

Hübler, Olaf

Working Paper

Health and weight

SOEPPapers on Multidisciplinary Panel Data Research, No. 914

Provided in Cooperation with:

German Institute for Economic Research (DIW Berlin)

Suggested Citation: Hübler, Olaf (2017) : Health and weight, SOEPPapers on Multidisciplinary Panel Data Research, No. 914, Deutsches Institut für Wirtschaftsforschung (DIW), Berlin

This Version is available at:

<https://hdl.handle.net/10419/162588>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

SOEPpapers
on Multidisciplinary Panel Data Research

SOEP – The German Socio-Economic Panel study at DIW Berlin

914-2017

Health and weight

Olaf Hübler

SOEPpapers on Multidisciplinary Panel Data Research at DIW Berlin

This series presents research findings based either directly on data from the German Socio-Economic Panel study (SOEP) or using SOEP data as part of an internationally comparable data set (e.g. CNEF, ECHP, LIS, LWS, CHER/PACO). SOEP is a truly multidisciplinary household panel study covering a wide range of social and behavioral sciences: economics, sociology, psychology, survey methodology, econometrics and applied statistics, educational science, political science, public health, behavioral genetics, demography, geography, and sport science.

The decision to publish a submission in SOEPpapers is made by a board of editors chosen by the DIW Berlin to represent the wide range of disciplines covered by SOEP. There is no external referee process and papers are either accepted or rejected without revision. Papers appear in this series as works in progress and may also appear elsewhere. They often represent preliminary studies and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be requested from the author directly.

Any opinions expressed in this series are those of the author(s) and not those of DIW Berlin. Research disseminated by DIW Berlin may include views on public policy issues, but the institute itself takes no institutional policy positions.

The SOEPpapers are available at
<http://www.diw.de/soeppapers>

Editors:

Jan **Goebel** (Spatial Economics)
Martin **Kroh** (Political Science, Survey Methodology)
Carsten **Schröder** (Public Economics)
Jürgen **Schupp** (Sociology)

Conchita **D'Ambrosio** (Public Economics, DIW Research Fellow)
Denis **Gerstorff** (Psychology, DIW Research Director)
Elke **Holst** (Gender Studies, DIW Research Director)
Frauke **Kreuter** (Survey Methodology, DIW Research Fellow)
Frieder R. **Lang** (Psychology, DIW Research Fellow)
Jörg-Peter **Schräpler** (Survey Methodology, DIW Research Fellow)
Thomas **Siedler** (Empirical Economics, DIW Research Fellow)
C. Katharina **Spieß** (Education and Family Economics)
Gert G. **Wagner** (Social Sciences)

ISSN: 1864-6689 (online)

German Socio-Economic Panel (SOEP)
DIW Berlin
Mohrenstrasse 58
10117 Berlin, Germany

Contact: soeppapers@diw.de



Health and weight

Olaf Hübler

Gottfried Wilhelm Leibniz University Hannover

May 11, 2017

Abstract: Many studies have shown that obesity is a serious health problem for our society. Empirical analyses often neglect a number of methodological issues and relevant influences on health. This paper investigates empirically whether neglecting these items leads to systematically different estimates. Based on data from the German Socio-Economic Panel, this study derives the following results. (1) Many combinations of weight and height lead to the same health status. (2) The relationship between health and body mass index is nonlinear. (3) Underweight strengthens individual health and severe obesity has a clear negative impact on health status. Underweight women are more affected than men but obese men are hit harder than women. (4) The hypothesis has to be rejected that weight has an exogenous influence on health. (5) A worse health status is linked with weight fluctuations and deviations between desired and actual working hours. (6) A healthy diet and long but not too long sleeping contribute to a good health status. Moreover, a good parental education and a high parental social status act favorably on health as does personal high income. (7) Four of the big five components of personality, namely openness, extraversion, conscientiousness and agreeableness, contribute to resilience against health problems.

Correspondence to: Olaf Hübler, Leibniz Universität Hannover, Wirtschaftswissenschaftliche Fakultät,

Empirische Wirtschaftsforschung, Königsworther Platz 1, 30167 Hannover

E-mail: huebler@ewifo.uni-hannover.de

1 Introduction

Overweight, especially obesity, but also underweight are globally discussed issues. The major studies on this topic come from the medical literature. However, economists also follow similar issues with a differentiated focus. In particular, the relationship between obesity and labor market outcome is analyzed (Johansson et al. 2009, Sabia/Rees 2012). Ensuing health problems and their economic consequences have been widely discussed, with various studies mentioning serious health implications, including higher mortality risk, and health problems such as cardiovascular disease, heart attack, stroke, diabetes, cancer, arthritis, gallstones, asthma, cataracts, infertility, snoring and sleep apnea (see, for example, Eliassen et al. (2006), Flegal et al. (2013) and Willett et al. (1995)). Bray (2004) argues that obesity is an epidemic disease that threatens to inundate health and that the effects of obesity come from two factors: the increased mass of adipose tissue and the increased secretion of pathogenetic products from enlarged fat cells. This concept of the pathogenesis of obesity as a disease allows an easy division of the disadvantages of obesity into those produced by the mass of fat and those produced by the metabolic effects of fat cells. The former category includes the social disabilities coming from the stigma associated with obesity, sleep apnea, which comes in part from increased parapharyngeal fat deposits, and osteoarthritis resulting from the wear and tear on joints from carrying an increased mass of fat. The second category includes the metabolic factors associated with distant effects of products released from enlarged fat cells.

In many developed countries the average weight of people has increased substantially in recent years, which has led to health problems. However, a declining mortality risk over time due to better control of risk factors for heart disease (Gregg et al. 2005) can be observed. Nevertheless, this development has not led to reduced disability risks (Alley and Chang 2007). Oswald and Powdthavee (2007) following Offer (2006) argue that economic prosperity undermines well-being. Happiness and mental health are worse among heavier people in Britain and Germany. For a given level of body mass index (BMI) they find that people who are educated or who have a high income are more likely to view themselves as overweight or even obese.

Soltoft et al. (2009) analyze the relationship between weight, measured by body mass index, and health-related quality of life using data from the Health Survey for England 2003. They find a nonlinear link. The best quality of life was reached at a BMI of 26.0 in men and 24.5 in women. BMI is negatively associated with quality of life for both underweight and obese individuals. At higher BMI values, men reported higher-quality perceptions than women. At lower BMI values, the reverse result is observed: Quality of life is lower in men than women.

Maclean et al. (2014) discuss the influence of personality disorders on body weight and find that women with these characteristics have a significantly higher BMI and are more likely to be obese than otherwise similar women. Only few statistically significant or economically meaningful effects are detected for men. Findings from unconditional quantile regressions demonstrate a positive gradient between personality disorders and BMI in that the effects are greater for higher-BMI respondents.

The aim of another study (Psouni et al. 2016) is to characterize the patterns of psychological and behavioral characteristics, in relation to body mass index. Cluster analysis identifies two distinct profiles. The first segment relates to more positive results in psychological variables. Interestingly, individuals in

the healthy segment correspond to a normal BMI. The second segment relates to more unhealthy behaviors including lower levels of exercise, unhealthy eating and negative psychological variables. As expected, individuals in the second segment had a mean overweight BMI. Furthermore, profiles from the unhealthy segment displayed higher levels of psychological distress and lower self-control.

Lissner et al. (1991) demonstrate that fluctuations in body weight have negative health consequences, independently of obesity and the trend of body weight over time. Preston et al. (2012) confirm this result. They find that higher volatility and increasing trends have large negative effects on mortality for obese people. Bhattacharya and Sood (2011) argue that body mass index should not be regarded as a medical diagnosis: Many classified as “obese” are physically fit as several studies make clear.

Most of these studies suffer from methodological shortcomings. They present their results with the help of descriptive statistics and draw far-reaching conclusions in relation to preventive measures. Normally, authors do not pay enough attention to weight differences between men and women and country-specific peculiarities when the effects on health are investigated. Underweight, normal weight, overweight and obesity should not be defined by the same BMI boundaries for men and women. Furthermore, many studies do not consider any or only consider a limited number of control variables and they also focus on linear relationships. They do not discuss interdependencies between health and weight and they neglect personal characteristics and behavior as determinants of the relationship between weight and health. Medical studies on the one hand, and socioeconomic investigations on the other, should not ignore the results from the other discipline. Specific individual behavior and characteristics, including those developed during youth, are important. All these aspects contribute to the outcome that current connections are insignificant and not stable, that some people or groups exhibit different results and that individual behavior can strengthen resilience against diseases or increase vulnerability. So far, it appears that a strong awareness campaign and public health policy against extreme body weight can avoid a lot of health problems. However, we find that we need specific measures for specific groups.

This empirical paper extends the literature in the following ways: (i) it consistently separates effects by gender; (ii) it takes into account heterogeneity and interdependencies; (iii) it extends the use of weight groups beyond the traditional international BMI classification and also uses quantiles of the German weight distribution; (iv) it analyzes the importance of personality characteristics and behavior, especially the link between the “Big5” factors, for resilience against various health problems.

One of our major objectives is to show that the influence of underweight and obesity on health status differs between men and women in that the effects depend on many other health determinants that also correlate with weight. The importance of these issues is revealed in several steps:

First, simple specifications are presented that are comparable with other studies.

Second, further health determinants and interdependencies between health and weight are taken into account. This enables the detection of whether weight effects change fundamentally. As health determinants, in particular mother’s education, father’s social status, schooling, gross wage and the difference between desired and actual working hours are added. Well-educated individuals are better informed than others about behavior that leads to good health, or this was learned from their parents (Case and Paxson 2002, Lindahl et al. 2016). More working hours than desired mean stress and this is not

good for health. And conversely, fewer working hours than desired mean dissatisfaction combined with psychological problems.

Third, we incorporate variables that are not collected every year in our data set but we guess that they increase resilience or vulnerability in relation to health. We distinguish whether the interviewed people had only slight weight variations in the past, whether they are nonsmokers, whether they eat healthily and whether they sleep longer or shorter than others. Furthermore, we consider characteristics that are inherent or developed during childhood or adolescence, namely whether they are undenominational as characterized by the parents and self-confident, and whether they actively played sports or music during their childhood. We expect that all these are positive influences on health and affect BMI (Alvarez and Ayas 2004, Ottenbach et al. 2016, Preston et al. 2012, Wehby et al. 2012).

Fourth, we are especially interested in the importance of personality traits. With one exception (Wehner et al. 2016) this topic was not analyzed in the health-weight context, although in other substantial economic relations personality was investigated (Becker et al. 2012, Bode et al. 2016). It also seems important for health and weight and may explain why some obese people are healthy while some normal-weight people have enormous problems. When these personality traits, together with all other characteristics, are jointly incorporated as control variables of a health function, strong multicollinearity and consequently insignificant influences have to be expected. As an alternative, a principal component analysis is conducted to bundle the influences and induce independent effects.

Fifth, heterogeneity is exemplarily investigated through the effects of personality traits on the probability of suffering from different diseases.

Sixth, robustness investigations are focused on extended interdependent estimates where resilience factors are incorporated.

2 Data and methods

2.1 Data

The data set used in this study, the Socio-Economic Panel (2015), is a representative annual household survey started in 1984 covering Western Germany at the time that was extended to Eastern Germany in 1990 (Wagner et al. 2007). Currently, more than 12,000 households and more than 23,000 individuals are interviewed per year. Our sample covers the years 2004, 2006, 2008, 2010, 2012 and partially 2013. Apart from in 2013, individuals do not report on specific diseases. Respondents are included in the respected sample if information about health and other necessary characteristics such as body mass index, wage, schooling, gender, age, height, mother's schooling and father's prestige score is available or can be calculated. Only unweighted data are used. For mothers and fathers, the SOEP group has generated the prestige score (Treiman score – Standard International Occupational Prestige Scale) based on three steps (Ganzeboom and Treiman 1996). This means an operational procedure is provided for coding internationally comparable measures of occupational status from the International Standard Classification of Occupation (ISCO) of the International Labor Office. First, for unit groups for which there is a one-to-one correspondence of two years of the ISCO, the scores are simply assigned to the more current categories. Second, for the remaining ISCO categories the occupational titles reported by Treiman (1977, Appendix A) are matched to these ISCO unit groups in the same way that Treiman had initially matched them to former ISCO unit groups. The scores for all occupation titles matching each

current ISCO unit group are then averaged to obtain a score for the unit group (usually the simple average is taken, but where occupational titles referred to rarely held jobs, weighted averages are taken - again in a manner analogous to the procedures used by Treiman in constructing unit group scores for the older ISCO). Third, where no occupational titles match a current ISCO unit group, scores are borrowed from similar unit groups for which scores are available.

Table A1 presents descriptive statistics for the utilized variables. The number of observations differs for different variables due to varying determinants of the estimates and missing values of the applied variables.

Seven health indicators are available for 2012: h1 – current health (=1 if very good, ... , =5 if bad); h2 – health satisfaction (=0 if completely dissatisfied, ... , =10 if completely satisfied); h3 – problems with climbing several flights of stairs on foot due to own health status (=1 if strong , =2 if some, =3 if not at all); h4 – problems with lifting something heavy or doing something requiring physical mobility (=1 if greatly, =2 if somewhat, =3 if not at all); h5 – suffering from any medical conditions or illnesses for at least one year or chronically (=0 if no, =1 if yes); h6 – total number of nights spent in hospital last year; h7 – number of days unable to work last year due to illness. A further health indicator, h8, is available based on 2013 data: Respondents were asked whether a doctor had found any chronic illness in the past (=0 if one or more diseases had been found; =1 if no disease had been detected). In Table A1 we show descriptive statistics, while Table 1 shows the simple Pearson correlation coefficients for men. The results for women are similar but usually a little bit larger, and are not presented in the paper.

Weight is mainly measured by body mass index ($BMI = \text{weight in kg} / (\text{height in cm} / 100)^2$). Subgroups are constructed following Costa (2015): “underweight” ($BMI < 18.5$ kg), “normal weight” ($18.5 \leq BMI < 25$), “overweight” ($25 \leq BMI < 30$), “obesity 1” ($30 \leq BMI < 35$) and “obesity 2” ($BMI \geq 35$). The WHO (2000) writes (p.9): “These BMI values are age-independent and the same for both sexes.” Alternatively, we split into light-, medium- and heavy-weight people, based on quantiles of German gender-specific BMI distributions.

Big five personality traits (“Big5”) are calculated from answers to question 151 in SOEP 2013.

Interviewees were questioned in relation to a total of 16 topic areas. On a scale from 1 to 7 (=1 if does not describe at all, ... , =7 if describes perfectly) they answered by giving their subjective assessment of individual personality. The same items were collected in 2005 and 2009 and the results are very similar. All “big five” factors – openness, extraversion, conscientiousness, agreeableness and neuroticism – are determined as the sum of the scores generated from answers to three questions. This means the minimum score for each factor is equal to three and the maximum score is equal to 21. Summary statistics can be seen in Table A1. Openness characterizes people who are original, have new ideas, who have artistic and aesthetic experiences and are imaginative. Extraversion describes people who are communicative, talkative, outgoing and sociable and who are not reserved. Typical traits for people with conscientiousness are that they are thorough workers, that they are not lazy and that they are effective and efficient in completing tasks. The fourth characteristic, agreeableness, expresses that people are not rude to others, that they can forgive and that they are considerate and kind to others. Individuals who are easily worried, who are nervous in many situations and who are not easily relaxed and cannot deal with stress strongly exhibit the fifth property, neuroticism.

All these characteristics of health, weight and personality traits are self-reported. Of course, this is a disadvantage. All such information may contain systematic measurement errors that may cause bias in the estimates. Interpersonal comparisons are not straightforward (Thomas and Frankenberg 2002). As long as no objective measurement is available we have to be careful in the interpretation of the results.

2.2 Empirical strategy

In the literature we observe different health and weight indicators. It is unclear whether they measure the same things and whether there are significant differences between men and women. First, we test whether the different available health indicators correlate significantly using simple correlation coefficients (see Table 1), or in other words whether we can restrict the following econometric analysis to one health indicator.

We also note that weight depends on height and both characteristics can induce serious health problems: Previous studies have not only observed a relation between weight and a range of disease outcomes but also between height and health (Batty et al. 2009). Therefore, it makes sense to combine both variables and/or to incorporate height as a determinant of weight. Most often, the body mass index (BMI) is used. In one sense, this measure introduces a first step of nonlinearity. In a second step, we test whether the incorporation of BMI² as well as BMI improves health estimation. A more differentiated analysis can be based on weight groups. This approach enables identification of the way the health status varies between weight groups and between men and women. Overweight individuals are chosen as the base group (Table 3). Besides OLS we use ordered probit estimates as the regressand health is measured by an ordinal variable.

A methodological problem follows if health is not only influenced by weight but reverse causality is also possible. Depending on a specific disease, bad health can lead to under- or overweight. On the one hand, most people with cancer experience weight and muscle losses during their treatment. On the other hand, thyroid insufficiency (hypothyroidism) or the so-called Cushing's syndrome may be responsible for a weight gain. This issue of mutual dependence can be resolved by simultaneous modeling. We address this issue in Tables 4 and 9 and in Appendix 2 where we use instrumental variables estimates. Good instruments strongly correlate with weight but not with the error term of the health function. As in our case there are no entirely convincing external instruments, we follow Lewbel (2012) for endogenous treatment effects. The decision to use IV estimates is based on theoretical explanations – see Section 3.2 - and the Durbin-Wu-Hausman test of endogeneity (Davidson and MacKinnon 2004) is applied. If the test statistic exceeds the critical value of $\chi^2(1)$, the null hypothesis of exogeneity has to be rejected. The null hypothesis of weak instruments is rejected if the Cragg-Donald (1997) F-statistic is larger than Stock-Yogo's critical value (Stock and Yogo 2005) if we are willing to tolerate distortion for a 5% Wald test based on the 2SLS estimator, so that the true size can be at most 10%.

Lewbel's technique enables the identification of structural parameters in fully simultaneous linear models

$$Y_1 = X'\beta_1 + Y_2\gamma_1 + e_1$$

$$Y_2 = X'\beta_2 + Y_1\gamma_2 + e_2$$

under the assumptions that x and e are uncorrelated, that the error terms e are heteroskedastic and that the covariance between z and the product $e_1 e_2$ is zero. In our case, Y_1 is the health variable and Y_2 is the BMI variable. The vector z contains observed variables, can be discrete or continuous and can be a subset of x . In the latter case, no information outside the model specified above is required. If the covariance assumption is violated, then the parameters are still identified, when it is assumed that the correlation between z and $e_1 e_2$ is smaller than the correlation between z and e_2^2 . Identification comes from a heteroskedastic covariance restriction and is achieved by having regressors that are uncorrelated with the product of heteroskedastic errors. In the simplest version, instruments W can be generated by the product of the residuals from the reduced form (\hat{e}_2) and the mean centered values ($X - \text{mean}(X)$) of an element of vector z as a subset of x :

$$W = (X - \text{mean}(X))' \hat{e}_2.$$

In one sense, this approach is a generalization of Altonji and Shakotko (1987), where time-demeaned centered variables are used as instruments. The advantage of Lewbel's method is that the weighting with \hat{e}_2 reduces the risk of a correlation between instruments and the error term of the above Y_1 equation. The parameters β_1 and γ_1 are identified by the ordinary linear two-stage least squares estimation of Y_1 on x and Y_2 using x and $(Z - E(Z))' \hat{e}_2$ as instruments. The assumption that Z is uncorrelated with $e_1 e_2$ means that $(Z - E(Z))' \hat{e}_2$ is a valid instrument for Y_2 in the main equation since it is uncorrelated with e_1 , with the strength of the instrument (its correlation with Y_2 after controlling for the other instruments x) being proportional to the covariance of $(Z - E(Z))e_2$ with e_2 , which corresponds to the degree of heteroskedasticity of e_2 with respect to Z . The Breusch-Pagan (1979) approach may be applied to test this. The greater the degree of scale heteroskedasticity in the error process, the higher will be the correlation of the generated instruments with the included endogenous variable, which is the regressand in the first-stage regression. In our case, this is the BMI equation. Besides the generated instruments W , external instruments can be considered. We have experimented with this case and presented some results in Appendix 2.

The next step tries to detect reasons that are responsible for some underweight and obese people not having health problems in contrast to the illustration in Figure 1 in Section 3.1. In other words, we want to find explanations why some people outside the weight norm are more resilient to bad health than others. Conversely, individual behavior can also cause a tendency for poor health among people classified as normal weight. In Table 5 we show results that address this issue. Separate estimates are derived based on the empirical weight distribution in Germany. We construct three groups: i) people whose weight is within the lower quartile, referred to as light-weight, ii) those within the interquartile range (medium-weight), and iii) finally those within the upper quartile of the weight distribution (heavy-weight). This classification is based on the following presumption: The concepts of under-, normal- and overweight as well as obesity are not completely comparable across countries and longer time periods. Hence, we use the current German-specific alternative.

Since many factors are strongly correlated, collinearity can become a relevant concern. Principal component analyses can bundle the influences with the aim of deriving uncorrelated components. The respective results are shown in Tables 6 and 7. We place particular emphasis on the importance of the "Big5" factors, a concept developed by the psychology of personality (see, for example, McCrae and Costa 1997): openness, conscientiousness, extraversion, agreeableness, and neuroticism. These

dimensions are commonly held to be a complete description of personality and the acronym OCEAN is often used.

As health can also be described by diseases (see Section 2.1) it is of interest to experience whether personality effects on different illnesses are similar or whether heterogeneous effects are typical. The answer to this question is presented in a specific table (Table 8) where the estimated effects are split into different diseases. Furthermore, we present estimates that take into account a correction of a sample selection bias (Table 9, columns 1 and 2). The inverse of the Mills ratio $\lambda = \phi(\cdot) / \Phi(\cdot)$ is incorporated, where $\phi(\cdot)$ is the normal density function, $\Phi(\cdot)$ is its cumulative normal distribution function and the argument of the functions is the linear combination of the BMI determinants. Finally, the Lewbel approach is repeated for the extended specification (Table 9, columns 3 and 4) where influences are incorporated that contribute to resilience against health problems or that are detrimental to health status.

3 Results and discussion

3.1 Descriptive results and graphs

(i) Health indicators

Table 1 demonstrates that the health indicators for men are all significantly correlated. For women we find two insignificant exceptions, namely $\text{corr}(h4, h7)$ and $\text{corr}(h7, h8)$ – not presented in the tables. In most cases the correlation coefficients are larger for women – see Hübler (2017a). No other systematic gender differences can be detected. The most powerful relations are revealed between current health status (h1), health satisfaction (h2), problems with climbing several flights of stairs (h3) and problems with lifting something heavy (h4). We restrict the subsequent analyses to h1.

(ii) Distribution of body mass index for healthy and unhealthy people

Table 2 (Panel A) shows that in the normal weight class (unsurprisingly) people with good health are overrepresented. Diseased people are overrepresented in the obese classes. Hence we find that high BMI correlates with bad health. This result is confirmed in Panel B, where the micro census is used as an alternative and larger data source. The comparison with the United States in Panel C demonstrates that country-specific weight distributions exist. The U.S. population is more overweight than the German population.

(iii) Simple relationship between bad health, weight and height

Figure 1 demonstrates the relationship between bad health and weight, namely between current health status (h1) and BMI, based on a fractional polynomial (Royston/Altman 1994). Health dispersion is larger within the group of underweight and obese people than within the group of normal-weight people. This finding supports the assumption that, especially in the extreme ranges of the BMI, distribution-specific individual behavior exists that has an impact on current health status. Factors of resilience and vulnerability may have an impact so that not all underweight and not all obese people have health problems.

Among underweight men, health status varies widely, while women with a low BMI usually report that they feel fine. The obese segment is characterized by the fact that the same weight (expressed in BMI

terms) is linked with better health for women than for men. The shape of the curve in Figure 1 illustrates a clear nonlinear link between health and weight.

The usual consideration of the link between bad health and BMI assumes a specific nonlinear relationship between weight and height, namely weight in kg/(height in cm/100)². A priori, it is not clear whether this describes the correct interaction effects on health. Linear modeling

$$\text{bad health} = \beta_0 + \beta_1 \text{weight} + \beta_2 \text{height} + \varepsilon$$

excludes joint effects while a simple interaction model

$$\text{bad health} = \beta_0 + \beta_1 \text{weight} + \beta_2 \text{height} + \beta_3 \text{weight} \cdot \text{height} + \varepsilon$$

only takes into account one specific interaction. One way to capture a more general nonlinear link is to use generalized additive models (GAMs; see Hastie and Tibshirani 1997). This means using the following relation:

$$(\text{bad health} | \text{weight, height}) = s(\text{weight, height}) + \varepsilon,$$

where s stands for a smooth function, i.e. a continuous but unspecified function of weight and height that is estimated from the data. We use a P-spline smoother for performing our empirical analysis (Marx and Eilers 1998). In this paper, we estimate the smoothing parameter using a restricted maximum likelihood (REML) approach (Ruppert et al. 2003). For these models the `mgcv` library (Wood 2006) in R is applied. All other estimates are based on STATA.

A graphical representation of the GAM estimates for women can be found in Figure 2. The tendency is not new: The shorter and heavier a person, the worse the status of health, i.e. the larger the numerical values of h_1 . However, the iso-lines for the same health status demonstrate a wide range of weight and height combinations where we cannot discriminate the status of health. The heavier a person is, the less relevant the height component of health becomes for women. This outcome is not so clearly observed for men and is not presented in a graph – see Hübler (2017a). Here, the data show a stronger interaction between weight and height, especially for heavy men. Within a range of from 60 kg to 90 kg on the one hand and 180 to 200 cm on the other, the iso-lines look similar for both sexes. The shape within the light areas is uncertain because only a few observations determine the course of the health iso-lines. In contrast to this outcome we can be more confident in the dark areas with many observations.

3.2 Nonlinear and interdependent regressions of health

Regressions of health on weight are usually restricted to simple linear models with a limited number of control variables at best, e.g. age, sex, race, education and income (Tomiyama et al. 2016, Alley and Chang 2007). In column 1 of Table 3 we present such an estimate with BMI as the weight measure. Due to our measurement of bad health we expect a positive correlation between bad health and BMI. And indeed, we find that a BMI decrease of one unit improves the health status by 0.033 units. Alternatively, weight can be split into BMI classes, which leads to the estimation results in column 4. In both cases health status continues to deteriorate with increasing BMI, which is in line with earlier studies. As an alternative to Figure 1, nonlinearity between health and weight can be captured by adding a squared term of BMI (see Table 3, columns 2 and 3). For women we find a significant (negative) effect of the squared BMI term on bad health. However, up until BMI=57.42, health nonetheless continues to worsen

with increasing BMI. For men the estimates seem to support the hypothesis of a linear relationship between bad health and BMI. The results in columns 4, 5 and 6, where BMI is split into five classes with overweight people as the base class, do not confirm the linear shape. Nevertheless, health is better at lower BMI values than at higher ones. While, for example, underweight individuals have on average a better health status of 0.19 units compared with overweight people, obese II individuals have a worse health status of 0.56 units again compared with overweight people.

Waived interdependency between weight and health leads to biased estimates. Causality is not restricted from weight to health. Specifically chronic diseases and necessary drugs can increase or decrease weight. Studies have demonstrated that some people eat more when affected by depression, anxiety or other emotional disorders. In turn, overweight and obesity themselves can promote emotional disorders: If you repeatedly try to lose weight and fail, or if you succeed in losing weight only to gain it all back, the struggle can cause tremendous frustration over time, which can cause or worsen anxiety and depression. A cycle develops that leads to greater and greater obesity, associated with increasingly severe emotional difficulties. Detailed information about this relationship is only available in a small number of medical surveys where necessary personal details, characteristics and individual behavior are usually missing (Allay and Chang 2007, Bray 2004). As long as not all details on both sides of the equation are at hand we have to switch to instrumental variables estimates. In Table 4, Lewbel's approach is applied.

Here, we focus our analysis on extreme weight ranges. In columns 1 and 2 the statistical weight impact on bad health of the 10 percent of people with the lowest and highest BMI, respectively, is compared with the effect for people in the intermediate 80 percent of the BMI distribution. These percentiles are based on the German BMI distribution using SOEP data. Conversely, in columns 3 and 4 the continuous BMI variable is substituted by BMI classes. Underweight and obese people of class II, following the international classification of BMI, are contrasted with the others.

We find that the hypotheses of homoscedasticity and of exogeneity are rejected in three of the four estimates (see rows Breusch-Pagan and Durbin-Wu-Hausman). The hypothesis of weak instruments is rejected in all our estimates – compare row Cragg-Donald with row Stock-Yogo. On average, heavyweights display a significantly worse state of health while lightweight men do not differ from medium-weight men (see row for people with a BMI \geq 90 percentile and BMI \leq 10 percentile), respectively. It seems that the relationship between bad health and weight (BMI) varies across gender. The results for underweight men (see columns 2 and 4) differ from those in columns 1 and 3, respectively. Column 1 of Table 4 illustrates that women belonging to the group with the 10 percent highest BMI have a worse health status by 0.22 units than women within the intermediate group of the BMI distribution.

It also seems interesting to report the impact of control variables on health status: High income helps to improve health for men and women. Increasing deviations between desired and actual working hours have a negative effect on health. The absolute value is larger for men than for women. Moreover, we observe positive health effects if mother's education and father's social status rise. We have the following explanations for these three results: (i) the higher the income, the more resources there are available for health expenditure. This should be true for both men and women. A priori, one would even expect the effect to be stronger for women because they are generally perceived to be more health

conscious. (ii) Permanent involuntary overtime and working hours in excess of desired hours have negative long-run effects on health. However, recent research has shown that women have better opportunities for aligning actual and desired working hours via working time accounts (Hübler 2017) so that we expect less influence of deviations between desired and actual working hours on health status for women than for men. (iii) The social status and education of parents are characteristics that shape the health of their children during childhood and youth. Low social and economic status is usually linked with low health awareness (Wardle/Steptoe 2003).

3.3 Resilience and vulnerability

In Figure 1 we have seen that the spreads of health outcomes in the tails of the BMI ranges for men and women are substantially more pronounced than in the intermediate range. This means that not all men and women with low or high weight display health problems. In other words, influences exist that render some underweight and obese people resilient against chronic or mental diseases while others are particularly vulnerable. In this section we look for personal characteristics and behavioral patterns that can potentially explain such individual differences. For the former we focus on the Big5 factors (see Section 2.1). The latter field includes characteristics regarding whether a person is a smoker, whether his/her diet is unhealthy and how long he/she normally sleeps, but also activities during youth and whether weight fluctuates or whether it increases over time. Some further control variables that we have already used in estimates under 3.2 are included. In Table 5 the Big5 variables are not included; we focus on behavioral characteristics instead. We present separate estimates for individuals with a BMI below the lower quartile, above the upper weight quartile and within the interquartile range, broken down by gender. Quartiles are determined by the respective male and female BMI distributions of the SOEP data in 2012.

We confirm that smoking (0= not currently; 1= yes currently) is not good for health, a well-known fact. This effect is not confirmed for obese men and women but is particularly strong for underweight individuals. Women are less affected on the whole. Lower smoking intensity may explain this outcome.

An unhealthy diet (1= very healthy, ... , 4= not healthy at all) weakens resilience for individuals. Obese women do not appear to benefit from a healthy diet. Instead we observe a negative relation for this group, a totally unexpected and counterintuitive result. This result suggests that subjective and self-reported answers to the diet question may not necessarily be trusted. Perhaps obese people lack the appropriate information on what constitutes a healthy diet or they are deluding themselves about how well they adhere to one.

Poor sleep has bidirectional influences on health including obesity (Burt et al. 2014). Lack of sleep can make the difference between good health and poor health. However, spending an excess amount of time in bed is also linked with health hazards. In some ways, oversleeping itself appears to directly influence certain risk factors, and in other cases, it may be a symptom of other medical conditions. Sleeping for too long makes it hard for the metabolism to rise over the day. Alvarez and Ayas (2004) report that both increased (>8 h/d) and reduced (<7 h/d) sleep duration lead to a modestly increased risk of all-cause mortality and cardiovascular disease, and of developing symptomatic diabetes. This brings us to a nonlinear modeling of sleep duration effects on health. Descriptive statistics for average health values (men: 2.61 if sleep duration<7h/d, 2.42 if sleep duration>=7 & <=8 h/d, 2.71 if sleep

duration > 8h/d; women: 2.79 if sleep duration < 7h/d, 2.40 if sleep duration ≥ 7 & ≤ 8 h/d; 2.69 if sleep duration > 8h/d) and Figure A1 in Appendix 1 support the hypothesis that neither short nor long sleeping is good for health. Our investigation in Table 5 shows that long sleeping hours contribute to resilience in general. For men this effect is only observed when they are underweight. However, the impact of sleep duration appears to be nonlinear: The positive impact is reversed or weakened when sleep duration is too long.

Sporting activities (=0 if no sports activities; =1 if sports activities) during youth also lead to better health outcomes that only partially continue into adulthood. Therefore, the effects are statistically insignificant. The group of underweight women is a clear exception. Here we find negative effects on health status. The influence of musical activities (0= no musical activities during youth; 1= musical activities) is mixed. Men, but not women, benefit from music.

The higher mother's education and father's social status, the better the child's chances are of good health during adolescence and later as adults. Parents with this status usually have the financial resources and knowledge that are prerequisites for good health, e.g. a varied and balanced diet. Self-confidence has a positive effect on health and increases resilience against diseases. Of course, the older an individual is, the more likely is a worsening of his/her health status.

We obtain no clear results when we consider the effect of weight fluctuations measured by the variation coefficient. In particular, we observe negative health effects for underweight women. Perhaps, women are more likely to start a weight loss program, gain body weight again and repeat the program, thereby creating an unhealthy yo-yo effect.

In the next step we analyze the impact of Big5 variables capturing personality traits. We extend the model in Table 5 by including the generated variables openness, extraversion, conscientiousness, agreeableness and neuroticism. As several control variables are strongly correlated, this inclusion leads to insignificant estimates. Therefore, we aggregate the variables using principal component analysis with "varimax" rotation, which maximizes the variance of the squared loadings within factors. First, we have to decide how many components to extract. Eight eigenvalues are larger than 1. This Kaiser criterion supports an eight-component model. The two smallest eigenvalues among these are 1.02, the next largest is 1.12 and the largest 2.04. Therefore, a six-component model also makes sense. Following the scree plot criterion that presents a graph from the largest to the smallest eigenvalue, we should choose a four-component model. In the scree plot, the eigenvalues after the fourth decline only slightly. We decide on the six-component model. The six-component model allows for the best interpretation (see Table 6) and explains roughly 50 percent of the total variance. Based on factor loadings larger than 0.3, we interpret the components in the following way:

- (1) OECA factor – four of the Big5 variables (openness, extraversion, conscientiousness and agreeableness – OECA) load high on this component.
- (2) Factor of early influences – parents' influence, musical and sporting activities during adolescence and undenominational are the most important variables for this component.
- (3) Self-confidence and success factor – the third component is mainly determined by low neuroticism, the fifth Big5 variable, a strong deviation of the actual from the desired working time, high self-confidence and high income (gross wage).

- (4) Health behavior factor – the fourth component loads high on the variables smoker, unhealthy diet, low sleep duration and large absolute difference between desired and actual working hours.
- (5) Mixed, nonmarket factor – religious commitment, sporting activities, public sector employment and low weight fluctuations load high on the fifth component.
- (6) Weight variation factor – the final component includes only two variables with high factor loadings, namely high weight fluctuations and weight growth rate.

We incorporate the scores of the six orthogonal components into ordered probit regressions of health (Table 7). The main objective is to clarify which components best explain bad health for individuals with a BMI in the lower and upper quartile of the distribution. Factor (5) does not allow a clear interpretation. Characteristics bundled in factor (4) undermine health. Furthermore, our estimates show that factors (1)–(3) are advantageous where the intensity of the effects varies between weight groups. For example, the OECA component with openness, extraversion, conscientiousness and agreeableness contributes to resilience against health problems. Obese men are the only group for whom this link is not confirmed.

3.4 Heterogeneity and robustness

When we use more specific estimations, namely ones where the endogenous variable is a dummy for a specific disease, the coefficients and explanatory power vary from disease to disease (see Table 8). For example, OECA contributes to resilience against depressive disorder for men but does the opposite for women. In the case of heart disease, women that are characterized by factor (1) incur a low risk, while we do not find the same for men. It is not so surprising that men in contrast to women with OECA do not suffer from migraine. Finally, it should be stressed that obese men and women benefit from increased resilience against asthma when displaying OECA characteristics, while underweight men with these personality traits are more vulnerable.

In the following we present some robustness checks. We test whether the results of Table 4 hold if resilience influences – discussed under Section 3.3 – are incorporated and alternative methods are applied. We use Heckman’s sample selection approach and Lewbel’s instrumental variables regression. These two estimations split by gender are compared. The results are presented in Table 9. Columns 1 and 2 are in accord with the hypothesis of individual self-selection concerning BMI. The influence of the inverse of the Mills ratios (λ) is statistically significant. This means that the fact that an individual is underweight or obese is not exogenous but endogenous. Sample selection problems would be a consequence if we did not correct for sample selection bias. Men that have a BMI within the range of the 10 percent smallest BMI have a better health by 0.19 units than those within the intermediate BMI range, while men with a BMI within the 10 largest BMI have a worse health, namely by 0.48 units. The analogous figures for women are 0.13 positive and 0.36 negative health units.

In columns 3 and 4 Lewbel’s method is applied where control variables are extended in comparison to Table 4. Namely, the Big5 personal traits are incorporated as well as behavioral variables that moderate or enlarge the statistical health effects of weight. Here we find significant coefficients as in columns 1 and 2. The signs are the same as in columns 1 and 2 but the absolute values are larger for underweight people. The negative health effect for obese women compared with normal or slightly overweight women amounts to only 0.18 units while Heckman’s approach signals 0.36 units. In comparison with the results of columns 1 and 2 of Table 4, we have to note similar but in absolute values somewhat larger

effects for underweight women and definitely smaller coefficients for obese women. While for underweight men insignificant effects resulted in Table 4, we find here, as for underweight women, significantly positive health effects. For obese men the negative effect in column 4 is smaller than the figure determined in Table 4.

As an alternative to the specification in columns 3 and 4 we have estimated a specification where instead of using a wide range of health determinants we incorporate, besides schooling, height and age as control variables, the six principal components derived in Table 6, measured by the factor scores. The results are presented in Table A4. Now, the difference of the negative health effects between obese men and women is larger than in columns 3 and 4 of Table 9 and in all cases significant. Finally, we should stress that no health effects for underweight men are manifest. This is in accord with the result of Table 4.

4 Conclusions

The main results of this study are the following. The relationship between health and weight differs for men and women in many aspects. A wide range of weight-height combinations is consistent with the same health status. Theoretical arguments and statistical tests support the hypothesis that the influence of weight on health status is not exogenous. Interdependent estimations reduce the negative health effects of obesity compared with OLS estimates. Strong obesity, but not slight overweight and not underweight, is usually connected with bad health. Smoking is clearly not good for health. This result is less pronounced for women and not confirmed for obese individuals. In general, a healthy diet strengthens resilience against diseases but for heavy-weight people this link is not so obvious. Parents have an influence on health status of their children during adolescence and when they are adult. Good parental education and a high social status as well as own self-confidence encourage health over a wide weight range. Personality traits measured by the Big5 model taken from personality psychology can be shown to be important for explaining vulnerability and resilience against health problems. Openness, extraversion, conscientiousness and agreeableness contribute to resilience against health problems. Neuroticism enhances the vulnerability to diseases. However, the influence of weight on health problems varies from disease to disease.

Two caveats have to be expressed. First, self-reported measures of bad health, weight and personality traits may cause bias in the estimates. Second, although our tests for exogeneity speak in favor of endogeneity of weight with respect to bad health, we should be cautious about the interpretation. Alternative reasons such as misspecification may be responsible for the rejection of the null hypothesis. Our investigation has shown that “differences in weight and adjustments to one’s weight can only partially explain and affect health problems.” A lot of innate and acquired characteristics determine whether a move towards one’s “optimal” weight is successful in improving health outcomes. We need a more group-specific health policy and the individual willingness to go against the laziness eliminating vulnerabilities like smoking. Health policy should not be the same for all population groups. Gender-specific and disease-specific measures are necessary. Health policy has to start during early adolescence. Screening in childhood is helpful for detecting personality traits and disease risks. Such preliminary investigations allow early statements on the probability of risk factors and diseases in future life. Prevention is more important than cure, or in other words, “*principiis obsta, sero medicina paratur.*”

Future studies with better data should take into account unobserved family and genetic information as Böckerman/Maczulskij (2016) have done for the analysis of the education-health nexus. Investigations of effects on specific diseases should not only focus on personality traits. Typical individual behaviors such as smoking, unhealthy diet and extreme sleeping are important. Feelings such as anxiety, stress, sadness and happiness should be taken into account. Furthermore, the incorporation of individual medical histories can be helpful for uncovering determinants of vulnerability against health risks. Forthcoming analysis should focus on long-run effects of weight changes on health, on the importance of the development of the working life and its assessment, and on the simultaneous impact of changes in behavior on multiple disease risks.

Acknowledgements

The author thanks Jörg Baten, Dominik Hübler and two anonymous referees for helpful comments and suggestions that have greatly improved the paper.

References

- Altonji, J. G. and R. A. Shakotko (1987). Do Wages Rise with Job Seniority? *The Review of Economic Studies*, 54:3, 437-459.
- Alley, D. E. and V. W. Chang (2007). The changing relationship of obesity and disability 1988-2004, *Journal of American Medical Association*, 298:17, 2020-2027.
- Alvarez, G.G. and N. T. Ayas, N. T. (2004). The Impact of Daily Sleep Duration on Health: a Review of the Literature, *Progress in Cardiovascular Nursing*, 19:2, 56–59.
- Batty, G. D., Shipley, M. J., Gunnell, D., Huxley, R., Kivimaki, M., Woodward, M., Lee, C. M. Y. and G. D. Smith (2009). Height, wealth, and health: an overview with new data from three longitudinal studies, *Economics and Human Biology*, 7:2, 137-152.
- Becker, A., Deckers, T., Dohmen, T., Falk, A. and F. Kosse (2012). The relationship between economic preferences and psychological personality measures, *Annual Review of Economics*, 4, 453–478.
- Bhalotra, S., Karlsson, M., Nilsson, T. and N. Schwarz (2016). Infant health, cognitive performance and earnings: evidence from inception of the welfare state in Sweden, IZA DP No. 10339.
- Bhattacharya, J. and N. Sood (2011). Who pays for obesity? *Journal of Economic Perspectives*, 25:1, 139-158.
- Bode, E., Brunow, S., Ott, I. and A. Sorgner (2016). Worker Personality: Another Skill Bias beyond Education in the Digital Age, *SOEP Discussion Paper* 875.
- Böckerman, P. and T. Maczulskij (2016). The Education-health Nexus: Fact and fiction, *Social Science & Medicine*, 150, 112-116.
- Bray, G. A. (2004). Medical consequences of obesity, *Journal of Clinical Endocrinology and Metabolism*, 89:6, 2583-2589.
- Breusch, T. and A. Pagan (1979). A simple test for heteroscedasticity and random coefficient variation, *Econometrica*, 47, 1287-1294.
- Burt, J., Dube, L., Thibault, L. and R. Gruber (2014). Sleep and eating in childhood: A potential behavioral mechanism underlying the relationship between poor sleep and obesity, *Sleep Medicine*, 15:1, 71–75.
- Cameron, A. C. and P. K. Trivedi (2009). *Microeconometrics using Stata*, Stata Press, Lakeway Drive.
- Case, A. and C. Paxson (2002). Parental behavior and child health, *Health Affairs*, 21:2, 164-178.
- Costa, D. L. (2015). Health and the economy in the United States from 1750 to the present, *Journal of Economic Literature*, 53, 503-570.
- Cragg, J. and S. Donald (1997). Inferring the Rank of a Matrix, *Journal of Econometrics*, 76, 223-250.
- Davidson, R. and J. G. MacKinnon (2004). *Econometric theory and methods*, Oxford University Press.
- Eliassen, A. H., Colditz, G. A., Rosner, B., Willett, W. C. and S. E. Hankinson (2006). Adult weight change and risk of postmenopausal breast cancer, *Journal of American Medical Association*, 296:2, 193-201.

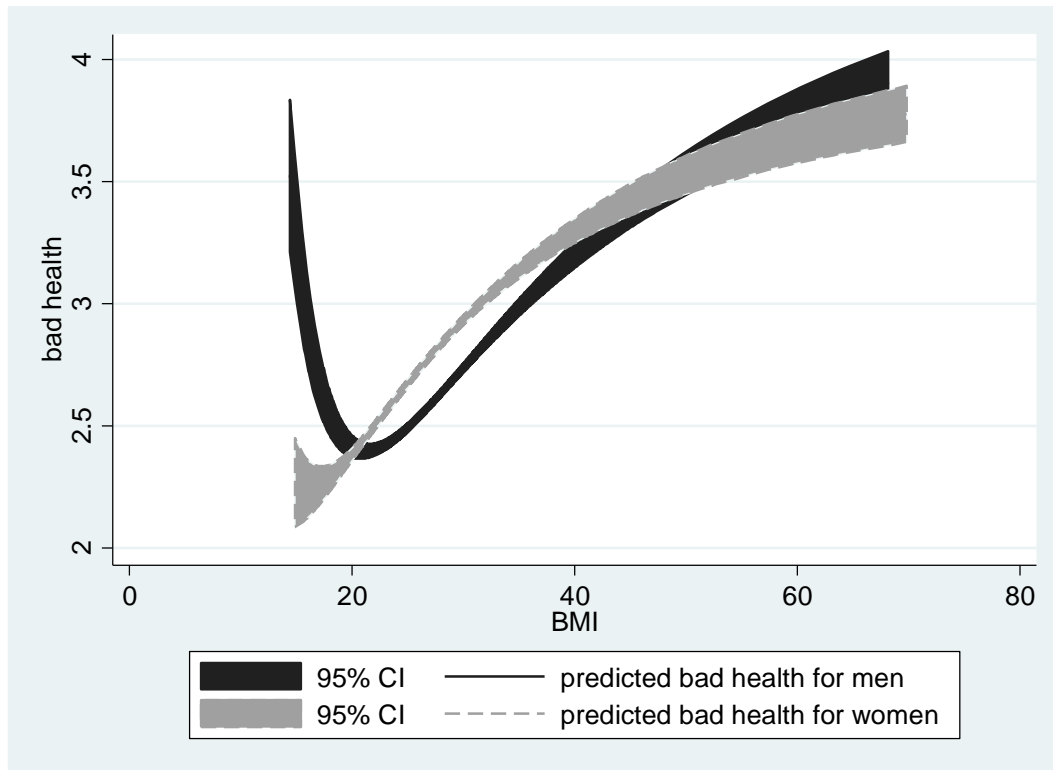
- Flegal, K. M., Kit, B. K., Orpana, H. and B. I. Graubard (2013). Association of all-cause mortality with overweight and obesity using standard body mass index categories: a systematic review and meta-analysis, *Journal of American Medical Association*, 309:1, 71-82.
- Ganzeboom, H. B. G. and D. J. Treiman (1996). Internationally Comparable Measures of Occupational Status for the 1988 International Standard Classification of Occupations, *Social Science Research*, 25, 201–239.
- Gregg, E. W., Cheng, Y. J., Cadwell, B. L., Imperatore, G., Williams, D. E., Flegal, K. M., Narajan, K. M. V. and D. F. Williamson (2005). Secular trends in cardiovascular disease risk factors according to body mass index in US adults, *Journal of the American Medical Association*, 293:15, 1868-1874.
- Hastie, T. and R. Tibshirani (1997). Generalized additive models, London: Chapman & Hall.
- Hosmer, D. W. jr. and S. Lemeshow (1980). Goodness of fit tests for the multiple logistic regression model, *Communications in Statistics*, A9:10, 1043-1069.
- Hübler, O. (2017). Men earn less under working time accounts: a puzzle? *Applied Economics Letters*, 24:1, 4-7.
- Hübler, O. (2017a). Health and body mass index – no simple relationship, unpublished paper.
- Johansson, E., Böckerman, P., Kiiskinen, U. and M. Heliövaara (2009). Obesity and labour market success in Finland: the difference between having a high BMI and being fat, *Economics and Human Biology*, 7, 36-45.
- Lewbel, A. (2012). Using heteroscedasticity to identify and estimate mismeasured and endogenous regressor models, *Journal of Business and Economic Statistics*, 30:1, 67-80.
- Lindahl, M., Lundberg, E., Palme, M. and E. Simeonova (2016). Parental influences on health and longevity: lessons from a large sample of adoptees, *IZA Discussion Paper No. 9688*.
- Lissner, L., Odell, P. M., D’Agostino, R. B., Stokes III, J., Kreger, B. E., Belanger, A. J. and Brownell, K. D. (1991). Variability of body weight and health outcomes in the framingham population, *The New England Journal of Medicine*, 324, 1839-1844.
- Maclean, J. C., Xu, H., French, M. T. and S. L. Ettner (2014). Personality disorders and body weight, *Economics and Human Biology*, 12, 153-171.
- Marx, B. D. and P. H. C. Eilers (1998). Direct generalized additive modeling with penalized likelihood, *Computational Statistics & Data Analysis*, 128, 193-209.
- McCrae, R. R. and Costa, P. T., Jr. (1997). Personality trait structure as a human universal, *American Psychologist*, 52, 509-516.
- Offer, A. (2006). The challenge of affluence: self-control and well-being in the United States and Britain since 1950, Oxford University Press, Oxford.
- Oswald, A. J. and N. Powdthavee (2007). Obesity, unhappiness, and the challenge of affluence: theory and evidence, *IZA Discussion Paper No. 2717*.

- Otterbach, S., Wooden, M. and Y. K. Fok (2016). Working-time mismatch and mental health, Melbourne Institute Working Paper No. 11/16.
- Preston, S. H., Stokes, A., Mehta, N. K. and B. Cao (2012): Projecting the effect of changes in smoking and obesity on future life expectancy in the United States, *NBER Working Paper* 18407.
- Psouni, S., Hassandra, M. and Y. Theodorakis (2016). Patterns of eating and physical activity attitudes and behaviors in relation to body mass index, *Psychology*, 7, 180-192.
- Royston, P. and D. G. Altman (1994). Regression using fractional polynomials of continuous covariates: parsimonious parametric modelling, *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, 43, 429-467.
- Ruppert, D., Wand, M.P. and R. J. Carroll (2003). Semiparametric regression, Cambridge University Press.
- Sabia, J. J. and D. I. Rees (2012). Body weight and wages: evidence from add health, *Economics and Human Biology*, 10, 14-19.
- Socio-Economic Panel (2015). Data for years 1984-2014, version 31, doi: 10.5684/soep.v31.
- Soltoft, F., Hammer, M. and N. Kragh (2009). The association of body mass index and health-related quality life of life in the general population: data from the 2003 Health Survey of England, *Quality of Life Research*, 18, 1293-1299.
- Statistisches Bundesamt (2014). Mikrozensus – Fragen zur Gesundheit, Wiesbaden.
- Stock, J.H. and M. Yogo (2005). Testing for weak instruments in linear IV regression. In: Identification and inference for econometric models: Essays in Honor of Thomas Rothenberg, ed. by D. W. K. Andrews and J.H. Stock, Cambridge University Press, 80-108.
- Thomas, D. and E. Frankenberg (2002). The measurement and interpretation of health in social surveys. In: C. Murray, J. Salomon, C. Mathers and A. Lopez (eds.). Summary measures of population health: concepts, ethics, measurement and applications, Geneva, World Health Organization, 387-420.
- Tomiyaama, A. J., Hunger, J. M., Nguyen-Cuu, J. and C. Wells (2016). Misclassification of cardiometabolic health when using body mass index categories in NHANES 2005–2012, *International Journal of Obesity*, preview 4 February 2016; doi: 10.1038/ijo.2016.17.
- Treiman, D. J. (1977). Occupational prestige in comparative perspective, Academic Press.
- Wagner, G., Frick, J. and J. Schupp (2007). The German Socio-Economic Panel Study – scope, evaluation, and enhancements, *Schmollers Jahrbuch*, 127, 139-169.
- Wardle, J. and A. Steptoe (2003). Socioeconomic differences in attitudes and beliefs about healthy lifestyles, *Journal of Epidemiology and Community Health*, 57, 440-443.
- Wehby, G. L., Murray, J. C., Wilcox, A. and R. T. Lie (2012). Smoking and body weight: Evidence using genetic instruments, *Economics and Human Biology*, 10, 113-126.
- Wehner, C., Schils, T. and L. Borghans (2016). Personality and Mental Health: The Role and Substitution Effect of Emotional Stability and Conscientiousness, IZA DP No. 10337.

WHO (2000). Obesity: preventing and managing the global epidemic. Report of a WHO Consultation. *WHO Technical Report Series 894*. Geneva: World Health Organization.

Willett, W. C., Manson, J. E., Stampfer, M. J., Colditz, G. A., Rosner, B., Spelzer, F. E. and C. H. Hennekens (1995). Weight, weight change, and coronary heart disease in women. Risk within the 'normal' weight range, *Journal of American Medical Association*, 273(6), 461-465.

Wood, S. N. (2006). *Generalized additive models: an introduction with R*, Chapman and Hall/CRC.



Notes: health=1 if very good, ... , =5 if bad, number of observations: $N_{men}=31,813$, $N_{women}=36,969$. Source: SOEP 2004, 2006, 2008, 2010 and 2012.

Fig. 1. Weight effects on bad health (h1), measured by BMI and subjective status of health, broken by gender and based on fractional polynomial estimates

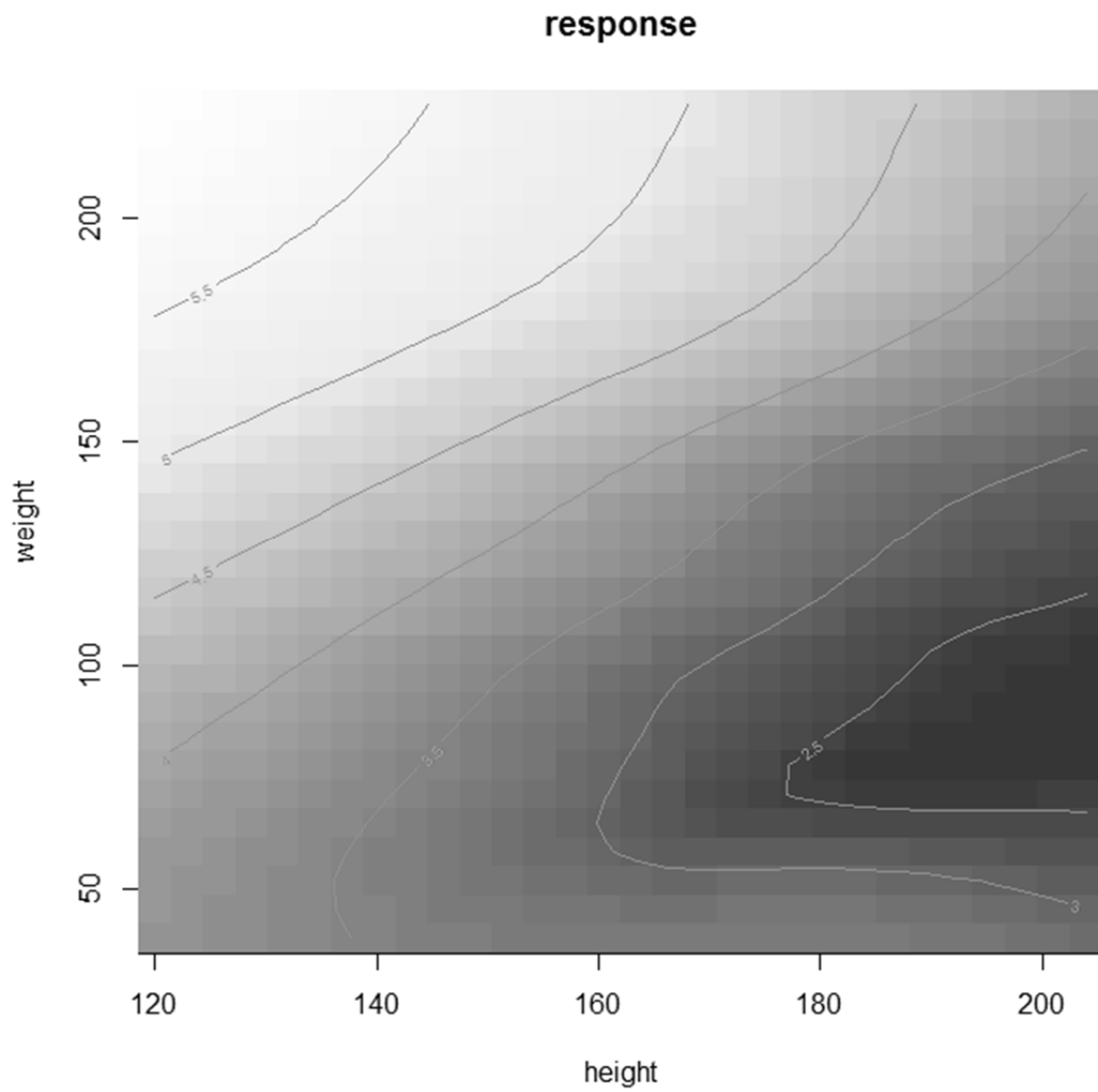


Fig. 2. Weight and height effects on bad health, measured by current subjective status of health – women -based on a generalized additive model

Table 1: correlations of health indicators – men

	h1	h2	h3	h4	h5	h6	h7	h8
h1	1.0000							
h2	-0.7379*	1.0000						
h3	-0.4634*	0.4354*	1.0000					
h4	-0.4815*	0.4852*	0.7014*	1.0000				
h5	0.4274*	-0.3920*	-0.4048*	-0.3994*	1.0000			
h6	0.2183*	-0.2340*	-0.1689*	-0.2149*	0.0888*	1.0000		
h7	0.1527*	-0.1390*	-0.0526*	-0.0437*	0.0331*	0.2018*	1.0000	
h8	-0.2975*	0.2954*	0.3293*	0.3454*	-0.4132*	-0.1224*	-0.0318*	1.0000

Notes: * $p \leq 0.05$; h1 – current health status (=1 if very good, ..., = 5 if bad); h2 – health satisfaction (=0 if completely dissatisfied, ..., =10 if completely satisfied); h3 – problems at climbing several flights of stairs on foot due to the own health status (=1 if greatly, =2 if somewhat, =3 if not at all); h4 – problems at lifting something heavy or doing something requiring physical mobility (=1 if greatly, =2 if somewhat, =3 if not at all) ; h5 – suffering from any conditions or illnesses for at least one year or chronically (=0 if no, =1 if yes) ; h6 – total number of nights spent to the hospital last year ; h7 – number of days unable to work last year due to illness; h8 - no chronic disease. Source: SOEP 2012/2013, author’s calculations.

Table 2: Body mass index distribution in percent among diseased and healthy people in Germany and the United States

	Underweight	normal weight	overweight	obese class I	obese class II
Germany					
BMI group	<=18.5	18.5-<25	25-<30	30-<40	>=40
A. German Socio-Economic Panel 2004-2012; age>=18					
All	1.4	41.7	38.0	14.3	4.5
Health: not good	1.0	34.6	40.1	17.7	6.6
Health: good	1.9	49.4	35.7	10.7	2.3
B. Microcensus 2013; age>=18					
All	2.0	45.5	36.7	14.7	1.0
Diseased	2.3	39.5	36.5	19.7	2.0
Healthy	1.9	46.7	36.8	13.7	0.8
United States					
BMI group	<=18.5	18.5-<25	25-<30	30-<35	>=35
C. NHANES 2005-2012; age>=18, metabolic status					
Unhealthy	0.9	18.0	33.2	38.3	9.6
Healthy	3.1	44.8	33.1	17.0	2.0

Sources: Statistisches Bundesamt (2014), Wagner et al. (2007), Costa (2015) and Tomiyama et al. (2016).

Table 3: OLS regressions and ordered probit estimates of bad health on weight

Method	→	OLS			ordered probit		
		Demographic group →	all	women	men	all	women
BMI		0.033*** (0.00)	0.096*** (0.01)	0.024** (0.01)			
BMI ² /1000			-0.836*** (0.16)	0.293 (0.18)			
Underweight					-0.187*** (0.05)	-0.535*** (0.06)	-0.195* (0.11)
Normal weight					-0.163*** (0.01)	-0.353*** (0.02)	-0.231*** (0.02)
Obesity I					0.224*** (0.02)	0.214*** (0.03)	0.218*** (0.02)
Obesity II					0.558*** (0.03)	0.430*** (0.04)	0.544*** (0.04)
Constant		0.944*** (0.04)	0.799*** (0.13)	1.743*** (0.16)			
R ²		0.147	0.060	0.034			
N		32,255	17,300	15,335	32,255	17,300	15,335

Notes: * p<0.10, ** p<0.05, *** p<0.01; standard errors in parentheses. Bad health is measured by a classified variable (=1, if very good health, ...,= 5, if bad health). Control variables in columns (1) and (4) are gender, age and schooling. In columns (4)-(6) overweight people are the base group. Source: SOEP 2004, 2006, 2008, 2010 and 2012, author's calculations.

Table 4: Lewbel's instrumental variables estimates of bad health

	women	men	women	men
BMI <=10 percentile	-0.3318*** (0.087)	0.068 (0.069)		
BMI >=90 percentile	0.222** (0.089)	0.726*** (0.110)		
Underweight			-0.138 (0.124)	-0.247* (0.146)
Obesity II			0.168* (0.103)	1.035*** (0.156)
Schooling	-0.009 (0.009)	-0.020** (0.008)	-0.016* (0.008)	-0.024*** (0.008)
Height	0.007*** (0.002)	-0.276*** (0.050)	0.006*** (0.001)	-0.292*** (0.050)
Height ²		0.001*** (0.000)		0.001*** (0.000)
Age	0.013*** (0.001)	0.020*** (0.001)	0.014*** (0.001)	0.020*** (0.001)
Gross wage/10000	-0.284*** (0.007)	-0.158*** (0.003)	-0.294*** (0.006)	-0.173*** (0.003)
Desired-actual working hours /1000	0.304** (0.142)	0.655*** (0.117)	0.326** (0.142)	0.673*** (0.117)
Mother's education	-0.014* (0.009)	-0.022*** (0.008)	-0.012 (0.009)	-0.021** (0.008)
Father's Treiman score	-0.004*** (0.001)	-0.001* (0.001)	-0.005*** (0.001)	-0.002* (0.001)
Constant	0.974*** (0.289)	26.537*** (4.523)	1.069*** (0.288)	28.012*** (4.489)

N	6,728	6,945	6,728	6,945
Durbin-Wu-Hausman	2.75*	5.33***	1.55	5.08***
Cragg-Donald	73.6	44.22	170.50	43.31
Stock-Yogo critical value (10% bias)	19.83	20.12	19.83	20.12
Breusch-Pagan	2.50	14.00***	5.11***	10.10***

Notes: * p<0.10, ** p<0.05, *** p<0.01, standard errors are in parentheses; Durbin-Wu-Hausman - endogeneity test (H_0 : exogeneity); Cragg-Donald - weak instrument test (H_0 : weak instruments), the Stock-Yogo critical values are based on the assumption that we are willing to tolerate distortion for a 5% Wald test based on the 2SLS estimator so that the true size can be at most 10%. We reject the null hypothesis if the Cragg-Donald test statistic exceeds the Stock-Yogo critical value. Breusch-Pagan – heteroskedasticity test (H_0 : homoskedasticity). BMI > 10 percentile & BMI < 90 percentile is the base group in columns 1 and 2. Normal weight, overweight and adipose I people are the base group in columns 3 and 4. There, 5 year dummies and 59 industries are further control variables. In columns 1-4 Lewbel’s instrumental variables estimator is used, where no additionally exogenous instruments are incorporated. Source: SOEP 2004, 2006, 2008, 2010 and 2012, author’s calculations.

Table 5: Resilience influences of people with different weight against bad health

Weight range →	lower quartile		interquartile		upper quartile	
	women	men	women	men	women	men
Smoker	0.223** (2.56)	0.839*** (8.79)	0.205*** (2.92)	0.384*** (6.24)	-0.382*** (-3.04)	-0.166* (-1.73)
Unhealthy diet	0.570*** (9.33)	0.254*** (5.14)	0.104** (2.56)	0.527*** (13.18)	-0.214*** (-3.82)	0.047 (0.81)
Sleep duration	-1.789*** (-5.52)	-2.817*** (-7.52)	-0.320* (-1.66)	0.383* (1.84)	-5.195*** (-12.63)	2.798*** (6.26)
Sleep duration ²	0.100*** (4.56)	0.217*** (7.56)	0.019 (1.33)	-0.040*** (-2.69)	0.412*** (13.21)	-0.242*** (-7.48)
Musical activities	0.563*** (7.05)	-0.327*** (-3.54)	0.139** (2.20)	-0.004 (-0.07)	0.275*** (2.88)	-0.711** (-6.82)
Sporting activities	0.246*** (3.07)	-0.010 (-0.11)	-0.041 (-0.68)	0.051 (0.68)	-0.072 (-0.71)	-0.059 (-0.66)
Undenominational	-0.618*** (-4.15)	-0.965*** (-6.21)	-0.294*** (-3.44)	0.402*** (3.53)	-1.283*** (-7.51)	0.092 (0.77)
Mother's education	-0.172*** (-4.90)	-0.345*** (-8.25)	-0.089*** (-4.24)	-0.136*** (-5.84)	0.136*** (3.76)	-0.215*** (-4.30)
Father's Treiman score	-0.004 (-1.09)	0.006* (1.67)	-0.010*** (-3.85)	-0.005* (-1.85)	-0.010*** (-2.72)	-0.011*** (-2.89)
Self-confidence	-0.054 (-0.86)	-0.467*** (-6.04)	-0.451*** (-9.52)	-0.079* (-1.74)	-0.807*** (-10.78)	-0.822*** (-12.28)
Age	0.068*** (15.43)	0.055*** (10.69)	0.030*** (9.19)	0.031*** (10.07)	0.011* (1.66)	0.040*** (9.34)
Schooling	-0.032 (-0.89)	0.222*** (6.55)	-0.085*** (-3.46)	-0.135*** (-5.03)	0.466*** (11.04)	-0.022 (-0.61)
Height	-0.001 (-0.13)	0.007 (1.17)	0.032*** (7.98)	-0.026*** (-5.90)	0.018** (2.52)	-0.002 (-0.37)

Gross wage/10000	-0.295*	0.013	0.456**	-0.472***	0.112	0.473***
	(-1.69)	(0.06)	(2.36)	(-4.56)	(0.35)	(4.30)
Public sector	-0.012	-2.591**	0.407**	-0.347*	1.261***	
	(-0.04)	(-2.95)	(2.32)	(-1.70)	(5.46)	
Desired-actual WT /100	-0.681***	-0.001	-0.086*	0.326***	-0.082	0.146**
	(-8.90)	(-0.02)	(-1.94)	(9.06)	(-1.43)	(2.55)
Weight_cv	11.184***	-2.594	-2.680***	12.346***	0.298	2.869
	(7.19)	(-1.28)	(-2.98)	(8.16)	(0.17)	(1.06)
Weight growth	3.431***	-2.913***	1.342***	-1.313***	-0.779	3.773***
	(5.51)	(-4.14)	(3.96)	(-2.98)	(-1.41)	(4.72)

N	1082	958	1918	1037	912	1045

Notes: * p<0.10, ** p<0.05, *** p<0.01, ordered probit estimates, z statistics in parentheses, cv – coefficient of variation, WT- working time in hours. Udenominational is measured by a dummy (=1, if the father of the interviewed person is no member of a Christian or another religious community; =0 otherwise). Source: SOEP 2012, author's calculations.

Table 6: Factor loadings of a six-components model after orthogonal varimax rotation

	comp 1	comp 2	comp 3	comp 4	comp 5	comp 6
Openness	0.4985	0.1261	0.1565	-0.0332	-0.0761	0.0652
Extraversion	0.5106	-0.0371	-0.0736	0.0241	0.1145	-0.0199
Conscientiousness	0.3726	-0.0918	0.2421	0.0957	-0.1636	0.0065
Agreeableness	0.4181	0.0533	-0.2539	-0.0863	0.0832	-0.0181
Neuroticism	0.2468	-0.1446	-0.3800	-0.0598	0.1129	0.0169
Smoker	-0.0429	-0.0083	0.0257	0.5964	-0.0148	-0.0958
Unhealthy diet	-0.2460	0.0941	-0.0872	0.3939	0.1383	0.1597
Sleep duration	-0.1571	-0.0909	0.0197	-0.4123	-0.0410	0.1100
Mother's education	0.0329	0.3871	-0.0908	0.1300	0.1191	0.1656
Father's Treiman score	0.0058	0.5525	0.1095	-0.1094	-0.0652	0.1162
Musical activities	0.0337	0.4224	0.0209	0.1700	-0.0208	-0.1181
Sporting activities	0.0176	0.3049	0.0198	-0.0680	0.3240	0.0444
Desired-actual WT	0.1294	-0.2189	0.4146	0.3402	-0.0714	0.0932
Public sector	0.0459	-0.0896	0.0003	0.1146	0.5973	0.1809
Undenominational	0.0480	0.3412	-0.2008	0.1130	-0.3433	-0.1532
Self-confidence	-0.0403	0.1541	0.4290	-0.2879	-0.0343	-0.0572
Gross wage	0.0525	0.0367	0.5252	-0.0088	0.2100	-0.0447
Weight_cv	0.0350	-0.0856	-0.0280	0.0897	-0.5112	0.4550
Weight growth	-0.0009	0.0454	-0.0025	-0.0479	0.0587	0.7835

Notes: Factor loadings larger than 0.3 determine the interpretation of the components; cv-coefficient of variation, WT-working time in hours. Source: SOEP 2012, author's calculations.

Table 7: Resilience influences of people with low and high weight (BMI) against bad health including Big5 of personality using principal component analysis – six-components model

	BMI <=		BMI >=	
	lower quartile		upper quartile	
	women	men	women	men
Component 1 – OECA	-0.084*** (-3.90)	-0.133*** (-5.18)	-0.136*** (-4.48)	0.140*** (5.74)
Component 2 – early influences	-0.159*** (-6.54)	-0.003 (-0.07)	-0.145*** (-3.76)	-0.038 (-1.35)
Component 3 – confidence & success	-0.174*** (-8.88)	-0.212 (-6.26)	-0.066* (-1.77)	-0.195*** (-5.45)
Component 4 – unhealthy behavior	0.079*** (3.57)	0.064** (2.24)	0.232*** (5.90)	0.125*** (3.83)
Component 5 – non-market influences	0.014 (0.70)	-0.073 (1.26)	0.029 (0.91)	-0.014 (-0.30)
Component 6 – weight variability	-0.150*** (-5.86)	-0.227*** (-5.20)	-0.004 (-0.15)	0.225*** (5.46)
N	1866	886	728	947

Notes: For the variables of the principal components and the factor loadings see Table 6, * p<0.10, ** p<0.05, *** p<0.01, ordered probit estimates, z statistics in parentheses. Source: SOEP 2012, author’s calculations.

Table 8: Influence of BIG4 via PCA on different diseases

Type of disease ↓	BMI<=		BMI>=	
	lower quartile		upper quartile	
	women	men	women	men
Sleep disturbance	-0.0815** (-2.20)	-0.1561* (-1.64)	-0.1603** (-2.06)	0.1331*** (2.59)
Asthma	0.0645 (1.14)	0.2609*** (5.87)	-0.3348*** (-5.32)	-0.1359** (-2.36)
Heart disease	-0.2543*** (-4.61)	-0.1393* (-1.77)	-0.6380*** (-5.63)	0.0467 (1.14)
Cancer	-0.2224*** (-4.08)	0.0988 (0.21)	0.1913 (1.38)	0.2350 (1.26)
Migraine	0.0254 (0.67)	-0.4050*** (-3.70)	0.1132** (2.07)	-0.2567*** (-2.82)
High blood pressure	-0.0734** (-2.35)	-0.3210*** (-6.71)	-0.1493*** (-3.92)	0.1428*** (4.70)
Depressive disorder	0.0874** (2.08)	-0.2173*** (-3.36)	0.5330*** (4.20)	-0.1508*** (-3.49)
Joint disease	-0.0613* (-1.92)	-0.2116** (2.18)	-0.1960*** (-4.94)	0.3207*** (7.91)
Chronical back complaints	0.2098*** (6.90)	0.1759** (1.82)	0.0086 (1.96)	-0.0868** (-5.43)
Other diseases	-0.3260*** (-8.54)	-0.1842*** (-3.84)	0.4800*** (7.43)	0.4390*** (8.82)
No diseases	-0.1211*** (-4.80)	0.0787** (2.48)	0.0537 (1.20)	-0.1233 (-3.77)

Notes: The modeling is the same as in Table 7 but the regressands are dummies of specific diseases.* p<0.10, ** p<0.05, *** p<0.01; probit estimates, z statistics in parentheses. Source: SOEP 2012/2013, author's calculations.

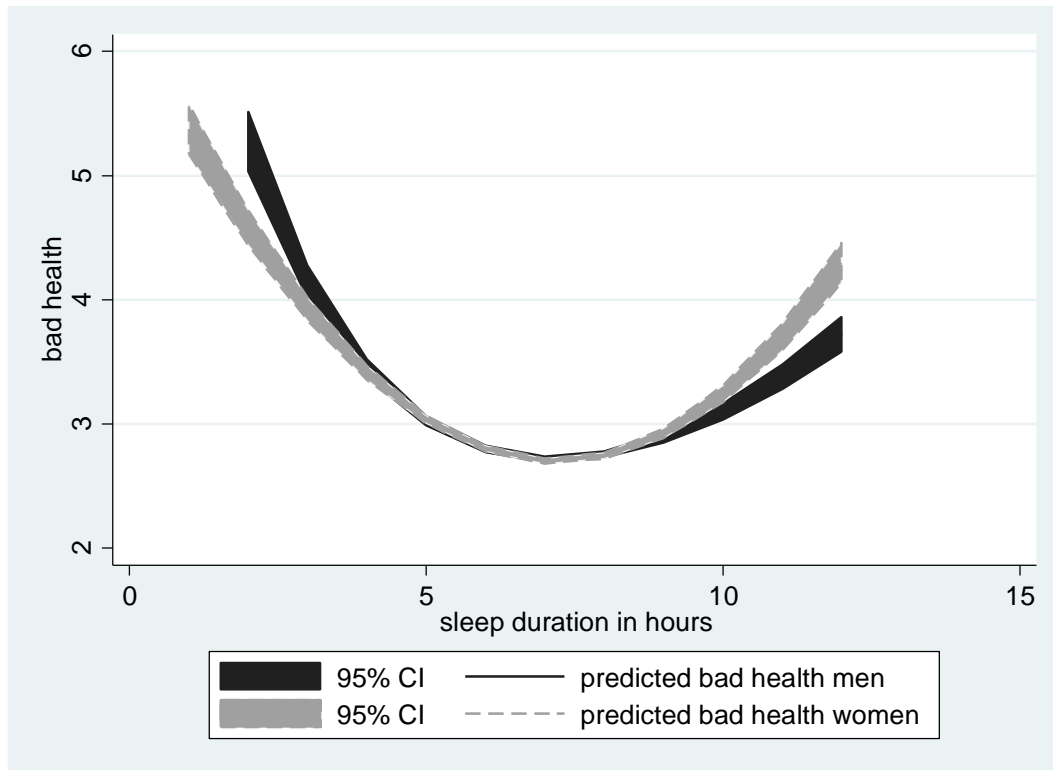
Table 9: Heckman's sample selection and Lewbel's instrumental variables estimates of bad health

Method	→	Heckman		Lewbel's IV	
		women	men	women	men
BMI <=10 percentile		-0.130*** (0.041)	-0.187*** (0.045)	-0.383*** (0.054)	-0.246*** (0.050)
BMI >=90 percentile		0.364*** (0.046)	0.479*** (0.042)	0.182*** (0.051)	0.503*** (0.046)
Schooling		-0.017 (0.012)	-0.004 (0.010)	-0.006 (0.012)	-0.011 (0.010)
Height		0.010*** (0.002)	-0.179** (0.072)	0.011*** (0.002)	-0.220*** (0.071)
Height ² /100			0.046** (0.020)		0.058*** (0.020)
Age		0.019*** (0.002)	0.019*** (0.001)	0.024*** (0.002)	0.018*** (0.001)
Gross wage/10000		-0.349*** (0.080)	0.253*** (0.050)	-0.247*** (0.073)	-0.132*** (0.044)
Desired-actual working hours /1000		0.145 (0.204)	0.656*** (0.157)	-0.079 (0.200)	0.825*** (0.153)
Mother's education		0.013 (0.011)	-0.099*** (0.011)	-0.023** (0.010)	-0.105*** (0.011)
Father's Treiman score		-0.009*** (0.001)	0.005*** (0.001)	-0.010*** (0.001)	-0.002* (0.001)
$\lambda_{\text{BMI} \leq 10 \text{ percentile}}$		-0.123** (0.053)	-0.321*** (0.048)		
$\lambda_{\text{BMI} \geq 90 \text{ percentile}}$		-0.415*** (0.084)	-1.195*** (0.070)		
Constant		1.358*** (0.426)	21.240*** (6.414)	2.159*** (0.443)	22.283*** (6.378)
N		3,677	3,625	3,677	3,625

Durbin-Wu_Hausman			114.302***	11.336***
Cragg-Donald			97,419	305.464
Stock-Yogo critical value (10% bias)			21.00	21.01
Breusch-Pagan	63.68***	3.37*	79.28***	2.32

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$, standard errors are in parentheses; Durbin-Wu-Hausman-endogeneity test (H_0 : exogeneity); Cragg-Donald -weak instrument test (H_0 : weak instruments), the Stock-Yogo critical values are based on the assumption that we are willing to tolerate distortion for a 5% Wald test based on the 2SLS estimator so that the true size can be at most 10%. We reject the null hypothesis if the Cragg-Donald test statistic exceeds the Stock-Yogo critical value (Cameron and Trivedi 2009, p.193). Breusch-Pagan – heteroskedasticity test (H_0 : homoskedasticity). BMI > 10 percentile & BMI < 90 percentile is the base group. Further control variables in columns 3 and 4 are: smoker, unhealthy diet, sleep duration, musical and sporting activities, openness, extraversion, conscientiousness, agreeableness, neuroticism, public sector employment, self-confidence, weight_vc and growth rate weight. In Lewbel’s instrumental variables estimator only the generated but no additional exogenous instruments are incorporated. Source: SOEP 2012, author’s calculations.

Appendix 1 – Graphs and tables



Notes: bad health=1, if health is very good; ... ; =5, if health is bad.

Fig. A1. Sleep effects on bad health where sleep duration is measured in hours on a normal working day and broken by gender. The estimates are based on fractional polynomials.

Table A1: Descriptive statistics of individual characteristics

Variable	N	Mean	Std. Dev.	Min	Max
Health indicators					
Bad health (h1)	68,782	2.672618	.9676437	1	5
Health satisfaction (h2)	66,745	6.658282	2.156527	0	10
Healthy (1, if (very) good)	68,817	.4694189	.4990675	0	1
No. of absent days (h7)	68,819	4.258359	19.71382	0	365
Unhealthy diet	33,224	2.383518	.7650056	1	4
Smoker (1, if yes)	33,265	.3421314	.4744305	0	1
Sleep duration in hours	18,313	6.995413	1.276908	1	12
Weight indicators					
Weight	32,762	77.33151	16.13296	32	225
Height	33,173	171.084	9.234607	117	205
BMI	32,647	26.31632	4.64416	14.45	69.86
Underweight	32,647	.0143045	.1187449	0	1
Normal weight	32,647	.4170368	.4930766	0	1
Overweight	32,647	.38028	.485463	0	1
Obesity I	32,647	.1430453	.3501244	0	1
Obesity II	32,647	.0452722	.2079036	0	1
Weight_cv	18,080	.0392647	.0307621	0	.715
Weight growth	17,897	.0152379	.0902887	-.558	.393
Socio-economic characteristics and schooling					
Male	68,819	.4625903	.4986022	0	1
Age	68,819	54.72436	16.32878	17	103
Schooling	68,103	2.384036	1.438804	1	7
Working characteristics					
Gross wage	36,017	2778.382	2664.009	0	70000
desired-actual WT	31,036	6.49963	7.71719	0	90

Characteristics as adolescents and adults

Musical activities	68,198	.3145987	.4643592	0	1
Sporting activities	68,819	.5535681	.4971258	0	1
Self-confidence	68,509	2.190836	.7166314	1	3

Parents: schooling and socio-economic status

Mother's schooling	66,271	1.554903	1.278308	0	6
Father's Treiman score	58,032	41.16718	12.33288	13	78

Big5

Openness	18,090	13.42941	3.664901	3	21
Extraversion	18,276	14.99141	2.299392	5	21
Conscientiousness	18,131	14.23953	1.967002	3	21
Agreeableness	18,292	14.30746	2.207225	6	21
Neuroticism	18,255	12.28129	2.609186	3	21

Notes: Weight and height and in consequence BMI are collected in 2004, 2006, 2008, 2010 and 2012, author's calculations. Bad health is measured by a classified variable (=1, if very good health; ... ; =5, if bad health), health satisfaction by 0 to 10 (=0, if completely dissatisfied; ... ; =10, if completely satisfied). Unhealthy diet describes the answer to the question, how much attention is paid to maintaining a healthy diet by 1 to 4 (=1, if a lot; ... ; =4, if none). Treiman score is explained in section 2.1. The BIG5 variables are only determined in 2013. The weight growth rate is calculated from 2004 to 2012. Sleep duration is surveyed on 2012. The subdivision into underweight, normal weight, overweight, obesity 1 and obesity 2 follows Costa (2015). Weight_cv measures the weight's coefficient of variation. WT is working time in hours per week. Musical and sporting activities are measured by dummies (=1, if activities during adolescence).

Appendix 2 – Lewbel’s estimator with generated and external instruments

Besides the generated instruments W (see Section 2.2) external instruments can be considered. The incorporation of exogenous instruments is sometimes necessary, namely when the number of generated instruments is not sufficient. Furthermore, this extension possibly reduces the problem that the instruments correlate with the squared error term. Allowing a comparison of the results with and without external instruments, we present the first-stage and the main equation estimates in Table A2 and A3 where two external instruments are included, namely conflicts with father during childhood and grade in mathematics. The reasons why these two variables are used as external instruments are as follows.

BMI as an adult cannot have an influence on these two variables that were observed in childhood. A substantial justification in relation to conflicts with father can be twofold with converse effects. On the one hand, it is possible that fathers have clumsily tried convincing their children that a healthy lifestyle is important but the children have done exactly the opposite and they are not standard compliant. On the other hand, it is possible that fathers have not convinced their children in the dispute but they have learned as adults.

A good grade in mathematics is an indicator of stringent and rational thinking and this determines own behavior. From this one can conclude that such people pay more attention to normal weight combined with good health while other people act less rationally and more emotionally. An opposite effect is also possible. If a student has to work hard for a good grade, he/she has to sit at a desk longer than others and does not move. This is detrimental for their health.

Although these two variables are statistically relevant BMI determinants in some cases (see Table A2) they are suboptimal as instruments because BMI and the health of a child may have statistical influence on conflicts with father and grade in mathematics on the one hand, and on health and weight as an adult on the other. Another channel of infant health effects on cognitive performance and earnings is described by Bhalotra et al. (2016) and this may have effects on the health-weight relationship of adults. We test in the first-stage regressions (BMI functions) whether the influence of generated and possibly additional exogenous instruments on BMI is significant and whether the specification fits well, where the Hosmer-Lemeshow (1980) goodness-of-fit test is applied (see Table A2).

If the applied external instruments are optimal, Lewbel’s approach is not required. However, we should emphasize that the estimates are mainly determined by the generated instruments. The addition of the two instruments leads only to small changes in the coefficients – compare columns 1 to 4 of Table A3 with columns 1 to 4 of Table 4.

Table A2: First-stage probit estimates of Lewbel’s instrumental variables approach

Demographic group	→	women	men	women	men
BMI percentile	→	<=10	<=10	>=90	>=90
W_schooling		0.082*** (0.008)	0.055*** (0.008)	-0.069*** (0.007)	-0.076*** (0.007)
W_height		0.005*** (0.002)	-0.001 (0.001)	-0.004*** (0.001)	0.002 (0.001)
W_age		-0.010*** (0.001)	-0.015*** (0.001)	0.009*** (0.001)	0.006*** (0.001)
W_gross wage/10000		0.106* (0.062)	-0.474*** (0.043)	0.001 (0.051)	0.061*** (0.021)
W_ desired-actual WT /1000		-0.288** (0.126)	0.200* (0.105)	0.234** (0.107)	0.031 (0.098)
W_mother’s education		-0.040*** (0.008)	-0.016** (0.007)	-0.031*** (0.007)	-0.010 (0.007)
W_father’s Treiman score		0.006*** (0.01)	0.001 (0.001)	-0.005*** (0.001)	0.001 (0.001)
Conflicts with father		0.007* (0.004)	0.000 (0.003)	0.005* (0.003)	-0.005 (0.003)
Grade in mathematics		-0.015*** (0.004)	0.001 (0.004)	0.002 (0.003)	-0.015*** (0.003)
N		6,523	6685	6,523	6,685
Pseudo_R ²		0.109	0.187	0.155	0.066
Goodness-of-fit-test		236,899***	8,593,189***	1,711,544***	18,806***

Notes: * p<0.10, ** p<0.05, *** p<0.01; marginal coefficients are presented; standard errors in parentheses. W – generated instruments following Lewbel – see section 2.2. The goodness-of-fit test statistic is chi-square distributed and follows Hosmer-Lemeshow. Source: SOEP 2004, 2006, 2008, 2010 and 2012, author’s calculations.

Table A3: Lewbel's instrumental variables estimates of bad health

	women	men	women	men
BMI <=10 percentile	-0.347*** (0.085)	0.083 (0.069)		
BMI >=90 percentile	0.271*** (0.084)	0.672*** (0.107)		
Underweight			-0.218*** (0.071)	-0.321** (0.134)
Obesity II			0.381*** (0.054)	0.575*** (0.047)
Schooling	-0.021** (0.009)	-0.0218*** (0.008)	-0.030*** (0.009)	-0.029*** (0.009)
Height	0.007*** (0.002)	-0.286*** (0.053)	0.007*** (0.002)	-0.332*** (0.052)
Height ²		0.001*** (0.000)		0.001*** (0.000)
Age	0.0134*** (0.001)	0.020*** (0.001)	0.014*** (0.001)	0.020*** (0.001)
Gross wage/10000	0.002*** (0.007)	-0.156*** (0.003)	0.009*** (0.001)	-0.211*** (0.003)
Desired-actual working hours /1000	0.327** (0.150)	0.695*** (0.120)	0.044*** (0.143)	0.377*** (0.112)
Mother's education	-0.018** (0.009)	-0.022*** (0.009)	-0.013 (0.009)	-0.027*** (0.009)
Father's Treiman score	-0.003*** (0.001)	-0.001* (0.001)	-0.004*** (0.001)	-0.001 (0.001)
Constant	0.922*** (0.293)	27.490*** (4.732)	0.861*** (0.297)	31.70*** (4.665)

N	6,523	6,685	6,523	6,685
Durbin-Wu-Hausman	34.22***	12.59***	47.00***	4.00
Cragg-Donald	67.39	42.82	3.37	1.95
Stock-Yogo critical value (10% bias)	20.12	20.33	7.03	7.03
Breusch-Pagan	3.14*	14.00***	11.59***	17.66***

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$, standard errors are in parentheses; Durbin-Wu-Hausman-endogeneity test (H_0 : exogeneity); Cragg-Donald -weak instrument test (H_0 : weak instruments), the Stock-Yogo critical values are based on the assumption that we are willing to tolerate distortion for a 5% Wald test based on the 2SLS estimator, so that the true size can be at most 10%. We reject the null hypothesis if the Cragg-Donald test statistic exceeds the Stock-Yogo critical value (Cameron and Trivedi 2009, p.193). Breusch-Pagan – heteroskedasticity test (H_0 : homoskedasticity). BMI > 10 percentile & BMI < 90 percentile is the base group in columns 1 and 2. Normal weight, overweight and adipose I people are the base group in columns 3 and 4. There, 5 year dummies and 59 industries are further control variables. Lewbel’s instrumental variables estimator is used, where additionally two exogenous instruments are incorporated (a dummy whether the individual had conflicts with his/her father during the adolescence, the grade in mathematics). Source: SOEP 2004, 2006, 2008, 2010 and 2012, author’s calculations.

Table A4: Lewbel’s instrumental variables estimates of bad health without external instruments and with principal components measured by factor scores

	women	men
BMI<=10 percentile	-0.430*** (0.06)	0.014 (0.05)
BMI>=90 percentile	0.165*** (0.07)	0.861*** (0.06)
Component 1 – OECA	-0.071*** (0.01)	-0.014* (0.01)
Component 2 – success & confidence	-0.051*** (0.01)	-0.126*** (0.01)
Component 3 – early influences	-0.075*** (0.01)	-0.032*** (0.01)
Component 4 – health behavior	0.181*** (0.01)	0.161*** (0.01)
Component 5 – non-market influences	0.028*** (0.01)	0.007 (0.02)
Component 6 – weight variability	0.037*** (0.01)	0.066*** (0.01)
Schooling	-0.024** (0.01)	-0.006 (0.01)
Height	0.009*** (0.00)	-0.010*** (0.00)
Age	0.025*** (0.00)	0.021*** (0.00)
Constant	-0.047 (0.36)	3.219*** (0.35)
N	3677	3625

Notes: For the variables of the principal components and the factor loadings see Table 6, * p<0.10, ** p<0.05, *** p<0.01, standard errors in parentheses. Source: SOEP 2012, author’s calculations.