(0.083)	(0.168)	(0.157)
	25 ≤ BMI < 30	-0.049**	-0.063	-0.070	-0.018	-0.004
		(0.021)	(0.126)	(0.081)	(0.165)	(0.154)
	BMI ≥ 30	-0.042*	-0.052	-0.069	-0.026	-0.027
		(0.022)	(0.126)	(0.082)	(0.168)	(0.156)
	N (person-years)	11,522	2,614	3,784	3,596	1,528

Note:
1. Combined (individual- and state-level) instruments include siblings' BMI, cost of beer, cost of fast food, and sales in restaurants (full-service and limited services) in 1,000 dollars per 100 state populations by states for both men and women sample.

Table 7.2 Summary of results for hourly wages by occupations group: 2SRI with the individual fixed-effects model for the combined sample[1]

Occupation group	Independent variable	Gender	
		Men	Women
Requiring social interactions	BMI < 25	0.010	-0.007
		(0.037)	(0.029)
	25 ≤ BMI < 30	-0.005	-0.012
		(0.037)	(0.029)
	BMI ≥ 30	-0.009	-0.010
		(0.036)	(0.029)
	N (person-years)	9,164	7,063
Not requiring social interactions	BMI < 25	0.029	-0.032
		(0.050)	(0.039)
	25 ≤ BMI < 30	0.030	-0.047
		(0.050)	(0.037)
	BMI ≥ 30	0.021	-0.040
		(0.050)	(0.038)
	N (person-years)	7,130	4,459

Note:

1. Combined (individual- and state-level) instruments include siblings' BMI, cost of beer, cost of fast food, and sales in restaurants (full-service and limited services) in 1,000 dollars per 100 state populations by states for both men and women sample**.**

APPENDIX 1

Monte Carlo Simulation

A Monte Carlo simulation was run to establish that the 2SRI model provides unbiased estimators when it is used for controlling for the endogeneity of BMI for the models on labor market outcomes when BMI was splined. The simulated results confirmed that inserting residuals from the linear estimation of the BMI in the first stage estimation into the second stage equation with splined BMI provides unbiased estimates.

For this simulation, random sets of 5,000 observations were drawn from the normal distribution in order to generate a dependent variable, an endogenous variable and four instruments for the endogenous variable. Equation (1) defines an endogenous variable. When Z is omitted in the estimation of equation (2), two splines of X_{endog} and interactions of those two splines with X_2 would be correlated with the error term in equation (2). In such a case, W_1 to W_4 can be used as the instruments for X_{endog} (or splines of X_{endog}) for identifying X_{endog} in estimating equation (2). The Y_2 is a dichotomous variable generated from a continuous dependent variable Y_1:

$$(1) X_{endo} = 0.5X_1 + W_1 + 2W_2 + 3W_3 + 4W_4 + (Z + \varepsilon_1)$$

$$(2) Y_1 = 0.2X_{endo}SPLINE_1 - 0.4X_{endo}SPLINE_2 - 0.1X_{endo}SPLINE_1 \times X_2 \\ + 0.2X_{endo}SPLINE_2 \times X_2 + (Z + \varepsilon_2)$$

$$(3) Y_2 = 1 \text{ if } Y_1 > 0, \text{else } Y_2 = 0$$

The simulation results are listed in Table 6.1. The second column of Table 6.1 displays unbiased estimators when equation (2) was estimated with Z included. The estimated coefficients were overall same as the coefficients for each variable in equation (2) as it should be. Estimators in the 2SRI were in the third column. In the 2SRI model, predicted residual from the estimation of equation (1) with four instruments (W_1 to W_4) and without Z were inserted in estimating equation (2) with omitting variable Z. In general, the 2SRI estimation resulted in almost similar estimators of the unbiased coefficients from the true model. For results for the Probit model, the magnitude of the estimators was roughly same as the unbiased coefficients when the standard errors were adjusted.

Based on the simulation results, the 2SRI was used for the two-stage instrumental variable estimation.

Table A.1 Monte Carlo Simulation results

Y_1	Linear regression (SE) True values	Linear regression (SE) 2SRI	Probit (SE) True values	Probit (SE) 2SRI
$X_{endog}\_spline_1$	0.198 **	0.196 **	0.186 **	0.155 **
	(0.005)	(0.006)	(0.011)	(0.010)
$X_{endog}\_spline_2$	- 0.396 **	-0.401 **	-0.412 **	-0.340 **
	(0.005)	(0.006)	(0.016)	(0.014)
$X_{endog}\_spline_1 \times X_2$	- 0.099 **	-0.098 **	-0.089 **	-0.077 **
	(0.007)	(0.008)	(0.013)	(0.012)
$X_{endog}\_spline_2 \times X_2$	0.201 **	0.202 **	0.213 **	0.173 *
	(0.007)	(0.008)	(0.018)	(0.016)
X_2	0.504 **	0.500 **	0.524 **	0.411 **
	(0.033)	(0.009)	(0.055)	(0.012)
Z	0.987 **		0.989 **	
	(0.010)		(0.022)	
γ		0.466 **		0.395 **
		(0.040)		(0.047)
Constant	0.003	0.032	-0.014	0.020
	(0.023)	(0.028)	(0.041)	(0.037)

Notes:

1. γ stands for predicted residual from the first stage regression: $X_{endo} = f(W_1, W_2, W_3, W_4)$.

Standard errors are in parentheses.

REFERENCES

Arellano M and Bond S. 1991. Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. *The Review of Economic Studies* 58(2): 277-297.

American Obesity Association. Accessed at March 13, 2005 from http://www.obesity.org/discrimination/employment.shtml.

Anderson PM, Butcher KF, and Levine FB. 2003. Maternal employment and overweight children. *Journal of Health Economics* 22: 477–504.

Angrist J and Krueger A. 1991. Does compulsory schooling affect schooling and earnings? *Quarterly Journal of Economics* 106(4): 979-1014.

Averett S and Korenman S. 1999. Black-white differences in social and economic consequences of obesity. *International Journal of Obesity Related Metabolism Disorder* 23(2):166-73.

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

Baum CL and Ford WF. 2004. The wage effects of obesity: a longitudinal study. *Health Economics* 13: 885-899.

Bhattacharya J and Bundorf MK. 2005. The incidence of the health care costs of obesity. NBER working paper No. 11303.

Bhattacharya J and Sood N. 2005. Health insurance and the obesity externality. NBER working paper No. 11529.

Becker GS. 1971. The economics of discrimination. Second edition. The University of Chicago Press.

Becker GS. 1973. A theory of marriage. *The Journal of Political Economy* 81(4): 813-846.

Becker GS, Grossman M, and Murphy KM. 1994. An empirical analysis of cigarette addiction. *The American Economic Review* 84(3): 396-418.

Becker GS and Murphy KM. 1988. A theory of rational addiction. *The Journal of Political Economy* 96(4): 675-700.

Behrman JR and Rosenzweig MR 2001. The Returns to Increasing Body Weight. PIER Working Paper 01-052.

Biddle JE and Hamermesh DS. 1998. Beauty, productivity, and discrimination: Lawyers' looks and lucre. *Journal of Labor Economics* 16(1): 172-201.

Bollen KA, Guilkey DK, and Mroz TA. 1995. Binary outcomes and endogenous explanatory variables: tests and solutions with an application to the demand for contraceptive use in Tunisia. *Demography* 32(1): 111-131.

Bond S. 2002. Dynamic panel data models: A guide to micro data methods and practice. *Portuguese Economic Journal* 1: 141-162

Borjas GJ and Bronars SG. 1989. Consumer discrimination and self-employment. *The Journal of Political Economy* 97(3): 581-605.

Buffum D and Whaples R. 1995. Fear and lathing in the Michigan furniture industry: employee-based discrimination a century ago. *Economic Inquiry* 33: 234-252.

Bureau of Labor Statistics. Department of Labor. Labor Force Statistics from the Current Population. Accessed at March 18, 2005 from http://www.bls.gov/cps/#data.

Card D. 1994. Earnings, schooling, and ability revisited. NBER working paper No. 4832.

Cawley J. 2000. Body weight and women's labor market outcomes. NBER working paper 7841.

Cawley J. 2000. An instrumental variables approach to measuring the effect of body weight on employment disability. *Health Services Research* 35(5): 1159-1179.

Cawley J. 2004. The impact of obesity on wages. *Journal of Human Resources* 39(2): 451-474.

Cawley J. 2004. An economic framework for understanding physical activity and eating behaviors. *American Journal of Preventive Medicine* 27(3S): 117-125.

Cawley J and Danziger S. 2004. Obesity as a barrier to the transition from welfare to work. NBER working paper 10508.

Chou SY, Grossman M, and Saffer H. 2002. An economic analysis of adult obesity: results from the Behavioral Risk Factor Surveillances System. NBER working paper 9247.

Chou SY, Grossman M, and Saffer H. 2004. An economic analysis of adult obesity: results from the Behavioral Risk Factor Surveillances System. *Journal of Health Economics* 23: 565-587.

Conley D and Glauber R. Gender. 2005. Body mass and economic status. NBER working paper 11343.

Cutler DM, Glaeser EL, and Shapiro JM. 2003. Why have Americans become more obese? *Journal of Economic Perspectives* 17(3): 93-118.

Evans WD et al. 2005. Public perceptions of childhood obesity. *American Journal of Preventive Medicine* 28(1): 26-32.

Dow WH and Norton EC. 2003. "Choosing Between and Interpreting the Heckit and Two-Part Models for Corner Solutions." *Health Services & Outcomes Research Methodology* 4: 5−18.

Gortmaker SL et al. 2001. Social and economic consequences of overweight in adolescence and young adulthood. *The New England Journal of Medicine* 329: 1008-1012.

Greenberg BS et al. 2003. Portrayals of overweight and obese individuals on commercial television. *American Journal of Public Health* 93(8):1342-1348.

Greene WH. 2000. Econometric Analysis. 4th edition. Prentice Hall, Inc.

Gruber J and Frakes M. 2006. Does falling smoking lead to rising obesity? *Journal of Health Economics* 25(2): 183-197.

Hausman JA and Taylor WE. 1981. Panel data and unobservable individual effects. *Econometrica* 49 (6); 1377-1398.

Hamermesh DS, and Biddle JE. 1994. Beauty and the labor market. *The American Journal or Economic Review* 84(5): 1174-1194.

Hamermesh DS, Meng Xin, and Zhang Junsen. 2002. Dress for success – does primping pay? *Labour Economics* 9: 361-373.

Harper B. 2000. Beauty, stature and the labour market: a British cohort study. *Oxford Bulletin of Economics and Statistics* 62: 771-800.

Heckman JJ. 1993. What has been learned about labor supply in the past twenty years? *American Economic Review* 83(2): 116-121.

Ioannides YM. 2002. Residential Neighborhood Effects. *Regional Science and Urban Economics* 32(2): 145–165.

Juhn C. 1992. Decline of male labor market participation: The role of declining market opportunities. *The Quarterly Journal of Economics* 107(1): 79-121.

Juhn C et al. 1991. Why has the natural rate of unemployment increased over time? *Brookings Papers on Economic Activity* 2: 75-142.

Kaufman PR et al. Do the poor pay more for food? Item selection and price differences affect low-income household food costs. Food and Rural Economics Division, Economic Research Service, U.S. Department of Agriculture. Agricultural Economic Report No. 759.

Lakadawalla D and Philipson T. 2002. The growth of obesity and technological change: A theoretical and empirical examinations. NBER working paper 8946.

Loh ES. The economic effects of physical appearances. 1993. *Social Science Quarterly* 74: 420-438.

Maclean JC, Norton EC, and French M. 2006. Alcohol consumption and body weight. University of Miami working paper.

MaCurday T, Green D, and Paarsch H. 1990. Assessing empirical approaches for analyzing taxes and labor supply. *The Journal of Human Resources* 25(3): 415-490.

Martin CJ. 1994. Protecting overweight workers against discrimination: is disability or appearance the real issue? *Employee Relations Law Journal* 20(1):133-142.

Martin E. 1990. Let an overweight person call on your best customers? Fat chance. *Sales and Marketing Management* 142: 66-70.

National Heart, Lung, and Blood Institute, National Institute of Health. 1998. Clinical guidelines on the identification, evaluation, and treatment of overweight and obesity in adults: the evidence of report. *Obesity Research* 6(2): 51S-215S.

National Organization for Research at the University of Chicago. Occupational Prestige Studies /summary. Accessed at July 24, 2005 from: http://cloud9.norc.uchicago.edu/faqs/prestige.htm.

Nielsen SJ and Popkin BM. 2004. Changes in beverage intake between 1977 and 2001. *American Journal of Preventive Medicine* 27(3): 205-210.

Nielsen SJ and Popkin BM. 2003. Patterns and trends in food portion sizes, 1977-1998. *JAMA* 289(4): 450-453.

Office of Administrative Law Judges Law Library. U.S. Department of Labor. Dictionary of Occupational Titles. Fourth edition, revised 1991. Accessed at March 11, 2005 from http://www.oalj.dol.gov/libdot.htm.

Office of the Surgeon General. Public Health Service. U.S. Department of Health and Human Services. The Surgeon General's call to action to prevent and decrease overweight and obesity. 2001. Accessed at Feb 28, 2005 from: http://www.surgeongeneral.gov/topics/obesity/calltoaction/CalltoAction.pdf.

Pagan JA and Davila A. 1997. Obesity, occupational attainment, and earnings. *Social Science Quarterly* 78(3).

Pencavel J. 1998. Assortative mating by schooling and the work behaviors of wives and husbands. *The American Economic Review* 88(2): 326-329.

Popkin, BM, SIEGA-RIZ AM, HAINES PS. 1996. A Comparison of dietary trends among racial and socioeconomic groups in the US. *The New England Journal of Medicine* 335(10): 716-722.

Popkin, BM et al. 2001. Where's the Fat? Trends in U.S. Diets 1965–1996. *Preventive Medicine* 32: 245–254.

Puhani PA. 2000. The Heckman correction for sample selection and its critique. *Journal of Economic Surveys* 14(1): 53-68.

Puhl R and Brownell KD. 2001. Bias, discrimination, and obesity. *Obesity Research* 9:788-805.

Puhl R and Brownell KD. 2003. Psychological origins of obesity stigma: toward changing a powerful and pervasive bias. *Obesity Research* 4:213-227.

Philipson T. 2001. The world-wide growth in obesity: an economic research agenda. *Health Economics* 10: 1-7.

Register CA and Williams DR. 1992. Labor market effects of Marijuana and cocaine use among young men. *Industrial Labor Relations Review* 45: 435-448.

Roehling M. 1999. Weight-Based Discrimination in Employment: Psychological and Legal Aspects. *Personnel Psychology* 52: 969-1016.

Rosen S. 1974. Hedonic prices and implicit markets: product differentiation in pure competition. *The Journal of Political Economy* 82(1): 34-55.

Ruhm CJ. 2004. Maternal employment and adolescent development. NBER working paper 10691.

Sargent JD and Blanchflower DG. 1994. Obesity and stature in adolescence and earnings in young adulthood. Analysis of a British birth cohort. *Archives of Pediatrics & Adolescent Medicine* 148(7): 681-687.

Sobal J, Nicolopoulos V, and Lee J. 1995. Attitudes about overweight and dating among secondary school students. *Internationl Journal of Obesity Related to Metabolic Disorders* 9(6): 376-81.

Sobal J and Stunkard AJ. 1989. Socioeconomic Status and Obesity: A Review of the Literature. *Psychological Bulletin* 105(2): 260-275.

Staiger D and Stock JH. 1997. Instrumental variables regression with weak instruments. *Econometrica* 65(3): 557-586.

Strotz RH. 1955-1956. Myopia and Inconsistency in Dynamic Utility Maximization. *The Review of Economic Studies* 23(3): 165-180.

Stevens J et al. 1998. The effects of age on the association between body-mass index and mortality. *New England Journal of Medicine* 338(1): 1-7.

Stewart H and Davis DE. 2005. Price dispersion and accessibility: a case study of fast food. *Southern Economic Journal* 71(14): 784-716.

Strum R. 2002. The effects of obesity, smoking, and drinking on medical problems and costs. *Health Affairs* 21: 245-253.

Stunkard AJ. 1996. Socioeconomic Status and Obesity. *Origins and Consequences of Obesity* 201: 174-187.

Terza JV. 2005. Endogeneity in nonlinear parametric models: A guide for applied researchers in health economics. Medical University of South Carolina working paper.

Toussaint-Comeau M and Rhine SLW. 2004. Tenure Choice with Location Selection: The Case of Hispanic Neighborhoods in Chicago. *Contemporary Economic Policy* 22(1): 95-110.

UCLA Stat Computing Portal. Accessed at March 9, 2005 from http://statcomp.ats.ucla.edu/.

U.S. Census Bureau, Economics and Statistics Administration, U.S. Department of Commerce. Annual benchmark report for retail trade and food services: January 1992 Through December 2000. A detailed summary of retail sales, purchases, accounts receivable, inventories, and food services sales. 2004. Access at March 12, 2005 from: http://www.census.gov/prod/2001pubs/br-00a.pdf.

U.S. Census Bureau, Economics and Statistics Administration, U.S. Department of Commerce. Annual benchmark report for retail trade: A detailed summary of Retail Sales, Purchases, Accounts Receivable, and Inventories. January 1987 Through December 1996. Access at March 27, 2005 from http://www.census.gov/prod/2/bus/retail/br-96rv.pdf.

U.S. Census Bureau, Historical CPS Income Tables. Access at March 27, 2005 from http://www.census.gov/hhes/income/histinc/p01ar.html

U.S. Department of Labor. NLSY79 users' guide: A guide to the 1979-2000 National Longitudinal Survey of Youth. 2001. Accessed at March 11, 2005 from http://www.bls.gov/nls/79guide/2001/nls79g0.pdf.

Woodridge JM. 2000. Introductory Econometrics: A modern Approach. South-Western College Publishing. 1st edition.