Department of Economics
Lund University
Sweden

Masters Thesis, 15 Credits

August 2008

The Impact of Obesity on Earnings – Evidence from Swedish Panel Data

Author: Tanjima Pervin

Supervisors:
Professor Inga Persson
Associate Professor Dan-Olof Rooth

Acknowledgements

I would like to thanks my supervisors Professor Inga Persson and Associate Professor Dan-Olof Rooth for their valuable comments and suggestions throughout the process of writing this thesis.

My thanks go to my fellow students at the Department of Economics for excellent cooperation during the year of studies.

Tanjima Pervin
Lund, August 2008

Abstract

Obesity is one of the most important public health concerns around the world. Research suggests that obesity has potentially important effects on labour market outcomes. Using longitudinal data, the thesis aims to study the effect of obesity on wage earnings in Sweden. The data set used in this study is the Swedish level-of-living Survey (LNU) from the year 1991 and 2000. Different estimation procedures, namely Pooled Ordinary Least Squares (POLS), Fixed Effects (FE) Model, and Instrumental Variable (IV) approach are used to examine the effect of obesity on wages in Sweden for all individuals, and for males and females separately. The empirical results suggest a strong negative correlation of Body Mass Index (BMI) and wages for females ($p<0.01$), but not for males. Though statistical significance reduces considerably, however, accounting for unobserved individual effects, the FE estimation also shows that a one unit increase in BMI score is associated with more than 0.6% lower females wage earnings, and an obese woman's wage earnings is about 6% lower than a normal weight woman's ($p<0.10$). Based on the IV estimation, I further find a strong statistically significant wage penalty for an obese female over a normal weight female that is causally about 8% ($p<0.01$). The findings are robust to alternative specifications and sub-samples examined in the sensitivity analysis. The thesis concludes that higher BMI score or obesity may cause a wage penalty for females but not for males in Sweden.

Keywords: Wage earnings; Body Mass Index (BMI); Obesity; Pooled OLS; Fixed-effects; Instrumental variable; Sweden

Table of contents

1 Introduction

Obesity is one of the most important public health concerns around the world. It is estimated that 1 billion people around the world are over weight, of whom around 300 million are clinically obese (WHO, 2003). Obesity is a risk factor for numerous health problems and many chronic diseases (WHO, 2002). Health experts argue that obesity levels will continue to rise in the early 21st century - with severe health consequences in the absence of quick and directed intervention as its prevalence has increased by 10-40% in most European countries over the last decade (WHO, 2003). The United States have become increasing alarmed about the growing incidence of obesity because recent research indicated that societal costs of obesity exceed the cost of cigarette smoking and alcoholism (Sturm, 2002). Moreover, obesity affects not only adults but also teenagers and children, which also justifies the importance of assessing both the determinants and the consequences of obesity (Philipson, 2001).

Research suggests that along with health problems obesity has potentially important effects on labour market outcomes. It is argued that obese people may be discriminated against by consumers or employers due to their distaste for obese people. A considerable empirical literature on the effects of obesity on labour market outcomes in the USA has been assessed (Cawley, 2004). One of their most robust findings is that obese women tend to earn less than their non-obese counterparts and that there are differences by ethnicity and/or race (Cawley, 2005). Nevertheless, the question is raised whether the inverse relationship between obesity and a person's wage is because of low wages causing obesity, obesity causing low wages, or a third factor causing both low wages and obesity is yet ambiguous. Recently, Rooth (2007) uses a field experimental method to investigate whether there is discrimination (or differential treatment) against obese individuals in the Swedish labour market. He finds that obese people (applicants' weight manipulated by their photos) get a significantly lower call back rate for an interview, being 6% lower for males and 8% lower for females.

Studying the effects of obesity on wage earnings between gender and ethnic groups might show diverse level of effects within and between countries. The empirical evidence suggests that culture and labour market institutions may be relevant for understanding the relationships between obesity and labour market outcomes, particularly on overall wages, differential wage effects on gender or ethnic and occupational groups (D'Hombres and Brunello, 2005; Lundborg et al., 2007). Recent studies based on groups of countries (e.g. Nordic, Central and Southern Europe), find that the effects of obesity on labour market outcomes differ across Europe (D'Hombres and Brunello, 2005; and Lundborg et al., 2007). Unfortunately, none of the studies provide a fully comparable country-by-country European analysis, which seems necessary according to the empirical evidence provided by D'Hombres and Brunello (2005), and Lundborg and colleagues (2007).

Furthermore, the question could be asked: has the labour market impact of obesity changed as the prevalence of obesity has risen? For example, the prevalence of obesity is 8.3% and 9.7% in Norway and Sweden respectively, whereas, the figures are 23% and 30.6% in the UK and the USA respectively. Consequently, the question could also be raised whether cross-country comparisons yield insight into the relationship between health behaviour and labour market outcomes? If it is presumed that cultural norms for thin body types are inversely related to the obesity prevalence in a society, its prevalence can be thought of as being a crude indicator for the social degree of acceptance of obesity in that country (D'Hombres and Brunello, 2005). Therefore, one could hypothesise that in societies with high obesity rates, one should expect to find low labour market discrimination or penalties associated with obesity. Hence, it would be interesting to analyse the wage impact of obesity in low obese prevalence countries, such as Sweden.

Nevertheless, the main challenge of measuring the effects of obesity on wages or earnings is to disentangle causation from mere statistical correlation. A number of channels have been proposed that explain why obesity may causally affect earnings or wages, such as lower productivity due to bad health and higher rates of absenteeism from work. But empirically, a correlation may also arise because obese people are different from non-obese, i.e. selection into the group of obese person is non-random with respect to characteristics relevant in the labour market To the extent that obese people are different from non-obese people in observed characteristics (e.g. education, age etc.) this can be controlled for by including them as covariates in a standard Ordinary Least Squares (OLS) regression. However, if selection takes place on unobservable characteristics of the individual (e.g. motivation or self-discipline/ self-confidence), estimation of the causal effects of obesity on wages becomes more complicated.

The overall aim of this thesis is to study whether obesity affects labour market outcomes, particularly individuals' wage earnings. Using the longitudinal data, the thesis investigates whether obesity affects individuals' wage earnings in Sweden. The Swedish Level of Living survey (LNU) data set is used in this work which has self-assessed information on height and weight for the years 1991 and 2000. To my best knowledge, this is the first study which investigates the obesity impact on earnings in Sweden using longitudinal data. Three different empirical strategies have been employed to examine the wage-weight relationships. The first strategy consists of the Pooled Ordinary Least Square (POLS) method, the second strategy uses a Fixed Effects (FE) approach to control for unobserved individual effects and, finally, to identify the causal effects of obesity on earnings the method of Instrumental variables (IV) is applied.

The remainder of the thesis is organized as follows. The next section provides the theoretical underpinnings as to why and how obesity affects a person's wage earnings. It also summarizes the previous literature concerning the data, econometric techniques used in the studies as well as their findings. I then describe the data and variables in my study, and the different empirical strategies that are used in the thesis as to examine the obesity and wages relationships. The estimation results follow next, and the final section presents the overall discussion and conclusions.

2 Theoretical Underpinnings and Previous Work

2.1 Theoretical Underpinnings

According to economic theory, the wage of an individual is related to his or her marginal productivity, i.e., a low wage implies a lower marginal productivity of labour compared to the marginal productivity for a worker with a higher wage and vice versa. The marginal productivity is also related to the level of education and how long the worker has been employed, i.e., the marginal productivity increases with skills or experience. The relationship between the marginal productivity of labour and the level of education was first introduced by Becker (1965) with the human capital model.

Economic models assume that consumers recognize changes in technology and relative prices and understand how these changes affect the optimality of tradeoffs between short-term pleasure and long-term health. In particular, most economic models assume that individuals are rational and forward looking (Becker and Murphy, 1988). For instance, lower food prices may or may not induce individuals to raise their food consumption, depending on the importance they place on future health hazards associated with higher weight. Therefore, it is important to know whether individuals fully consider the long-term costs of excessive obesity when making their current decisions or despite such knowledge whether individuals do face a self-control problem in making appropriate choices.

The theory of rational addiction states that addictions are rational in the sense that addictions are forward looking utility maximizations with stable preferences (Becker and Murphy, 1988). The theory also says that present consumption of a certain good is significantly increased by past consumption, i.e. consumption of the good increases the marginal utility of the good to a greater extent than the total cost of the good, including harmful effects such as a reduction of the capital stock of health and thus a reduction of future income. Goods giving risk to such behaviour are referred to as addictive. The theory of rational addiction has been applied to various goods and conceivably most importantly to tobacco and cigarette smoking. Becker and Murphy (1998) include eating as a rational addictive behaviour as well. The increase in overweight and obesity in the western world is well documented (Philipson et al., 2004) and could be an area where the rational addictive theory could be applied.

To assess the effects of obesity, two basic types of behavioural factors need to be noticed. Firstly, the factors that are internal to the workers and, secondly, the external forces impacting them in their work environments (Baum and Ford, 2004). In general, Baum and Ford (2004) assume that obese workers may earn lower wages because they are less productive because of health problems. Alternatively, obese workers may suffer a wage penalty because they place a higher premium on present consumption (Baum and Ford, 2004). Baum and Ford (2004) further argue that economically myopic workers, who have higher marginal rates of time preference, may be less concerned about the possible long term health effects of obesity and consequently, may be more likely to be obese. Further, such myopic workers may also be less likely to engage in training, resulting in a flatter career earnings profile (Baum and Ford, 2004).

In addition, Baum and Ford (2004) recognise that a number of external behavioural forces in the obese worker's environment may also help to explain the observed earnings differences. Employers may discriminate in hiring, promoting, compensating, or firing such workers. Other possible forms of external discrimination against obese workers are due to cultural and customer discrimination. For example, obese workers in occupations involving direct public contact may experience a wage penalty through their interaction with customers. This would occur if customers are averse to interacting with those who are obese (Baum and Ford, 2004).

2.2 Previous Work

The economics literature has given relatively little attention to the effects of obesity. Nevertheless, there are a number of economic studies mostly based on US data; the available empirical evidence for Europe is rather limited. There are some studies for particular European countries: UK (Sargent and Blanchflower, 1994; Morris 2005, 2006), Finland (Sarlio-Lahteenkorva and Lahelma, 1999), Germany (Cawley et al., 2005), and Denmark (Greve, 2005). A summary of some seminal works on the obesity-wage relationship is given in Table 1.

Register and Williams (1990) use cross-sectional data from the 1982 National Longitudinal Study of Youth (NLSY) to examine the effect of obesity on wages (in wage level models) by gender. Their results indicate that obesity reduces females' wages by 12 percent but has no significant effects for males. Averett and Korenman (1996) use 1988 NLSY data to examine the effects of obesity on wages. Their results are mixed: wage level models indicate that women and, to a lesser extent, men both suffer obesity wage penalties, but their sibling–differenced fixed-effects models indicate no significant obesity wage penalties. In particular, in their wage level models, obese women suffer a 10-24 percent wage penalty and men suffer an obesity wage penalty of about 8 percent.

Pelkowski and Berger (2004) estimate the long term impact of health problems on labour market decisions and outcomes. They distinguish between temporary and permanent health problems (due to obesity) to find the effect on wage earned and hours worked, and find that poor health has effects on employment outcomes and that these effects differ between genders. Permanent health conditions have a significant negative effect on wages for both males and females where the effect is slightly larger for women.

Table 1: Summary of some previous economic studies on obesity-wage nexus

Studies	Data	Econometric methods used	Overall findings
Register and Williams (1990)	Cross-sectional data from the 1982 NLSY in USA	Ordinary least squares(OLS) regression	Obesity reduces female wages 12% but has no significant effect for males
Sargent and Blanchflower (1995)	A birth cohort at age 23 years from the National Child Development Study, which consists of all children born in England, Scotland, and Wales between March 3 and 9, 1958.	A series of ordinary least-squares (OLS) regression	The wage of young females is lower if they have been overweight or obese in the past and found little evidence of a difference in wage for male.
Averett and Korenman (1996)	Use 1988 NLSY data	Ordinary least squares(OLS) regression and fixed-effect (FE) model	Mixed findings: OLS indicates that women and men both suffer obesity wage penalties, but sibling-differenced fixed-effects model indicate no significant obesity wage penalties.
Baum and Frod (2004)	Use NLSY data from 1979 to 1994	Various models including fixed-effects model	Obese workers (male & female) suffer a wage penalty.
Cawley (2004)	Use NLSY data for the years 1981-2000	OLS, Fixed effects and Instrument variable approach (IV)	Wage penalty for obese white female workers of 9 percent and finds that the wage penalty is equivalent to the wage effect of roughly one and a half years of education or three years of work experience.
D'Hombres and Brunello (2005)	Use the European community Household Panel (ECHP) data	Instrument variable approach (IV)	Association between BMI and wages is negative for women and positive for men. But got different results from biological family member (negative effect for both men & women)
Cawley et al. (2005)	Panel data in 1979 from NLSY survey.	OLS regression, probit model	Negative correlation between weight & wages for other gender-ethnic groups
Garcia and Quintana-Domeque (2006)	Use the European community Household Panel (ECHP) data in 1998	Multinomial logit & OLS	Obese people are more unemployed & find to difficulty to detect relationship between obesity & wage
Lundborg et al. (2007)	Survey of Health, Ageing and Retirement in Europe(SHARE) data	Fixed-effect model and Instrument variable (IV) approach	Obesity is negatively associated with men &women aged 50 plus people.
Attella, et al. (2007)	Use the European community Household Panel (ECHP) data 1998-2001.	Quantile Regression Approach	Association between BMI and wages is negative all over the distribution for women, and negative and significant only in the bottom part of the distribution for men.

The forgoing literature has some shortcomings, and these shortcomings probably account for their contrasting results. First, the literature uses small and unrepresentative samples; in particular, the NLSY survey includes respondents between the ages of 17 and 24. These respondents are less likely to be obese because they are still quite young. Second, the estimates found in the literature fail to control sufficiently for unobserved heterogeneity. Unobserved heterogeneity could bias the results if the same unmeasured traits that determine obesity simultaneously affect wages.

Baum and Ford (2004) use the NLSY data to examine the effects of obesity on wages by gender. Their study is distinctive in that they use 15 separate statistical models and a very large sample size. They attempt to test whether obese workers receive lower wages because they are: 1) less motivated, 2) they are more costly for employers to insure, 3) they are discriminated against by customers, or 4) they are limited in their jobs due to their obesity. They find that obese men and women earn lifetime wages that are on average 3.4% and 6.3% lower than those of other workers with similar educational backgrounds, work experience and socio-economic characteristics. They conclude that the standard socioeconomic covariates do not explain why obese workers experience persistent wage penalties which implies that the other factors - including job discrimination, health-related factors and/or obese workers' behaviour patterns - may be the channels through which obesity adversely affects wages.

Using the same data set (i.e. the NLSY); Cawley (2004) finds an even larger wage penalty for obese white female workers, of 9 percent. The author also shows that the wage penalty is equivalent to the wage effect of roughly one and a half years of education or three years of work experience. However, it is deduced that negative correlations between weight and wages observed for other gender-ethnic groups appear to be due to unobserved heterogeneity.

Using pooled cross-sectional Health Survey Data for England for the years 1997 and 1998, Morris (2006) finds that obesity has a positive and significant effect on mean hourly occupational earnings for males and a negative and significant effect for females, though the association for males is not robust across different specifications. However, after using the mean BMI (and/or the prevalence of obesity) across individuals living in the same health authority area as an instrument for individual BMI, the author finds no statistically significant effect, either for men or for women.

In Finland, obese females are found to have lower income levels than non-obese ones, but that it is not the case for males (Sarlio-Lahteenkorva and Lahelma, 1999). For Denmark, using information on whether the individuals' parents have ever taken medication related to obesity or obesity related diseases (namely hypertension and Type 2 diabetes) and their mortality cause, Greve (2005) finds a negative and significant relationship between BMI and probability to be employed for women and an insignificant relationship for men.

The empirical evidence for Germany shows that obesity is negatively associated with wages, both for men and for women (Cawley et al, 2005). Moreover, once the authors control for endogeneity using genetic factors, they conclude that there is no significant relationship between weight and wages.

Using the Survey of Health, Ageing and Retirement in Europe (SHARE), a recent study by Lundborg et al. (2007) analyze the effect of obesity on employment, hours worked and hourly wages in 10 European countries for people aged 50 and above. Pooling all countries, they find that obesity is negatively associated with being employed for both men and women and with female hourly wages.

European wide analyses have been conducted using pooled data from the European Community Household panel (ECHP) data for nine European countries by Sousa (2005), Garcia and Quintana-Domeque (2007), Brunello and D' Hombres (2007). Sousa (2005) focuses on the impact of the BMI on labour force participation. She finds that being overweight decreases labour force participation for women, but it increases labour force participation for men. However, she is not able to estimate the obesity effect for each country separately, because using the propensity score matching approach reduces enormously the sample size. Brunello and D' Hombres (2007) find a negative and statistically significant impact of obesity on wages independently of gender for the pooled sample of countries. Furthermore, they also find that the negative relationship between wage and obesity is higher in southern Europe than in northern Europe.

Employing European Community Household panel (ECHP) data and using a multinomial logit model, recently Garcia and Quintana-Domeque (2007) find weak evidence that the obese workers are more likely to be unemployed or tend to be more segregated in self employment than their non-obese counterparts. For example, this study finds statistically significant association with a higher relative probability of being unemployed in Belgium. Obese women in Belgium are more than twice likely to be unemployed rather than working as employees, and for men this ratio is even higher, more than three times. On the other hand, obese Belgium men tend to earn a higher hourly wage (8%) than their counterparts.

Most of the studies that explain the relationship between wages and obesity have been based on a mean regression approach, which looks only at the role of obesity at the mean level of wages, which ignores individual wage heterogeneity. Hoverer, it could be that obesity affects individuals differently across the wage distribution. By using ECHP data, recently Atella and colleagues (2007) try to fill this gap by adopting a quintile regression approach. By employing the quintile pooled regression model, the authors find that the obesity-wage relationship seems to be negative and significant all over the distribution for women. Their results also show that men in the bottom part of the wage distribution seem to suffer from a wage penalty due to obesity (-3.1% and -1.3% at the 15th and 25th percentile, respectively) while the effect is not statistically significant in the remaining quintiles (50th, 75th, 85th).

3 Data and Methods

3.1 The Data

The empirical analysis is based on data from the Swedish Level of Living Survey (LNU), a data set designed and coordinated by the Swedish Institute for Social Research, Stockholm University, Sweden. The LNU provides a longitudinal panel data from 1968 to 2000, with a focus on household income, living conditions, individual health and access to care, education, employment and working conditions, family and social integration, political security etc. The Swedish Level-of-Living survey (LNU) is one of the longest running longitudinal social science surveys in the world. It was first conducted in 1968. Thereafter, it has been replicated in 1974, 1981, 1991, and

2000. The basis for LNU was a random sample of 1/1000 of the Swedish population between 15 and 75 years of age. In 1991, the lower age limit was raised to 18 years. The same respondents have been interviewed again at later waves, and 1,750 respondents have in fact contributed to all five waves. In addition to the interview data, register information has been added, mainly in order to calculate household income. Due to the lack of information about weight and height for the earlier years, the present thesis uses data for the years 1991 and 2000.

In total, 6,773 and 5,141 individuals responded in years 1991 and 2000 respectively of which 4,160 individuals participated in both years/waves. After dropping individuals aged less than 18 and aged above 65 (working age group), I come up with a total of 5,771 and 4, 552 individuals in 1991 and 2000 respectively. Finally, after dropping individuals with missing values on the items of interest, the pooled data comprised 4,325 individuals (2,211males and 2,116 females)[1] in two years (in total, 6,277 observations for all waves). The balanced panel was a sub-set of the 1,952 individuals (3,904 observations) present in all two waves. Table 2 summarises the definitions of the variables that have been used in the analyses. Table 3 provides a description of the final data considered in the analyses, and summary statistics for the males and females, separately and together.

3.1.1 Dependent variable

Log hourly wage

This thesis considers hourly wage for the respondent's current job as the dependent variable. I have converted nominal hourly wage into real wage using the consumer price index (CPI) by using 2000 as the base year. It should be mentioned that hourly wages are preferred over total earnings for two reasons. First, the theoretical arguments centre on the effects on productivity, and differences in a worker's productivity are expected to be reflected in wages per hour. Second, total wages may overstate a worker's earnings capacity if he works many hours, and the number of hours worked is potentially endogenous as a control variable as it is likely be jointly determined with total wages. The log form accounts for the right-skewed distribution of wages and restricts the model to non-negative values for the dependent variable.

[1] It seems that one individual has reported different sex in the alternative years.

3.1.2 Independent variables

Body mass index (BMI)

In the LNU data respondents are asked for their height in meters and weight in kilograms in the years 1991 and 2000. Based on height and weight, I have calculated the individuals' body mass index (BMI) score. [2] Body Mass Index (BMI) is calculated as the person's weight (in Kilograms) divided by the square of height (in meters). Following the earlier literature, based on their BMI, I have divided individuals into four categories: BMI less than 18.5 is classified as under weight for both genders. For males, BMI more than 18.5 but less than 25 ($18.5 \leq \text{BMI} < 25$) is considered as normal weight, BMI 25 or above but less than 30 ($25 \leq \text{BMI} < 30$) is categorised as over weight, and BMI 30 or above ($\text{BMI} \geq 30$) is classified as obese. The corresponding numbers for women is BMI between 23.8 and 28.6 ($23.8 \leq \text{BMI} < 28.6$), and BMI over 28.6 ($\text{BMI} \geq 28.6$) are categorized as overweight and obese respectively (Swedish National Institute of Public Health, 2006; Rooth, 2007).

Education Qualification

The education level of individuals usually indicates the level of human capital accumulation of people. Moreover, for more educated people (and especially for women) education may have a negative influence on weight due to higher frequency of weight monitoring (Wardle & Griffith, 2001), different life styles, lower inter-temporal discount rates. Therefore, in the analysis one needs to control for individuals' education level. Educational status is defined as level of educational attainment or years of schooling. For a few individuals the years of schooling is found to be more than 20 years, for those individuals years of schooling are top coded to 20 years.

Experience and Experience squared

In a typical wage equation, it is common to include level of work experience as a covariate (e.g. Mincer, 1974). Quadratic form of the variable also is often used in applied labour economics to capture decreasing or increasing marginal effects of experience on individuals' wage earnings.

---

[2] As the other researchers, I rely on self reported measures of height and weight and calculate the BMI. With this procedure, there is a possibility that the BMI can be measured with errors. However, Cawley (2004) shows that this does not seem to be a major problem. He finds that even if women tend to under-report their weight but not their heights, using reported BMI instead of corrected BMI does not alter significantly the estimates.

Table 2: Description of the variables used in the analysis

Variables	Description
hwage	Hourly wage
lhwage	Log of hourly wage
bmi	Body Mass Index calculated as the person's weight (in Kilograms) divided by the square of his/her height (in meters)
underw	Dummy variable = 1 if BMI is less than 18.5 (both for males and females = 0 if otherwise
normalw (omitted category)	Dummy variable = 1 if BMI is $18.5 \leq \text{BMI} < 25$ for males and if BMI is $18.5 \leq \text{BMI} < 23.8$ for females = 0 if otherwise
overw	Dummy variable = 1 if BMI is $25 \leq \text{BMI} < 30$ for males and if BMI is $23.8 \leq \text{BMI} < 28.6$ for females = 0 if otherwise
obese	Dummy variable = 1 if BMI ≥ 30 for males and if BMI ≥ 28.6 for females = 0 if otherwise
age	Age in years
education	Number of years of schooling
exp	Number of years of Experience, calculated as: Age-years of schooling- 6
exp2	Experience squared
male	Dummy variable = 1 if male = 0 if female
alone	Dummy variables = 1 if respondents are single/ unmarried/ widowed = 0 if respondents are married/ cohabiting
ghealth	Dummy variable = 1 if individual's self-assessed health is good = 0 if individual's self-assessed health is bad or neither bad or good health
fhealth	Dummy variable = 1 if individual's self-assessed health is neither bad or good health = 0 if individual's self-assessed health is bad or good health
Smoking	Dummy variable = 1 if individual currently smoke (smokes less than 10 cigarettes, or 10 and more cigarettes in a day) = 0 if individual never smoked or quit smoking
year2000	Dummy variable = 1 if individual responded in the year 2000 = 0 if individual responded in the year 1991

It is expected that the wage increases with experience, but experience has decreasing effect on wage earnings. This implies that the coefficient for the variable experience is assumed to be positive; while the coefficient for experience squared should be negative (a parabolic shape is expected for the wage-experience relationship). Since, in the current LNU data set, I do not have any direct information on the individuals' job experience for the year 2000, therefore, I have constructed the variable as: experience= age - years of schooling 6.[3]

Gender

Wage differences between genders are well documented. Therefore, in the wage regression analyses it is important to control for gender. In this study, when I run the regressions for all individuals, females are used as a base case.

Marital Status

Marital status defined as whether individual lived alone or otherwise, with widows included in the 'single' category. A dummy variable = 0 if people are married / cohabiting and = 1 if people are single/ widowed /un-married.

Health

According to the existing literature, it is observed that health problems are more frequent in obese workers than non-obese people and such problems may also affect labour market performance. Based on the question in the LNU data on health status, I have classified individuals into three categories: good health, bad health, and neither good nor bad health (fair health) and where bad health has been considered as the omitted category.

Smoking habits

Smoking habit is widely used in wage models in order to control for systematic differences in observed characteristics between individuals, as some of them may affect simultaneously weight and wages and their effects need to be netted out (Baum, 2008). Many studies have found that smoking is negatively correlated with labour productivity (Brune, 2007). In this study, based on

[3] Which seems to be reasonable for men, but maybe less so for women, however, for seek of parsimony, I use the same construction for both genders.

the individuals' smoking status, I have classified people into two categories: current smokers (whether smokes less than 10 or 10 and more cigarettes in a day) and non-smokers (never smoked or quit smoking); non-smokers are considered as the base case.

3.2 Descriptive analysis

Table 3 reports descriptive statistics of the variables by all individuals, and by males and females. As seen in table 3, the respondents' mean hourly wage has been reported to be about SEK 104 for all individuals. As expected, males' hourly wages (SEK 115) appear to be higher than for females (SEK 93). A similar difference is also evident for the BMI: the average BMI score is higher for males (24.86) than females (23.54). However, when the individuals' BMI score are categorised into dummy variables, it seems that males are overweight to be a larger extent than their counterpart females. The figures are 36.7% and 19.4% for males and females respectively. Nevertheless, females are more likely to be obese than males. About 9% of the females included in the sample are categorised as obese, whereas the corresponding figure for the males is 6.4%.

For other included covariates, no significant differences are observed for the males and females. Average age of the sample is about 40 years for both genders. An average year of schooling is reported around 12 years for both genders and the figure for the covariates 'experience' is estimated to about 22 years for both the groups. About 30% and 27% of males and females respectively report that they live alone. More than 83% and 82% of males and females respectively state that their self-assessed health is good. The corresponding figures for fair health are about 14% and 15% for the males and females respectively. The smoking status is slightly higher for the females; more than 28% of the females in the sample are smokers, whereas the proportion for the males is about 26%.

Table 3: Descriptive statistics of the variables used in the analyses by all individuals, males and females

Variable	All Individual Number of observations=6,277				Males Number of observations=3,181				Females Number of observations=3,096			
	Mean/ Propor	Std. Dev.	Min	Max	Mean/ Propor	Std. Dev.	Min	Max	Mean/ Propor	Std. Dev.	Min	Max
hwage	103.73	46.491	10.33	1732.10	114.56	56.290	20.654	1732.10	92.615	29.718	10.33	444.13
lhwage	4.583	0.323	2.335	7.457	4.674	0.340	3.028	7.457	4.488	0.274	2.335	6.096
bmi	24.211	3.506	15.242	44.82	24.860	3.195	17.433	44.818	23.543	3.683	15.24	43.91
underw	0.019	0.136	0	1	0.004	0.061	0	1	0.034	0.182	0	1
overw	0.282	0.450	0	1	0.367	0.482	0	1	0.194	0.395	0	1
obese	0.076	0.266	0	1	0.064	0.244	0	1	0.089	0.285	0	1
education	12.096	3.094	0	20	12.097	3.181	0	20	12.094	3.001	4	20
exp	22.279	12.931	0	53	22.098	13.009	0	53	22.465	12.849	0	52
exp2	663.54	632.72	0	2809	657.51	637.42	0	2809	669.745	627.90	0	2704
age	40.375	12.034	18	65	40.196	12.126	18	65	40.559	11.939	18	65
male	0.507	0.500	0	1								
alone	0.281	0.450	0	1	0.296	0.456	0	1	0.267	0.442	0	1
ghealth	0.829	0.377	0	1	0.834	0.372	0	1	0.823	0.381	0	1
fhealth	0.148	0.355	0	1	0.143	0.350	0	1	0.154	0.361	0	1
smoking	0.273	0.445	0	1	0.264	0.441	0	1	0.281	0.450	0	1
year2000	0.470	0.499	0	1	0.476	0.499	0	1	0.464	0.499	0	1

3.3 Empirical Strategy

3.3.1 Pooled Ordinary Least Squares (POLS)

Given the panel nature of the data used in the thesis and to make full use of this data, cross-section pooled across time is a standard starting point of the analysis. Suppose wages for the *ith* individual randomly drawn from the population at time *t* are determined by weight or obesity, other human capital factors (such as education), and other controls (e.g. experience, gender, marital status etc). One may estimate the model to estimate the effect of obesity on wage earnings and the relationship can be specified as:

$$W_{it} = \beta X_{it} + \gamma O_{it} + \varepsilon_{it} \tag{1}$$

Where, the dependent variable, log of hourly wages, for individual *i* at the survey year ("wave") *t* is denoted as $W_{it.,}$ X is a vector of control variables and O is overweight/obesity (as measured by individual's BMI score). Following the standard human capital model of wage determination (e.g. Mincer, 1974), X_{it} contains the control variables, namely education, experience and experience squared, gender, marital status, self-assessed health and smoking habits. ε_{it} is the error term for observation *i* in time *t* which can assumed to have a mean of zero due to the inclusion of a constant: $E(\varepsilon_{it}) = 0$ for all *i* and *t*. If individuals are randomly sampled from the population then, in any particular wave, the ε_{it} are independent and identically distributed (i.i.d), so $E(\varepsilon_{it,}\ \varepsilon_{jt}) = 0$ for $i \neq j$ at any one t (Wooldridge, 2002) .

However, independence of disturbances of an individual across time is not presumed in the POLS approach as in a longitudinal data set, hence the repeated observations for the same individual (s) over time are not independent. Moreover, if the assumption that ε_{it} and ε_{is} come from the same distribution for $s \neq t$ is not valid, then their respective variances would be differed across the waves of the panel. While violating of these two conditions do not prevent to get consistent estimates through OLS, however, the resulting serial correlation and heteroskedasticity respectively would mean that the usual OLS standard errors would not be consistent (i.e. estimations would not be efficient). Hence, a robust covariance structure is required to derive correct standard errors for statistical inference (Wooldridge, 2002).

The main interest of this thesis is to estimate the effect of obesity/BMI on wage earnings. In order to derive the consistent estimates of the coefficients of the above model via pooled ordinary least squares (POLS), one would have to assume that the regressions are contemporaneously exogenous in the sense that, at a minimum, $E(X_{it}\ \varepsilon_{it}) = 0$ and $E(O_{it}\ \varepsilon_{it}) = 0$ for t=1,2….Ti. This assumption might be violated if covariates are omitted from the model's specifications that are correlated with one or more of the regressors and the dependent variable (Wooldridge, 2002). In particular, if O is strictly exogenous then an OLS estimate of β can be interpreted as a consistent estimate of the true effect of weight or BMI score on log wages. However, there may be unobserved factors that affect wages and that are correlated with the X.

OLS may be biased estimates for at least three reasons. Firstly, unobservable individual effects related to genetic and non-genetic factors (e.g., ability, self control, parental background), might be associated both with wages and the individual's BMI. Secondly, a reverse causality problem may also exist. For instance, the quality and the quantity of food intake might determine by the individual's earning potentialities, but, simultaneously, individual's wages might influence her quality and quantity of food intake. Finally, since researchers rely on self reported measures of weight and height, therefore the BMI score can be measured with errors, which in turn produce inconsistent estimates. The inconsistency of the estimates can be produced due to the correlation between the error term and the BMI (Cawley, 2004).

3.3.2 The Fixed Effects Model

To present a specification which can address the problem with endogenity caused by unobserved heterogeneity are presented in the following section. At a more intuitive level, panel data may provide more information because the same individuals are repeatedly observed. In the one hand, an important advantage of panel data compared to time series or cross-sectional data sets is that it allows identification of certain parameters without the need to make restrictive assumptions (Verbeek, 2004). For example, panel data make it possible to analyze changes on an individual level. On the other hand, having the same individuals rather than different ones may imply less variation in the explanatory variables and thus relatively inefficient estimators. Most panel data models are estimated under either the fixed effects (FE) or the random effects (RE) assumptions. The parameters should be estimated using the FE methods if the RE model fails a Hausman test.

The appropriate FE estimator explicitly uses the panel structure of the data and may eliminates endogeneity resulting from unobserved time invariant individual effects. By assuming that all the slope coefficients (βs) are constant for all individuals (i) and over time (t) and that each individual has a different intercept term, we may specify the individual-specific effects linear model for log wage of individual i at the time t. In the fixed effect approach, we need to assume that the wage model contains an individual fixed component, α_i which represents individual-specific unobserved heterogeneity. Thus, the wage model can be written as follows:

$$W_{it} = \mu_i + \alpha_i + \beta X_{it} + \gamma O_{it} + \varepsilon_{it} \tag{2}$$

Where, μ_t is a mean intercept, α_i are the time-invariant random variables that capture unobserved individual-level heterogeneity; β is a vector of coefficients representing the partial effects of the exogenous regressors X_{it} and γ is the coefficient of O_{it}. The mean intercept μ_t is only identifiable from the fixed effect α_i by imposing the following restriction $\Sigma\ \alpha_i = 0$. The individual fixed effect can be seen as the *i*th individual's deviation from the common mean (Hsiao, 2003).

To obtain the FE estimator, α_i is eliminated by performing the so called "within transformation" and subsequently applying OLS to the transformed model. The within transformation consists of averaging each individual's observation over time and then subtracting the individual's average over time from each of the individual's observation. The unobserved fixed effects are thus eliminated from the model and hence identification of the coefficients of the remaining (time varying) variables does not depend on the statistical properties of α_i; particularly, correlation with any of the regressors is no longer a problem. In other words, by the nature of the FE estimator, identification of the model's parameters works through within-group variation, i.e. variation across individuals. The estimation equation can then be expressed as:

$$(W_{it} - \overline{W}_i) = (\mu_i - \overline{\mu}) + \beta(X_{it} - \overline{X}_i)\beta + \gamma(O_{it} - \overline{O}_i) + (\varepsilon_{it} - \overline{\varepsilon}_i) \tag{3}$$

i= 1, 2… ….N, t= 1, 2, …….., Ti

Equation (3) is estimated by using the covariance estimator, which removes α_i by transforming the data into mean deviations. OLS estimation of Equation (3) leads to consistent and unbiased estimates of the β's as N and /or $\rightarrow$ to infinite (Verbeek, 2004).

Notice that, though POLS estimation necessitates contemporaneous exogeneity of the predictors, for the fixed effect (FE) approach to produce consistent estimates it is required to assume a form of strict exogeneity that is, at least, $E(X_{is}\varepsilon_{it}) = 0$ and $E(O_{is}\varepsilon_{it}) = 0$ for all i and s, t= 1, 2, …….., Ti. The major difference between POLS and FE estimation is that, though correlation between a regressor and α_i produce inconsistent estimates for the POLS, however, any such correlation is allowed for FE estimation.

Despite the FE estimator's ability for consistent estimation in the presence of unobserved heterogeneity, other sources of endogenity are assumed not to be present. For example, the present model assumes that there is a one way relationship between obesity (O_{it}) and wages (W_{it}) in the sense that obesity influences wages but not the other way around. If higher wages cause a higher or lower incidence of obesity then in a single equation model γ is no longer consistently estimated through FE. Therefore, two-way causation may exist on theoretical grounds. But the magnitude of the bias may be assumed to be small empirically.

In early studies, though Baum and Ford (2004), Cawley (2004), Cawley and DanZiger (2005) use the fixed effects estimators to control for unobservable individual effects. However, this identification strategy does show some drawbacks. In particular, as also noted by Garcia and Quintana-Domeque (2007), a FE strategy does not solve the reverse causality problem. In addition, there is a clear trade-off-between consistency of the estimates obtained with longer panel and plausibility of unobservable time invariance. However, all these identification strategies are somewhat unsatisfactory, since the identification strategies are based on very strong assumptions.

Nevertheless, fixed-effects estimators would simultaneously control for sample selection bias assuming any unobserved factors that determine employment are fixed. However, if the unobserved variables vary over time and the error differences are correlated with the differenced covariates, then the fixed effects estimates will also be biased. In general, the fixed-effects approach necessitates the regressors to be strictly exogenous and that all the omitted relevant variables and unobserved individual characteristics remain constant over time.

3.3.3 Instrumental Variables Approach

To break the correlation between the observed right-side variables and the compound disturbance terms that include unobserved determinants in addition to stochastic terms, one estimation strategy is to use instrumental variables (IV) techniques. In IV estimates the endogenous right-side variables are replaced by their predicted values that depend on "instruments" that do not appear directly in the relation of interest.

However, it is not an easy task to find a good instrument. Good instruments must (1) predict well the variable being instrumented (i.e. corr (Z_i, O_i) $\neq 0$ implies that the instrument is relevant and a more relevant instrument produces more accurate estimator), and (2) not be correlated with the error term [i.e. corr($Z_i \varepsilon_i = 0$), then the instrument is called to be exogenous, i.e. the instrument is valid]. Where, Z_i indicates the instrument of O_i. If a variable satisfies these conditions, then the estimated coefficients [i.e. β's and particularly γ in equation (1)] can be consistently estimated through the two-Stage Least Square (2SLS) approach.

The first requirement of a good instrument can be tested easily. However, one can not directly test the validity of an instrument but rather test the over-identifying restrictions in the model. Suppose we have only one endogenous explanatory variable, then we need to assume that at least one IV is exogenous. Then one can test the over-identifying restrictions that are used in 2SLS. This allows one to test the validity of the instruments, i.e. whether the correlation of the instruments with the errors is zero. In other words, if we have more than one instrumental variable, we can effectively test whether some of them are uncorrelated with the structural error. In applied research, it is usual to use the Hansen's J test statistics, as the test statistics is the robust and equivalent of the usual Sargan (1958) test of over-identifying restrictions (Baum *et al.*, 2003).

To disentangle causality from correlation in the relationship between BMI and labour market outcomes (i.e. the endogeneity problem), in the literature several empirical strategies are tried with alternative identification strategies. For example, Cawley (2000 & 2004) uses the BMI of 'biological' family members that is including parents', siblings' and children' BMI as instrument of individuals BMI. In another work, Cawley *et al*. (2005) use sibling weight to instrument of individual weight. Morris (2006) adopts the average BMI and prevalence of obesity across individuals living in the same health authority area as instruments. Greve (2005) uses information on whether the individuals' parents have ever taken medication related to obesity or obesity related diseases (namely hypertension and Type 2 diabetes). Lundborg *et al.* (2007) choose as instruments the presence of the other obese persons in the household, being an oldest child, and having sisters only. To solve the endogeneity problem, D'Hombres and Brunello (2007) consider the biological BMI (computed as average of all household member's BMI) as an instrument of individual's BMI.

To overcome the difficulty of finding suitable instruments, Sousa (2005) uses a propensity score matching approach. However, since this procedure implies to find comparable individuals within the same dataset it might lead to reduce enormously the sample size. A similar problem is found by Behrmen and Rosenzweig (2001) and Conley and Glauber (2007) when using information on siblings and twins to remove the common household effect due to both genetic and non-genetic factors, given that the number of households with at least two children living in is limited and, therefore, it may create problems of representativeness. In recent works, Cawley (2004, 2005) acknowledges the problem that there exists the possibility that a substantial part of the genes responsible for obesity are also responsible for other factors that affect labour market outcomes or other kind of unobserved characteristics. Since the current knowledge on which particular genes are responsible for obesity and other factors related with wages is too scant, researchers are rather suspicious about the validity of these instruments.

Sargent and Blanchflower (1996), Gortmaker et al. (1993), Averett and Korenman, (1996) address reverse causality by replacing the contemporaneous BMI with its lagged value. However, the validity of this strategy relies on the hypothesis of the independence between the lagged BMI and the residual. Though the independence between the lagged BMI and the residual assumption is rather strict, however, in the present study, due to lack of the genetic information (e.g. sibling information), in particular, in the data set, I also follow such a strategy, that is, I use the individual's own BMI of 1991 as an instrument of his or her BMI of 2000. Based on the BMI scores the categorical variables are also generated accordingly. Since there is a nine-year lag period, one may deduce that the dependency between the lagged BMI and the residual may not be so problematic.

Within the 2SLS procedure, in the first stage, the individuals' own BMI of 2000 (i.e. labelled as BMI2000) is regressed on all other exogenous variables including the instrumental variable (i.e. BMI91) via OLS. The second stage involves regressing again by OLS, the original dependent variable (i.e. in our case log of hourly wage) on the predicted dependent variable of the first stage plus the selected control variables. The predicted values of the first stage regression are a linear projection of exogenous regressors. Employing the predicted values in the second stage then basically has the effect of "purging" the correlation of the BMI variable with the disturbance, which thus allows consistent estimation of γ in equation (1).

4 Estimation Results

In this section I report the results of the empirical analyses I have undertaken. For different specifications (i.e. POLS, FE, and IV), I estimate two different models: with the same covariates, the first model considers weight (i.e. BMI) as a continuous variable and the second model includes weight as a categorical variable. Models are also estimated for all individuals and for males and females separately.

4.1 Pooled Ordinary Least Square (POLS) Results

Table 4 reports the results of the BMI and other included covariates obtained from OLS regressions for the pooled sample for all individuals and for males and females separately.[4] Notice that, employing Ramsey RESET test, I have tested for functional form misspecifications for alternative specifications. It seems that all separate models, that is, for males and females, pass the RESET test.[5] The adjusted R^2 indicates a reasonable fit of the models.

[4] To account for dependence of repeated observations of an individual, and to get standard errors robust to heteroskedaticity, the STATA option 'cluster' and 'robust' are used respectively.

[5] For separate specifications for males (F=1.73; p=0.1768) and females (F=2.25; p=0.106) the models pass the RESET test at the 1% level. When considering BMI as categorical variables the test statistics are: F=1.723; p=0.180

As seen in Table 4, all included variables are highly statistically significant. The estimated coefficient of BMI is negative and statistically significant at the 10% level which implies that for all individuals on average a one unit increase in the BMI score is associated with a decrease in hourly wages of 0.2%. The BMI coefficient is negative and statistically significant at the 1% level for female but not for men thus suggesting the existence of a wage penalty only for women (of around 0.5%) in Sweden.

Table 4: Pooled Ordinary Least Square Estimates: weight (BMI score) considered as a continuous variable

Predictors	All Individuals		Males		Females	
	Coefficient	p-value	Coefficient	p-value	Coefficient	p-value
BMI	-0.0020* (0.0011)	0.054	0.0016 (0.0018)	0.384	-0.0046*** (0.0012)	0.000
Education	0.0411*** (0.0016)	0.000	0.0444*** (0.0024)	0.000	0.0364*** (0.0019)	0.000
Experience	0.0176*** (0.0010)	0.000	0.0208*** (0.0016)	0.000	0.0143*** (0.0012)	0.000
Experience squared	-0.0002*** (0.0000)	0.000	-0.0003*** (0.0000)	0.000	-0.0002*** (0.0000)	0.000
Male	0.1909*** (0.0074)	0.000	-----		-----	
Alone	-0.0521*** (0.0078)	0.000	-0.0698*** 0.0120	0.000	-0.0248** (0.0101)	0.014
Good health	0.0512** (0.0203)	0.012	0.0900*** (0.0322)	0.005	0.0208 (0.0237)	0.380
Fair health	0.0118 (0.0214)	0.581	0.0358 (0.0339)	0.290	-0.0099 (0.0251)	0.692
Smoking	-0.0188** (0.0078)	0.017	-0.0381*** (0.0122)	0.002	-0.0063 (0.0098)	0.520
Year2000	0.1724*** (0.0055)	0.000	0.1622*** (0.0084)	0.000	0.1853*** (0.0071)	0.000
Number of individuals (n)	4,325		2,211		2,116	
Number of observation (N)	6,277		3,181		3,096	
Adjusted R^2	0.396		0.344		0.359	

Notes: Robust standard errors are given in parentheses.
*, ** and *** represent p<0.10, p<0.05 and p<0.01 respectively.

for males and F=1.66; p=0.190 for females. Therefore, it can be deduced that there is no evidence of functional misspecifications while considering males and females separately.

The coefficient values of the other included covariates also show expected signs. For example, the coefficient of education is positive and significant at the 1% level for all individuals (0.041) and for both males and females.[6] A one year increase in schooling is associated with an increase in wages of about 4.4% for males and 3.6% for females. It seems that level of experience is significantly associated with higher wages with decreasing marginal returns to experience (the coefficient of experience squared is negative and significant) for both the males and the females. Controlling for other covariates in the model, a one year increase in experience is associated with an increase in hourly wages of 2.2% (0.0208 - 2*0.0003) for males and about 1.4% for females.[7]

A single person earns about 5% lower hourly wages than those who are married or live with their partner. The coefficient value of male is 0.1909 which implies that males earn on average about 20% higher hourly wages than women. The coefficient value of good health is positive (0.0512) and statistically significant at the 5% level for all individuals. It implies that healthy individuals earn about 5% higher wages per hour compared to individuals with bad health. The coefficient value of good health is positive for both men and women but statistically significant only for men at 1% level. It is evident that a good healthy man earns 9% higher wages than a man with bad health. The coefficient value of smoking is negative (-0.0188) and statistically significant at the 5% level for all, which implies that the smokers on average have 2% lower wages per hour compared to non-smokers. Smoking status seems to produce a significant wage penalty for the males (at the 1% level) but not for the females. A male smoker on average earns about 4% lower wages than the non-smokers.

Table 5 presents the results of the POLS estimates where I categorised individuals' BMI score into four categorical variables and where normal weight is considered as the base case. As

[6] I have also classified the education variable into four categories: Pre-secondary education (9 years of education or less); Short secondary education (10-11 years education); Secondary education (12-13 years); University education (more than 13 years of education). Including these categorical variable in the wage equation, I have also run the regressions, however, the coefficient of the prime predictor, BMI shows the same sign with similar magnitudes. The results are not reported here.

[7] To see whether the association between wage and obesity measure varies with age, instead of experience and experience squared, I have included age dummies in the models. I have generated five age dummies (i.e. age between 18-24, age between 25-34, age between 35-44, age between 45-55, and age between 56-65) and consider age between 18-24 as the base case. However, regarding the coefficient of obesity measure, for all specifications, the conclusions are the similar as the previous (i.e. with experience and experience squared).

seen in table 5, comparing with the normal weight (base case) the coefficients on underweight are negative but not significant for both the genders. There is no wage penalty observed for the overweight or obese males, however, seems to be negatively and significantly associated for the females. On average, an overweight female's wage earnings is 2.5% lower than a normal weight women. The wage penalty is significantly higher for obese females: an obese woman earns more than 6% lower wages than a normal weight woman.

Table 5: Pooled Ordinary Least Square Estimates: weight considered as categorical variables

Predictors	All Individuals		Males		Females	
	Coefficient	p-value	Coefficient	p-value	Coefficient	p-value
Under weight	-0.0153 (0.0252)	0.054	-0.0425 (0.0908)	0.640	-0.0328 (0.0265)	0.217
Over weight	0.0003 (0.0081)	0.972	0.0153 (0.0115)	0.182	-0.0249** (0.0104)	0.017
Obese	-0.0366** (0.0131)	0.005	0.0001 (0.0226)	0.995	-0.0618*** (0.0146)	0.000
Education	0.0411*** (0.0016)	0.000	0.0445*** (0.0024)	0.000	0.0365*** (0.0019)	0.000
Experience	0.0175*** (0.0010)	0.000	0.0208*** (0.0015)	0.000	0.0139*** (0.0012)	0.000
Experience squared	-0.0002*** (0.0000)	0.000	-0.0003*** (0.0000)	0.000	-0.0002*** (0.0000)	0.000
Male	0.1867*** (0.0074)	0.000	-----		-----	
Alone	-0.0516*** (0.0078)	-6.58	-0.0691*** (0.0120)	0.000	-0.0245** (0.0101)	0.015
Good health	0.0499** (0.0202)	2.47	0.0894** (0.0322)	0.006	0.0188 (0.0234)	0.421
Fair health	0.0109 (0.0213)	0.51	0.0360 (0.0339)	0.288	-0.0110 (0.0248)	0.658
Smoking	-0.0184** (0.0079)	-2.34	-0.0381*** (0.0122)	0.002	-0.0059 (0.0098)	0.548
Year2000	0.1716*** (0.0055)	31.11	0.1618*** (0.0080)	0.000	0.1838*** (0.0070)	0.000
Number of individuals (n)	4,325		2,211		2,116	
Number of observation (N)	6,277		3,181		3,096	
Adjusted R^2	0.396		0.344		0.360	

Notes: Robust standard errors are given in parentheses.

*, ** and *** represent p<0.10, p<0.05 and p<0.01 respectively.

Notice that, as seen in table 5, the sign and the magnitudes of other covariates in the models are found to be similar as to those for the model that considered BMI as a continuous variable (see table 4).

It could be interesting to study whether the obesity penalty has developed over time. To consider the issue, I run the regressions including interaction of the 2000 year dummy and the BMI variable, however, the interaction term is highly insignificant with negative value (while BMI use as continuous variable) and is positive for obesity coefficient (categorical variable). The magnitudes of coefficients are also very low. These results suggest that there is no significant difference in wage penalty in 2000 compared to 1991. Moreover, including this interaction term in the models does not seem to affect the magnitude and significance level of the individual obesity measure considerably.

Overall, the POLS estimates suggest that a more educated, experienced, healthy and male worker in 2000 and with non-smoking status earns significantly higher wages than their counterparts. After controlling for other covariates, it seems that a higher BMI score is negatively and significantly associated with the female worker's wage earnings and that an overweight or obese woman earns significantly lower wages than a normal weight woman.

4.2 Fixed Effects (FE) Estimates

Based on the POLS estimates, despite the inclusion of controls, there remains a significant wage gap between obese and non-obese females. To isolate the variation that is causally due to weight, unobserved heterogeneity needs to be taken into account. The parameters have been estimated using the fixed-effects methods as the random-effects models fails a Hausman test[8], probably due to the correlation between the independent variables and the error term.

[8] Models for all individuals $\chi 2$ =78.49 (p=0.000), for males $\chi 2$ =43.89 (p=0.000) and for females $\chi 2$ =32.40 (p=0.000).

Tables 6 (considers BMI as a continuous variable) and 7 (considers BMI as a categorical variables) report the results from the fixed effects (FE) models for all individuals, and for males and females separately.

As observed in table 6, controlling for unobserved individual effects, the BMI variable is no longer significant for all individuals and as with the POLS models not significant for the males as well. Though the BMI score is not highly significantly associated with females wage as was the case with POLS, it is, however, still significant at the 10% level. Notice that, controlling for observable covariates and unobserved individual effects, the absolute magnitude of the coefficient is, in fact, slightly increased. On average, a one unit increase in the BMI score is associated with more than 0.6% lower female wage earnings (for POLS the figure is 0.46%, see table 4).

Table 6: Fixed Effects Estimates: weight considered as a continuous variable (i.e. BMI score)

Predictors	All Individuals		Males		Females	
	Coefficient	p-value	Coefficient	p-value	Coefficient	p-value
BMI	-0.0040 (0.0026)	0.127	-0.0005 (0.0042)	0.905	-0.0063* (0.0032)	0.053
Education	0.0056 (0.0480)	0.907	0.0520 (0.0826)	0.529	-0.0207 (0.0571)	0.717
Experience	0.0098 (0.0479)	0.838	0.0550 (0.0826)	0.505	-0.0163 (0.0567)	0.775
Experience squared	-0.0004*** (0.0000)	0.000	-0.0005*** (0.0000)	0.000	-0.0003*** (0.0000)	0.000
Alone	-0.0233* (0.0123)	0.057	-0.0063 (0.0182)	0.729	-0.0387** (0.0164)	0.018
Good health	0.0358 (0.0280)	0.200	0.0725* (0.0405)	0.074	-0.0062 (0.0384)	0.871
Fair health	0.0519* (0.0281)	0.065	0.1044** (0.0409	0.011	-0.0020 (0.0382)	0.959
Smoking	-0.0019 (0.0149)	0.897	-0.0184 (0.0223)	0.408	0.0131 (0.0197)	0.506
Year2000	0.3387 (0.4307)	0.432	-0.0361 (0.7432)	0.961	0.5376 (0.5110)	0.293
Number of individuals (n)	4,325		2,211		2,116	
Number of Observation (N)	6,277		3,181		3,096	

Note: *, ** and *** represent p<0.10, p<0.05 and p<0.01 respectively.

As seen in table 7, similar findings are also observed when BMI is considered as a categorical variable. The effect of obesity is similar for females (negative and significant at the 10% level). The absolute magnitude of the obesity coefficient is found to be rather lower for the FE model. Compared with a woman with normal weight, an obese woman is wage earnings is slightly lower than 6% (the magnitude of the coefficient is found to be -0.0618 and significant at the 1% level for POLS, see table 5).

Table 7: Fixed Effects Estimates: weight considered as categorical variables

Predictors	All Individuals		Males		Females	
	Coefficient	p-value	Coefficient	p-value	Coefficient	p-value
Under weight	0.0278 (0.0344)	0.420	0.1877* (0.1113)	0.092	-0.0003 (0.0344)	0.993
Over weight	0.0009 (0.0121)	0.939	0.0163 (0.0172)	0.342	-0.0234 (0.0171)	0.172
Obese	-0.0210 (0.0233)	0.368	0.0148 (0.0359)	0.681	-0.0589* (0.0302)	0.051
Education	0.0082 (0.0481)	0.864	0.0552 (0.0827)	0.504	-0.0237 (0.0573)	0.679
Experience	0.0121 (0.0480)	0.800	0.0575 (0.0826)	0.487	-0.0195 (0.0570)	0.732
Experience squared	-0.0004*** (0.0000)	0.000	-0.0005*** (0.0000)	0.000	-0.0003*** (0.0000)	0.000
Alone	-0.0227* (0.0123)	0.064	-0.0052 (0.0182)	0.776	-0.0380** (0.0164)	0.021
Good health	0.0367 (0.0280)	0.190	0.0731* (0.0405)	0.071	-0.0055 (0.0384)	0.886
Fair health	0.0522* (0.0281)	0.063	0.1041** (0.0409)	0.011	-0.0021 (0.0383)	0.957
Smoking	-0.0011 (0.0149)	0.939	-0.0189 (0.0222)	0.394	0.0148 (0.0197)	0.453
Year2000	0.3109 (0.4317)	0.471	-0.0621 (0.7435)	0.933	0.5594 (0.5130)	0.276
Number of individuals (n)	4,325		2,211		2,116	
Number of Observation (N)	6,277		3,181		3,096	

Note: *, ** and *** represent p<0.10, p<0.05 and p<0.01 respectively.

After controlling for unobserved individual heterogeneity, though, it seems that the under weight males earn around 19% higher wages (significant at the 10% level) compared to the normal weight males. However, one needs to be cautious about the findings as very few males belong to this category (i.e. in the selected sample only 0.004% of the males are underweight).

As seen in tables 6 and 7, in the FE models, most of the control variables are found to be insignificant. In particular, controlling for unobserved individual level heterogeneity, though the absolute magnitude of the education variable is found to be comparable for the males, however, the variable no longer significantly affects wages for both genders. [9] A similar finding is also observed for experience, but experience squared is negatively and significantly associated with wages for all FE specifications. The wage penalty (around 4%) is still significantly higher for single women but not for single men. Good or fair health status positively and significantly affects male wage earnings, but do not significantly influence females' wages. Smoking habit seems no longer significantly associated with neither male nor female wage earnings.

Based on the FE results, it appears that unobserved heterogeneity plays an important role in getting statistically significant effects of different covariates on individuals' wage earnings. Nevertheless, even after controlling for unobserved heterogeneity, it seems that females' weight is an important factor in influencing their wage earnings.

4.3 Instrumental Variable (IV) Estimates

We have discussed the estimation procedure of the instrumental variables estimator earlier and highlighted that an IV must satisfy two requirements: it must be correlated with the endogenous explanatory variable and uncorrelated with the error. I have considered individuals' earlier year BMI (i.e. BMI1991) as an instrument of his/her current BMI (i.e. BMI2000). The first condition of my chosen instrument has been tested by employing OLS regression where BMI2000 has been considered as a dependent variable and BMI1991 as an independent variable with all other control variables. The full estimated results are given in Appendix A. It can be seen that BMI1991 is positively and highly significantly associated with BMI2000 (t-value=41.50). This result may suggest that my chosen instrument passes the relevancy condition of the instrument.

[9] Since I have a very short panel of only two observations, it seems that the FE models do not work properly, probably because, for instance, few individuals change education between 1991 and 2000.

Since I have used a single exogenous variable as instrument of BMI, therefore the model is found to be exactly identified, hence it is not possible to test for the over-identification restriction (to examine indirectly whether the instrument is uncorrelated with the error).

Tables 8 and 9 show the IV estimation results. As seen in table 8, I get similar results as with the POLS estimation approach. In particular, an increase in the BMI score significantly causes a wage penalty for all individuals and for females, but not for males. The absolute magnitude of the coefficients seems to be comparatively higher for the IV estimates than for the POLS, but rather similar to the FE estimates. For example, according to the IV estimates the wage penalty is 0.6% for females, whereas the figure is 0.46% for POLS. As expected the standard errors are rather higher for the IV estimates, however, the BMI coefficient is still significant at the 1% level for females.

Table 8: Instrumental Variable Estimates: weight considered as a continuous variable (i.e. BMI)

Predictors	All Individuals		Males		Females	
	Coefficient	p-value	Coefficient	p-value	Coefficient	p-value
BMI 2000	-0.0039** (0.0020)	0.048	-0.0008 (0.0035)	0.820	-0.0059*** (0.0021)	0.005
Education	0.0430*** (0.0025)	0.000	0.0520*** (0.0038)	0.000	0.0328*** (0.0028)	0.000
Experience	0.0077*** (0.0029)	0.007	0.0082* (0.0048)	0.087	0.0072** (0.0032)	0.021
Experience squared	-0.0000 (0.0000)	0.163	0.0000 (0.0001)	0.609	-0.0001* (0.0000)	0.070
Male	0.2100*** (0.0118)	0.000	-----		----	
Alone	-0.0530*** (0.0140)	0.000	-0.0575** (0.0218)	0.008	-0.0380** (0.0170)	0.026
Good health	0.0617* (0.0320)	0.054	0.1251*** (0.0486)	0.010	0.0053 (0.0362)	0.882
Fair health	0.0264 (0.0340)	0.043	0.0725 (0.0524)	0.167	-0.0194 (0.0382)	0.631
Smoking	-0.0246* (0.0138)	0.075	-0.0470** (0.0225)	0.037	-0.0133 (0.0164)	0.418
Number of individuals (n)	1,952		971		981	

Notes: Robust standard errors are given in parentheses.
*, ** and *** represent p<0.10, p<0.05 and p<0.01 respectively.

The sign and absolute size of coefficients of the control variables are also similar magnitude in the IV estimates compared with the POLS estimation. For example, education and experience positively and significantly influence individuals' wages and the associations are higher for males than for females. Living alone is significantly associated with lower wage earnings for both males and females. Good health status has a positively, but smoking has a negative influence on males wage earnings, but the corresponding coefficients for females are not significant.

Table 9: Instrumental Variable Estimates: weight considered as a categorical variable

Predictors	All Individuals		Males		Females	
	Coefficient	p-value	Coefficient	p-value	Coefficient	p-value
Underweight 2000	-0.2946 (0.2257)	0.192	-0.9058 (1.160)	0.435	-0.3357 (0.2355)	0.154
Overweight 2000	0.0204 (0.0417)	0.624	0.0559 (0.0550)	0.310	-0.0345 (0.0593)	0.558
Obese 2000	-0.0711*** (0.0292)	0.013	-0.0749 (0.0597)	0.210	-0.0763*** (0.0266)	0.004
Education	0.0434*** (0.0025)	0.000	0.0524*** (0.0040)	0.000	0.0332*** (0.0030)	0.000
Experience	0.0079*** (0.0029)	0.006	0.0091* (0.0047)	0.055	0.0071** (0.0030)	0.022
Experience squared	-0.0000 (0.0000)	0.135	-0.0000 (0.0000)	0.471	-0.0000* (0.000)	0.075
Male	0.1965*** (0.0144)	0.000				
Alone	-0.0520 (0.0140)	0.000	-0.0460 (0.0224)	0.040	-0.0400 (0.0171)	0.020
Good health	0.0616* (0.0320)	0.054	0.1141** (0.0498)	0.022	0.0085 (0.0370)	0.818
Fair health	0.0270 (0.0341)	0.429	0.0581 (0.0540)	0.282	-0.0118 (0.0393)	0.764
Smoking	-0.0203* (0.0140)	0.147	-0.0411** (0.0228)	0.071	-0.0094 (0.0170)	0.581
Number of individuals (n)	1,952		971		981	

Notes: Robust standard errors are given in parentheses.
*, ** and *** represent p<0.10, p<0.05 and p<0.01 respectively.

Table 9 shows the IV estimation results when BMI is considered as a categorical variable. As seen in table 9, for this specification, I also get similar results as with the POLS estimation approach. In particular, the negative effect of obesity on wage earnings is similar for females. The absolute magnitude of the obesity coefficient is found to be rather higher for the IV estimate than for the POLS or FE model. Compared with a woman with normal weight, an obese woman is wage earnings is around 8% (the magnitude of the coefficient is found to be -0.0618 and significant at the 1% level for POLS, see table 5) lower. The absolute magnitude of the coefficients seems to be comparatively higher for the IV estimates than for the FE estimates. As observed in the previous specification, standard errors are rather higher for IV estimates. However, the obesity dummy is still significant at the 1% level for females.

It is important to note that the IV estimate in fact gives an estimate for men that is the same as for women, even if not statistically significant. It seems that the obesity penalty for men is great in an economic sense but is estimated with a bad precision.

Notice that, as seen in table 9, the sign and the magnitudes of other covariates in the models are found to be similar to those obtained that with the model considered BMI to be a continuous variable (see table 8).

4.4 Sensitivity analysis

Based on the three alternative estimation procedures, a number of sensitivity analysis results pertaining to the different selected samples are presented in Table 10. In the sensitivity analysis, I first consider a restricted sub-sample of ages between 25 and 55 years. This restricted age group focuses the analysis on the more homogenous group of prime age workers that have for most parts completed their education. This could be so important as to avoid complications concerned with issues of life time labour supply such as full time education and early retirement. This restricted analysis could facilitate the comparison of my results with other studies that generally employ a restricted sample to prime age individuals both for males and females separately (Atella et al., 2007). As seen in table 10, for this sub-sample, in general, for every alternative estimation approach, I get similar findings as earlier, i.e. an increase in the BMI score negatively and significantly affect females' wages earnings and obese females significantly earn less than normal weight females. Nevertheless, it is worth mentioning that compared with the respective full sample analyses (i.e. the base line analyses) the absolute magnitudes of the BMI and obesity coefficients seemed to be rather higher for all estimation procedures. Moreover, according to the IV estimates, it is shown that the wage penalty is around 11% for obese males and it is significant at the 10% level.

In the second sensitivity analysis, I consider another restricted sample, where I drop individuals with extremely low and high values of the hourly wage. Excluding individuals with extreme wage earnings could also be important as they may produce noise in the estimation results. It is observed that the estimated coefficients for females (both BMI and obesity dummy) are now significant at the 5% level for the FE approach (the coefficients have been significant at the 10% level for full sample); however, as seen in table 10, I do not obtain any considerable divergences in the magnitudes of the estimated coefficients.

Finally, I also attempt to re-estimate the baseline models considering balanced panel individuals (same individuals responded both years). Since I have considered individuals' BMI scores in the year 1991 as instruments of their BMI in the year 2000, therefore, the IV estimation procedure has selected only the balanced sample individuals. Therefore, this sub-set of the data may also be important for the precise comparisons of the estimates obtained different alternative estimation approaches. However, again as observed in table 10, I do not find any notable differences in the absolute magnitudes (or the significance levels) of the estimated coefficients for the balanced sample as well.

Table 10: Sensitivity analysis of the association between BMI (and also categorical variable) and log hourly wages by different specifications

	Pooled OLS	Fixed-Effects	IV
Exclude individuals with age below 25 years and above 55 years			
All individuals: BMI	-0.0031***(0.0012)	-0.0024 (0.0031)	-0.0044** (0.0021)
Males: BMI	-0.0008 (0.0020)	-0.0044 (0.0047)	-0.0035 (0.0039)
Females: BMI	-0.0058*** (0.0013)	-0.0081** (0.0041)	-0.0048 ** (0.0023)
All			
Under weight	-0.0040 (0.0286)	-0.0156 (0.0400)	-0.2512 (0.2508)
Over weight	-0.0052 (0.0286)	0.0028 (0.0141)	0.0211 (0.0508)
Obese	-0.0388 *** (0.0141)	0.0200 (0.0279)	-0.0852***(0.0315)
Males			
Under weight	-0.1549** (0.0718)	-0.0221 (0.1346)	-0.9 994 (1.2349)
Over weight	-0.0138 (0.0130)	0.0200 (0.0198)	0.0265 (0.0647)
Obese	-0.0027 (0.0250)	0.0400 (0.0419)	-0.1145* (0.0679)
Females			
Under weight	0.0033 (0.0410)	0.0065 (0.0410)	-0.1590 (0.2642)
Over weight	-0.0342*** (0.0202)	0.0242 (0.0202)	0.0193 (0.0827)
Obese	-0.0662*** (0.0371)	0.0416 (0.0371)	-0.0642** (0.0307)
Exclude individuals with the hourly wage less than SEK 50 and above SEK 500			
All individuals: BMI	-0.0021** (0.0010)	-0.0039 (0.0026)	-0.0040** (0.0019)
Males: BMI	0.0013 (0.0017)	0.0000 (0.0042)	-0.0006 (0.0035)
Females: BMI	-0.0046*** (0.0011)	-0.0066** (0.0031)	-0.0063*** (0.0021)
All			
Under weight	0.0053 (0.0215)	0.0472 (0.0339)	-0.2913 (0.2232)
Over weight	0.0002 (0.0078)	-0.0043 (0.0119)	0.0278 (0.0416)
Obese	-0.0340*** (0.0123)	-0.0208 (0.0228)	-0.0709** (0.0287)
Males			
Under weight	-0.0470 (0.0896)	0.1808 (0.1093)	-0. 8714 (1.1338)
Over weight	0.0124 (0.0110)	0.0112 (0.0170)	0.0610 (0.0548)
Obese	0.0008 (0.0223)	0.0202 (0.0356)	-0.0704 (0.060)
Females			
Under weight	-0.0072 (0.0217)	0.0225 (0.0337)	-0.3356 (0.2332)
Over weight	-0.0212** (0.0099)	-0.0283* (0.0167)	-0.0259 (0.0584)
Obese	-0.0586*** (0.0130)	-0.0627** (0.0292)	-0.0773*** (0.0264)
Consider individuals who responded in both years (balanced panel)$^{\phi}$			
All individuals: BMI	-0.0030** (0.0014)	-0.0040 (0.0026)	-0.0039** (0.0020)
Males: BMI	0.0017 (0.0025)	-0.0005 (0.0042)	-0.0008 (0.0036)
Females: BMI	-0.0061***(0.0015)	-0.0063* (0.0032)	-0.0059*** (0.0021)
All individuals			
Under weight	-0.0032 (0.0308)	0.0278 (0.0344)	-0.2947 (0.2256)
Over weight	0.0017 (0.0103)	0.0009 (0.0121)	0.0205 (0.0417)
Obese	-0.0367** (0.0159)	-0.0210 (0.0233)	-0.0711*** (0.0288)
Males			
Under weight	0.0242 (0.1355)	0.1877* (0.1113)	-0.9059 (1.1611)
Over weight	0.0291* (0.0149)	0.0163 (0.0172)	0.0559 (0.0550)
Obese	0.0085 (0.0288)	0.0148 (0.0359)	-0.0749 (0.0597)
Females			
Under weight	-0.0263 (0.0304)	-0.0003 (0.0344)	-0.3357 (0.2356)
Over weight	-0.0375***(0.0127)	-0.0234 (0.0171)	-0.0346 (0.0589)
Obese	-0.0675*** 0.0166)	-0.0589* (0.0302)	-0.0763***(0.0266)

Notes: All models are also controlled for other covariates reported in the earlier tables.
Robust standard errors are given in parentheses.
*, ** and *** represent p<0.10, p<0.05 and p<0.01 respectively.
$^{\phi}$ For this sub-set of the data the IV estimation results are the same as given in tables 9 and 10. Yet, for the sake of the comparisons of the results between POLS and FE estimation approaches, I further provide the IV results.

5 Discussion and Conclusions

5.1 Discussion

Research suggests that obesity has potentially important effects on labour market outcomes, particularly on individuals' wage earnings. There are several theoretical underpinnings that explain why might have obesity effects on wages. One explanation could be that by lowing productivity or because of workplace discrimination obesity lowers wages. In contrast, the second reason would be that low wages cause obesity and the last explanation would be that unobserved variables cause both obesity and low wages (Cawley, 2004). In general, though existing empirical studies have linked obesity to wages, the validity of their estimations and findings yet remain questionable, particularly, due to potential weaknesses in the strategies employed to control for the endogeneity of obesity. Besides, major problems may exist which introduce biases in the estimated effects, e.g. unobserved variables, measurement errors.

To shed some light on the question whether a society with a low prevalence of obesity has wage penalties for obese people, in this thesis, I have investigated the relationship between obesity and wage earnings based on Swedish longitudinal data. To estimate the effects empirically, I have tried alternative econometric specifications, namely pooled OLS, fixed-effects and instrumental variable techniques.

My estimated results suggest, firstly, an individuals' weight negatively and significantly affects female but not male wages, implying that a wage earnings impact of obesity is evident only for females in Sweden. Secondly, on average a one unit increase in BMI score is associated with around 0.5% lower females wages, and after controlling for individual level heterogeneity (via FE estimation), the magnitude of the association seems rather higher in that the females' wage penalty is more than 0.6%. This significant negative effect (with similar magnitude) of an increase in BMI is also evident for females in IV estimation approach. Thirdly, with different estimation approaches, the results show that an obese woman's wage earnings is around 6-8% lower than a normal weight woman. Fourthly, the significant negative wage penalties for obese females are also robust for different selected sub samples (e.g. prime working age group) as evidenced by the sensitivity analysis.

It should be noticed that the estimated results are quite comparable with those of other studies, conducted for other societies. For example, using cross sectional data from the 1982 National Longitudinal Study of Youth (NLSY), Register and Williams (1990) find that obesity reduces females' wages by 12% but has no significant effects for males. Using the same data, Cawley (2004) also finds a wage penalty for obese white female workers, of 9% in the USA.

Overall, though a wage penalty for obese male is not statistically significant in any estimation approaches, however, the IV estimate indeed gives an estimate for male that is the same as for female. Moreover, in the sensitivity analysis, when I consider a subset of the sample ages between 25 and 55 (i.e. the more homogenous group of prime age workers), I find that the wage penalty is more than 11% for obese males and it is significant at the 10% level. Hence, given that we have faith in the IV estimates (as I have) then, at least, I would say that I can not exclude the possibility that men and women have a similar obesity penalty, which is also found in a recent field experimental study conducted by Rooth (2007) in Sweden.

Regarding other covariates, in particular, as the returns to education, the POLS models receive estimates that seem to be trustworthy. In particular, using the Swedish Level of Living Survey data for the year 1991 and including a set of control variables, Isacsson (1999) finds that the estimate of the return to schooling is 4.5% in the population at large in Sweden. Furthermore, using the Swedish Twin Registry data, the author also shows that the estimate of the return to schooling is found to be 4.6% in the sample of MZ twins and 4.7% in the sample of DZ twins.

My study is not without its limitations. First, though I have included most of the controls in my estimated wage specifications that are also considered in other studies, there may be other covariates (e.g. information on training, tenure, household composition, sick leave etc), those also need to be controlled for. Due to lack of information of some of these covariates in the current data set and/or complications in calculating some variables (e.g. to estimate tenure, I need to have information on individual's full work histories) means that I am unable to adjust for these covariates.[10] Secondly, wage penalties might be different in different parts of the wage

[10] However, even if the omitted covariates are correlated with BMI, excluding them would still be picked up by the BMI variable and seems not to be problematic, since in IV estimates, I instrumenting BMI2000 with own BMI in 1991.

distribution and exploring the effect of obesity on wages in different parts of the wage distribution might provide better information on the relations. Therefore, it may be crucial to explore the role of obesity at different points of the wage distribution, as it could be that obesity is related to individual wages differently at the bottom or at the top of the wage distribution (Atella et al., 2007). Due to time limitation, I have not considered the wage distributions in my analyses, however, it could be a worth while exercise for the future. Moreover, since I have used a short panel of only two observations, therefore, care should be needed in interpreting the findings of the FE estimations. Finally, regarding IV estimates, due to using a single instrument, I am not able to confirm that the instrument I have chosen is orthogonal with the error term in the wage equation (in applied research it is usual to test the validity condition of the instruments indirectly via testing the over-identification restriction), hence I may need to be cautious in interpreting the relationship between obesity and wage as causal.

5.2 Conclusions

Using the different estimation techniques, the thesis finds a strong statistically significant wage penalty for an obese female over a normal weight female. The findings are robust to alternative specifications and sub-samples examined in the sensitivity analysis. The thesis concludes that higher weight or obesity may cause a wage penalty for females but not for males in a low prevalence obese society, such as Sweden. Nonetheless, since the effect of obesity on earnings is important to consider for policy and decision making both at the individual and at the societal level further exploration is needed. In particular, questions could be raised: why does a weight effect appear to lower wages for women (or white women) but not for other groups? What factors influence the differential effects of obesity on wages across race-ethnic and sex categories? It seems that the relationship between obesity and wages yet remains elusive and future research is needed to explore the issues further. Understanding the pathways and the true effect of obesity on labour market outcomes may be helpful in formulating better public policy for combating the obesity epidemic.

References

Attella, V., Pace, N. and Vuri, D. 2007. Wages and weight in Europe: Evidence using quantile regression model. Centre for Household, Income, Labour and Demographic Economics (CHILD) working paper n.23/2007. University of Rome, Tor Vergata.

Averett, S and Korenman, S. 1996. The economic reality of the beauty myth. *Journal of Human Resources* 31(2): 304-30.

Baum C.L, and Ford, W. 2004. The wage effects of obesity: a longitudinal study. *Health* Economics, 13: 885-899.

Baum C. L. 2008. The effects of cigarette costs on BMI and obesity. *Health Economics,* DOI: 10. 1002/hec. 1340.

Becker, G. S. 1965. *Human Capital: A Theoretical and Empirical Analysis with Special Reference to Education*. New York: National Bureau of Economic Research.

Becker, G. and Murphy, K. 1988. A theory of rational addiction. *The Journal of Political Economy*, 96: 675-700.

Behrman, J., and Rosenzweig, M. 2001. The returns to increasing body weight. PIER Working Paper, No. 11303.

Brune, L. 2007. The smoker's wage penalty puzzle- Evidence from Britain. Institute For Social & Economic Research Working Paper 2007-31. Colchester: University of Essex.

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

Cawley, J. Grabka, M. and Lillard, D. 2005. A comparison of the relationship between obesity and earnings in the U.S. and Germany. *Journal of Applied Social Science Studies* (*Schmollers Jahrbuch*), 125: 119-129.

Cawley, J. 2005. The labour market impact of obesity. in *Obesity, Business, and Public Policy*, Zoltan Acs and Alan Lyles (eds.).

Cawley, J. and DanZiger, S. 2005. Obesity as a barrier to employment and earnings for current and former welfare recipients. *Journal of Policy Analysis and Management*, 24: 727-743.

Conley, D. and Glauber, R. 2007. Gender, body mass, and socioeconomic status. New Evidence from PSID, *Advances in Health Economics and Health Services Research*, Vol. 17. The Economics of Obesity, K. Bolin and J. Cawley, (eds.) Amsterdam: Elsevier.

D'Hombres, B. and Brunello, G. 2005. Does obesity hurt your wages more in Dublin than in Madrid? Evidence from the ECHP. IZA, Discussion Paper No. 1704.

Garcia, J. and Quintana-Domeque, C. 2007. Obesity, employment and wages in Europe. *Advances in Health Economics and Health Services Research*, Vol.17. The Economics of Obesity, K. Bolin and J. Cawley, (eds.). Amsterdam: Elsevier.

Gortmaker, S.L., Must, A., Perrin, J.M., Sobol, A.M. and Dietz, W. H. 1993. Social and economic consequences of overweight among adolescents and young adults. *New England Journal of Medicine*, 329: 1008-1012.

Greve, J. 2005. Obesity and labour market outcomes: new Danish evidence. Unpublished manuscript, Cornell University and Aahus School of Business, Denmark.

Hsiao, C. 2003. *Analysis of Panel Data.* 2nd edition, Cambridge University Press, Cambridge.

Isacsson, G. 1999. Estimates of the return to schooling in Sweden from a large sample of twins. *Labour Economics,* 6: 471–489.

Lundborg, P., Bolin, K. Höjgård, S. and Lindgren, B. 2007. Obesity and occupational attainment among the 50+ of Europe. *Advances in Health Economics and Health Services Research*, vol. 17, The Economics of Obesity, K. Bolin and J. Cawley (eds. Amsterdam: Elsevier.

Mincer, J. 1974. *Schooling, Experience and Earnings*. New York: National Bureau of Economic Research.

Morris, S. 2006. Body mass index and occupational attainment. *Journal of Health Economics*, Vol; 25: 347-364.

Pelkowski, J. and Berger, M. 2004. The impact of health on employment, wages, and hours worked over the life cycle. *The Quarterly review of Economics and Finance*, 44: 102-121.

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

Philipson, T., Dai, C., Helmchen, L. and Variyam, J. 2004. The economics of obesity. A report on the workshop held at USDA's economic research service, E-FAN'.

Register C.A. and Williams D. R. 1990. Wage effects of obesity among young workers. *Social Science Quarterly*, 71: 130-141.

Rooth, D. 2007. Obesity, attraction and differential treatment in hiring- Field experiment. IZA working paper, IZA, Bonn.

Sargent, J. and D. Blanchflower. 1994. Obesity and stature in adolescence and earnings in young adulthood. Analysis of a British birth cohort. *Archives of Pediatrics and Adolescent Medicine*, 148: 681-687.

Sarlio-Lahteenkorva, S. and E. Lahelma. 1999. The association of body mass index with social and economic disadvantage in women and men. *International Journal of Epidemiology*, 28: 445-449.

Sousa, S. 2005. Does size matter? A propensity score approach to the BMI on labour market outcomes, paper presented at ESPE 2005, Paris.

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

Verbeek, M. 2004. *A Guide to Modern Econometrics*, 2nd edition, John Wiley & Sons.

Wardle, J. and Griffth, J. 2001. Socioeconomic status and weight control practices in British adults. *Journal of Epidemiology Community Health*, 55: 185-190.

WHO. 2002. The World Health Report 2002 - Reducing risks, promoting healthy life.

WHO. 2003. Diet, nutrition, and the prevention of chronic diseases. WHO Technical Report 916.

Wooldridge, J. M. 2002. *Econometric Analysis of Cross Section and Panel Data*. The MIT Press, Cambridge, Massachusetts.

APPENDIX

Table A: Results of the first stage IV regression: Dependent variable individual BMI in the year2000 (i.e. BMI2000)

bmi2000	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
bmi91	.9425682	.0227098	41.50	0.000	.89803	.9871064
education	-.0469978	.0167333	-2.81	0.005	-.0798149	-.0141807
exp	-.0155592	.0246529	-0.63	0.528	-.0639081	.0327896
exp2	-.0003056	.0004266	-0.72	0.474	-.0011422	.0005311
male	-.1437388	.0999769	-1.44	0.151	-.3398121	.0523346
alone	-.1430057	.1240994	-1.15	0.249	-.3863877	.1003763
ghealth	-.6759202	.2670524	-2.53	0.011	-1.19966	-.1521806
fhealth	-.2377273	.2904044	-0.82	0.413	-.8072644	.3318097
smoking	-.1817746	.1054797	-1.72	0.085	-.3886398	.0250907
_cons	4.892221	.6715691	7.28	0.000	3.575149	6.209293

```
. test bmi91

 ( 1)  bmi91 = 0

       F(  1,  1942) = 1722.65
            Prob > F = 0.0000
```