



Munich Personal RePEc Archive

The Impact of Obesity on Wages: the Role of Personal Interactions and Job Selection

Moro, Andrea and Tello-Trillo, D. Sebastian and Tempesti, Tommaso

3 January 2018

Online at <https://mpra.ub.uni-muenchen.de/90690/>

MPRA Paper No. 90690, posted 26 Dec 2018 09:05 UTC

The Impact of Obesity on Wages: the Role of Personal Interactions and Job Selection

Andrea Moro, D. Sebastian Tello-Trillo, Tommaso Tempesti*

December 20, 2018

Abstract

We estimate the effects of obesity on wages accounting for the endogenous selection of workers into jobs requiring different levels of personal interactions in the workplace. Using data from the National Longitudinal Survey of Youth 1979 combined with detailed information about occupation characteristics from O*Net, we confirm the results from the literature finding a wage penalty for obese white women. This penalty is higher in jobs that require a high level of personal interactions. Accounting for job selection does not significantly change the estimated wage penalty.

Keywords: Obesity, Wages, Personal interactions, Selection

JEL Codes: J24, J31, J70, I19

*Moro : Department of Economics, Vanderbilt University. Tello-Trillo: Frank Batten School of Leadership and Public Policy, University of Virginia. Tempesti: Department of Economics, University of Massachusetts at Lowell. Emails: andrea@andreamoro.net, sebastian.tello@virginia.edu, TommasoTempesti@uml.edu. We thank the editor, an anonymous referee, and Ekaterini Kiriazidou for providing software code that we adapted to estimate one of our model specifications.

1 Introduction

This paper investigates how the relationship between obesity and wages is affected by the level of personal interactions required in the workplace while accounting for selection into these types of jobs. A vast literature studies how obesity affects labor market outcomes.¹ In a seminal paper, Cawley (2004) finds a negative effect of obesity on wages for white women.² A possible explanation of this effect is that obesity results in social stigma, affecting men and women differently (the so-called “beauty premium”).³ To the extent that body weight affects a person’s appearance, it may also affect other people’s perception of his or her labor-market relevant traits.⁴ Other papers have documented how the wage penalty varies according to the degree of personal interactions required in the workplace, however these papers have not accounted for the possibility of how weight affects selection into these types of jobs. If the effect of stigma on wages is more prominent for workers employed in jobs requiring frequent personal interactions, then obese workers may be more likely to self-select into jobs requiring fewer interactions. We study how this self-selection affects the observed job-specific relationship between obesity and wages.

There exist at least two channels that may generate a weight penalty in jobs requiring personal interactions. One possibility is that overweight individuals, being

¹See Averett and Korenman (1996), Cawley (2004), Pagan and Davila (1997), Register and Williams (1990) and references therein.

²Using more recent data, DePasquale and Macis (2016) find that this penalty has declined after year 2000.

³A growing literature (see Hamermesh and Biddle (1994)) documents the relationship between wages and physical appearance, which is at the root of the “beauty premium” conjecture. Bhattacharya and Bundorf (2009) offer an alternative explanation for the obesity wage penalty. Given that obese workers have higher expected medical expenditures, they also have higher health insurance costs. If these workers obtain health insurance from their employer, their employer will pass some of these costs to the employee in the form of lower wages. They find supporting evidence for this type of mechanism

⁴See Baum and Chou (2011) for an analysis of the socioeconomic causes of obesity. Another explanation from the sociological literature is that obesity has a more adverse impact on the self-esteem of white female than on that of other groups, but Averett and Korenman (1996) do not find empirical support for such conjecture.

perceived as less attractive in the workplace, are discriminated against by co-workers, customers, or employers who have a taste against interacting with individuals with above normal weight (as in Becker (1957)'s theory of discrimination). In this case, jobs that require more interactions display a stronger relationship between weight and labor market outcomes. A second theory is that co-workers, customers, or employers *statistically* discriminate against overweight individuals.⁵ According to this theory employers hold a (possibly biased) belief that overweight workers are on average less productive, for example because they are less capable of performing the productive tasks requiring interactions. If productivity is imperfectly observed, body weight is used by employers as a proxy in order to improve the workers' job allocation. Therefore, individuals carrying a different body weight, but otherwise identical, are treated differently. This (statistical) discrimination will be larger in jobs requiring personal interactions. Both channels carry the same implication: workers in jobs that require more interactions with customers or co-workers will be more strongly affected by their appearance.⁶

When the effects of obesity on wages are economically relevant, workers may, depending on their body mass, self-select into jobs requiring different levels of personal interactions. Our contribution is to account for the endogenous selection of workers into job types by adopting a job selection model to correct for the potential bias occurring when selection is ignored.⁷

To this end, we merge two sources of data. The first one is the National Longitudinal Survey of the Youth 1979, a nationally representative sample of men and

⁵The theory was originally proposed by Phelps (1972); see Fang and Moro (2011) for a survey.

⁶With experimental data, Rooth (2009) shows that employers discriminate, at the point of hiring, against obese job applicants, but without distinguishing between applications to jobs that require personal interactions and those that do not.

⁷Previous research has noted that people with different body mass choose different jobs. Han et al. (2011) and Morris (2006) examined the relationship between individuals' weight and their employment decisions over the life cycle, but do not account explicitly for the endogenous selection of workers with different obesity into different types of jobs.

women aged between 14 and 22 when first interviewed in 1979. This dataset contains detailed information about the respondents, including their weight, height, employment status, occupation and wages. The second source of data is the O*Net database, which classifies occupations according to hundreds of standardized descriptors illustrating each occupations' characteristics and the workers' required skills.⁸ From this information we use factor analysis to construct a variable measuring the level of personal interactions required by each occupation. We classify occupations in one of two types, depending on whether the job's level of personal interactions value is above or below the median.

The literature on the effects of obesity on wages typically regresses wages on obesity status, using different statistical methods to identify the causal effect of weight (usually fixed effects or instrumental variables techniques), to account for the possibility of reverse causality (i.e poorer individuals have less time and monetary resources to eat healthy food, exercise, etc.). We complement this approach by adopting a Roy model of self-selection.⁹ In this framework, workers of different body mass may have a comparative advantage in performing different types of jobs. We assume each person chooses between two types of jobs, requiring a high or low level of personal interactions. Each job-type requires specific skills, which are distributed differently among workers of different body mass. Workers select the job that gives them the highest expected earnings. In the standard OLS regression of wages on Body Mass Index (BMI), ignoring this selection biases the coefficient on BMI, even after controlling for job type. For example, obese individuals may disproportionately

⁸O*Net provides for each occupation values to descriptors such as "Contact with Others" (*How much does this job require the worker to be in contact with others, face-to-face, by telephone, or otherwise in order to perform it?*), Face-to-Face Discussions (*How often do you have to have face-to-face discussions with individuals or teams in this job?*), or Work With Work Group or Team (*How important is it to work with others in a group or team in this job?*). The full set of descriptors used in our analysis is listed in Appendix B.

⁹See Roy (1951) and Heckman (1990). Seminal papers in this empirical literature are Willis and Rosen (1979), Heckman and Sedlacek (1985), and Borjas (1987).

find it more advantageous to seek employment in jobs requiring fewer interactions, where obesity has a smaller impact. Because we do not observe the wages that obese workers would obtain in jobs requiring interpersonal interactions, a standard OLS regression of wages on BMI would return a biased coefficient.

The standard specification used in the literature uses a wage equation with individual fixed effects to account for reverse-causality between obesity and wages, that is, the possibility that low-wages cause obesity (perhaps because poorer families have easier access to fattening food), and the possibility that unobserved covariates affect both obesity and wages.¹⁰ To correct for the selection bias, we model job selection with an equation capturing the discrete choice between two job-types. Estimates from this equation provide information about the selection bias in the wage equation. To unbiasedly estimate the wage equation including fixed effects in both the selection and treatment equations, we adopt in our preferred specification the two-step semi-parametric estimation procedure developed in Kyriazidou (1997) that does not rely on specifying a parametric structure of the residuals. This approach allows us to control for time-invariant unobservables, such as personality traits, that may affect both selection into job type and be correlated with BMI.¹¹ This represents a substantial improvement over previous work that has not included fixed effects when studying occupational choice as a function of weight status. In fact, we are able to show that without including for the fixed effects in the selection equation could lead to a different conclusion from the one we find in this paper.¹² As a robustness check we verify our results by adopting the two-step parametric method introduced by Willis and Rosen (1979) that does not include fixed effects in the selection equation,

¹⁰This approach is adopted in Cawley (2004).

¹¹For example, extroverted individuals may be more likely to work in jobs with high social interactions. Personality traits may also be correlated with obesity (Elfhag and Morey (2008); Sutin et al. (2011))

¹²See e.g. Shinall (2014), discussed below.

and a standard fixed effects model that ignores the endogenous job selections.

To identify the selection bias correction, we use the respondent's closest siblings' job type as an exclusion restriction. We argue that this variable affects the relationship between obesity and wages only through job choice. Individuals may find it easier, regardless of their obesity, to find jobs requiring skills that are similar to the jobs of their family members, either because of direct referrals, or because family members correlate on other skills required in the workplace.¹³

Our findings confirm, using up-to-date information and modeling for selection, the literature's result that the negative relationship between obesity and wages is significant only in some demographic groups, notably white women. We find that the negative relationship between obesity and wages is mostly coming from the subset of workers in jobs needing a high level of personal interactions. In such jobs, the relationship is stronger than average. In jobs that require a lower level of personal interactions, the relationship between BMI and wages is smaller and not statistically significant. This effects are consistent even when accounting for selection into different types of jobs. This reinforces findings in this literature that have no accounted for selection.

Our analysis is most related to the research in Cawley (2004), Han et al. (2009) and Shinall (2014). Cawley (2004) focuses on the effect of obesity on wages. To account for sources of bias, such as omitted variables and reverse causality, he follows different strategies: controlling for lagged values of BMI, using siblings' weight as an instrument for the respondent's weight, and, in his preferred specification, including individual fixed effects. However, he does not control for job type. Han et al. (2009) introduce the use of job characteristics. They use data from the Dictionary

¹³We note that the two-step selection models do not need an exclusion restriction, they can be identified out of the functional form. However, it has been documented that identification from the functional form may be sensitive to specification Dow and Norton (2003); Puhani (2000)

of Occupational Titles (DOT - a precursor of O*Net, the database we use) to obtain information about the jobs the individuals are performing. They find that the association with BMI and wages is stronger for jobs that require more interpersonal skills, but do not account for endogenous job selection.¹⁴ Finally, Shinall (2014) uses job characteristics to compare the effects of obesity on wages for jobs that require physical skills versus jobs that require social skills. However, in that paper the data does not offer the possibility of controlling for individual fixed effects, which was the preferred identification strategy in Cawley (2004), nor endogenous job selection is accounted for.

Our paper builds on these three papers and extends their analysis by explicitly accounting for the endogenous selection of workers into jobs to verify if job sorting affects the magnitude of the effects of obesity on wages. To account for such selection, we adopt two different methodological approaches and base our conclusions on the comparison of these results with the results from the literature that does not account for job selection.

In another related paper, Harris (2017) studies the relationship between body mass and occupational choice. He uses a dynamic model along the lines of Keane and Wolpin (1997) and he estimates it using maximum likelihood. He also uses random effects in order to account for the within-individual serial correlation in wages. We consider our work complementary to Harris (2017): even though our model is not dynamic, by using a specification with fixed effects, we are able to account for time-invariant factors affecting both wages and (for our preferred specification) occupational choice.

¹⁴The importance of job choice is also suggested by Lundborg et al. (2014), where it is found that teenage overweight status, which affects skill acquisition, affects adult labor market outcomes.

2 Data

The National Longitudinal Survey of the Youth 1979 is a nationally representative sample of the American youth. Respondents were sampled first in 1979 when they were between 14 and 22 years of age, every year until 1994, and every two years since. We use 15 years of data ranging from 1982 thru 2006.¹⁵ This is a widely used survey in research on employment and obesity, allowing for comparability of our results with other studies. The survey contains detailed questions about employment and tenure, household environment and structure, and personal health information like height and weight. For the empirical analysis we use information on gender, race, education status, marital status, number of children in the household, Armed Forces Qualification Test (AFQT), age of youngest child, highest grade achieved by mother and father, work tenure, years of experience, and region dummies. We also follow the literature and generate dummies representing missing values for each of the variables to account for different types of misreporting. The estimation sample includes 41,589 person-years (19,175 person-years for women and 22,414 person-years for men) after excluding women who were pregnant at the time of interview, individuals under the age of 18, and individuals who did not have full employment or weight information. About 91 percent of our sample is employed, with on average 141 weeks of tenure at their current job.

In order to construct the main dependent outcome we obtain information on the hourly wage of each individual and then, as in Cawley (2004), we top-code the hourly wage at \$500. The hourly wage variable is a computed hourly wage from NLSY, which bases the hourly wage on questions about last-week employment. In order to compute the hourly wage, the NLSY uses information on earnings, time unit

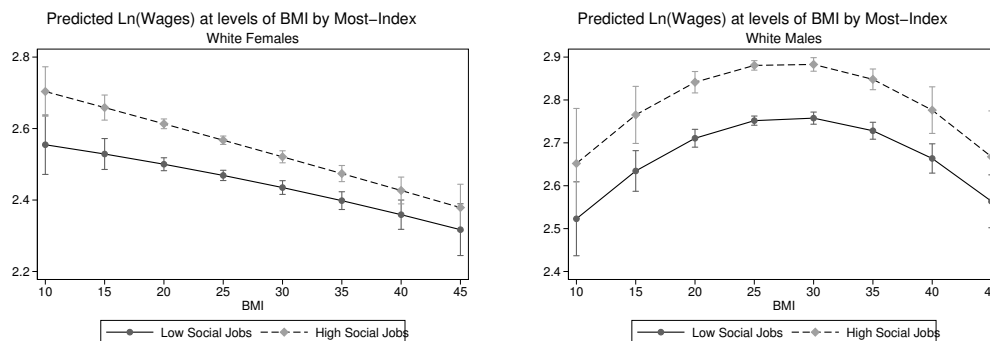
¹⁵We cannot use all of the survey years because the survey does not include body weight information in these years. The years that we end up using for our sample are: 1982, 1985, 1986, 1988, 1989, 1990, 1992, 1993, 1994, 1996, 1998, 2000, 2002, 2004 and 2006.

of compensation (hourly, weekly), and usual hours worked. We normalize wages to 2010, using the Consumer Price Index for Urban Consumers. Our left hand side variable is the log of real hourly wages. Our estimation sample shows that the average wage for men is \$14.57 while the average wage for women is \$11.58. When focusing on the difference between normal weight category and obese category, we estimate a wage differential of -\$0.77 for women and a \$3.01 for men, i.e. obese men earn - on average - higher hourly wages than normal weight men.

The Body Mass Index (BMI) is computed as the ratio of a person’s weight in Kilograms and the square of height in meters. As standard practice in the literature, we look at the effect on wages of categorical values of BMI: individuals are defined underweight if their BMI is below 18.5, overweight if their BMI is between 25 and 30, and obese if their BMI is greater than 30. In order to calculate the Body Mass Index (BMI), we pool the responses for all years that recorded self-reported weight. Height was assumed to be equal to the height recorded in 1985, when respondents were between 20 and 27 years of age. We adopt the standard practice in the literature to correct for self-reporting error of weight and height by using the procedure proposed by Burkhauser and Cawley (2008). This procedure exploits information on the relationship between true and reported height and weight collected in the National Health and Nutrition Survey. They regress accurate measures of height and weight on self-reported measures of height and weight, age, and age square separately by race and sex. We use the coefficients from these regressions with our sample to produce a predicted “Y” which is the corrected self-reported weight or height.¹⁶ In our estimation sample the average BMI is around 25.69 for women and 26.68 for males. Appendix Table A1 provides more information on the summary statistics of our sample.

¹⁶See also Lee and Sepanski (1995) and Bound et al. (2001) for a technical discussion of this method.

Figure 2: Wages and BMI for different groups and job types



may be equally *important* for secretaries and journalists, but a professional writer may need a higher *level* of writing skills. The database contains a total of 277 job descriptors. Figure 1 reports, as an example, one of the questions which is relevant to infer the importance and level of personal interactions required to perform the respondent’s job.

We use information from the categories “skills” and “abilities” in O*Net to calculate how important personal interactions are to perform any given job. We chose 17 work activities and job skills that, to our judgement, are most important for discerning the importance of personal interactions. These questions assess how relevant such interactions are to perform a job, to communicate with others (supervisors, co-workers, or customers), to resolve conflicts or coordinate other individuals, to speak, negotiate, coordinate others, etc. The full list of job descriptors we selected can be found in Appendix B.

We used the “level” values for each task and skill; to reduce the dimensionality of the information we used factor analysis to extract a single variable measuring the degree of personal interactions required by each job.¹⁷ We use the median of

¹⁷We replicated the analysis using the “importance” indicators, which lead to similar results. In our main specifications we use the “level” measure because it displays more variation than the “importance” measure.

this standardized score for the entire sample as a threshold to categorize jobs in two categories: those requiring high and low levels of personal interaction. In our estimation sample we find that 58% of women are occupied in jobs requiring a high level of personal interactions, whereas only 44% of males are employed in these job. In Figure 2, we plot the predicted logarithm wage across the different levels of BMI for jobs requiring a high level of interactions versus jobs requiring low levels of personal interactions for two groups, white women and white males.¹⁸

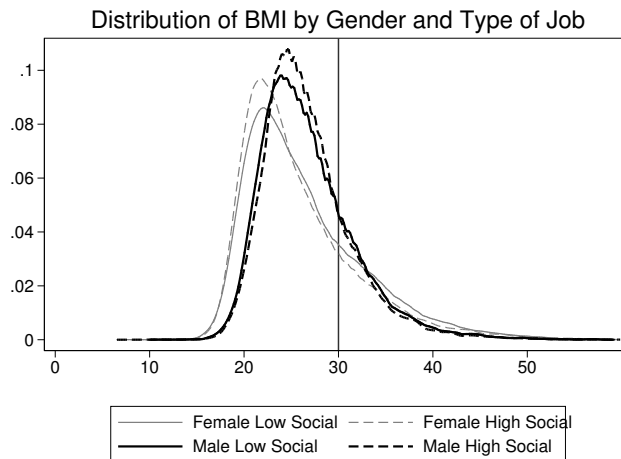
These figures illustrate several facts. First, jobs requiring interactions usually pay higher wages than other jobs. Second, the relationship between BMI, wages, and job-type is different between males and women. The relationship between wages and BMI is close to linear for women, and decreasing in BMI; the same relationship for males is non-monotonic. Finally, the wage-gap between jobs with high and low levels of personal interactions for white women is decreasing in BMI, suggesting that there is a higher penalty of BMI on wages for women in jobs requiring interactions, presumably because weight and appearance is more important in such jobs. The same penalty is not as evident for males' wages as for women' wages.

Figure 3 displays the distribution of BMI by type of job and gender. The BMI distribution appears to display a smaller variance in jobs requiring more interactions. Harris (2017) also documents how the share of obese workers in a given occupation varies across age categories. These facts suggests the possibility of job selection which we will explore in the empirical analysis.

Our empirical specification will use information about the individual's sibling, hence the sample will be restricted to individuals who have siblings (and that have non-missing information on the sibling). This sample restriction means that we are using around 32 percent of the individuals in the full sample. Table A4 in

¹⁸The other race-gender groups can be found in Appendix C

Figure 3:



the Appendix A compares our sample to the excluded sample in key observable characteristics showing that the sample we use in our estimation is very similar to the excluded sample.¹⁹ While we do not see a strong evidence of sample selection, it is possible that our results are not robust to extending the analysis to a larger sample, for example because families with one child, which belong to wealthier backgrounds, may have different job opportunities.

3 Empirical framework

We now turn to describing our empirical strategy to account for the selection into the different types of jobs. We assume that individuals can be employed in two types of jobs j , requiring ($j = 1$) or not requiring ($j = 2$) a high level of personal interactions. Let w_{itj} be the log wage of individual i at time t employed in job $j \in \{1, 2\}$, which we assume to depend on a set of covariates X_{it} , including her or his obesity at time t , BMI_{it} , and, depending on the adopted specification, individual and time effects to

¹⁹There are, however, differences in average tenure and experience between the two samples

account for the possibility of reverse causality from wages to BMI:

$$w_{itj} = \alpha_j X_{it} + \zeta_i + \epsilon_{jit} \quad (1)$$

Let J_{it} be a latent variable determining the job-type choice of the individual and let Z_{it} denote a set of observed variables that influence the job choice of individual i at time t with some elements of Z_{it} not affecting wages directly. We assume the following job selection model:

$$J_{it} = \beta Z_{it} + \eta_i - \epsilon_{3it} \quad (2)$$

with job choice $j_{it} = 1$ if $J_{it} \geq 0$, and $j_{it} = 2$ otherwise. Hence,

$$\Pr(j_{it} = 1) = \Pr(\epsilon_{3it} \leq \beta Z_{it} + \eta_i) \quad (3)$$

Because individuals select into different jobs according to their possibly time-varying characteristics, wages are not observed from a random sample of the population and an OLS regression of the wage equation with fixed effects delivers biased coefficients because:

$$\begin{aligned} E(w_{it1}|J_{it} = 1) &= \alpha_1 X_{it} + \zeta_i + E(\epsilon_{1it}|J = 1) = \alpha_1 X_{it} + \zeta_i + E(\epsilon_{1it}|\epsilon_{3it} \leq \beta Z_{it} + \eta_i) \\ E(w_{it2}|J_{it} = 2) &= \alpha_2 X_{it} + \zeta_i + E(\epsilon_{2it}|J_{it} = 2) = \alpha_2 X_{it} + \zeta_i + E(\epsilon_{2it}|\epsilon_{3it} > \beta Z_{it} + \eta_i) \end{aligned}$$

Procedures to account for the selection bias in linear models have been developed since the seminal paper by Heckman (1979).

It is worth elaborating on the possible source of this bias. The wage equation (1) includes individual fixed effects. If the factors that affect the probability of an individual selecting a job with high personal interactions do not change over time

(i.e. the Z_{it} variables in (2)), then the expected value of such probability will not change over time and will be absorbed by the individual fixed effects. On the other hand, if there is within-individual variation over time in the factors affecting the selection equation, then the OLS regression may be biased, even if it includes fixed effects. This will occur if the error term ϵ_{3it} in the job selection equation (2) is correlated with the error term ϵ_{jit} in the wage equation (1). This type of concerns motivates our use of a selection model.

Fixed effects may also enter the selection equation. For example, unobservable personality traits likely affect the probability of working in a job with high personal interactions. Moreover, such traits are also likely correlated with BMI.²⁰ In order to alleviate these problems, in our preferred specification we include fixed effects not only in the wage equation, but also in the selection equation. The standard procedure in estimating fixed effects models with panel data is to compute time-differences of the equation of interest, which eliminates the individual fixed effects ζ_i . However, in the presence of selection with fixed effects, sample selectivity creates nonlinearities in the wage equation that cannot be differenced out. We therefore adopt the method proposed in Kyriazidou (1997) and label this as specification K in our results.

Kyriazidou (1997)'s procedure follows a two-step approach in the spirit of Heckman (1979), in which the unknown coefficients of the selection equation (2) consistently estimated in the first step are used to estimate the equation of interest (1) in the second step. This estimator matches observations within individuals that have the same selection effect in two time periods. It then differences out both the individual heterogeneity term, and the selection term. Identification of the fixed effects specification of the selection equation requires individuals transitioning between job

²⁰The psychology literature has found a correlation between personality traits, such as extraversion and impulsiveness, and obesity (Elfhag and Morey (2008); Sutin et al. (2011)). In the economic literature, Courtemanche et al. (2014) find that greater impatience increases BMI.

t \ $t + 1$	Low	High
	0.78	0.22
	0.17	0.83

Table 1: Transition probabilities from jobs with Low/High levels of social interactions at time t to jobs with Low/High levels of social interactions at time $t + 1$.

types. Table 1 displays the transition probabilities across jobs with different levels of social interactions, showing that job-type transitions are quite frequent.

This approach makes no parametric assumptions about the correlation structure about the unobservables in the model. This is an advantage over previous traditional selection models such as Heckman (1979) that impose a bivariate normal distribution on the joint probability of unobservables.²¹

The wage equation is estimated with the usual procedure of taking time differences of the observed variables as in standard linear panel data models, which eliminates the fixed effects ζ_i . To account for sample selectivity, each observation is assigned a weight computed in the first step. The procedure assumes that, for an individual choosing job j in two consecutive periods, the selection effect remains the same if the variables that determine the job choice do not change over time, in which case the selection effect is completely eliminated by the time differencing in the second stage. A kernel weight is therefore computed as a function of the magnitude of the estimated differences $|\hat{\beta}Z_{i,t+1} - \hat{\beta}Z_{i,t}|$ so that a larger difference corresponds to a smaller weight, where $\hat{\beta}$ is consistently estimated using a logistic model. In other words, for a given individual, if we observe a switch to a different job type and no change in BMI, then this observation will not receive too much weight in the wage equation, as opposed to individuals who experience changes in both jobs and BMI.

In this approach identification of model parameters is guaranteed by nonlineari-

²¹For a discussion of selection methods and their differences see Dustmann and Rochina-Barrachina (2007)

ties in the first stage. However, non-parametric identification is usually preferable if a variable is available that affects the selection equation without affecting the wage equation directly. To this end, we include in Z_{it} the closest sibling's job type.²² This choice is supported by a large body of evidence documenting that jobs are often found through social networks (family, friends etc.), via, for example, job referrals (Topa (2011)). We hypothesize that the closest sibling's job type provides information regarding the type of jobs a worker may be referred to. As to the validity of our instrument, we rely on results from the literature that find, at least for the U.S., more mixed effects of job referrals on wages, with effects sometimes being positive and other times, especially for the U.S., close to zero (Topa (2011); Pellizzari (2010); Loury (2006)). This exclusion restriction will improve identification of the parameters of interest, the coefficients on BMI categories.²³

As a robustness check we also estimate the model (1-2) using a traditional parametric approach without fixed effects in the selection equation. We assume joint normality of the error term vector $[\epsilon_1, \epsilon_2, \epsilon_3]$, with zero means and variance-covariance matrix $\Sigma = [\sigma_{ij}]$ and adopt the two-step procedure by Willis and Rosen (1979) who use the probit job selection model to compute the predicted values of the inverse Mills ratios and, in a second stage, estimate the wage equations (1) by Ordinary Least Squares, including among the regressors the inverse Mills ratios in order to correct for selection bias.²⁴ We label this specification WR in our results. Finally we compare our results to the preferred specification in Cawley (2004), who includes

²²In order to identify the "closest" sibling of each individual we create a measure of distance between each individual and all of their siblings. This measure is the sum of the squared difference between the siblings' age, race and gender. The "closest" possible sibling is one that has the same age, race and gender.

²³The use of sibling's information as a possible instrument is a common strategy in the literature: for example, Cawley (2004), who does not control for selection, uses a sibling's BMI as an instrument for an individual's BMI.

²⁴We bootstrap with 400 replications the standard errors of the second stage, in order to account for the fact that the inverse Mills ratios are estimated.

		High personal interactions			Low personal interactions		
		$N \times T$	N	$\overline{\ln(w)}$	$N \times T$	N	$\overline{\ln(w)}$
Whites							
	Women	7,082	1,213	2.73	4,223	1,120	2.36
	Men	6,865	1,171	3.02	6,287	1,306	2.61
Blacks							
	Women	2,705	549	2.56	2,903	630	2.28
	Men	1,880	473	2.81	4,549	767	2.46
Hispanics							
	Women	1,987	373	2.64	1,291	317	2.32
	Men	1,808	352	2.91	2,486	445	2.58

Table 2: Sample sizes and mean log wages

fixed effects in the wage equation. We estimate this specification for both job types and label its result OLS-FE.

4 Results

To allow for appropriate comparisons, we use the same sample in all specifications. Sample size and average wages for all specifications are reported in Table 2, and other descriptive statistics in Appendix A. The main results are reported in tables 3 (Whites), 4 (Blacks), and 5 (Hispanics). In each table, columns 1, 2, and 3 report the results respectively for the specifications that include fixed effects in both the selection and wage equation (K), fixed effects in only the wage equation (WR), and the fixed effect specification that ignores job selection (OLS-FE), as described in the previous section. The first two columns represent the specifications that account for selection, and the difference between the first and second column portray the difference between including and excluding individual FE from the selection equation. The tables report the effects of the categorical BMI levels (normal BMI is the omitted category) on wages. In the wage and selection equations we also control include for the regressors typically used in this literature: age, number of children, age of

the youngest child, the respondent and the respondent parent’s highest completed grade, job tenure, enrollment in school status, experience, marriage status, region dummies, and a time trend. The full set of coefficient estimates, including results from the first stage regressions for specifications (OLS-FE) and (WR) are reported in the Online Appendix. Standard errors are clustered at the individual level. The top panel of each table reports the estimates for women and the bottom panel for men. In each panel we report the results for individuals in jobs that require high and low levels of personal interactions. In parenthesis we report robust standard errors clustered by individual to account for correlations in the error terms of each individual over time.

Taking a bird-eye view by looking for traction on the coefficients of the categorical BMI variables across different groups and specifications, we note that above normal BMI has consistent statistically and economically significant effects on white women working in jobs requiring a high level of personal interactions. This group suffers a wage penalty from being obese between 8 and 11 percent, depending on the specification.²⁵ The wage penalty associated from being simply overweight is smaller in size, but not statistically significant at the 5% level. The effect of abnormal BMIs on black and hispanic women in jobs with a high level of personal interactions is in almost all specifications smaller in magnitude and noisier, and not statistically significant.

²⁵In the WR specification, the coefficient is statistically significant only at the 10% level, not the 5% level which is the only one reported in the table.

Table 3: Effects of BMI on log wages for whites

Women	(1) K	(2) WR	(3) OLS-FE
<i>High personal interactions</i>			
Underweight	-0.125 (0.110)	-0.098* (0.049)	-0.070 (0.039)
Overweight	-0.044 (0.037)	-0.052 (0.028)	-0.050 (0.029)
Obese	-0.108* (0.049)	-0.082 (0.049)	-0.103* (0.045)
Inverse Mills Ratio		-0.438 (0.282)	
<i>Low personal interactions</i>			
Under weight	-0.013 (0.044)	-0.103 (0.053)	-0.067 (0.046)
Overweight	-0.017 (0.066)	0.009 (0.028)	0.012 (0.023)
Obese	-0.046 (0.079)	0.009 (0.042)	-0.026 (0.039)
Inverse Mills Ratio		-0.503* (0.191)	
Men	(1) K	(2) WR	(3) OLS-FE
<i>High personal interactions</i>			
Underweight	0.539* (0.137)	0.106 (0.089)	0.094 (0.079)
Overweight	-0.057 (0.088)	0.032 (0.027)	0.037 (0.024)
Obese	0.347 (0.427)	0.041 (0.044)	0.047 (0.040)
Mills Ratio		0.215 (0.181)	
<i>Low personal interactions</i>			
Underweight	-0.010 (0.088)	-0.093 (0.118)	-0.065 (0.103)
Overweight	0.026 (0.077)	0.063* (0.023)	0.024 (0.024)
Obese	-0.052 (0.096)	0.036 (0.04)	0.018 (0.030)
Mills Ratio		-0.326 (0.168)	

Robust standard errors in parenthesis. *Significant at 5% or less

Table 4: Effects of BMI on log wages for blacks

Women	(1) K	(2) WR	(3) OLS-FE
<i>High personal interactions</i>			
Underweight	-0.01 (0.104)	-0.108 (0.104)	-0.101 (0.092)
Overweight	0.036 (0.114)	0.007 (0.066)	0.047 (0.039)
Obese	-0.097 (0.144)	0.019 (0.068)	0.042 (0.056)
Inverse Mills Ratio		-0.429 (0.409)	
<i>Low personal interactions</i>			
Under weight	-0.193* (0.067)	-0.101 (0.064)	-0.099 (0.061)
Overweight	0.128 (0.085)	0.052 (0.038)	0.061* (0.030)
Obese	-0.018 (0.088)	0.088* (0.041)	0.093* (0.041)
Inverse Mills Ratio		-0.097 (0.248)	
Men	(1) K	(2) WR	(3) OLS-FE
<i>High personal interactions</i>			
Underweight	0.016* (0.000)	0.540 (0.302)	-0.308* (0.047)
Overweight	0.002 (0.066)	0.075 (0.058)	0.017 (0.038)
Obese	0.039 (0.064)	0.110 (0.083)	0.042 (0.060)
Mills Ratio		0.681 (0.413)	
<i>Low personal interactions</i>			
Underweight	0.278 (0.174)	0.056 (0.073)	-0.057 (0.067)
Overweight	-0.055 (0.047)	0.024 (0.030)	0.024 (0.024)
Obese	-0.065 (0.065)	0.058 (0.045)	0.059 (0.039)
Mills Ratio		-0.010 (0.310)	

Robust standard errors in parenthesis. *Significant at 5% or less

Table 5: Effects of BMI on log wages for hispanics

Women	(1) K	(2) WR	(3) OLS-FE
<i>High personal interactions</i>			
Underweight	-0.068 (0.061)	-0.288 (0.180)	-0.168* (0.082)
Overweight	0.042 (0.107)	0.066 (0.072)	0.005 (0.053)
Obese	0.038 (0.130)	0.052 (0.097)	-0.063 (0.066)
Inverse Mills Ratio		-1.629* (0.549)	
<i>Low personal interactions</i>			
Under weight	0.15 (0.105)	-0.272 (0.149)	-0.183 (0.108)
Overweight	0.105 (0.057)	0.077 (0.059)	0.040 (0.039)
Obese	0.156 (0.083)	0.103 (0.088)	0.019 (0.058)
Inverse Mills Ratio		-0.755 (0.517)	
Men	(1) K	(2) WR	(3) OLS-FE
<i>High personal interactions</i>			
Underweight	0.42 (0.563)	0.749 (0.466)	0.744 (0.444)
Overweight	-0.378* (0.118)	-0.000 (0.058)	0.000 (0.054)
Obese	-0.337* (0.140)	0.027 (0.094)	0.038 (0.082)
Mills Ratio		0.217 (0.400)	
<i>Low personal interactions</i>			
Underweight	-0.174 (0.100)	-0.150 (0.082)	-0.147 * (0.068)
Overweight	-0.057 (0.040)	-0.028 (0.029)	-0.028 (0.027)
Obese	-0.090 (0.057)	0.038 (0.048)	0.035 (0.045)
Mills Ratio		-0.070 (0.328)	

Robust standard errors in parenthesis. *Significant at 5% or less

The effects of abnormal BMI for women with jobs with low levels of personal interactions are usually smaller in magnitude for all racial groups. In the rare cases when they are statistically significant, the results are not robust to different specifications. For example, obese black women receive a wage premium of about 9 percent in these jobs, significant at the 5 percent level according to specifications (WR) and (OLS-FE). This result is reversed, but noisy, in our preferred specification (K). In the Hispanic women sample, in our favorite specification (K), overweight women earn higher wages, with a premium of respectively 11% and 16%, barely missing the 5% statistical significance threshold. Both the size of these effects and statistical significance of the estimates are lower in the other specifications.

The effect of abnormal BMI on men is rarely significant, and not robust to changing specifications. For example, we find a negative effect of abnormal BMI on hispanic men’s wages, but only in our preferred specification. We note however that our exclusion restriction variable is not significant in this specification’s first stage (see the Online Appendix for details), therefore this result may be the outcome of some of this noise and should be interpreted with caution.

5 Discussion

The replication of the results with the OLS-FE specification (but added years of data) is consistent with the findings in Cawley (2004) and Han et al. (2009) (see Table 6 for a comparison of their main results with our selection specifications K and OLS-FE).²⁶ Both of these papers find significant negative effects of obesity on wages for white women, and the magnitude of these estimates is within the confidence interval

²⁶There are several differences across the different specifications: Cawley (2004) and Han et al. (2009) do not include year 2006. Han et al. (2009) wage equations are conditioned on people who are employed. Cawley (2004) and Han et al. (2009) use a blue collar dummy and several state and local economic characteristics. Finally, our sample only includes individuals with siblings.

Table 6: Comparison of our results with Cawley (2004) and Han et al. (2009)

	Cawley (2004)	Han et al. (2009)	Our results			
			(K)	(OLS-FE)		
			High	Low	High	Low
White women	-0.087* (0.015)	-0.075* (0.021)	-0.108* (0.049)	-0.046 (0.079)	-0.103* (0.045)	-0.026 (0.039)
Black women	0.002 (0.017)	-0.049* (0.025)	-0.097 (0.144)	-0.018 (0.088)	0.042 (0.056)	0.093* (0.041)
Hispanic women	-0.020 (0.024)	0.032 (0.035)	0.038 (0.130)	0.156 (0.083)	-0.063 (0.066)	0.019 (0.058)
White men	0.013 (0.015)	-0.001 (0.017)	0.347 (0.427)	-0.052 (0.096)	0.047 (0.040)	0.018 (0.030)
Black men	0.031 (0.019)	0.014 (0.027)	0.039 (0.064)	-0.065 (0.065)	0.042 (0.060)	0.059 (0.039)
Hispanic men	0.023 (0.025)	0.008 (0.032)	-0.337 (0.140)	-0.090 (0.057)	0.038 (0.082)	0.035 (0.045)

*Significant at 5% or less. Coefficients on the category for “Obese” in each paper. Robust standard errors are in parentheses.

of our unadjusted estimates. Similarly, these papers do not find strong evidence for the effects of obesity in males or black and hispanic minorities. This reinforces the evidence for an obesity wage penalty occurring for white women, but not other race-gender groups even when adding more years to the analysis.

The second finding is that most of the negative effect of obesity for white women is due to women working in jobs requiring a high level of personal interactions, confirming Han et al. (2011)’s results. These patterns are consistent with the evidence displayed in Figure 2 and suggest that white women’s wage penalty for obesity may be driven driven, in part, by social stigma, consistently with the hypothesis that

appearance matters more in these types of jobs.

More importantly, we do not find evidence of drastic changes in these patterns when accounting for selection. The estimated coefficients with and without individual fixed-effects in the selection equation are only slightly different relative to the OLS-FE results. These results help validate previous findings on the estimation of BMI on wages because most of those studies do not correct for selection into these types of job. Nevertheless, the precision of the results change if individual fixed effects are not included in the selection equation, hence we think it is non-trivial to raise the relevance of including individual fixed-effects in the selection equation.

This result could have some implications. In other settings, it has been shown that selection into jobs may have an important role, but in this case it does not seem to have a relevant impact, given our results. This could be because the other observable characteristics and the individual fixed-effects are already capturing enough of the time-invariant heterogeneity in the choice of jobs. An alternative hypothesis is that there isn't much selection occurring in the first place, and people's choice of job-type may depend on a set of other characteristics such that weight becomes less trivial. Evidence from the first stage (the BMI coefficients) suggest that there is a role of weight explaining job choices, which weakens support to this hypothesis.

It is also important to note some of the limitations of our methodology. First, even though we are accounting for selection, there could be some *unobservable* time-varying components across individuals that would be biasing the effects of obesity on wages. For example, if higher weight is correlated with lower health and in turn health affects productivity, our approach, as most others in the literature, would confound the effect of weight with the one of health.²⁷ Second, our exclusion restriction may fail: given that, like Cawley (2004), we have a single instrument, we

²⁷Unfortunately, the NLSY does not collect data on health of individuals in all waves. Currie and Madrian (1999) discuss the evidence of the effect of health on wages.

cannot test the validity of such restriction. Finally, our selection equation, even when fixed effects are included, may also suffer from omitted variable bias. For example, inter-temporal preferences may change over time, even conditional on the observables. This time-varying heterogeneity may in turn affect both labor market outcomes and be correlated with BMI. Despite these limitations, our results extend the literature of obesity and wages by providing evidence that the impact of BMI on wages depends on job types but not on sorting.

6 Conclusion

We jointly model the relationship between wages and weight status and the selection of workers into jobs with different levels of personal interactions by merging the NLSY, a longitudinal data set with information on individual’s employment and wages, with O*NET, a data set on job characteristics that help us identify the level of interpersonal interactions that are associated with a given occupation.

We estimate the effect of obesity on wages using a semi-parametric specification that accounts for time-invariant unobservable factors affecting selection, and for potential endogeneity of body mass by including fixed effects. In our results we find a wage penalty of about 10% for obese white women in jobs that require a high level of personal interactions. These results therefore complement the literature finding an impact of obesity on white women suggesting that this effect stems from the need for personal interactions required by the job. Given the potential concern for selection in these estimations, we find no strong evidence that selection plays an important role in the estimated effects of obesity on wages.

A Appendix: descriptive statistics

Table A1: Summary Statistics of NLSY 1987-2006

	Women			Men			
	White	Black	Hispanic	White	Black	Hispanic	Total
<i>Employment Variables</i>							
Employed	0.96	0.92	0.94	0.96	0.91	0.9	0.95
Hourly Wage in Dollars in 2010 \$	11.99	9.42	10.71	15.25	10.80	13.29	13.19
Ln(Wage) in 2010 \$	2.63	2.43	2.53	2.87	2.57	2.74	2.72
High personal interactions job	0.65	0.48	0.61	0.55	0.30	0.44	0.56
Tenure (weeks)	215.72	212.01	197.17	245.88	194.44	209.69	226.63
Currently in School	0.11	0.07	0.09	0.08	0.05	0.07	0.09
Experience (1000s hours)	20.51	19.27	19.86	26.08	22.68	24.46	23.14
<i>Body Mass Index</i>							
BMI Adjusted	25.10	28.81	26.85	26.61	26.80	27.40	26.22
Underweight (BMI < 18.5)	0.04	0.02	0.02	0.00	0.01	0.01	0.02
Overweight (25 < BMI < 30)	0.22	0.29	0.31	0.39	0.35	0.41	0.32
Obese (BMI: >30)	0.17	0.36	0.25	0.19	0.22	0.24	0.20
<i>Demographics</i>							
Age	31.42	32.40	31.88	31.60	31.90	31.45	31.60
Highest Grade	13.83	13.24	12.65	13.48	12.52	12.34	13.48
Mother Highest Grade	11.78	10.34	7.95	11.65	10.15	7.64	11.30
Father Highest Grade	12.01	7.95	7.26	11.96	7.74	7.27	11.18
Number of Children	1.04	1.52	1.44	0.96	1.38	1.28	1.08
Age of Youngest Child	3.08	4.90	4.26	2.18	2.11	2.27	2.73
Never Married	0.28	0.46	0.26	0.36	0.51	0.39	0.34
Married	0.58	0.33	0.55	0.54	0.34	0.49	0.53
Separated	0.02	0.10	0.07	0.02	0.06	0.04	0.03
Divorced	0.12	0.11	0.11	0.08	0.09	0.07	0.10
Widowed	0.00	0.01	0.01	0.00	0.01	0.00	0.00

Table A2: Summary Statistics of NLSY 1987-2006 for Females

	Underweight	Normal	Overweight	Obese	Total
BMI Adjusted	17.64	21.90	27.19	35.66	25.69
<i>Employment Variables</i>					
Employed	0.92	0.96	0.96	0.95	0.96
Hourly Wage in Dollars in 2010 \$	10.31	11.82	11.65	11.05	11.58
Ln(Wage) in 2010 \$	2.54	2.63	2.59	2.52	2.60
High personal interactions job	0.63	0.64	0.63	0.58	0.62
Tenure (weeks)	135.30	189.89	225.59	281.06	214.27
Currently in School	0.16	0.13	0.08	0.06	0.10
Experience (1000s hours)	13.09	18.12	21.94	25.60	20.31
<i>Demographics</i>					
Age	27.90	30.20	32.57	34.77	31.57
Highest Grade	13.61	13.87	13.64	13.30	13.69
Mother Highest Grade	12.13	11.73	11.13	10.65	11.39
Father Highest Grade	11.64	11.73	10.96	10.08	11.22
Number of Children	0.84	0.95	1.32	1.43	1.13
Age of Youngest Child	2.43	2.72	3.97	4.68	3.39
Never Married	0.36	0.31	0.27	0.31	0.30
Married	0.41	0.54	0.59	0.52	0.54
Separated	0.04	0.03	0.04	0.05	0.04
Divorced	0.18	0.12	0.10	0.12	0.12
Widowed	0.00	0.00	0.00	0.00	0.00

Table A3: Summary Statistics of NLSY 1987-2006 for Males

	Underweight	Normal	Overweight	Obese	Total
BMI Adjusted	17.56	22.82	27.18	33.87	26.68
<i>Employment Variables</i>					
Employed	0.83	0.94	0.96	0.96	0.95
Hourly Wage in Dollars in 2010 \$	8.75	12.80	15.86	15.81	14.57
Ln(Wage) in 2010 \$	2.34	2.75	2.90	2.85	2.83
High personal interactions job	0.36	0.49	0.54	0.51	0.51
Tenure (weeks)	100.74	183.62	260.98	304.17	237.17
Currently in School	0.11	0.11	0.06	0.04	0.08
Experience (1000s hours)	11.24	19.78	27.84	33.30	25.55
<i>Demographics</i>					
Age	25.89	29.15	32.66	34.84	31.63
Highest Grade	11.98	13.34	13.41	13.01	13.29
Mother Highest Grade	9.91	11.33	11.21	11.03	11.22
Father Highest Grade	9.36	11.27	11.25	10.70	11.14
Number of Children	0.37	0.78	1.14	1.35	1.03
Age of Youngest Child	0.34	1.34	2.47	3.36	2.18
Never Married	0.72	0.49	0.31	0.28	0.38
Married	0.22	0.41	0.58	0.60	0.51
Separated	0.01	0.03	0.02	0.02	0.02
Divorced	0.04	0.07	0.08	0.10	0.08
Widowed	0.01	0.00	0.00	0.00	0.00

Table A4: Summary Statistics Comparing Estimation and Excluded Sample NLSY 1987-2006

	Estimation Sample	Non-Sibling Sample
<i>Employment Variables</i>		
Employed	0.95	0.91
Hourly Wage in Dollars in 2010 \$	13.19	11.83
Ln(Wage) in 2010 \$	2.72	2.64
High personal interactions job	0.56	0.55
Tenure (weeks)	226.63	129.66
Currently in School	0.09	0.13
Experience (1000s hours)	23.14	16.89
<i>Body Mass Index</i>		
BMI Adjusted	26.22	26.33
Underweight (BMI < 18.5)	0.02	0.02
Overweight (25 < BMI < 30)	0.32	0.31
Obese (BMI: >30)	0.20	0.21
<i>Demographics</i>		
Age	31.60	30.39
Highest Grade	13.48	9.79
Mother Highest Grade	11.30	10.84
Father Highest Grade	11.18	10.39
Number of Children	1.08	0.81
Age of Youngest Child	2.73	2.24
Never Married	0.34	0.37
Married	0.53	0.49
Separated	0.03	0.03
Divorced	0.10	0.10
Widowed	0.00	0.01
N × T	44,066	218,241

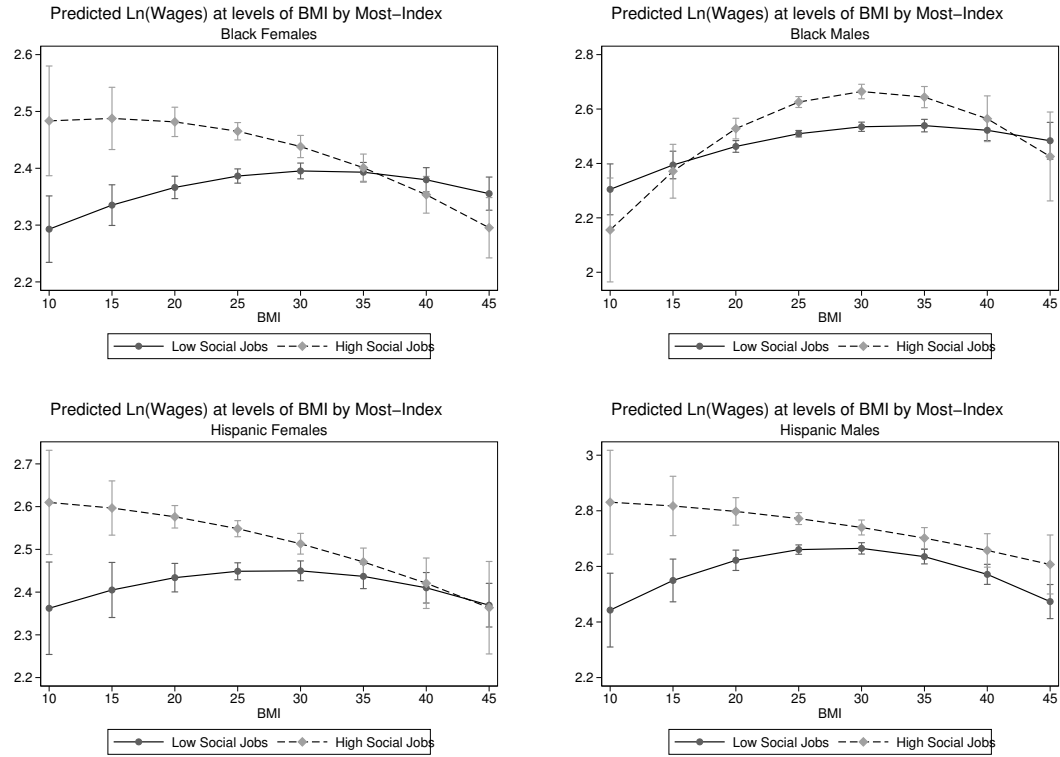
B Appendix: job descriptors

We report below the set of abilities, skills and work activities we used to construct the index that encompasses the level of personal interaction required on the job.

- Work Activities:
 - Communicating with Supervisors, Peers, or Subordinates
 - Communicating with Persons Outside Organization
 - Establishing and Maintaining Interpersonal Relationships
 - Assisting and Caring for Others
 - Selling or Influencing Others
 - Resolving Conflicts and Negotiating with Others
 - Coordinating the Work and Activities of Others
- Skills:
 - Speaking
 - Social Perceptiveness
 - Coordination
 - Persuasion
 - Negotiation
 - Instructing
 - Service Orientation

C Relationship between BMI and wages, other groups

Figure C1: Relationship between BMI and wages across race and gender groups



References

Averett, Susan and Saners Korenman, "The economic reality of The Beauty Myth," *The Journal of Human Resources*, 1996, 31 (2), 304–30.

Baum, Charles L and Shin-Yi Chou, "The Socio-economic causes of obesity," Technical Report, National Bureau of Economic Research 2011.

- Becker, Gary S.**, *The Economics of Discrimination*, University of Chicago Press, Chicago, 1957.
- Bhattacharya, Jay and M Kate Bundorf**, “The incidence of the healthcare costs of obesity,” *Journal of health economics*, 2009, 28 (3), 649–658.
- Borjas, George J.**, “Self-Selection and the Earnings of immigrants,” *American Economic Review*, September 1987, 77 (4), 531–553.
- Bound, John, Charles Brown, and Nancy Mathiowetz**, “Measurement error in survey data,” in “Handbook of econometrics,” Vol. 5, Elsevier, 2001, pp. 3705–3843.
- Burkhauser, Richard V and John Cawley**, “Beyond BMI: the value of more accurate measures of fatness and obesity in social science research,” *Journal of health economics*, 2008, 27 (2), 519–529.
- Cawley, John**, “The impact of obesity on wages,” *Journal of Human Resources*, 2004, 39 (2), 451–474.
- Courtemanche, Charles, Garth Heutel, and Patrick McAlvanah**, “Impatience, incentives and obesity,” *The Economic Journal*, 2014, 125 (582), 1–31.
- Currie, Janet and Brigitte C Madrian**, “Health, health insurance and the labor market,” *Handbook of labor economics*, 1999, 3, 3309–3416.
- DePasquale, Christina and Mario Macis**, “Have the obesity employment and wage penalties disappeared? Evidence from the NLSY, 1989-2008,” 2016. Unpublished typescript.

- Dow, William H and Edward C Norton**, “Choosing between and interpreting the Heckit and two-part models for corner solutions,” *Health Services and Outcomes Research Methodology*, 2003, 4 (1), 5–18.
- Dustmann, Christian and María Engracia Rochina-Barrachina**, “Selection correction in panel data models: An application to the estimation of females’ wage equations,” *The Econometrics Journal*, 2007, 10 (2), 263–293.
- Elfhag, Kristina and Lesley C Morey**, “Personality traits and eating behavior in the obese: poor self-control in emotional and external eating but personality assets in restrained eating,” *Eating behaviors*, 2008, 9 (3), 285–293.
- Fang, Hanming and Andrea Moro**, “Theories of Statistical Discrimination and Affirmative Action: A Survey,” in Jess Benhabib, Matthew O. Jackson, and Alberto Bisin, eds., *Handbook of Social Economics, Volume 1A*, Vol. 1A, The Netherlands: North Holland, 2011, chapter V, pp. 133–200.
- Hamermesh, Daniel and Jeff E. Biddle**, “Beauty and the Labor Market,” *American Economic Review*, 1994, 84 (5), 1174–94.
- Han, Euna, Edward C Norton, and Lisa M Powell**, “Direct and indirect effects of body weight on adult wages,” *Economics & Human Biology*, 2011, 9 (4), 381–392.
- , **Edward Norton, and Sally Stearns**, “Weight and Wages: Fat Versus Lean Paychecks,” *Health Economics*, 2009, (18), 535–548.
- Harris, Matthew**, “The Impact of Body Weight on Occupational Mobility and Career Development,” *Working Paper*, 2017.
- Heckman, James J.**, “Sample Selection Bias as a Specification Error,” *Econometrica*, 1979, 47 (1), 153–161.

—, “The Empirical Content of the Roy Model,” *Econometrica*, September 1990, *58* (5), 1121–1149.

Heckman, James J and Guilherme Sedlacek, “Heterogeneity, aggregation, and market wage functions: an empirical model of self-selection in the labor market,” *The Journal of Political Economy*, December 1985, *93* (6), 1077–1125.

Keane, Michael P and Kenneth I Wolpin, “The career decisions of young men,” *Journal of political Economy*, 1997, *105* (3), 473–522.

Kyriazidou, Ekaterini, “Estimation of a panel data sample selection model,” *Econometrica*, 1997, pp. 1335–1364.

Lee, Lung-Fei and Jungsywan H. Sepanski, “Estimation of linear and non-linear errors-in-variables models using validation data,” *Journal of the American Statistical Association*, 1995, *90* (429), 130–140.

Loury, Linda Datcher, “Some contacts are more equal than others: Informal networks, job tenure, and wages,” *Journal of Labor Economics*, 2006, *24* (2), 299–318.

Lundborg, Petter, Paul Nystedt, and Dan-Olof Rooth, “Body size, skills, and income: evidence from 150,000 teenage siblings,” *Demography*, 2014, *51* (5), 1573–1596.

Morris, Stephen, “Body mass index and occupational attainment,” *Journal of Health Economics*, 2006, *25* (2), 347–364.

Pagan, Jose A and Alberto Davila, “Obesity, occupational attainment, and earnings: Consequences of obesity,” *Social Science Quarterly*, 1997, *78* (3), 756–770.

- Pellizzari, Michele**, “Do friends and relatives really help in getting a good job?,” *ILR Review*, 2010, *63* (3), 494–510.
- Phelps, Edmund S**, “The Statistical Theory of Racism and Sexism,” *The American Economic Review*, 1972, *62* (4), 659–661.
- Puhani, Patrick**, “The Heckman correction for sample selection and its critique,” *Journal of economic surveys*, 2000, *14* (1), 53–68.
- Register, Charles A and Donald R Williams**, “Wage effects of obesity among young workers,” *Social Science Quarterly*, 1990, *71* (1), 130–141.
- Rooth, Dan-Olof**, “Obesity, attractiveness, and differential treatment in hiring a field experiment,” *Journal of human resources*, 2009, *44* (3), 710–735.
- Roy, Andrew Donald**, “Some thoughts on the distribution of earnings,” *Oxford economic papers*, 1951, *3* (2), 135–146.
- Shinall, Jennifer B.**, “Why Obese Workers Earn Less: Occupational Sorting and its implications for the legal system,” *Social Science Research Network*, 2014.
- Sutin, Angelina R, Luigi Ferrucci, Alan B Zonderman, and Antonio Terracciano**, “Personality and obesity across the adult life span,” *Journal of personality and social psychology*, 2011, *101* (3), 579.
- Topa, Giorgio**, “Labor markets and referrals,” in “Handbook of social economics,” Vol. 1, Elsevier, 2011, pp. 1193–1221.
- Willis, Robert J. and Sherwin Rosen**, “Education and Self-Selection,” *Journal of Political Economy*, October 1979, *87* (5), S7–S36.