

Obesity with labor market outcome : Evidence in U.S.

Hsiang-en, Ho

Abstract

This paper investigates the causation between obesity and incomes in the labor market in order to explore the discrimination against obese. The dataset comes from IPUM's National Health Interview Survey, which is a nationally representative sample investigates the health issue in U.S. since 1963. Results show the positive association between body size and wages, no matter how the models group by personal or household characteristics such as race, working class or industry. The Instrumental Variable method is used for disentangling the causation from the correlation, however, the results present the positive effect whereas passing the tests of endogenous and explanatory power. In conclusion, there might be some other confounding factors mitigate the effect, hence requiring future studies investigating the better control.

1 Introduction

Obesity is one of the most important health issues in the whole world, since it often leads to higher risk of suffering from cardiovascular disease according to World Health Organization. The prevalence of obesity in the United States continues to rise, exceeding 30% in most sex and age groups (Flegal et al., 2010). Moreover, since obesity potentially exacerbates many health problems, such as type 2 diabetes, coronary heart disease and rising incidence of certain forms of cancer (Kopelman, 2000). Each year, approximately 300,000 US adults die due to causes correlated to obesity and diabetes (Mokdad et al., 2003).

Many empirical studies have investigated the effect of obesity on labor market outcome. There are many researches found that obesity has a negative effect on incomes, particularly for women (Pagan and Davila, 1997; Cawley, 2004; Norton and Han, 2008), while others tested the influence on possibility of employment (Garcia and Quintana-Domeque, 2007; Bolin et al., 2008; Frayling et al., 2007). In the previous work, body mass index (BMI) were generally used to measure obesity. Moreover, to evaluate the accurate fatness without calculating muscle or bone, Burkhauser and Cawley (2006) considered the medical literature to better estimate the fatness. On the other hands, body composition such as body fat (BF) and fat-free mass (FFM) are used to disentangle between body fat and fat-free body (Burkhauser and Cawley, 2006).

Obesity affects labor outcome by different reasons. Obesity might reduce one's ability or flexibility in workplace, further deteriorating his or her productivity (Greve, 2008). That is to say obesity limits the amount or quality they can perform, hence leading to the wage penalty due to health problems (Baum and Ford, 2004). Furthermore, the salaries penalty is aroused because the health care might be expensive. As the higher risk of suffering from severe disease, the employers might worry about providing extra health insurance or further assistance. Burkhauser and Cawley (2008) tested the hypothesis that obesity leads to employment disability, and the mixed results conclude that obesity increases the chance of health-related limitations and receive disability-related earnings.

Discrimination against obese people may be one other significant factor attributed to the wages penalty. The discrimination might be caused by compensating wages for higher health costs as previous literatures (Baum and Ford, 2004) mentioned, or just pure discrimination (Han et al., 2011). Discrimination can be interpreted as the category of preference against people with specific characteristics (Becker, 1971). These preferences lead employers to regard these workers "as more expensive than they truly are" (Charles and Guryan, 2008), hence deciding not to hire them rather than contemplating to the abilities or qualifications. That is their personal features other than those really relevant enter into the decision (Arrow et al., 1973). The common types of workplace discrimination including discrimination against racism or ethnicity (Deitch et al., 2003; Elliott and Smith, 2004; Fox and Stallworth, 2005), gender (Badgett, 1995; Weichselbaumer, 2003; Badgett, 2009) or disability (Chan et al., 2005). Specifically, obese people may meet the distaste from the employers in the labor market, since obesity seems to be the signal of lack of self-control and laziness (Sobal and Stunkard, 1989; Puhl and Latner, 2007) which employers may not like.

This paper investigates the causal effect between body size and labour market outcome. To disentangle the causation from correlation, previous studies provided several approaches to cope with this issue, including lagged variable, fixed effect or instrumental variable. This research uses family member's BMI to serve as IV due to the convenience of dataset, allowing to estimate the genetic variation. Controlling the relative covariates helps to eliminate effect from observed characteristic or heterogeneous.

Empirical strategies such as ordinary Least Regression is conducted to evaluate the relationships between the outcomes of interest - body size and the salaries. Besides, there are many factors might influence the income. For instance, the academic degree might affect the labor market outcome significantly. Other personal and individual characteristics are also included in the model such as region, seniority, race, etc.

This paper would use the Instrumental variable strategy to distinguish the causal effect from the correlation since the lagged BMI has insufficient observations in the dataset. Nevertheless, it provides necessary information of characteristic of interviewers' family members, including the outcome of interest - body size. After considering the genetic factor, the earning residual might be composed of genetic factor, nongenetic factor and the residual.

The dataset is collected from IPUM's National Health Interview Survey. The survey is a nationally representative sample which has investigated health issue since 1963. NHIS has included the information of BMI since 1974, but it began to cover personal earning after 1996. Accordingly, this paper adopts the dataset after 1997 to satisfy the requirement. From 1997 to 2018, it interviewed 4,423,611 people in U.S., while both genders keep approximately half of the sample. Furthermore, the BMI is recorded as continuous variable, which would be classified to different categories in order to interpret different groups of body size.

One other main variable of the labor market outcome is the personal total earnings, which is presented by the previous calendar years. In addition, further characteristics of the individuals in labor market outcome also been considered, including occupation, working class and years on main or longest or last job. Furthermore, the BMI is also classified to the different categories. Following the classification from WHO, this study converts the BMI into 8 groups: Very severely underweight(15), Severely underweight(15-16), Underweight(16-18.5), Normal(18.5-25), Overweight(25-30), Moderately obese(30-35), Severely obese(35-40) and Very severely obese(40).

The relationship between earnings and BMI is an inverted U-shaped curve. The trends illustrate that the salaries do not always drop when weight elevates. On the other hand, the relationship between earnings and BMI for women is much smoother than the curve for men, elucidating that the women's wages are much invariant relative men's. On the whole, the inverted U-shaped curve presents the heterogeneity within different classifications of body size.

The results show that the body size has strong correlation with wages, causing the purpose of distinguishing causation from the association to be more difficult. Several models which consideration for different covariates regress incomes on the outcome of interest, the body size, separately. The body size is represents by Body Mass Index (BMI), and is converted to the dummy of whether the employee is obese or not in order to help better estimation. However, the positive association in the result illustrates that income may affect body size of the workers reversely, bringing ambiguous result concluded by the regressions.

This paper also explores the association between these two by grouping by gender and different characteristics such as race, working class or industry, but the correlation can't provide much information of the causation between the real labour market outcome. To buttress the causation between the outcomes of interest, this study chooses IV to be the empirical strategy. As a result, the Two-Stage Least Squares result doesn't lend support to the supposition of the discrimination against specific body size in the workplace. Much regressions pass the tests of endogenous and explanatory power, but the results of positive effect can't strengthen the assumption of wages penalty.

The causation isn't strengthened due to the lack of finding great instrumental variable. Present IV doesn't achieve the requirement to show significantly negative associating. The dataset doesn't include additional available information can be used as IV. There might be some other confounding factors to mitigate the effect, hence require future studies investigating better control.

2 Empirical strategy

Body size is a reflection of lifestyle, which is attributed to several aspects. Pressure might lead one to eat more, hence elevating his or her weight. On the other hand, one in poverty may have no enough money to buy healthy food, while the unhealthy food cause obesity. On the whole, these factors also impact the BMI and make the analysis more difficult.

Previous studies had discussed the food issue. Workers with lower paid or live in poorer regions could only eat cheap fast foods, acquiring little nutrients while getting much fat. That is to say that the body size perhaps has the correlation rather than causal effect, it's important to causality from correlation in the relationship between BMI and labor market outcomes.

Consider the very basic model, including the primary variables only:

$$\log(W_i) = \gamma X_i + \epsilon_i$$

Where the W represents the wages, and X_i implies the primary variable, specifically, the body size. ϵ represents the error term for different individuals. Accordingly, the ordinary least squares (OLS) regression is used for exploring the association.

Besides, there are many factors might influence the income. Most directly, the academic degree might affect the labor market outcome significantly. Other personal and individual characteristics are also included in the model such as region, seniority, race, working class, industry, age, gender, etc. Thus, the model would be like:

$$\log(W_i) = \gamma X_i + \sum_{j=1} \beta_j v_i^j + \epsilon_i$$

Several empirical approaches have been used: lagged BMI (Conley and Glauber, 2005; Norton and Han, 2008), fixed-effects strategies to control the individual differences (Cawley, 2004; Norton and Toohey, 2011) or instrumental variable (Cawley, 2004; Brunello and d’Hombres, 2007; Frayling et al., 2007; Greve, 2008)

First, the lagged specification strategy refers to using interviewers’ BMI when they were younger. The methods depend the past body size to distinguish the nowadays BMI from labor market outcome.

Second, fixed effect helps to control unobserved heterogeneity. In estimating indirect pathways of BMI on wages by education and occupation choice, the measurement for the variable of interest (late teen BMI) precedes the measurement of the dependent variables (Norton and Toohey, 2011).

Third, instrumental variable needs to satisfy exogenous condition in order to serve as an appropriate strategy. The general use of instrumental variable is to utilize family members’ BMI to control the genetic variation. For instance, some study used father’s or mother’s body size as IV. However, one’s specific characteristic might also derived from his or her mother, that is exogenous requirement cannot be satisfied. In more recent work, Cawley (2004) discussed the other factors which may attribute the labour market outcome, such as time discount rate or other unobserved characteristics.

This paper would use the third strategy to distinguish the causal effect from the correlation since the lagged BMI has too many missing value in the dataset. Nevertheless, the dataset provides information of characteristic of interviewers’ family members, including my outcome of interest – body size. After considering the genetic factor, the residual might be composed of genetic factor, nongenetic factor and the residual:

$$X_i = \phi W_i + G_i + NG_i + \xi_i$$

This study would estimate the effect of obesity and BMI in the OLS regression models first, while including the household and personal characteristics. The category variables are encoded to several dummy variables, and the educational degree may be represented as continuous variable or dummy variables in different models to estimate its influence. To better investigate the effect of body size on wages, the model includes different categories of body size to distinguish heterogeneous effect.

3 Data

The data is collected from IPUM’s National Health Interview Survey. The survey is a nationally representative sample which has investigated the health, health care and behaviors of the citizens since 1963.

NHIS has included the information of BMI since 1974, but it began to include personal earning after 1996. Accordingly, this research considers the dataset after 1997 to cover the required variables. From 1997 to 2018, it covered 4,423,611 people in U.S., while both genders are approximately half of the sample.

The main variable of the labor market outcome is the personal total earnings, which is presented by the previous calendar years. Persons age 18 and older who had worked for pay during the previous calendar year were asked to report the best estimate of their earnings before taxes and deductions from all jobs and businesses in last calendar year. Specifically, the definition of earnings follows Field Representative’s Manuals for 1997, which provides detailed instructions for what to include and exclude as earnings:” The types of income included and excluded as earnings varied across time, and several types of income included as earnings in some years were excluded in other years (e.g., Veteran’s payments).” In addition, the subjects were asked to report their earnings in non-intervalled amounts amount, but the dataset encoded the information into interval-type in the original NHIS public use files. The interval is ranging from “\$01 to \$4999” group to “\$75000 and over” group, while the increment is about \$5000 for lower groups and \$10000 for higher groups. Table 1 shows that men and women earns on \$35875.09 and \$25807.28 in the previous year respectively.

In addition, further characteristic of the individuals in labor market outcome also been considered, including occupation, working class and years on main or longest or last job. Occupation used a modified version of the 1995 revised Standard Occupational Classification-based codes (SOC), which provides the constant long-term category of different occupation in the labor market; Table 1 presents that the considered occupations including Managerial and Professional Specialty; Technical, Sales, and Administrative Support; Service; and Operators, Fabricators and Laborers account for over 10% of the sample separately. The majority of male workers with positive wages have Managerial and Professional Specialty occupation (30%), while the female workers are in Managerial and Professional Specialty occupation (35%) and Technical, Sales, and Administrative Support (33%).

Working class indicates the person is a self-employed, work in private industry or public sector, or worked without pay in a family business. If the individual worked more than one job, the question applied to the job where longer hours were worked. The public sector includes employers in Federal government, State government and Local government. Table 1 shows that nearly three quarters men worked in private company, when 44% paid women worked in private company only and 42% in public sectors.

Table 1
Summary statistics for variables

	male			female		
	mean	sd	range	mean	sd	range
Dependent variable						
personal earnings	35875.09	23716.33	[1, 75000]	25807.28	20846.62	[1, 75000]
Independent variable						
BMI	27.43	4.42	[15.8, 51.4]	26.77	5.75	[15.1, 55.3]
obesity	0.24	0.43	[0, 1]	0.25	0.43	[0, 1]
age	41.65	13.73	[18, 85]	41.49	13.78	[18, 85]
number of persons in family	2.47	1.49	[1, 18]	2.54	1.42	[1, 21]
years on main or longest or last job	8.39	9.26	[0, 35]	7.32	8.33	[0, 35]
education	13.37	3.12	[0, 22]	13.53	2.89	[0, 22]
number of persons in family	2.47	1.49	[1, 18]	2.54	1.42	[1, 21]
region						
Northeast	0.16	0.37	[0, 1]	0.17	0.37	[0, 1]
North central or Midwest	0.23	0.42	[0, 1]	0.23	0.42	[0, 1]
South	0.35	0.48	[0, 1]	0.36	0.48	[0, 1]
legal marital status						
married	0.33	0.47	[0, 1]	0.28	0.45	[0, 1]
divorced	0.14	0.34	[0, 1]	0.19	0.39	[0, 1]
never married	0.30	0.46	[0, 1]	0.28	0.45	[0, 1]
race						
white	0.81	0.39	[0, 1]	0.77	0.42	[0, 1]
black	0.11	0.32	[0, 1]	0.16	0.37	[0, 1]
occupation						
Managerial and Professional Specialty	0.30	0.46	[0, 1]	0.35	0.48	[0, 1]
Technical, Sales, and Administrative Support	0.18	0.38	[0, 1]	0.33	0.47	[0, 1]
Service	0.11	0.31	[0, 1]	0.19	0.39	[0, 1]
Operators, Fabricators and Laborers	0.16	0.37	[0, 1]	0.06	0.23	[0, 1]
class of worker						
private sector	0.74	0.44	[0, 1]	0.71	0.45	[0, 1]
public sector	0.24	0.42	[0, 1]	0.26	0.44	[0, 1]
self employed	0.10	0.30	[0, 1]	0.06	0.24	[0, 1]
N	162771			167864		

Note: The intervals of personal earnings are (\$1 - \$ 4,999), (\$5,000 - \$9,999), (\$10,000 - \$14,999), (\$15,000 - \$19,999), (\$20,000 - \$24,999), (\$25,000 - \$34,999), (\$35,000 - \$44,999), (\$45,000 - \$54,999), (\$55,000 - \$64,999), (\$65,000 - \$74,999), (\$75,000 - over). The (\$1 - \$4,999) group is represented by \$1 instead of \$0 to disentangle itself from missing values.

Years on job presents the amount of years the person worked at his or her job for all the adults satisfy the condition that aged equal or over 18 and are currently working or have ever worked. The information is ranging from less than 1 year to 35 or more years with 1-year increment. Table 1 illustrates that men and women respectively have 8.39 and 7.32 years on their main or longest or last jobs.

The outcome of interest, characteristic of body size, has been comprehensively investigated by NHIS. The dataset offers the information of height, weight and BMI calculated by previous two variables. Height is defined as height in inches without wearing shoes by a self-reported methodology by interviewers. Weight is defined as the weight in pounds without wearing clothes or shoes by a self-reported methodology by interviewers, too. BMI is calculated by height and weight in NHIS. Table 1 shows that the average BMI is 27.43 for men and 26.77 for women.

Furthermore, the BMI is also classified to the different categories. Following the classification from WHO, this study converts the BMI into 8 groups: Very severely underweight(15), Severely underweight(15-16), Underweight(16-18.5), Normal(18.5-25), Overweight(25-30), Moderately obese(30-35), Severely obese(35-40) and Very severely obese(40). Figure 1 presents that the majority of the men with positive salaries is overweight, and the highest ratio of paid women is normal. In the general case, I define the last three categories as the obese type. Accordingly, Table 1 indicates that the percentage of the obese people is approximately one quarter, which is slightly lower than the citing literature previously.

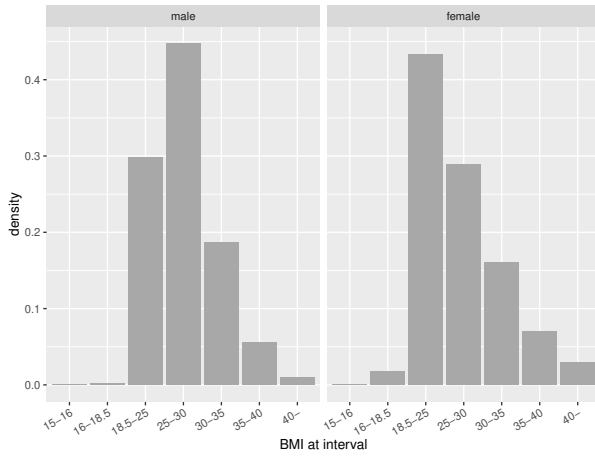


Figure 1: Distribution of BMI by sex

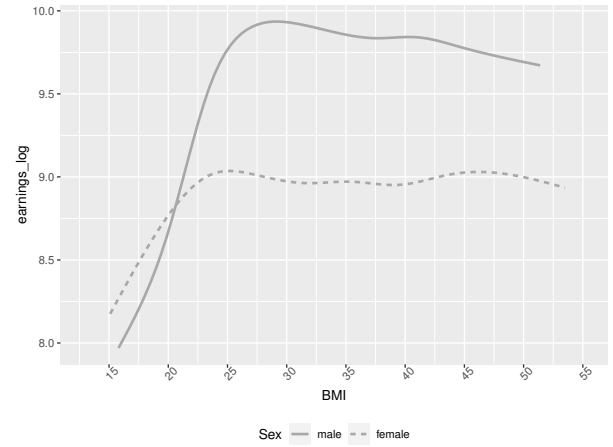


Figure 2: Fitting plot for BMI and earnings

Academic degree presents the highest level of schooling an individual has completed, in terms of completed grades for persons with less than a high school degree, and in terms of degrees attained for high school graduates and those with higher education. It distinguishes between those who completed twelfth grade but did not attain a diploma (e.g., those who failed to pass state-mandated high school exit examinations), those who graduated from high school, and those passed the GED (General Equivalence Degree) Test (which certifies that the test taker has attained high school-level academic skills). That is a guideline consistent with the emphasizing on degrees attained rather than years spent in the classroom. In addition, for a person attending special education classes or a school for persons with mental, physical, or developmental difficulties, interviewers were to ask which grade in a regular school most closely matched the individual's education level. The category variable is encoded into continuous form in this paper to better interpret how many years the interviewers attained in their school life. Table 1 presents that men and women have attained 13.37 and 13.53 years of schooling in average respectively.

The dataset also provides a variety of characteristics of household or individual, and these features serve as the covariates. The variables include age (the average is 41.65 for men, and 41.49 for women); number of persons in family (the average is 2.47 for men, and 2.54 for women); region (Northwest; North central or Midwest; South or West); legal marital status (married; spouse; widowed; divorced; separated or never married); race (White; Black; Chinese; Filipino; Asian Indian or other race). The variety of variables are collected in order to better interpret the circumstance of labor market.

4 Result

Figure 2 shows that the relationship between earnings and BMI is an inverted U-shaped curve. The wages increase as BMI rises, reaching the peak when BMI is near 30 around. The figure also presents that the fitted values of earnings remain high for obese people, since it only drops slightly after passing the boundary between overweight and obese. The trends illustrate that the salaries do not always drop when BMI elevates, but increases as BMI rises reversely. On the other hand, the relationship between earnings and BMI for women is much smoother than the curve for men, elucidating the women's wages are much invariant relative men's. The fitted value for obese women elevated much lower than fitted value for obese men. On the whole, the inverted U-shaped curve presents the heterogeneity within different classification of body size.

Table 2 presents regression result of personal earnings from previous calendar year on BMI and other control variables. In models (1-4), the regression control different covariates separately. Model 1 shows that obesity is associated with an 27.2% rise on the wages for men, while it correlates with a 16.5% increasing for women. Model 2 use BMI to estimate the effect of body size. It shows that one unit of BMI is associated with 5.27% increasing for men and with 2.27% for women. Model 3 controls many other dummy variables including household and family characteristics, etc. The covariates used for control in model 3 is selected by the Double-Lasso, hence might be different by gender. The result shows that obesity is correlated with 13.9% and 8.8 % for men and women, respectively. Model 4 uses the category of BMI to classify different body size states. It shows that obesity is associated with 4.75% rise for men and 1.32% increasing for women separately.

If obesity leads to discrimination in the workplace, bring penalty to the wages, then the race might be other factor which exacerbate this salaries erosion. That is to say about the obesity risk, the relationship differs greatly by individual race/ethnicity (Kirby et al., 2012). They mentioned that previous studies suggest that the availability and accessibility of the activities such as sidewalks, parks, recreational facilities sand food selection are related to obesity independent of individual characteristics.

On the other hand, the discrimination in the workplace is often accompanied with discrimination against specific race. To distinguish these two effect, the separate regressions are run for black and white which account for nearly the entire sample size by gender. Table A1, in models (1-4), the regression control different covariates separately. Model 1 shows that obesity is associated with an 25.7% rise on the wages for white men, while it correlates with a 16.1% increasing for white women. The rising is lower than for black, especially for men. Model 2 shows that one unit of BMI is associated with 5.14% increasing for white men and with 2.27% for white women, which is both lower than for black. The result in model 3 presents that the coefficients of obesity for white are much lower than for black. Model 4 shows that obesity correlated to higher wages for white men than for black men, but is associated with lower salaries for black men than for black women.

Table 2
OLS regression results of earnings in the previous calendar on body size, by sex

	male				female			
	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
Obesity	0.272*** (0.0134)		0.159*** (0.0142)	0.0533*** (0.014)	0.165*** (0.0172)		0.0962*** (0.0181)	0.006 (0.0199)
BMI		0.0527*** (0.0015)				0.0227*** (0.0013)		
education	0.152*** (0.002)	0.154*** (0.002)	0.111*** (0.003)	0.153*** (0.002)	0.247*** (0.003)	0.250*** (0.003)	0.160*** (0.003)	0.251*** (0.003)
Northeast			-0.0009 (0.0205)				0.0248 (0.0249)	
North Central or Midwest			-0.144*** (0.0191)				-0.201*** (0.0237)	
South			-0.0915*** (0.0168)				-0.0842*** (0.0211)	
married			0.120*** (0.015)				0.228*** (0.024)	
never married			-0.860*** (0.019)				-0.284*** (0.027)	
White			0.143*** (0.018)				-0.158*** (0.020)	
Managerial and Professional Speciality			0.409*** (0.020)				1.872*** (0.050)	
Technical, Sales and Administrative support			0.0564 ** (0.022)				1.448*** (0.049)	
Service			-0.346*** (0.028)				(0.858)*** (0.052)	
Operators, Fabricators or Laborers			0.080*** (0.021)				1.414*** (0.058)	
private sector			0.240*** (0.021)				0.304*** (0.020)	
underweight				-1.582*** (0.189)				-0.641*** (0.064)
normalweight				-0.545*** (0.020)				-0.253*** (0.024)
N	162,541	162,541	130,543	162,541	167,695	167,695	133,698	167,695

note: The models have controlled the covariates, including age, seniority, region, etc. These covariates are controlled but are not reported by the table. The dummy variables are encoded if the features account for high enough ratio of the sample size. The exact variables used in the model 3 for both genders are chosen by the Post-Double Selection method, and checked by the Variance Inflation Fraction test (VIF). Specifically, if the variable is reported over 10 largely, it will be eliminated from the original model. The test helps to obviate the multicollinearity problem. The model 3 includes fewer observations than other three models due to the missing values while considering more control variables for both genders.

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

The workplace and employers might influence the wages. Biddle and Hamermesh(1998) distinguished between employees and the self-employed and between the public and private sectors. This study only consider the public and private sectors because the small number of self-employed workers in the sample. To examine whether the discrimination against to obesity shows different trend across sectors, the regressions are conducts separately by sector and gender. The result of personal earnings from previous calendar year on BMI and other control variables. Table A2, in models (1-4), the regression outcomes are appended with different covariates gradually.

Table 3
OLS regression results of earnings in the previous calendar on body size, by sex and industry

	male				female			
	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
Service								
Obesity	0.421*** (0.0474)		0.156*** (0.0491)	0.0281 (0.0497)	0.115*** (0.0433)		0.0337*** (0.0456)	-0.0858 * (0.0500)
BMI		0.0808*** (0.005)				0.0213*** (0.003)		
education	0.155*** (0.008)	0.154*** (0.008)	0.133*** (0.009)	0.154*** (0.008)	0.147*** (0.008)	0.150*** (0.008)	0.141*** (0.008)	0.153*** (0.008)
underweight				-1.841*** (0.506)				-0.857*** (0.172)
normalweight				-0.933*** (0.0561)				-0.331*** (0.0473)
N	17,720	17,720	14,949	17,720	31,390	31,390	26,329	31,390
Sales								
Obesity	0.292*** (0.0461)		0.206*** (0.0491)	0.0453 (0.0477)	0.0487 (0.0642)		-0.0456 (0.0680)	-0.0123* (0.0736)
BMI		0.0583*** (0.005)				0.0139*** (0.005)		
education	0.214*** (0.009)	0.215*** (0.009)	0.190*** (0.010)	0.211*** (0.009)	0.291*** (0.012)	0.294*** (0.012)	0.249*** (0.013)	0.294*** (0.012)
underweight				-1.334** (0.585)				-0.706*** (0.215)
normalweight				-0.629*** (0.0525)				-0.262*** (0.0628)
N	14,625	14,625	11,728	14,625	16,902	16,902	13,633	16,902
Other industries								
Obesity	0.252*** (0.0145)		0.143*** (0.0154)	0.0622*** (0.0152)	0.197*** (0.0190)		0.129*** (0.0201)	0.0563** (0.0220)
BMI		0.0475*** (0.002)				0.0241*** (0.001)		
education	0.139*** (0.002)	0.142*** (0.002)	0.136*** (0.002)	0.141*** (0.002)	0.223*** (0.003)	0.226*** (0.003)	0.204*** (0.003)	0.226*** (0.003)
underweight				-1.508*** (0.215)				-0.546*** (0.0713)
normalweight				-0.474*** (0.0165)				-0.222*** (0.0199)
N	130,196	130,196	103,866	130,196	119,403	119,403	93,736	119,403

note: The models have controlled the covariates, including age, seniority, region, etc. These covariates are controlled but are not reported by the table. The dummy variables are encoded if the features account for high enough ratio of the sample size. The exact variables used in the model 3 for both genders are chosen by the Post-Double Selection method, and checked by the Variance Inflation Factor test (VIF). Specifically, if the variable is reported over 10 largely, it will be eliminated from the original model. The test helps to obviate the multicollinearity problem. The model 3 includes fewer observations than other three models due to the missing values while considering more control variables for both genders.

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

In Table A2, model 1 shows that obesity is associated with an 25.8% rise on the wages for men in private sector, while it correlates with a 12.2% increasing for women in private sector. The increasing for men in private sector is higher than for men in public sector, but the rise for women in private sector is lower than for women in public sector. Model 2 shows that one unit of BMI is associated with higher increasing for men in private sector than in public sector, and is correlated to lower rise for women in private sector than for in public sector. Model 4 shows that obesity has negative correlation with salaries for women in private sector, but the effect is not significant.

As the employers who might pose effect on salaries as mentioned before, the customers may be responsible, either. Hübler (2009) implied that if employers see that (other things being equal) customers prefer to interact with moderately tall salesmen and with moderately short saleswomen. Then these favoured employees will generate higher earnings not only for themselves but also for their employers. Everett (1990) and Puhl and Brownell (2001) demonstrated that employers regard obese persons as unfit for public sales positions and as more proper for telephone sales without face-to-face contact. Carr and Friedman (2005) also reported that workers whose BMI exceed 35 are more likely to report job-related discrimination in the workplace due to their weight and appearance.

If the discrimination against obesity is truly heterogeneous across industries, it may be severer within occupations where workers need to high-frequently interact with the customers. Specifically, in these kinds of jobs, the productively is defined as the service they are able to provide rather than selling some merchandises. How consumers regard these "salesperson" directly affects how they tend to accept the service, hence the physical appearance of the workers are much more important in these industries. Accordingly, if obesity leads to higher wages reduction to the workers, it lends support to the supposition that the pure discrimination is attributed to the discrimination against obesity in the workplace.

As mentioned above, the pure discrimination disentangled from the salaries penalty implies the distaste against to people who have higher BMI, making the consumers to be more reluctant to accept the service. It erodes the chance for the obese workers to provide their service, having their service to be more expensive than they truly are. On the other hands, the distaste may also come from the employers. The employers might afraid of the compensating wages potential health cost, or reckon that the obese employees are lazy and lack of self-control, therefore are much reluctant to hire somebody who is obese.

To examine whether the discrimination against to obesity shows different trend across industries. the regressions are conducts separately by sector and industry. The model distinguish the service industry from others since the workers in the service industry require to spend much more time interacting with consumers. The service industry category includes private household occupations, food service and health service, etc. To better capture all the industries highly-correlated with interacting with the consumers, the sales occupation are also separated from the Technical, Sales, and Administrative category.

In Table 3, model 1 shows that obesity is correlated to the lower rise on the wages for women in service and sales industry than for women in other industries. Model 2 also reports that one unit of BMI is associated with the highest increasing in earnings for women in other industries. model 3 presents that obesity has negative correlation with earnings for women in sales industry, but it's not significant. Model 4 illustrates that obesity is associated with the reduction on the salaries for women in service industry and sales industry, and the effects are significant at 10% level.

The instrumental variable method

In the developing countries, nutritious foods is not affordable for the majority of citizens to purchase. In addition to the monetary reason, people in these countries could not consume health foods because lack of time or having too many family members. By and large, people who don't earn much income are easier to eat unhealthy foods which lead them to obese (Mullan et al., 2017). In other words, low food security is correlated to obese and derivative health problems. USDA defined it as "reports of reduced quality, variety, or desirability of diet. Little or no indication of reduced food intake" (Dhurandhar, 2016). People who live under the poverty line can only purchase the high-calories food without selection. On the other hand, the lack of knowledge of health is also the important reason: People who have taught the knowledge of the relative fields know how to keep themselves from intake the unhealthy foods than people who don't.

Correlation between body size and earnings may be attributed to the mental health. For example, Carpenter et al. (2000) and Onyike et al. (2003) had explored the association between mental health and obesity for women in America, finding evidence that obesity is truly associated with anxiety disorder. If people who earn higher salaries suffer from severer mental health problem, then is possible that these people tend to eat lots of foods to mollify their anxiety. It may bring obesity to these high-income workers as the labor market outcome.

There are more empirical studies provide evidence to buttress the positive correlation between salaries and the body size. Specifically, the socioeconomic status (SES) have positive association with these two simultaneously. Nguyen et al. (2007); Dinsa et al. (2012); Minh et al. (2007) had offered evidence to support the positive associations. However, the phenomenon is more general in the developed countries rather than the developed countries, and the positive trend is also aroused by the changing diet culture and sedentary lifestyle in these nations. Inequality distribution in the society causes people who have higher SES acquire power and prestige than others. The power and prestige lend monetary and other kinds of support for them to purchase the food as supplement. On the other hand, the resources centralized in the important region also allow people in the high-class to obtain the food much easier.

To overcome the issue, the approach to disentangle the causation from the correlation is highly needed in order to buttress the assumption of discrimination against obese workers. This paper use Instrumental variable (IV) method to investigate the causation between these two, while the fixed effect model can't be used because the dataset isn't panel data. In previous studies, IV is often been used to cope with similar issue. [Cawley \(2004\)](#) discussed the other factors which may attribute the labourmarket outcome, such as time discount rate or other observable characteristic. [Greve \(2008\)](#) also used IV to explored the effect of obesity on labour market outcome.

Several individual and characteristic have been considered as an instrumental variable, but the majority of them can't been used due to too many missing values. Some features aren't interviewed in the survey in constant. Specifically, some questions are only asked in one or two years, impeding them from been used for the IV. By and large, this paper used family member's BMI as IV in order to control the genetic variation. Besides, a dummy variable has also been generated to represent the family member is obese or not.

Table 4

2SLS regression results of earnings in the previous calendar on body size, by sex

	male				female			
	(1) Obesity	(2) BMI	(3) Obesity	(4) BMI	(1) Obesity	(2) BMI	(3) Obesity	(4) BMI
First-stage test								
Family member's BMI	11.4811* (6.0233)	1.365* (0.758)	20.421** (9.173)	2.422** (1.0685)	-0.928 (6.598)	-0.909 (0.654)	18.840 (14.472)	1.579* (0.8767)
F-test	9.78116 F(1,51864)	4.1361 F(1,51864)	8.59155 F(1,65566)	12.2688 F(1,65566)	4.84273 F(1,52038)	4.1414 F(1,52038)	2.60564 F(1,52048)	6.26912 F(1,52048)
R^2	0.0108	0.0249	0.1286	0.4517	0.0117	0.0203	0.2657	0.5905
Family member is obese or not	58.810 (38.41)	1.878*** (0.406)	31.761*** (5.594)	2.814*** (0.354)	16.68*** (2.34)	0.987*** (0.113)	16.816*** (2.629)	1.619*** (0.243)
F-test	2.35685 F(1,130697)	22.5436 F(1,130697)	33.3865 F(1,162766)	66.974 F(1,162766)	60.298 F(1,167849)	99.9707 F(1,167849)	52.5291 F(1,167859)	58.0207 F(1,167859)
R^2	0.0109	0.025	0.1824	0.4780	0.0132	0.0223	0.2755	0.5887

note: The covariates such as age, seniority and region are hold constant but not reported. The exact variables used in the model 3 for both genders are chosen by the Post-Double Selection method, and checked by the Variance Inflation Faction test (VIF). Specifically, if the variable is reported over 10 largely, it will be eliminated from the original model. The test helps to obviate the multicollinearity problem. The model 3 includes fewer observations than other three models due to the missing values while considering more control variables for both genders.

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table 5

Wald test of endogenous results of earnings in the previous calendar on body size, by sex

	male				female			
	(1) Obesity	(2) BMI	(3) Obesity	(4) BMI	(1) Obesity	(2) BMI	(3) Obesity	(4) BMI
Wald test of Endogenous								
Family member's BMI								
$\chi^2(1)$	3.62168	3.60085	11.6731	11.5961	0.021548	0.024492	4.76075	4.73226
p -value	(0.0570)	(0.0578)	(0.0006)	(0.0007)	(0.8833)	(0.8756)	(0.0291)	(0.0296)
Family member is obese or not								
$\chi^2(1)$	273.367	266.914	963.289	954.127	303.03	298.868	192.505	190.832
p -value	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)

Table 4 presents the result of the first stage. Column 1 and 2 are from model 3 in the last section, which meaning control many personal and household characteristic selected by Post-double selection. Column 3 and 4 are from model 4 in the last section, which meaning control different interval of BMI. Obesity and BMI are endogenous variables, respectively.

The upper side reports the case when family member's BMI as IV, and the bottom side shows the case while IV is dummy for the family member is obese or not. In the upper one, the F-statistics is under 10 in column 1, 2 and 3 for men, and all columns for women, either. Column 4 for male is the only case that F-statistics is significant. In the bottom one, only column 1 for men reports insignificant F-statistics. The F-statistics are significant for the other regressions.

Table 5 presents the Wald test of endogenous. As above table, the upper side reports the case when family member's BMI as IV, and the bottom side shows the case while IV is dummy for the family member is obese or not. In the upper one, column 3 and column 4 for men reject the null hypothesis at 1% significance level, when column 3, 4 for women and column 1, 2 for men reject the null hypothesis at 5% and 10 % significance level, respectively. In the bottom one, all 8 columns reject the null hypothesis at 1% significance level.

The negative effect from the Instrumental variable method doesn't pass the test of endogenous or explanatory power, as shown in column 1 and 2 in table 5. Dummy for the family member is obese or not as IV pass both test, but the results present a positive effect between body size and earnings.

Only case when the effect between body size and wages is significant negative is pooled the sample together without distinguishing gender. The result is shown as table. When dummy for the family member is obese or not serves as IV, it passes the test of endogenous or explanatory power, while the effects are negative at 1 % significant level. However, the circumstance that disentangling the gender can't lend support to the hypothesis of discrimination against obese workers. There may exist heterogeneous characteristic across gender, hence pooling these two together will lead to some problems.

5 Conclusion

Obesity is one of the most important issues due to its risk of suffering from cardiovascular disease or type 2 diabetes, etc. Besides, obesity might also arouse the discrimination from the employers or the consumers, causing wages penalty in the workplace. This paper use IPUM's National Health Interview Survey to conduct the discrimination against obese workers in U.S. The survey is a national representative sample which explore the health issue of the residents since 1963.

The result shows that the body size has strong correlation with wages, causing the purpose of distinguishing causation from the association to be more difficult. Several models which consider different covariates regress the incomes on the outcome of interest, the body size, separately. The body size is represents by Body Mass Index (BMI), and is converted to the dummy of whether the employee is obese or not in order to help better estimation. However, the positive association in the result illustrates that income may affect body size of the workers reversely, bringing ambiguous result concluded by the regressions. This paper also explored the association between these two by grouping by gender and different characteristics such as race, working class or industry, but the correlation can't provide much information of the causation between the real labour market outcome. The prediction curve illustrates that the correlation between wages and BMI is close to the inverted U-shaped curve. That is to say the salaries do not always drop when weight elevates.

The positive correlation between obesity and income are supported by abundant past studies, while the association is more significant in the developing nations rather than in the developed nations due to changing lifestyle and diet culture. Inequality distribution in the countries cause people who have higher SES obtain power and prestige than people with lower SES. The power and prestige lend monetary and other kinds of support for them to purchase the food as supplement. On the other hand, the trend that resources centralized in the important region also allows people with higher SES to obtain the resources much easier.

To buttress the causation between the outcomes of interest, this paper used Instrumental Variable method as previous studies. Past literature has made effort to disentangle the causation from the correlation, including fixed-effects model or IV strategies. This study choose IV to be the empirical strategy due to the availability of the dataset, and the instrumental variable here is the family member's BMI of the interviewers in the survey, in order to control the genetic variation.

As a result, the Two-Stage Least Squares result doesn't lend support to the supposition of the discrimination against specific body size in the workplace. Much regression pass the tests of endogenous and explanatory power, but the positive effect presented from the result can't strengthen the assumption. Only case that shows the significantly negative effect of obesity on earnings id the condition that pooled both gender together, but the manipulation may blur the distinction of the two genders. The conclusion can't provide sufficient evidence due to omitting the heterogeneous.

The causation isn't demonstrated due to the lack of finding great instrumental variable. Present IV which control the genetic factor pass both tests, but it doesn't fulfill the requirement to show significantly negative associating. The dataset doesn't include more available information to be used for serving as IV. This paper considers some other approach to estimate the effect, but can't find the appropriate dataset to include all the needed variables. Hence there might be some other confounding factors to mitigate the effect, hence require future studies investigating better control. In spite of the limitations, it's believed that this research provide contribution to economics literature in understanding the causation between obesity and labour market outcome.

Reference

- Kenneth Arrow et al. The theory of discrimination. *Discrimination in labor markets*, 3(10):3–33, 1973.
- MV Lee Badgett. The wage effects of sexual orientation discrimination. *ILR Review*, 48(4):726–739, 1995.
- MV Lee Badgett. *When gay people get married: What happens when societies legalize same-sex marriage*. NYU Press, 2009.
- Charles L Baum and William F Ford. The wage effects of obesity: a longitudinal study. *Health economics*, 13(9):885–899, 2004.
- Howard Saul Becker. *Sociological work*. Transaction publishers, 1971.
- Kristian Bolin, Bjorn Lindgren, and Petter Lundborg. Informal and formal care among single-living elderly in europe. *Health economics*, 17(3):393–409, 2008.
- Giorgio Brunello and Béatrice d’Hombres. Does body weight affect wages?: Evidence from europe. *Economics & Human Biology*, 5(1):1–19, 2007.
- Richard V Burkhauser and John Cawley. The importance of objective health measures in predicting early receipt of social security benefits: The case of fatness. *Michigan Retirement Research Center Research Paper No. WP*, 148, 2006.
- Richard V Burkhauser and John Cawley. Beyond bmi: the value of more accurate measures of fatness and obesity in social science research. *Journal of health economics*, 27(2):519–529, 2008.
- Kenneth M Carpenter, Deborah S Hasin, David B Allison, and Myles S Faith. Relationships between obesity and dsm-iv major depressive disorder, suicide ideation, and suicide attempts: results from a general population study. *American journal of public health*, 90(2):251, 2000.
- Deborah Carr and Michael A Friedman. Is obesity stigmatizing? body weight, perceived discrimination, and psychological well-being in the united states. *Journal of health and social behavior*, 46(3):244–259, 2005.
- John Cawley. The impact of obesity on wages. *Journal of Human resources*, 39(2):451–474, 2004.
- F Chan, BT McMahon, G Cheing, DA Rosenthal, and J Bezyak. Drivers of workplace discrimination against people with disabilities: Physical vs. mental-behavioral genesis. *WORK: Journal of Assessment, Prevention, and Rehabilitation*, 25(1):77–88, 2005.
- Kerwin Kofi Charles and Jonathan Guryan. Prejudice and wages: an empirical assessment of becker’s the economics of discrimination. *Journal of political economy*, 116(5):773–809, 2008.
- Dalton Conley and Rebecca Glauber. Gender, body mass and economic status. Technical report, National Bureau of Economic Research, 2005.
- Elizabeth A Deitch, Adam Barsky, Rebecca M Butz, Suzanne Chan, Arthur P Brief, and Jill C Bradley. Subtle yet significant: The existence and impact of everyday racial discrimination in the workplace. *Human Relations*, 56(11):1299–1324, 2003.
- Emily J Dhurandhar. The food-insecurity obesity paradox: A resource scarcity hypothesis. *Physiology & behavior*, 162:88–92, 2016.
- G Deye Dinsa, Yevgeniy Goryakin, Elena Fumagalli, and Mark Suhrcke. Obesity and socioeconomic status in developing countries: a systematic review. *Obesity reviews*, 13(11):1067–1079, 2012.
- James R Elliott and Ryan A Smith. Race, gender, and workplace power. *American sociological review*, 69(3):365–386, 2004.
- Martin Everett. Let an overweight person call on your best customers? fat chance. *Sales and Marketing Management*, 142(4):66–70, 1990.
- Katherine M Flegal, Margaret D Carroll, Cynthia L Ogden, and Lester R Curtin. Prevalence and trends in obesity among us adults, 1999–2008. *Jama*, 303(3):235–241, 2010.
- Suzy Fox and Lamont E Stallworth. Racial/ethnic bullying: Exploring links between bullying and racism in the us workplace. *Journal of vocational behavior*, 66(3):438–456, 2005.
- Timothy M Frayling, Nicholas J Timpson, Michael N Weedon, Eleftheria Zeggini, Rachel M Freathy, Cecilia M Lindgren, John RB Perry, Katherine S Elliott, Hana Lango, Nigel W Rayner, et al. A common variant in the fto gene is associated with body mass index and predisposes to childhood and adult obesity. *Science*, 316(5826):889–894, 2007.

- Jaume Garcia and Climent Quintana-Domeque. The evolution of adult height in europe: a brief note. *Economics & Human Biology*, 5(2):340–349, 2007.
- Jane Greve. Obesity and labor market outcomes in denmark. *Economics & Human Biology*, 6(3):350–362, 2008.
- Euna Han, Edward C Norton, and Lisa M Powell. Direct and indirect effects of body weight on adult wages. *Economics & Human Biology*, 9(4):381–392, 2011.
- Olaf Hübler. The nonlinear link between height and wages in germany, 1985–2004. *Economics & Human Biology*, 7(2):191–199, 2009.
- James B Kirby, Lan Liang, Hsin-Jen Chen, and Youfa Wang. Race, place, and obesity: the complex relationships among community racial/ethnic composition, individual race/ethnicity, and obesity in the united states. *American journal of public health*, 102(8): 1572–1578, 2012.
- Peter G Kopelman. Obesity as a medical problem. *Nature*, 404(6778):635–643, 2000.
- Hoang Van Minh, Peter Byass, Dao Lan Huong, Nguyen Thi Kim Chuc, and Stig Wall. Risk factors for chronic disease among rural vietnamese adults and the association of these factors with sociodemographic variables: findings from the who steps survey in rural vietnam, 2005. 2007.
- Ali H Mokdad, Earl S Ford, Barbara A Bowman, William H Dietz, Frank Vinicor, Virginia S Bales, and James S Marks. Prevalence of obesity, diabetes, and obesity-related health risk factors, 2001. *Jama*, 289(1):76–79, 2003.
- Barbara Mullan, Nikos Ntoumanis, Cecilie Thøgersen-Ntoumani, and Ottmar V Lipp. ” it’s a bit more complicated than that”: A broader perspective on determinants of obesity. *Behavioral and Brain Sciences*, 40, 2017.
- Minh Duc Nguyen, Shirley AA Beresford, and Adam Drewnowski. Trends in overweight by socio-economic status in vietnam: 1992 to 2002. *Public health nutrition*, 10(2):115–121, 2007.
- Bonny Norton and Kelleen Toohey. Identity, language learning, and social change. *Language teaching*, 44(4):412–446, 2011.
- Edward C Norton and Euna Han. Genetic information, obesity, and labor market outcomes. *Health economics*, 17(9):1089–1104, 2008.
- Chiadi U Onyike, Rosa M Crum, Hochang B Lee, Constantine G Lyketsos, and William W Eaton. Is obesity associated with major depression? results from the third national health and nutrition examination survey. *American journal of epidemiology*, 158(12): 1139–1147, 2003.
- Jose A Pagan and Alberto Davila. Obesity, occupational attainment, and earnings. *Social Science Quarterly*, pages 756–770, 1997.
- Rebecca Puhl and Kelly D Brownell. Bias, discrimination, and obesity. *Obesity research*, 9(12):788–805, 2001.
- Rebecca M Puhl and Janet D Latner. Stigma, obesity, and the health of the nation’s children. *Psychological bulletin*, 133(4):557, 2007.
- Jeffery Sobal and Albert J Stunkard. Socioeconomic status and obesity: a review of the literature. *Psychological bulletin*, 105(2):260, 1989.
- Doris Weichselbaumer. Sexual orientation discrimination in hiring. *Labour Economics*, 10(6):629–642, 2003.

Appendix

Table A1

OLS regression results of earnings in the previous calendar on body size, by sex and race

	male				female			
	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
White								
Obesity	0.257*** (0.0146)		0.139*** (0.0155)	0.0475*** (0.0152)	0.161*** (0.0201)		0.086*** (0.0210)	0.0132*** (0.0232)
BMI		0.0514*** (0.002)				0.0222*** (0.002)		
education	0.143*** (0.002)	0.145*** (0.002)	0.103*** (0.003)	0.144*** (0.002)	0.239*** (0.003)	0.242*** (0.003)	0.149*** (0.004)	0.242*** (0.003)
underweight				-1.595*** (0.207)				-0.529*** (0.0704)
normalweight				-0.531*** (0.0169)				-0.231*** (0.0203)
N	132,261	132,261	105,952	132,446	129,725	129,725	103,232	129,725
Black								
Obesity	0.418*** (0.0404)		0.302*** (0.0424)	0.137*** (0.0425)	0.187*** (0.0382)		0.144*** (0.0404)	0.0090 (0.0430)
BMI		0.0664*** (0.004)				0.0246*** (0.003)		
education	0.216*** (0.007)	0.214*** (0.007)	0.165*** (0.009)	0.211*** (0.007)	0.304*** (0.007)	0.306*** (0.007)	0.213*** (0.009)	0.305*** (0.007)
underweight				-1.317** (0.0545)				-1.225*** (0.0240)
normalweight				-0.712*** (0.0525)				-0.358*** (0.0488)
N	18,461	18,461	15,049	18,461	26,644	26,644	21,431	26,644

note: The models have controlled the covariates, including age, seniority, region, etc. These covariates are controlled but are not reported by the table. The dummy variables are encoded if the features account for high enough ratio of the sample size. The exact variables used in the model 3 for both genders are chosen by the Post-Double Selection method, and checked by the Variance Inflation Fraction test (VIF). Specifically, if the variable is reported over 10 largely, it will be eliminated from the original model. The test helps to obviate the multicollinearity problem. The model 3 includes fewer observations than other three models due to the missing values while considering more control variables for both genders.

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table A2

OLS regression results of earnings in the previous calendar on body size, by sex and working class

	male				female			
	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
private sector								
Obesity	0.258*** (0.0148)		0.142*** (0.0158)	0.0423*** (0.0155)	0.122*** (0.0197)		0.0872*** (0.0208)	-0.407 (0.0277)
BMI		0.0503*** (0.002)				0.0190*** (0.002)		
education	0.154*** (0.002)	0.156*** (0.002)	0.0964*** (0.007)	0.156*** (0.002)	0.236*** (0.003)	0.239*** (0.003)	0.116*** (0.009)	0.204*** (0.003)
underweight				-1.565*** (0.205)				-0.576*** (0.0708)
normalweight				-0.528*** (0.0169)				-0.260*** (0.0204)
N	121,000	121,000	98,910	121,000	119,303	119,303	97,528	119,303
Public sector								
Obesity	0.222*** (0.0274)		0.185*** (0.0299)	0.0602*** (0.0286)	0.158*** (0.0328)		0.0638*** (0.0354)	0.387 (0.0379)
BMI		0.0424*** (0.003)				0.0209*** (0.003)		
education	0.126*** (0.004)	0.129*** (0.004)	0.113*** (0.012)	0.128*** (0.004)	0.241*** (0.005)	0.244*** (0.005)	0.179*** (0.007)	0.244*** (0.005)
underweight				0.762** (0.382)				-0.755*** (0.139)
normalweight				-0.455*** (0.0323)				-0.185*** (0.0341)
N	38,394	38,394	31,012	38,394	43,151	43,151	35,090	43,151

note: The models have controlled the covariates, including age, seniority, region, etc. These covariates are controlled but are not reported by the table. The dummy variables are encoded if the features account for high enough ratio of the sample size. The exact variables used in the model 3 for both genders are chosen by the Post-Double Selection method, and checked by the Variance Inflation Faction test (VIF). Specifically, if the variable is reported over 10 largely, it will be eliminated from the original model. The test helps to obviate the multicollinearity problem. The model 3 includes fewer observations than other three models due to the missing values while considering more control variables for both genders.

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$