International Journal of Health Economics and Management
https://doi.org/10.1007/s10754-021-09315-4

RESEARCH ARTICLE

Temporary and persistent overweight and long-term labor market outcomes

Liisa T. Laine[1,3] · **Ari Hyytinen**[2]

Received: 16 February 2021 / Accepted: 12 October 2021

Abstract
We study how the duration of being overweight earlier in life is related to subsequent long-term labor market outcomes. Our data on fraternal and identical twins born and raised in the same household contain weight measurements of the twins during their early adulthood measured in 1975, 1981, and 1990 and is linked to register-based administrative data on the earnings and employment from 1990 to 2009. When combined, these data enable an empirical strategy that controls for the family environment and genes shared by twins. We find that being persistently overweight during early adulthood is negatively associated with long-term earnings for both women and men. We find that for women, the association is driven by a decrease in labor market-attachment, whereas for men, the association is driven by lower annual earnings.

Keywords Overweight · Obesity · Long-term earnings · Labor market attachment · Genes

JEL Classification J21 · I10

Introduction

Obesity is a global public health problem. The prevalence of obesity has almost tripled between 1975 and 2016 worldwide; and in 2016, more than 1.9 billion adults were overweight, and of these, over 650 million adults were obese. (WHO, 2018) Being obese is

We thank Anirban Basu, Petri Böckerman, Norma B. Coe, Randall Ellis, Jaakko Kaprio, Ching-to Albert Ma, Terhi Maczulskij, Miikka Rokkanen, Benjamin Solow, Wenjia Zhu, and the participants of the ASHEcon 2014 in Los Angeles and the Boston University Department of Economics Health Economics reading group for their comments and suggestions. Dr Laine gratefully acknowledges funding from the the Yrjö Jahnsson Foundation (research Grant No. 6578) and the KAUTE Foundation.

✉ Liisa T. Laine
lainel@wharton.upenn.edu

1 The Wharton School, University of Pennsylvania, Philadelphia, USA

2 Hanken School of Economics and Helsinki GSE, Helsinki, Finland

3 Present Address: Leonard Davis Institute, University of Pennsylvania, 3641 Locust Walk, 308 CPC, Philadelphia 19104-6218, USA

Published online: 09 November 2021

a significant risk factor for various noncommunicable diseases, such as cardiovascular diseases. Obesity is also associated with higher medical costs (Cawley and Meyerhoefer, 2012) and worse labor market outcomes, such as lower earnings (wage penalty), weaker labor market attachment (Brunello and Beatrice, 2007; Morris, 2007; Averett, 2011; Cawley, 2015) and discrimination (Campos-Vazquez and Gonzalez, 2020).

Previous studies have shown that the weight penalty on earnings is more significant for individuals with a history of being chronically overweight or obese and that these effects are gender-specific (Han et al., 2011; Chen, 2012; Pinkston, 2017). A subtler point, however, is how unmeasured genetic factors shape this relationship. It is known that the shared home environment (Segal and Allison, 2002; Gregory and Ruhm, 2011; Datar, 2017) and other family-level peer effects (Gwozdz et al., 2015) affect weight development. Genes can have additional importance because the heritable component of weight is very high (Cawley, 2004; Farooqi and O'Rahilly, 2007; Cawley, 2015) and because labor market outcomes are also known to have such a component (Benjamin et al., 2012).

We study the association between different durations of being overweight and long-term labor market outcomes and focus, in particular, on the role unmeasured family environment and genetic factors. We use data from Finland containing multiple weight and height measures of fraternal and identical twins during their early adulthood over a period that covers 15 years. We link these data to comprehensive administrative data of twins' subsequent earnings and labor-market attachment information over a 20-year period.

Our empirical strategy exploits within-twin variation to control for the omitted variable bias arising from the family environment and genetic factors.[1] We estimate twin-differenced models using two samples. First, we use the full sample of fraternal and identical twins to control for family environment. Second, because fraternal twins share approximately half of their genes but identical twins share all of their genes (Zwijnenburg et al., 2010), we can use the sample of identical twins to control for genes and the differences in appearance (Jenq et al., 2015).

Several mechanisms, such as lower productivity and various forms of labor market discrimination (Campos-Vazquez and Gonzalez, 2020), have been proposed as explanations for the negative relationship between being overweight and labor market outcomes (Cawley, 2004). Many of the mechanisms, however, are more critical at specific points in life. To test the relative importance of the different types of overweight histories, we create measures of weight histories by dividing individuals into four exclusive groups measuring whether individuals are currently overweight or have been persistently overweight, whether they have lost or gained weight or whether they have never been overweight.

We use our data and econometric framework to explore the underlying mechanisms as follows. First, we study the link between being overweight earlier in life versus being overweight later in life and long-term labor market outcomes. Second, we exploit the feature of our data that we can measure an individual's weight histories at different points of time corresponding for twins being aged 17–31, 23–37, and 32–46 in each measurement year.[2]

[1] Extensive literature has used within-sibling variation to study, for instance, the effects of education on economic outcomes in different contexts (Griliches, 1979; Altonji and Dunn, 1996; Aaronson, 1998; Sacerdote, 2007; Abramitzky et al., 2012). Specifically, our approach is related to Behrman and Rosenzweig (2002) and Behrman et al. (1996) who estimate the returns to education using twin-data.

[2] The age range in our study is slightly wider than in prior literature: Han et al. (2011) study the effect of adolescent BMI on education and occupation. The cohorts in Chen (2012) are teens (age 16), young adults (age 23), and adults (age 33). In Pinkston (2017) ages vary between 16–30.

The benefit of using the long-term earnings instead of a cross-sectional measure of wages as an outcome is that the long-term earnings outcome captures the cumulative labor market effects of being overweight. We also study the association between being overweight and long-term average employment. This measure reflects the genuine labor market attachment better than cross-sectional measures of labor market status, such as temporary unemployment spells. We can thus distinguish whether being overweight is associated with lifetime earnings because of lower annual earnings or because of weaker labor market attachment.

We find that being persistently overweight is associated with adverse long-term labor market outcomes. Genetic factors shape the relationship between overweight history and labor market outcomes more for men than for women. For men, the associations are uncovered only after genetics are fully controlled for. For women, the results are robust across all specifications. This difference implies that the extent to which genetic factors are a source of omitted variable bias may differ across women and men. It also provides a potential explanation for the lack of consensus in the previous literature on the existence and magnitude of the earnings penalty for overweight men (Sargent and Blanchflower, 1994; Averett and Korenman, 1996; Behrman and Rosenzweig, 2001; Baum and Ford, 2004; Cawley, 2004; Brown and Routon, 2018).

Data and empirical strategy

Data

We use information from two data sets. The first is the Older Finnish Twin Cohort Study, a survey-based panel data that covers all Finnish twin pairs born before 1958, surveyed in 1975, 1981, and 1990 (Kaprio et al., 1979). These data provide the weight and height information for the measures for Body Mass Index (BMI), which we use to construct our weight history variables.

The second data is the Finnish Longitudinal Employer-Employee Data, a comprehensive individual-level longitudinal data on the education and labor market outcomes of the working-age population in Finland from 1990 to 2009. The data is collected and maintained by Statistics Finland and constructed from several different administrative registers. We use these data to create our long-term labor market outcomes. The two data sets were linked by using personal coded identifiers by Statistics Finland (see online Appendix and Hyytinen et al. 2013).[3]

Sample construction: We restrict our sample to same-sex twin pairs born after 1944 but before 1958 to measure our outcomes over a period that covers people's primary working age. Thus, our estimation sample consists of 2,364 women and 1,564 men who were from 32 to 46 years old in 1990 (the first year of our earnings data), and from 51 to 65 years old in 2009 (the last year of our earnings data). Over representation of women is common in

[3] Record linkages of the cohort study data conform to the Finnish Data Protection Act and were originally approved by the ethical committee of the Department of Public Health, University of Helsinki (Kaprio et al., 1979). All the data work of this paper was carried out at Statistics Finland, following its terms and conditions of confidentiality.

survey-based twin data (Silventoinen et al., 2015). 1,182 of twin pairs are women of which 747 are fraternal, and 435 identical, and 782 of twin pairs are male, of which 515 are fraternal and 267 identical.[4]

Outcome variables: We study the association between individual's overweight history and three long-term labor market outcome measures. Our first outcome variable is long-term earnings, which is the natural logarithm of the average annual gross earnings in year 2000 euros, measured over the period 1990– 2009. Our earnings measure includes wages, salaries, and entrepreneurial income.

Our second outcome variable is long-term average employment, which we calculate as the average annual employment months. We use it to distinguish if the association is caused by a direct link on earnings or because of it being linked to weaker labor market attachment. Our third outcome variable is the average monthly earnings (in year 2000 euros), calculated as the natural logarithm of the ratio of average annual earnings and average annual employment months measured over the period 1990–2009.

An advantage of our earnings and employment data is that it is constructed from several different administrative registers, including annual information on earnings and other income such as social benefits, employment, and education. Thus, unlike earnings and employment measures based on survey data sources, our earnings and employment measures do not suffer from common survey data problems such as recall error or over- or underreporting. The second benefit of our data is that the earnings data are not top-coded.

Constructing individual weight histories using BMI metrics: Following the prior literature studying the effects of overweight history (Han et al., 2011; Chen, 2012; Pinkston, 2017), we use BMI to construct our overweight history measures. We use the height and weight information of each individual to calculate the individual's BMI in 1975, 1981, and 1990.

The individuals in our sample were aged 17–31, 23–37, and 32–46, during the respective survey years. In comparison to the previous literature the age range here is slightly wider: Han et al. (2011) study the effect on adolescent BMI on education and occupation, the cohorts in Chen (2012) were teens (age 16), young adults (age 23), and adults (age 33), and in Pinkston (2017) ages vary between 16–30.

We then classify individuals with a $BMI \geq 25.00$ as being overweight. The threshold is the same for women and men.[5] Our focus on being overweight instead of obesity ($BMI \geq 30.00$) is motivated by the previous literature which suggests that the relevant BMI range for adverse labor market outcomes may be below the official WHO obesity threshold (Gregory and Ruhm, 2011), especially for women (Caliendo and Gehrsitz, 2016). This allows us to also to make qualitative comparison with the previous literature and provide a broader international contribution to the literature which has used mostly American data. To consider the sensitivity of our findings to non-linearities (Caliendo and Gehrsitz, 2016), we also show and discuss the results using *different* BMI thresholds in section Robustness.[6]

[4] We acknowledge that the number of twin pairs is relatively small. However, the problem and the possible consequences of small sample size, for example on the power of the study, are shared by all studies using twin data.

[5] As discussed below, these commonly used thresholds (in research and clinical work) are obviously proxies of excess body weight and have their limitations (see below). Anyhow, we use the official cutoffs determined by WHO for transparency instead of relying more "ad hoc" measures.

[6] Figure A1 plots the local polynomial regressions of each long-term outcome on BMI (average over time) and Figure A2 plots twin-differenced local polynomial regressions of each long-term outcome on twin-differenced BMI (average over time). The former provide support for our use of BMI = 25 as the threshold for being overweight.

Using BMI, we divide our sample into four exclusive groups called persistently overweight, previously overweight, currently overweight, and never overweight. We call an individual persistently overweight if the individual was overweight in either 1975 or 1981 (or both) and was also overweight in 1990. We call an individual previously overweight if the individual was overweight in 1975 or 1981 but was not overweight in 1990. We call an individual currently overweight if the individual was overweight during the year we started measuring our labor market outcomes in 1990 but was not overweight in 1975 or 1981. We use the term "current" because it refers to the first year of our earnings data. Last, we call an individual never overweight, if the individual was not overweight in any of the survey years 1975, 1981, and 1990. Besides the alternative values for the BMI thresholds, we have studied the robustness of our results by using models with a different threshold specification that divides BMI into 3-year and gender-specific terciles (section Robustness).

A critique against BMI as a measure of being overweight (an excess body fat) is that it does not take age, sex, or body composition into account (Burkhauser and Cawley, 2008). Because muscles weigh more than fat, BMI can underestimate the amount of body fat in people with low muscle mass, and it can overestimate the amount of body fat on lean people with high muscle mass.

We consider this critique to be much less of a concern to us because the body composition between twins can be assumed to be more similar than for two randomly chosen individuals. Even though body composition can differ between siblings (Price and Swigert, 2012), two randomly chosen individuals differ even more because of between-family variation in environment and genes. Moreover, the body composition of identical twins is as similar as it can biologically be because, in addition to their home environment, identical twins share nearly all of their genes (Zwijnenburg et al., 2010).

Using within-twin differencing to remove the effects of shared family background and genes also alleviates the concern that our results would be driven by cohort effects or the position of the individuals in the earnings distribution; see Brown and Routon (2018) who document the importance of these factors.

Descriptive statistics: Table 1 reports descriptive statistics. There are several differences between men and women. The mean of the logarihtm of long-term earnings is 9.37 for women (approximately 14,900 euros per year). The corresponding number for men is 9.70 (approximately 21,308 euros). Women (men) worked on average 7.9 (8.2) months per year, and their monthly mean earnings were 2300 (2600) euros.

We can also see that 10 (24) percent of women (men) were persistently overweight during the sample period. Also, 13 (17) percent of women (men) who were not overweight initially were overweight at the last measurement time in 1990. The percentage of those who had previously been overweight but were no longer overweight in 1990 is small for both genders, approximately 2–3 percent.

The average age in 1990 was 39 years both women and men. Women and men had, on average, 12 years of schooling in 1990.[7] The table also shows that men had slightly smaller number of diseases than women in 1975.

Table A1 in our online Appendix summarizes the number of observations (individuals) and the corresponding means of the absolute values of the twin-differenced variables for the fraternal and identical twins. The numbers indicate that there is quite a bit of variation in the twin-differenced data as well.

[7] Our education variable refers to the number of years in school, which is based on information on achieved degrees and standard degree times.

Table 1 Summary statistics: full, fraternal, and identical twin samples

Sample	Women			Men		
	Full	Fraternal	Identical	Full	Fraternal	Identical
Panel A: Dependent variables (1990-2009)						
Long-term earnings (in year 2000 euros, log)	9.37	9.36	9.38	9.70	9.72	9.69
Average employment months (months)	7.90	7.91	7.89	8.19	8.23	8.13
Monthly earnings (in year 2000 euros, log)	7.42	7.42	7.43	7.71	7.71	7.70
Panel B: Overweight history dummies						
Never overweight	0.75	0.74	0.76	0.56	0.52	0.63
Persistently overweight	0.10	0.11	0.09	0.24	0.26	0.19
Previously overweight	0.02	0.02	0.02	0.03	0.03	0.03
Currently overweight	0.13	0.13	0.13	0.17	0.19	0.15
Panel C: Control variables						
Age in 1990 (years)	38.70	38.61	38.85	39.18	39.20	39.12
Height in 1975 (cm)	163.48	163.65	163.20	176.55	176.71	176.24
Schooling (years)	12.21	12.18	12.26	12.24	12.12	12.46
Number of diseases 1975	0.50	0.49	0.52	0.39	0.38	0.42
Number of individuals	2364	1494	870	1,564	1030	534
Number of twin pairs	1,182	747	435	782	515	267

Variables in Panel A in addition to age and schooling variables in Panel C are from our education and earnings data. Variables in Panel B in addition to the height, number of diseases, and employment information in 1975 are from the twin data (Kaprio et al., 1979).

Table 2 displays the means of BMI measured at different points in time, age in 1990, and the outcome variables by the overweight history categories, as well as p-values from F-tests for equality of the group means. The table depicts a few patterns: first, average weight increases over time. For example, the average BMI of persistently overweight women (men) increases from 28.0 (27.0) in 1975 to 32.0 (28.9) in 1990.

Persistently overweight individuals are, on average, somewhat older than the rest of the individuals in the other groups. Average long-term earnings are lowest for those who are either persistently or currently overweight. Also, labor market attachment is weakest for those who were persistently overweight.

Empirical strategy

We estimate the following equation:

$$Y_{ij} = \beta_1 BMI_{ij}^{persistent} + \beta_2 BMI_{ij}^{previously} + \beta_3 BMI_{ij}^{currently} + \gamma' Z_{ij} + \delta_j + \epsilon_{ij}, \tag{1}$$

where Y_{ij} is one of our three outcome variables: long-term earnings, long-term average employment, or monthly earnings of twin i in pair j. Our weight history variables are

Table 2 Means in different overweight history categories

	Never overweight	Persistently overweight	Currently over-weight	Previously overweight	p-value of F-test
Panel A: Women					
BMI history dummy					
BMI in 1975	20.1	25.4	21.8	24.4	< 0.01
BMI in 1981	20.6	28.0	22.9	25.0	< 0.01
BMI in 1990	21.6	32.0	26.9	23.3	< 0.01
Age in 1990	38.5	39.8	39.1	39.2	< 0.01
Dependent variables (1990–2009)					
Long-term earnings (in 2000 euros, log)	9.42	9.12	9.30	9.06	< 0.01
Average employment months (months)	8.02	7.40	7.63	7.73	< 0.01
Monthly earnings (in 2000 euros, log)	7.46	7.26	7.39	7.22	< 0.01
Number of individuals	1,766	241	306	51	
Panel B: Men					
BMI history dummy					
BMI in 1975	21.2	25.3	22.4	24.3	< 0.01
BMI in 1981	21.8	26.9	23.5	25.4	< 0.01
BMI in 1990	22.6	28.2	26.4	24.2	< 0.01
Age in 1990	38.7	40.6	38.7	39.9	< 0.01
Dependent variables (1990–2009)					
Long-term earnings (in 2000 euros, log)	9.72	9.63	9.77	9.60	< 0.01
Average employment months (months)	8.31	7.89	8.29	7.90	< 0.01
Monthly earnings (in 2000 euros, log)	7.71	7.68	7.73	7.65	< 0.01
Number of individuals	874	369	272	49	

The first four columns report the average characteristics for the four overweight history categories. The last column reports p-values from F-tests for equality of the group means

$BMI_{ij}^{persistent}$, $BMI_{ij}^{previously}$, and $BMI_{ij}^{currently}$, referring to persistent, previous, and current overweight, respectively. Our omitted overweight history category is thus that of those who are never overweight. Z_{ij} includes the observed control variables, δ_j denotes the unobserved family environment and genes shared by twins, and ϵ_{ij} is the individual specific error term.

We begin by estimating equation (1) by ignoring δ_j and treating it as part of the error term, which is the same as estimating the equation in levels using variation both within and across twins (i.e., OLS regression; see Sect. 4.1). The consistency of this estimator requires that BMI is uncorrelated with both δ_j and ϵ_{ij}.

We then include twin fixed effects δ_j and estimate the full equation (1), which is the same as estimating the equation with within-twin differences and controls for family environment and genes (some or all) shared by twins. The consistency of this model does not require that BMI be uncorrelated with δ_j.

We estimate the twin-differenced model using two samples. We first estimate it using the full sample, which includes both fraternal and identical twins (Sect. 4.2). Using full sample allows us to control for the family environment shared by twins (but partially for the genetic differences). We also estimate the twin-differenced model using a sample of identical twins (Sect. 4.3). This specification allows us to study how unmeasured genetic factors are associated with the estimates (because identical twins share all of their genes).

We report standard errors clustered by twin pairs for models estimated in levels and heteroskedasticity-robust standard errors in twin-differences. We estimate all three specifications without control variables and then by using two sets of control variables.

Our main set of control variables includes exogenous control variables: age, age^2, and height. We also estimate the three specifications by using a more extensive set of controls: the exogenous controls (age, age^2, and height) and predetermined health (number of reported diseases in 1975) and education (years of schooling in 1990) as additional control variables. We control for (lagged) health because there is a well-documented connection between being overweight and weak health outcomes. We include schooling, since being obese in the late teenage years reduce the amount of acquired schooling and affects occupation outcomes (Han et al., 2011).

The results using the more extensive set of controls should be taken with a degree of caution. Adding education as a control can be problematic as it may be affected by being overweight in the past, thus giving rise to a bad control problem (Angrist and Pischke, 2009). We also do not have information of children in our data. Even if we had the data, including e.g. the number of children as a control would be problematic, because such a variable is an intermediate outcome (i.e., something that is affected by BMI and in turn affects long-term income).[8] Importantly, we view BMI as a proxy measure of several different weight-related factors that can be associated with long-term labor market outcomes.

A general challenge in the literature studying the economic consequences of being overweight and obese is to find plausibly exogenous variation in weight that allows one to identify the effect of being overweight or obese (Cawley, 2004). The earlier studies have used various methods and instruments to solve this problem. One of the more current approaches is to use genetic risk scores as an instrument (Böckerman et al., 2018). We cannot use this approach, because as far as we are aware, medical science and behavioral genetics have not identified genes that would predict individuals' weight in different parts of their lifecycle or, the duration of being overweight. Also, even if such instruments became available, it is not clear that they would satisfy the required exclusion restriction for instrumental variable estimation in our context.

[8] It is unlikely that our results are explained by children only. For example, Finland universal health coverage and all individuals are insured by the National Health Insurance scheme. All pregnant women also attend regular well-child visits prior and after labor. It is also likely that the existing family leave policies Finland alleviate the effects of children on labor market outcomes.

Conceptual framework

Our long-term earnings measure allows us to evaluate the overall association of the histories of being overweight with the individual's total lifetime earnings. If there is such an association, we can use our two other labor market outcomes, average employment months and monthly earnings, to explore the association in more detail.

On one hand, lower long-term earnings might be caused by a labor supply effect, which refers to a weaker labor market attachment. On the other hand, lower long-term earnings might also be caused by a wage penalty effect, which refers to lower earnings per amount worked. It is also possible that lower long-term earnings could be a combination of these two.

Many of the mechanisms proposed in the literature have a more significant role at specific points in life. One this kind of mechanism is skill formation in childhood. Being overweight as a child or in early adulthood can cause underdeveloped social skills, lower self-esteem, and poor communication skills. These adverse consequences of being overweight early in life can subsequently affect an individual's career path negatively and cause lower wages. (Heckman, 2006; Mobius and Rosenblat, 2006; Florin et al., 2011) This kind of mechanism is likely to be something that persists throughout an individual's life. Other mechanisms are more relevant in adulthood than in childhood. One is labor market discrimination. For example, promotions and salary raises could be granted less often to workers who are physically less attractive (Puhl and Brownell, 2001; Mobius and Rosenblat, 2006) or discrimination occurs during the interview and hiring stage (Rooth, 2009; Gregory and Ruhm, 2011; Han et al., 2009; Campos-Vazquez and Gonzalez, 2020). There is evidence that discrimination is worse against female workers (Caliendo and Lee, 2013; Campos-Vazquez and Gonzalez, 2020).

Another mechanism that is more relevant in adulthood is due to lower wages reflecting lower productivity. Currently overweight workers may, for example, be absent from work more often or less handy in performing certain types of physically demanding tasks. Their lower productivity may show up as lower wages.

Being overweight can also be a proxy for some unobservable characteristics. This mechanism persists throughout an individual's life. For instance, overweight individuals may have high discount rates and therefore invest less in education, health, and weight control (Falkner et al., 2001; Kristjansson et al., 2010; Sabia and Rees, 2012). These kinds of cumulative effects of being overweight on labor market outcomes are captured by persistent and previous overweight.

If the underlying mechanism is related to skill formation or permanent unobservable characteristics, we should see a non-negligible negative effect on long-term labor market outcomes for those who have been persistently or previously overweight. If the underlying mechanism is related more to labor market discrimination or lower productivity, we should see a non-negligible negative effect of being currently overweight on long-term labor market outcomes.

Firms can also impose a wage penalty for overweight individuals to compensate for their higher medical costs on their employer-provided health insurance (Bhattacharya and Bundorf, 2009). This mechanism is relevant in countries in which firms are required to provide employer-sponsored health insurance. This mechanism is not that relevant to us: all Finnish citizens have publicly funded National Health Insurance and thus Finnish employers have relatively weak incentives to internalize the medical costs of obesity.

We use the estimated coefficients for being persistently overweight (β_1), having been previously overweight (β_2), and being currently overweight (β_3) to test the relative importance of various weight histories. We can do so by focusing on the following five comparisons.

We test i) whether the effects of being overweight in different parts of life are different ($\beta_1 = \beta_2 = \beta_3$); ii) whether the effect of being persistently overweight is the same as the effect of having been previously overweight ($\beta_1 = \beta_2$); iii) whether the effect of persistently being overweight is the same as that of currently being overweight ($\beta_1 = \beta_3$); iv) whether the effect of being previously overweight is the same as that of being currently overweight ($\beta_2 = \beta_3$); and v) whether the effect of being persistently overweight is larger (or smaller) than the sum of the effects of having been overweight in the past and of being overweight currently ($\beta_1 = \beta_2 + \beta_3$).

The effect of being persistently overweight could be larger than just the sum of being previously and currently overweight if the adverse effects accumulate or increase over time. It could also be smaller if those who are persistently overweight learn to take compensatory measures to alleviate the adverse labor market consequences of being overweight.

Results

OLS regression

We begin by estimating equation (1) using variation both within and across twin pairs. Table 3 displays the results for our three outcome variables and for the three different regression specifications. Panel A of the table reports the results for women and panel B the results for men. The lower parts of both panels summarize the p-values for the test of the relative importance of the various patterns of overweight history on the labor market outcomes.

Based on the OLS regressions, we find different associations of being overweight for women and men. Our results show that persistently overweight women have 20 to 30 percent smaller long-term earnings than women who are never overweight. Persistently overweight women also work 0.4 to 0.6 months less per year and have 13 to 19 percent smaller monthly earnings than women who have never been overweight. The OLS coefficients for women become somewhat smaller (in absolute values) when education and health are included as controls. We do not find similar associations for men.

The results for transitory obesity measures, previously overweight and currently overweight, are mixed. The coefficients are consistently negative for women, but the estimated associations are no longer statistically significant at the 5 percent level when education and health are included as controls. For men, the coefficients are smaller and never significant in regressions that include controls.

Within-twin differences: the role of family environment

Previous studies have shown that family environment and genetic effects are potential sources of omitted variable bias because of their association with variation in overweight and long-term labor market outcomes. This implies that the OLS results should be taken with caution.

Table 3 OLS regressions, full sample by gender

	Long-term earnings			Average employment months			Monthly earnings		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Panel A: Women									
Overweight history categories									
Persistently overweight (β_1)	−0.306***	−0.277***	−0.207***	−0.628***	−0.521***	−0.443**	−0.192***	−0.188***	−0.130***
	(0.067)	(0.065)	(0.062)	(0.195)	(0.192)	(0.188)	(0.039)	(0.039)	(0.036)
Previously overweight (β_2)	−0.365**	−0.359**	−0.298	−0.289	−0.297	−0.236	−0.233**	−0.228**	−0.176
	(0.171)	(0.169)	(0.178)	(0.420)	(0.414)	(0.421)	(0.109)	(0.109)	(0.113)
Currently overweight (β_3)	−0.125**	−0.106**	−0.057	−0.389**	−0.332	−0.289	−0.063**	−0.058	−0.014
	(0.051)	(0.051)	(0.050)	(0.168)	(0.169)	(0.169)	(0.031)	(0.031)	(0.029)
Controls									
Age	N	Y	Y	N	Y	Y	N	Y	Y
Height	N	Y	Y	N	Y	Y	N	Y	Y
Education	N	N	Y	N	N	Y	N	N	Y
Health	N	N	Y	N	N	Y	N	N	Y
Tests (p−values)									
$\beta_1 = \beta_2 = \beta_3 = 0$	0.000	0.000	0.003	0.003	0.017	0.055	0.000	0.000	0.002
Joint test of controls	–	0.001	0.000	–	0.000	0.000	–	0.000	0.000
$\beta_1 = \beta_2 = \beta_3$	0.043	0.051	0.083	0.541	0.701	0.774	0.009	0.008	0.016
$\beta_1 = \beta_2 + \beta_3$	0.329	0.313	0.445	0.915	0.818	0.864	0.382	0.411	0.619
$\beta_1 = \beta_2$	0.745	0.646	0.624	0.444	0.610	0.640	0.721	0.727	0.697
$\beta_1 = \beta_3$	0.022	0.030	0.047	0.325	0.432	0.517	0.004	0.004	0.008
$\beta_2 = \beta_3$	0.170	0.146	0.183	0.820	0.937	0.904	0.126	0.127	0.162
Number of individuals	2,364	2,364	2,364	2,364	2,364	2,364	2,364	2,364	2,364
Panel B: Men									
Overweight history categories									

Table 3 (continued)

	Long-term earnings			Average employment months			Monthly earnings		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Persistently overweight (β_1)	−0.097	−0.077	0.036	−0.418**	−0.198	−0.058	−0.027	−0.044	0.038
	(0.062)	(0.061)	(0.058)	(0.170)	(0.162)	(0.159)	(0.037)	(0.036)	(0.034)
Previously overweight (β_2)	−0.122	−0.108	−0.015	−0.402	−0.288	−0.189	−0.064	−0.067	0.005
	(0.134)	(0.128)	(0.122)	(0.421)	(0.380)	(0.373)	(0.076)	(0.077)	(0.072)
Currently overweight (β_3)	0.048	0.048	0.060	−0.015	−0.011	0.008	0.018	0.017	0.024
	(0.060)	(0.060)	(0.055)	(0.169)	(0.168)	(0.165)	(0.038)	(0.038)	(0.033)
Controls									
Age	N	Y	Y	N	Y	Y	N	Y	Y
Height	N	Y	Y	N	Y	Y	N	Y	Y
Education	N	N	Y	N	N	Y	N	N	Y
Health	N	N	Y	N	N	Y	N	N	Y
Tests (p-values)									
$\beta_1 = \beta_2 = \beta_3 = 0$	0.161	0.284	0.718	0.072	0.577	0.943	0.644	0.443	0.697
Joint test of controls	–	0.000	0.000	–	0.000	0.000	–	0.000	0.000
$\beta_1 = \beta_2 = \beta_3$	0.088	0.161	0.813	0.122	0.571	0.859	0.460	0.319	0.879
$\beta_1 = \beta_2 + \beta_3$	0.878	0.906	0.947	0.998	0.812	0.768	0.828	0.941	0.913
$\beta_1 = \beta_2$	0.856	0.812	0.684	0.970	0.817	0.733	0.642	0.763	0.656
$\beta_1 = \beta_3$	0.037	0.075	0.707	0.045	0.340	0.733	0.305	0.169	0.729
$\beta_2 = \beta_3$	0.219	0.241	0.548	0.374	0.485	0.612	0.314	0.297	0.797
Number of individuals	1,564	1,564	1,564	1,564	1,564	1,564	1,564	1,564	1,564

The control variables are age and age squared in 1990, height in 1975 (in centimeters), schooling in 1975 (in years), and the number of diagnosed diseases in 1975. Standard errors are in parentheses (clustered by twin pair). $**p < 0.05$, $***p < 0.01$.

Table 4 Twin fixed effect regressions, full sample by gender

	ΔLong-term earnings			ΔAverage employment months			ΔMonthly earnings		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Panel A: Women									
Overweight history categories									
ΔPersistently overweight (β_1)	−0.298***	−0.290***	−0.270***	−0.780**	−0.785**	−0.769**	−0.170**	−0.162**	−0.148
	(0.106)	(0.105)	(0.104)	(0.318)	(0.319)	(0.319)	(0.082)	(0.082)	(0.081)
Δ Previously overweight (β_2)	−0.353	−0.365	−0.378	−0.388	−0.382	−0.349	−0.175	−0.185	−0.204
	(0.221)	(0.221)	(0.216)	(0.478)	(0.480)	(0.484)	(0.148)	(0.149)	(0.142)
ΔCurrently overweight (β_3)	−0.118	−0.109	−0.092	−0.658***	−0.663***	−0.653***	−0.029	−0.021	−0.008
	(0.067)	(0.067)	(0.066)	(0.217)	(0.218)	(0.219)	(0.044)	(0.044)	(0.042)
Controls									
ΔHeight	N	Y	Y	N	Y	Y	N	Y	Y
ΔEducation	N	N	Y	N	N	Y	N	N	Y
ΔHealth	N	N	Y	N	N	Y	N	N	Y
Tests (p-values)									
$\beta_1 = \beta_2 = \beta_3 = 0$	0.020	0.024	0.037	0.007	0.007	0.009	0.149	0.153	0.135
Joint test of controls	–	0.087	0.000	–	0.742	0.329	–	0.012	0.000
$\beta_1 = \beta_2 = \beta_3$	0.219	0.204	0.170	0.744	0.736	0.724	0.128	0.112	0.078
$\beta_1 = \beta_2 + \beta_3$	0.468	0.438	0.387	0.635	0.644	0.680	0.838	0.792	0.691
$\beta_1 = \beta_2$	0.803	0.733	0.613	0.444	0.435	0.422	0.976	0.889	0.722
$\beta_1 = \beta_3$	0.112	0.111	0.108	0.716	0.717	0.730	0.070	0.068	0.065
$\beta_2 = \beta_3$	0.277	0.240	0.177	0.578	0.565	0.538	0.322	0.269	0.167
Number of individuals	1,182	1,182	1,182	1,182	1,182	1,182	1,182	1,182	1,182
Panel B: Men									
Overweight history categories									
ΔPersistently overweight (β_1)	−0.021	−0.021	0.039	−0.012	−0.012	0.048	−0.040	−0.040	0.006
	(0.095)	(0.095)	(0.092)	(0.248)	(0.248)	(0.246)	(0.058)	(0.058)	(0.056)

Table 4 (continued)

	ΔLong-term earnings			ΔAverage employment months			ΔMonthly earnings		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
ΔPreviously overweight (β_2)	−0.311**	−0.307**	−0.251	−0.781	−0.783	−0.761	−0.153	−0.145	−0.094
	(0.148)	(0.146)	(0.140)	(0.436)	(0.435)	(0.443)	(0.107)	(0.104)	(0.096)
ΔCurrently overweight (β_3)	−0.047	−0.046	−0.037	−0.296	−0.297	−0.297	−0.035	−0.032	−0.023
	(0.082)	(0.082)	(0.081)	(0.219)	(0.220)	(0.220)	(0.051)	(0.051)	(0.050)
Controls									
ΔHeight	N	Y	Y	N	Y	Y	N	Y	Y
ΔEducation	N	N	Y	N	N	Y	N	N	Y
ΔHealth	N	N	Y	N	N	Y	N	N	Y
Tests (p-values)									
$\beta_1 = \beta_2 = \beta_3 = 0$	0.185	0.186	0.184	0.176	0.172	0.150	0.534	0.561	0.707
Joint test of controls	–	0.463	0.000	–	0.895	0.139	–	0.042	0.000
$\beta_1 = \beta_2 = \beta_3$	0.128	0.130	0.100	0.171	0.167	0.122	0.512	0.532	0.534
$\beta_1 = \beta_2 + \beta_3$	0.039	0.041	0.035	0.029	0.028	0.024	0.202	0.227	0.241
$\beta_1 = \beta_2$	0.045	0.046	0.033	0.078	0.077	0.067	0.269	0.291	0.275
$\beta_1 = \beta_3$	0.772	0.791	0.407	0.264	0.263	0.175	0.936	0.890	0.630
$\beta_2 = \beta_3$	0.076	0.076	0.136	0.272	0.271	0.302	0.269	0.278	0.469
Number of individuals	782	782	782	782	782	782	782	782	782

The control variables are height in 1975 (in centimeters), schooling in 1975 (in years), and the number of diagnosed diseases in 1975. All dependent, explanatory, and control variables are within-twin differenced (denoted by Δ). Standard errors are in parentheses (heteroscedasticity robust). $**p < 0.05$, $***p < 0.01$

We address these concerns by estimating models with twin fixed effects, using the full samples of women and men that include both fraternal and identical twins. In these regressions, we control for shared family background and for genetic differences partially, because fraternal twins share approximately half of their genes while identical twins share all of their genes.

Table 4 shows more and stronger associations for women after the effect of family environment and genes shared by twins are controlled for. Persistently overweight women have approximately 30 percent smaller long-term earnings than women who are never overweight. Persistently overweight women also work 0.8 months less and have 15–17 percent smaller monthly earnings than women who are never overweight. For men, there is no such association in regard to being persistently overweight.

Also, for currently overweight women, the association between being overweight and average employment months becomes stronger after the family environment and genes shared by twin siblings are controlled for. There is no such association for currently overweight men.

How do previously overweight individuals who have been able to lose weight differ from the others? The number of individuals who were able to lose weight is small (3 percent of people in our data, see also Table 1). The estimated effects of having been previously overweight on long-term labor market outcomes are insignificant and not very robust. This result is due to the small number of female and male twin pairs in which one, but not both, of the twins would belong to the overweight history category of previously overweight individuals. Thus, it seems that the effect of losing weight is difficult to quantify, especially if one wants to control for shared family environment and genes.

Our results suggest that for women being overweight matters for long-term earnings and labor market attachment. For these two outcome variables, the joint null hypothesis $\beta_1 = \beta_2 = \beta_3 = 0$ is rejected at better than the 5 percent level.

Our measure for being persistently overweight allows capturing the cumulative effects of being overweight on labor market outcomes. The underlying mechanism may, therefore, be related to skill formation or other permanent unobservable characteristics of individuals. For women, the negative relationship between being persistently overweight and long-term earnings is related to their weaker labor market attachment. For them, the mechanism can be related to something that erodes women's labor market attachment throughout life, for example, labor supply decisions or the decision to have children.

Within-twin differences: the role of genetics

We now consider the importance of unobserved genetic factors and estimate equation (1) with twin fixed effects, using the sample of identical twins.

The results in Table 5 suggest that genetic factors seem to shape labor market outcomes of men more than those of women. We now find that persistently overweight identical male twins have approximately 45–55 percent smaller long-term earnings than those who are never overweight. This result is driven by a wage effect because persistently overweight men have approximately 30 percent smaller monthly earnings.

The results for women are in line with our previous full sample results. We find that persistently overweight identical twin women have approximately 40 percent smaller long-term earnings and work approximately 1.1 months less than women who are never overweight. This result suggests that the wage penalty of overweight on long-term earnings is driven by persistently overweight women working less. The only difference compared to

Table 5 Twin fixed effect regressions, identical twin sample by gender

	ΔLong-term earnings			ΔAverage employment months			ΔMonthly earnings		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Panel A: Women									
Overweight history categories									
ΔPersistently overweight (β_1)	−0.397**	−0.391**	−0.401**	−1.159**	−1.120**	−1.107**	−0.296	−0.297	−0.309
	(0.185)	(0.189)	(0.178)	(0.545)	(0.553)	(0.553)	(0.218)	(0.216)	(0.214)
ΔPreviously overweight (β_2)	−0.318	−0.311	−0.259	−0.841	−0.798	−0.701	−0.196	−0.197	−0.174
	(0.269)	(0.272)	(0.288)	(0.994)	(1.004)	(1.031)	(0.102)	(0.100)	(0.105)
ΔCurrently overweight (β_3)	−0.080	−0.078	−0.085	−0.792**	−0.774**	−0.783**	−0.059	−0.060	−0.064
	(0.118)	(0.117)	(0.113)	(0.371)	(0.369)	(0.363)	(0.085)	(0.085)	(0.084)
Controls									
ΔHeight	N	Y	Y	N	Y	Y	N	Y	Y
ΔEducation	N	N	Y	N	N	Y	N	N	Y
ΔHealth	N	N	Y	N	N	Y	N	N	Y
Tests (p-values)									
$\beta_1 = \beta_2 = \beta_3 = 0$	0.187	0.216	0.163	0.067	0.081	0.079	0.258	0.238	0.398
Joint test of controls	–	0.526	0.002	–	0.222	0.118	–	0.853	0.007
$\beta_1 = \beta_2 = \beta_3$	0.291	0.309	0.248	0.795	0.814	0.808	0.238	0.226	0.304
$\beta_1 = \beta_2 + \beta_3$	0.997	0.994	0.852	0.645	0.663	0.722	0.742	0.747	0.579
$\beta_1 = \beta_2$	0.761	0.760	0.605	0.731	0.729	0.671	0.496	0.497	0.370
$\beta_1 = \beta_3$	0.121	0.130	0.095	0.520	0.548	0.566	0.178	0.175	0.155
$\beta_2 = \beta_3$	0.379	0.391	0.539	0.960	0.981	0.935	0.110	0.104	0.222
Number of individuals	435	435	435	435	435	435	435	435	435
Panel B: Men									
Overweight history categories									
ΔPersistently overweight (β_1)	−0.552***	−0.550***	−0.459***	−0.537	−0.533	−0.371	−0.333***	−0.331***	−0.292***
	(0.192)	(0.192)	(0.174)	(0.455)	(0.457)	(0.443)	(0.103)	(0.105)	(0.103)

Table 5 (continued)

	ΔLong-term earnings			ΔAverage employment months			ΔMonthly earnings		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
ΔPreviously overweight (β_2)	−0.228	−0.223	−0.212	−0.566	−0.559	−0.536	−0.074	−0.071	−0.060
	(0.137)	(0.139)	(0.143)	(0.345)	(0.350)	(0.350)	(0.083)	(0.086)	(0.088)
ΔCurrently overweight (β_3)	−0.193	−0.194	−0.185	−0.352	−0.354	−0.339	−0.079	−0.080	−0.078
	(0.152)	(0.152)	(0.151)	(0.402)	(0.402)	(0.404)	(0.091)	(0.092)	(0.093)
Controls									
ΔHeight	N	Y	Y	N	Y	Y	N	Y	Y
ΔEducation	N	N	Y	N	N	Y	N	N	Y
ΔHealth	N	N	Y	N	N	Y	N	N	Y
Tests (p-values)									
$\beta_1 = \beta_2 = \beta_3 = 0$	0.039	0.042	0.072	0.410	0.432	0.492	0.014	0.017	0.041
Joint test of controls	–	0.403	0.015	–	0.692	0.319	–	0.446	0.095
$\beta_1 = \beta_2 = \beta_3$	0.062	0.065	0.171	0.843	0.854	0.885	0.021	0.024	0.059
$\beta_1 = \beta_2 + \beta_3$	0.475	0.475	0.751	0.428	0.433	0.300	0.156	0.159	0.242
$\beta_1 = \beta_2$	0.050	0.051	0.135	0.945	0.952	0.693	0.010	0.011	0.026
$\beta_1 = \beta_3$	0.031	0.034	0.089	0.635	0.645	0.934	0.032	0.036	0.074
$\beta_2 = \beta_3$	0.829	0.860	0.874	0.604	0.620	0.636	0.966	0.935	0.872
Number of individuals	267	267	267	267	267	267	267	267	267

The control variables are height in 1975 (in centimeters), schooling in 1975 (in years), and the number of diagnosed diseases in 1975. All dependent, explanatory, and control variables are within-twin differenced (denoted by Δ). Standard errors are in parentheses (heteroscedasticity robust). **p < 0.05, ***p < 0.01

our findings using the full sample (Sect. Within-twin differences: the role of family environment.) is that the association between persistently being overweight and on women's monthly earnings is no longer statistically significant.

These findings suggest that the mechanisms explaining variation in labour market outcomes could be gender-specific. For persistently overweight men, the mechanism seems to be related to something that erodes their earnings power on the labor market but not their labor market attachment throughout their lifecycle. This result could reflect an unobservable characteristic, such as a high discount rate. It can also reflect characteristics related to skill formation earlier in life, such as underdeveloped social skills or lower self-esteem.

Unlike men, currently overweight women work less than women who are never overweight, suggesting additional explanatory mechanisms. Besides factors that affect their current productivity, one potential mechanism may be discrimination against overweight women and their job applications (Rooth, 2009; Campos-Vazquez and Gonzalez, 2020). However, we note that because the period of our analysis is long, we acknowledge that stereotypes and prejudices evolve and change as the general population becomes heavier, which can also affect the ability to generalize the obtained results across decades.

Our estimated coefficient sizes are larger than those reported in the previous literature. However, there are many reasons why our and the previous results do not have to be in conflict. First, unlike many previous papers, we have been careful not to add control variables that themselves may be affected by individuals' overweight history. The reason for this is that we are interested in capturing the total effect of having a history of being overweight on lifetime earnings. This requires that we allow the regression coefficients of the overweight variables also to capture the indirect effects, such as labor supply decisions, occupational choices, and health behaviors (Han et al., 2011).

Second, previous papers have used shorter-term measures, such as the hourly rate of pay, as an outcome variable. Comparing our estimates to those is not straightforward because our long-term outcome measures have a smaller measurement error than the short-term measures. Our outcome measures also reflect potential consequences that previous measures were unable to capture, such as the wage effects of repeated unemployment spells and switching jobs often.

Our results show that controlling for both genetics and family environment is important for men but not for women. Our findings indicate, in particular, that genetic factors shape the association of being overweight and labor market outcomes for men but not for women. For women, the results are fairly robust across all specifications. For men, they are not: persistently overweight men have smaller long-term earnings and lower monthly earnings after family environment and genetics are controlled for.

To summarize, our results suggest that whether and how genetic factors are a source of omitted variable bias differ across women and men. This interpretation is consistent with the view that genetically inherited traits, such as appearance, and other personality traits that are affected by genetics and environment, such as impatience, contribute to weight gains and shape an individual's labor market outcomes. Our results thus provide a potential explanation for the lack of consensus in the previous literature regarding the earnings penalty of overweight men (Averett and Korenman, 1996; Sargent and Blanchflower, 1994; Baum and Ford, 2004; Cawley, 2004) and, hence, for why the evidence for men has been more mixed (Cawley, 2015). Our gender-specific results may help interpreting studies using genetic risk scores as instruments for obesity. For example, Böckerman et al. (2018) suggests that the power of genetic instruments and the exclusion restrictions that their validity require may be gender-specific.

Robustness

Alternative characterization of weight history: We test the robustness of our results by using a stricter definition for being persistently overweight as a robustness check and define an individual persistently overweight if the individual is overweight in every survey year: 1975, 1981, and 1990. We also modify how we define the class of previously overweight individuals to keep the estimation sample unchanged: An individual is called previously overweight if the individual was overweight in either 1975 or 1981 but not in 1990. The BMI threshold is at 25.

Because of the stricter definition, the group of persistently overweight individuals becomes smaller, and the group of previously overweight individuals becomes larger because some from the group of persistently overweight individuals belong to the group of currently overweight individuals. If having some overweight history matters for those who were overweight in 1990, we expect our results for the persistently overweight individuals to weaken and for the currently overweight individuals to become stronger. If our results for the persistently overweight individuals are driven by those who were consistently overweight over the fifteen-year period from 1975 to 1990 during which the weight measurements took place, we should expect our results for the persistently overweight individuals to become stronger when the definition of persistently overweight is changed as described.

This robustness analysis shows that the adverse effects of being overweight on long-term labor market outcomes are related to having *some* overweight history in addition to being overweight in 1990. Characterizing weight history differently changes the results somewhat but in the expected fashion: The results for the currently overweight women become statistically stronger, whereas the results for the persistently overweight women become statistically weaker in models with twin fixed effects. For men, the associations exist when genetic differences are fully controlled for. As for women, the results for the currently overweight men become statistically stronger.

Alternative BMI-thresholds: Our main analysis uses the standard BMI threshold for being overweight ($BMI \geq 25.00$) when constructing our main explanatory variables. This threshold is assumed to be the same for both sexes and for all survey years. Below, we consider an alternative way of defining the threshold to see whether the threshold specification affects our results.

Our alternative definition for BMI thresholds is the same as the definition used in the main analysis in Chen (2012). We divide BMI distribution into 3-year and gender-specific terciles and use the lower bounds of the highest terciles as the thresholds. The resulting thresholds for women are 21.55 for 1975, 22.35 for 1981, and 24.90 for 1990. For men, the thresholds are 23.41, 24.49, and 25.68 for 1975, 1981, and 1990, respectively. This means that for women the thresholds for 1975 and 1981 are much lower than the standard threshold used in our main estimations. The thresholds are also lower than the standard for men in 1975 and 1981 but not as low as that of women.

Before we describe our results it is useful to briefly discuss the mechanism behind how these decreases in the thresholds could affect the results. First, a lower threshold adds thinner individuals to the overweight history category, thus changing the composition of each overweight history category. What is particularly important is that the change in the threshold(s) somewhat compromises our definition of being persistently overweight. If our results change and become weaker, this does not automatically imply that our baseline findings were fragile. Instead, it may just mean that keeping the threshold at a higher level also during the earlier measurement years is important for measuring the adverse labor market effects of being persistently overweight.

Our results show that using a lower BMI threshold for the definition of overweight matters. When we replicate the OLS estimations of Subsection 4.1. and the fixed effects estimations of Subsections Within-twin differences: the role of family environment. and 4.3, we find that our results become less robust across the models (Table A5). For example, when the tercile threshold specification is used we no longer find a consistent negative association between persistently being overweight and long-term labor market outcomes for women. For men, using the tercile approach also yields a relatively mixed set of results and weakens the findings for the persistently overweight history category.

These results are in line with the previous literature that suggests that it may be difficult to pin down exactly the point at which weight gains start to matter for labor market outcomes (see, for example, (Gregory and Ruhm, 2011). Within the context of our study, the standard BMI threshold for being overweight seems to capture the longer-term consequences of being overweight relatively well.

Alternative outcome measures: We also re-estimate our models using alternative measures for long-term earnings and monthly earnings in which we calculate the average of the natural logarithm of annual earnings first before taking the sum. Taking the logarithm before averaging can increase the number of missing observations because the logarithm is not defined at zero. Using these alternative measures reduces the sample size but has a negligible effect on our results.

Additional control for physical condition: Previous literature has suggested that being overweight and physical condition are related, and being physically active may be related to better long-term labor market outcomes (Hyytinen and Lahtonen, 2013). To check the robustness of our results, we have also run our estimations with an additional control for individuals' physical condition, measured in 1975. This variable is defined as 1 if the respondent felt short of breath with light physical effort and is zero otherwise. We find that adding control for the physical condition to the regression models does not change our results.

Limitations: A potential concern to our estimation strategy is that a twin's weight may deviate from the co-twin's weight for reasons that are correlated with the regression error term, such as a physical injury or a health shock could cause one twin to gain weight. A health shock can bias our results if it happened before the weight measurements took place and if it continues to have a labor market impact much later over the 20-year period.

We argue that such a bias is unlikely in our study because BMI measures are measured much earlier than the labor market outcomes. Predetermined BMI measures reduces the concern that the (contemporary) error term in the labor market outcome regressions is correlated with past weight measures. Importantly, our results are also robust to controlling for within-pair differences in initial health and physical conditions, both measured in 1975.

BMI is influenced by genetic factors (Yang et al., 2007). The rest of the variation in weight within a twin pair may be related to behavioral factors, such as nutrition, physical activity and smoking. Also, the fact that some of the BMI measures were taken place while some of the individuals were already at their prime working age is obviously a limitation. We acknowledge these econometric concerns and note here explicitly that if valid, our findings cannot be given a causal interpretation.[9]

[9] The descriptive nature of our study is also in line with the interpretation of BMI being a proxy of different factors that can be associated with long-term labor market outcomes.

One possible additional concern is whether the harmful effects of high BMI change over time. Unfortunately, our data do not allow evaluate this in a reliable fashion because our sample size is not large enough if the sample is split to too small subsamples. Moreover, the way how being overweight is associated with labor market outcomes may change over time for a number of reasons, such as changing weight penalty, the individuals in our sample getting older, and changing economic environments. Our graphical analysis (see Figures A4 - A10 in our Appendix) of level outcomes provides no evidence of consistent patterns. Because of this and because of the mentioned potential confounding factors, we refrain from taking a stance on whether the penalty changes over time, and leave it for further studies to explore.

Conclusions

We have studied how factors related to family environment and genes shape how the weight history of individuals, characterized by classes of being persistently, previously, and currently overweight, is associated with their long-term earnings, average employment months, and monthly earnings using data on Finnish twins. We find that (i) being persistently overweight in early adulthood is associated with lower subsequent long-term earnings for both women and men and (ii) that the mechanism is gender-specific.

Understanding these linkages is relevant for policy design. When being overweight reduces people's lifetime earnings, as compared to it reducing current earnings, the implications for lifetime consumption possibilities and health policy are profound. Interventions and policies that aim at reducing the number of people currently overweight or their contemporary discrimination are likely to have heterogeneous effects because the earnings penalty seems to be associated with persistently being overweight.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s10754-021-09315-4.

Funding Dr Laine received financial support from the Yrjö Jahnsson Foundation (research Grant No. 6578) and the KAUTE Foundation.

Availability of data and material Each dataset is as such proprietary, but researchers can independently obtain access to the datasets following the procedures provided in the supplemental details.

Code availability Stata-programs that allow replication of all of the results in the paper can be obtained from the authors.

Declarations

Conflict of interest The authors declare no conflicts of interest or competing interests.

Ethical approval The two data sets were linked by using personal coded identifiers by Statistics Finland. Record linkages of the cohort study data conform to the Data Protection Act and were originally approved by the ethical committee of the Department of Public Health, University of Helsinki (Kaprio et al., 1979). Statistics Finland has accepted the record linkages used for the matched data used in this paper. All the data work of this paper was carried out at Statistics Finland, following its terms and conditions of confidentiality.

References

Aaronson, D. (1998). Using sibling data to estimate the impact of neighborhoods on children's educational outcomes. *The Journal of Human Resources, 33*(4), 915–946.

Abramitzky, R., Boustan, L. P., & Eriksson, K. (2012). Europe's tired, poor, huddled masses: Self-selection and economic outcomes in the age of mass migration. *American Economic Review, 102*(5), 1832–1856.

Altonji, J. G., & Dunn, T. A. (1996). Using siblings to estimate the effect of school quality on wages. *The Review of Economics and Statistics, 78*(4), 665–671.

Angrist, J. D., & Pischke, J.-S. (2009). *Mostly harmless econometrics: An empiricist's companion*. Princeton: Princeton University Press.

Averett, S. (2011). Labor market consequences: Employment, wages, disability and absenteeism. In J. Cawley (Ed.), *The Oxford Handbook of the Social Science of Obesity, Volume 1 of Oxford Handbooks* (pp. 531–552). Oxford: Oxford University Press.

Averett, S., & Korenman, S. (1996). The economic reality of beauty myth. *Journal of Human Resources, 31*(2), 304–330.

Baum, C., & Ford, W. (2004). The wage effects of obesity: A longitudinal study. *Health Economics, 13*, 885–899.

Behrman, J., & Rosenzweig, M. (2001). *The returns to increasing body weight*. PIER Working Paper 01(052).

Behrman, J., & Rosenzweig, M. (2002). Does increasing women's schooling raise the schooling of the next generation? *American Economic Review, 92*(1), 323–334.

Behrman, J., Rosenzweig, M., & Taubman, P. (1996). College choice and wages: Estimates using data on female twins. *The Review of Economics and Statistics, 78*(4), 672–685.

Benjamin, D., Cesarini, D., Chabris, C. F., Glaeser, E. L., Laibson, D., Guonason, V., Harris, T. B., Launer, L. J., Purcell, S., Smith, A. V., Johannesson, M., Magnusson, P. K., Beauchamp, J. P., Christakis, N. A., Atwood, C. S., Hebert, B., Freese, J., Hauser, R. M., Hauser, T. S., & Lichtenstein, P. (2012). The promises and pitfalls of genoeconomics. *Annual Review of Economics, 4*(1), 627–662.

Bhattacharya, J., & Bundorf, K. (2009). The incidence of the healthcare costs of obesity. *Journal of Health Economics, 28*(3), 649–658.

Böckerman, P., Cawley, J., Viinikainen, J., Lehtimäki, T., Rovio, S., Seppälä, I., et al. (2018). The effect of weight on labor market outcomes: An application of genetic instrumental variables. *Health Economics.*,.

Brown, C., & Routon, P. W. (2018). On the distributional and evolutionary nature of the obesity wage penalty. *Economics & Human Biology, 28*, 160–172.

Brunello, G., & Beatrice, D. (2007). Does body weight affect wages?: Evidence from Europe. *Economics & Human Biology, 5*(1), 1–19.

Burkhauser, R. V., & Cawley, J. (2008). Beyond BMI: The value of more accurate measures of fatness and obesity in social science research. *Journal of Health Economics, 27*(2), 519–529.

Caliendo, M., & Gehrsitz, M. (2016). Obesity and the labor market: A fresh look at the weight penalty. *Economics & Human Biology, 23*, 209–225.

Caliendo, M., & Lee, W.-S. (2013). Fat chance obesity and the transition from unemployment to employment. *Economics & Human Biology, 11*(2), 121–133.

Campos-Vazquez, R. M., & Gonzalez, E. (2020). Obesity and hiring discrimination. *Economics & Human Biology, 37*, 100850.

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

Cawley, J. (2015). An economy of scales: A selective review of obesity's economic causes, consequences, and solutions. *Journal of Health Economics, 43*, 244–268.

Cawley, J., & Meyerhoefer, C. (2012). The medical care costs of obesity: An instrumental variables approach. *Journal of Health Economics, 31*(1), 219–230.

Chen, A. J. (2012). When does weight matter most? *Journal of Health Economics, 31*(1), 285–295.

Datar, A. (2017). The more the heavier? Family size and childhood obesity in the U.S. *Social Science & Medicine, 180*, 143–151.

Falkner, N. H., Neumark-Sztainer, D., Story, M., Jeffery, R. W., Beuhring, T., & Resnick, M. D. (2001). Social, educational, and psychological correlates of weight status in adolescents. *Obesity Research, 9*(1), 32–42.

Farooqi, I. S., & O'Rahilly, S. (2007). Genetic factors in human obesity. *Obesity Reviews, 8*, 37–40.

Florin, T. A., Shults, J., & Stettler, N. (2011). Perception of overweight is associated with poor academic performance in us adolescents. *Journal of School Health, 81*(11), 663–670.

Gregory, C. A., & Ruhm, C. J. (2011). *Where does the wage penalty bite?*, pp. 315 – 347. University of Chicago Press.
Griliches, Z. (1979). Sibling models and data in economics: Beginnings of a survey. *Journal of Political Economy, 87*(5), S37–S64.
Gwozdz, W., Sousa-Poza, A., Reisch, L. A., Bammann, K., Eiben, G., Kourides, Y., Kovacs, E., Lauria, F., Konstabel, K., Santaliestra-Pasias, A. M., Vyncke, K., & Pigeot, I. (2015). Peer effects on obesity in a sample of European children. IZA Discussion Papers 9051, Institute for the Study of Labor (IZA).
Han, E., Norton, E. C., & Powell, L. M. (2011). Direct and indirect effects of body weight on adult wages. *Economics & Human Biology, 9*(4), 381–392.
Han, E., Norton, E. C., & Stearns, S. C. (2009). Weight and wages: Fat versus lean paychecks. *Health Economics, 18*(5), 535–548.
Heckman, J. J. (2006). Skill formation and the economics of investing in disadvantaged children. *Science, 312*(5782), 1900–1902.
Hyytinen, A., Ilmakunnas, P., & Toivanen, O. (2013). The return-to-entrepreneurship puzzle. *Labour Economics, 20*, 57–67.
Hyytinen, A., & Lahtonen, J. (2013). The effect of physical activity on long-term income. *Social Science & Medicine, 96*, 129–137.
Jenq, C., Pan, J., & Theseira, W. (2015). Beauty, weight, and skin color in charitable giving. *Journal of Economic Behavior & Organization, 119*, 234–253.
Kaprio, J., Artimo, M, Sarna, S., & Rantasalo, I. (1979). The Finnish twin registry: Baseline characteristics. section i: Materials methods, representativeness and results for variables special to twin studies. *Kansanterveystieteen julkaisuja, M*.
Kristjansson, L. A., Sigfusdottir, I. D., & Allegrante, J. P. (2010). Health behavior and academic achievement among adolescents: The relative contribution of dietary habits, physical activity, body mass index, and self-esteem. *Health Education & Behavior, 37*(1), 51–64.
Mobius, M. M., & Rosenblat, T. S. (2006). Why beauty matters. *The American Economic Review, 96*(1), 222–235.
Morris, S. (2007). The impact of obesity on employment. *Labour Economics, 14*(3), 413–433.
Pinkston, J. C. (2017). The dynamic effects of obesity on the wages of young workers. *Economics & Human Biology, 27*, 154–166.
Price, J., & Swigert, J. (2012). Within-family variation in obesity. *Economics & Human Biology, 10*(4), 333–339.
Puhl, R., & Brownell, K. D. (2001). Bias, discrimination, and obesity. *Obesity Research, 9*(12), 788–805.
Rooth, D.-O. (2009). Obesity, attractiveness, and differential treatment in hiring: A field experiment. *Journal of Human Resources, 44*(3), 710–735.
Sabia, J. J., & Rees, D. I. (2012). Body weight and wages: Evidence from Add Health. *Economics & Human Biology, 10*(1), 14–19.
Sacerdote, B. (2007). How large are the effects from changes in family environment? A study of Korean American adoptees. *The Quarterly Journal of Economics, 122*(1), 119–157.
Sargent, J., & Blanchflower, D. (1994). Obesity and stature in adolescence and earnings in young adulthood: Analysis of a British birth cohort. *Archives of Pediatrics & Adolescent Medicine, 148*(7), 681–687.
Segal, N. L., & Allison, D. B. (2002). Twins and virtual twins: Bases of relative body weight revisited. *International Journal of Obesity and Related Disorders, 26*(4), 437–441.
Silventoinen, K., Jelenkovic, A., Sund, R., Honda, C., Aaltonen, S., Yokoyama, Y., Tarnoki, A. D., Tarnoki, D. L., Ning, F., Ji, F., et al. (2015). The codatwins project: The cohort description of collaborative project of development of anthropometrical measures in twins to study macro-environmental variation in genetic and environmental effects on anthropometric traits. *Twin Research and Human Genetics, 18*(4), 348–360.
WHO (2018). Obesity and overweight. https://www.who.int/en/news-room/fact-sheets/detail/obesity-and-overweight. Accessed February 12, 2020.
Yang, W., Kelly, T., & He, J. (2007). Genetic epidemiology of obesity. *Epidemiologic Reviews, 29*(1), 49–61.
Zwijnenburg, P. J., Meijers-Heijboer, H., & Boomsma, D. I. (2010). Identical but not the same: The value of discordant monozygotic twins in genetic research. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics, 153B*(6)