RUHR ECONOMIC PAPERS

Lucas Hafner
Harald Tauchmann
Ansgar Wübker

Does Moderate Weight Loss Affect Subjective Health Perception in Obese Individuals?

Evidence from Field Experimental Data

#730

Imprint

Ruhr Economic Papers

Published by

RWI – Leibniz-Institut für Wirtschaftsforschung
Hohenzollernstr. 1-3, 45128 Essen, Germany

Ruhr-Universität Bochum (RUB), Department of Economics
Universitätsstr. 150, 44801 Bochum, Germany

Technische Universität Dortmund, Department of Economic and Social Sciences
Vogelpothsweg 87, 44227 Dortmund, Germany

Universität Duisburg-Essen, Department of Economics
Universitätsstr. 12, 45117 Essen, Germany

Editors

Editorial Office

Ruhr Economic Papers #730

Responsible Editor: Jochen Kluve

ISSN 1864-4872 (online) – ISBN 978-3-86788-850-9

Ruhr Economic Papers #730

Lucas Hafner, Harald Tauchmann, and Ansgar Wübker

Does Moderate Weight Loss Affect Subjective Health Perception in Obese Individuals?

Evidence from Field Experimental Data

Bibliografische Informationen der Deutschen Nationalbibliothek

The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data are available on the Internet at http://dnb.dnb.de

RWI is funded by the Federal Government and the federal state of North Rhine-Westphalia.

http://dx.doi.org/10.4419/86788850
ISSN 1864-4872 (online)
ISBN 978-3-86788-850-9

Lucas Hafner, Harald Tauchmann, and Ansgar Wübker[1]

Does Moderate Weight Loss Affect Subjective Health Perception in Obese Individuals?

Evidence from Field Experimental Data

Abstract

This paper analyzes whether moderate weight reduction improves subjective health perception in obese individuals. To cure possible endogeneity bias in the regression analysis, we use randomized monetary weight loss incentives as instrument for weight change. In contrast to related earlier work that also employed instrumental variables estimation, identification does not rely on long-term, between-individuals weight variation, but on short-term, within-individual weight variation. This allows for identifying short-term effects of moderate reductions in body weight on subjective health. In qualitative terms, our results are in line with previous findings pointing to weight loss in obese individuals resulting in improved subjective health. Yet, in contrast to these, we establish genuine short-term effects. This finding may encourage obese individuals in their weight loss attempts, since they are likely to be immediately rewarded for their efforts by subjective health improvements.

JEL Classification: I12, C26, C93

Keywords: Self-rated health; BMI; obesity; randomized experiment; short-term effect; instrumental variable

December 2017

1 Lucas Hafner, Universität Erlangen-Nürnberg; Harald Tauchmann, Universität Erlangen-Nürnberg, RWI and CINCH; Ansgar Wübker, RWI, RUB, and Leibniz Science Campus Ruhr. *– We gratefully acknowledge the comments and suggestions of Martin Salm, John Cawley, Hendrik Jürges, Hendrik Schmitz, Thomas Siedler, Annika Herr, the participants of the annual meeting of the German Association of Health Economists (dggö) 2017, the annual congress of the German Economic Association 2017, and the Workshop on Risky Health Behaviors at the University of Hamburg, 2017. The authors are furthermore grateful to Franziska Valder for research assistance.–* *All correspondence to: Lucas Hafner, Professur für Gesundheitsökonomie, Findelgasse 7/9, 90402 Nürnberg, Germany, e-mail: lucas.hafner@fau.de*

1 Introduction

It is well documented in the literature that excessive accumulation of body fat (obesity) is associated with many undesirable health outcomes such as heart disease (Hubert et al., 1983), type 2 diabetes (Mokdad et al., 2003), and several forms of cancer (Calle et al., 2003). A recent meta-analysis (Di Angelantonio, 2016) even finds that obese individuals face a higher risk of all-cause mortality compared to their normal-weight counterparts. Although, at least in western societies, the general public seems meanwhile to be well aware of the health risks associated with obesity (Tompson et al., 2012), its prevalence is at an all-time high and further increasing worldwide (WHO, 2000; Ng et al., 2014).

Even for moderate weight loss (5-10% percent of body weight) in obese individuals, substantial benefits for objectively measurable health outcomes, such as blood lipid profiles or cardiovascular risk factors, have been established (Blackburn, 1995; Wing et al., 2011). However, despite likely health benefits from losing weight, many obese struggle with realizing even small, sustained reductions in body weight. This ubiquitous everyday experience is also well documented in the scientific empirical literature. In a systematic review of long-term weight management schemes Loveman et al. (2011), for instance, find that short-run reductions of bodyweight are commonly offset by subsequent weight regain. A better understanding of the mechanisms that make weight loss sustainable and the factors that let weight loss efforts fail is, hence, crucial for battling the obesity 'epidemic'.

One possible explanation is that moderate weight loss insufficiently induces short-term improvements in perceived health. Objective health measures, for which beneficial effects are well established, do not necessarily reflect patient's subjective health perception. Yet the latter is likely to matter much for health and obesity related behavior. If one realizes some weight loss under great efforts without feeling better, it may be tough to keep up the discipline to maintain or further reduce one's body weight.

In order to contribute to the discussion, we empirically address the question of whether moderate weight loss causally influences the subjective health perception of obese individuals. Several analyses have examined the relationship between self-rated health (*SRH*) and excess bodyweight. The vast majority of the existing literature find a significant negative association that is poor self-rated health accompanies obesity. Using a national survey with Americans Ferraro and Yu (1995) found that – even after controlling for morbidity and functional limitations – obese individuals have a higher probability of bad self-rated health compared to normal weight individuals. Okosun et al. (2001) approve this finding also analyzing a sample of Americans. Phillips et al. (2005), Prosper et al. (2009), and Baruth et al. (2014) are further, more recent examples for analyses yielding similar results based on US data.

This general pattern is not confined to studies using data from the US. Guallar-Castillón et al. (2002), for instance, analyze a sample of Spanish women and find that overweight and obese individuals are significantly more likely to report poor health compared to normal weight women. Molarius et al. (2007) found that overweight (body mass index $[BMI] \geq 25^{kg}/_{m^2}$) and obese ($BMI \geq 30^{kg}/_{m^2}$) Swedes have a higher probability to rate their health as poor, compared to normal weight survey respondents. Examining health surveys from Portugal and Switzerland Marques-Vidal et al. (2012) find that obese subjects rated their health significantly worse compared to their normal weight counterparts. This also holds for UK residents as shown in Ul-Haq et al. (2013b).

Only very few empirical studies yield mixed findings or do not find a significant association between self-rated health and obesity at all. Looking at American cross-sectional data over a time span of 30 years (1976-2006) Macmillan et al. (2011) confirms the above pattern for women. Yet, for men, the association between obesity and *SRH* is weaker and only significant in roughly half of the considered years. Imai et al. (2008) find that the association of *BMI* and *SRH* varies significantly across different ages and sexes. They generally confirm previous findings, stating that being underweight or severely obese is associated with bad *SRH*. However, they find no significant association for obese men older than 65. Darviri et al. (2012) find no significant association between *SRH* and *BMI* for a rural population in Greece, neither do Kepka et al. (2007) using a sample of Hispanic immigrants in the US.

Although a close association of excess body weight and *SRH* is very well documented in the literature, the question remains unsettled whether excess body weight causally affects self-rated health. Such effect is crucial for subjectively perceived health improvements encouraging obese individuals in their weight loss efforts. However, the mere correlation may just capture the influence of confounding third factors such as certain life styles that affect both body weight and self-perceived health. An example for a confounding third factor is sleep duration. Studies find short sleep duration to be associated with poor self-rated health (Frange et al., 2014) as well as obesity (Patel and Hu, 2008). Stress may serve as another example for such confounding factors. An increase in stress is likely to have detrimental effects on self-rated health. At the same time, stress may induce overeating (Zellner et al., 2006). Moreover, reverse causality may also be an issue. One may, for instance, think of individuals who feel well and healthy and are motivated by this to practice an active life-style that prevents them from becoming overweight.

The above mentioned studies analyze the relationship of inter-individual weight-variation and self-rated health in cross-sectional data sets. They spend little effort in establishing causality in the link between *SRH* and obesity. One notable exception in this literature is Cullinan and Gillespie (2015) who employ instrumental variables estimation to identify a causal link. Following several examples from the literature (Ali et al., 2014; Cawley and Meyerhoefer, 2012; Sabia and Rees, 2011; Kline and Tobias, 2008) they use body weight of biological relatives (children) as instrumental variable. This choice of instrument seems to be well justified by evidence from adoption (Vogler et al., 1995; Sacerdote, 2007) and twin studies (see Elks et al., 2012; Maes et al., 1997, for surveys of this literature), which suggests that shared genetics explain intra-family correlation of *BMI* much better than the shared social environment.[1] However, despite the major importance of genetic disposition, household level environmental conditions may still play some role for intra-family correlation of body weight. They may, in turn, contaminate biological relatives' body weight as instrument, since such conditions may also matter for health and subjective health perception. More importantly, even if close relatives' body weight is a valid and strong instrument for the level of *BMI* or overweight status in a cross-section of data, it can hardly be used as instrumental variable if the analysis is concerned with the effects of relatively small *changes* in body weight, which are observed over a relatively short period of time.

This is precisely the focus of the present analysis that aims at identifying subjective health effects of a moderate, short-term weight loss in obese individuals. Our contribution is to develop an empirical strategy that allows for identifying such intra-individual short-term effects. Following Cullinan and Gillespie (2015) and earlier work, we rely on instrumental variables estimation to establish a causal link. Yet, we do not adopt their instrument, which provides an exogenous source of variation in the long-term level of *BMI*. We rather make use of a randomized controlled experiment that exogenously induced short-term variation in body weight,[2] and hence provides a basis for identifying short-term effects attributable to moderate weight loss.

The remainder of this paper is structured as follows. In section 2 we introduce our data. In section 3 we describe our estimation procedure. In section 4, we show the results of our estimations. Finally in section 5 we summarize and discuss our main findings and present a conclusion.

2 Data

2.1 The field experiment

The data used in the present analysis originate from a field experiment that was conducted by RWI – Leibniz-Institut für Wirtschaftsforschung. Its prime objective was to test whether monetary incentives are an effective instrument for assisting obese individuals in losing body weight. Four medical rehabilitation clinics operated by the German Pension Insurance of the federal state of Baden-Württemberg and the association of pharmacists of Baden-Württemberg cooperated with RWI in this project. The *Pakt für Forschung und Innovation*, which is part of the excellence in research initiative of the German federal government, provided funding. The study protocol of the project was approved by the ethics commission of the Chamber of Medical Doctors of Baden-Württemberg. See Augurzky et al. (2012) and Augurzky et al. (2014) for a more detailed discussion of the project.

[1] A closely related identification strategy is to directly use genetic information as instrument for body weight. Few recent contributions (Norton and Han, 2008; Fletcher and Lehrer, 2011; von Hinke et al., 2016; Willage, 2017), which consider different outcomes than subjective health, have adopted this strategy.

[2] Reichert (2015) and Reichert et al. (2015) use the same source of exogenous weight-variation but consider different outcomes than health.

Upon admission to one of the four involved clinics, 695[3] obese individuals were recruited for participation in the experiment between March 2011 and August 2012. The medical staff in charge was advised to approach any new patient whose *BMI* exceeded 30[4] and to invite him or her to take part in the experiment. Yet, participation was entirely voluntary and had no consequence for any treatment or advice the patient received over their rehab stay, which usually takes three weeks. The prime objective of rehab stays in these clinics is to preserve, or to restore, patients' workableness. Our study population is hence biased towards the working population, which is however no challenge to the internal validity of our analysis. For the vast majority of participants, obesity was not the prime reason for being sent to rehabilitation. Yet, many suffered from health problems related to overweight such as chronical back pain. Hence, all obese patients, irrespective of participation in the experiment, were advised to reduce their body weight.

At rehab discharge participants' body weight was measured again and participants were set an individual weight loss target by the physician in charge, which they were prompted to realize within four months. Physicians were asked to choose a weight loss target of about 6 to 8 percent of current body weight. Yet they were in principal free to deviate from this guideline. Right after rehab discharge, the participants were randomly assigned to one control and two treatment/incentive groups, and subsequently informed about the result of the randomization by regular mail (intervention). While in this letter all participants were prompted to realize their weight loss target, treatment group members were informed about the monetary reward they could earn by being successful in losing weight.[5] For one treatment group the maximum reward was 150 €, for the other it was 300 €. If participants failed to realize at least 50 percent of the contractual weight loss they did not receive any money. If they were partially successfully, i.e. they lost more than 50 percent but less than 100 percent, they were rewarded proportionally to the degree of target achievement.

By the end of the four month weight loss period, all participants received another letter, by which they were prompted to visit a specified pharmacy in a specific week for a weigh-in. Body weight measured in the pharmacy served as basis for the cashout of rewards. Upon attending the weigh-in all participants, irrespective of the experimental group they were assigned to, received an expense allowance of 25 €. Each letter was accompanied by a questionnaire, which the participants had to answer. The questionnaires covered a wide range of questions regarding socio-economic characteristics and weight related behavior, such as exercising, eating habits, etc. Most importantly, participants were also asked about their health status. Two health questions addressed self-rated

[3]Originally 700 patients were recruited, yet five had to be excluded because of ex-post violation of the inclusion criteria (pregnancy, developing cancer) or missing documents.

[4]In addition to $BMI > 30$, a detailed list of inclusion criteria needed to be met; in detail: age between 18 and 75 years, resident of the federal state of Baden-Württemberg, sufficient German language skills, no pregnancy, no psychological and eating disorders, no substance abuse, no other serious illnesses.

[5]At recruitment, all participants where informed about the design of the experiment (randomization, monetary rewards) control group members, hence, knew that they missed the chance of financially benefitting from losing weight.

health and physical well-being, in a standard fashion.

The experiment included two further phases. A six month weight maintenance phase, which directly followed the weight reduction phase, and a subsequent twelve month follow-up phase. In the weight maintenance phase, participants who were at least partially successful in meeting their weight loss target were offered another monetary reward for not exceeding their target weight. In the follow-up phase participants were not exposed to any monetary incentives for weight loss. In both phases, the weigh-in procedure was the same as for the weight loss phase. The present analysis only uses information up to the end of the weight reduction phase. The reason for this is that in the weight reduction phase the exogenous source of weight variation, i.e. being member of the control or the treatment arm of the experiment, is clearly random by the design of the experiment. This applies less to the subsequent weight maintenance phase, since the second randomization was conditional on success in the previous phase.

The econometric analysis rests on information which was collected at rehab discharge and by the end of the weight loss phase. While the information regarding body weight is complete for the first time of measurement, this does not hold for the second, since roughly one-forth of the participants did not attend the weigh-in by the end of the weight loss phase. In consequence, weight-change information is available for only 517 participants. Augurzky et al. (2012) comprehensively discuss the issue of experiment drop-out and its possible implications. Using a battery of different econometric techniques, they find that the results are rather robust to correcting for selective drop-out. Unlike body weight, which was measured in the clinic or the pharmacy, the information regarding self-rated health and physical well-being was collected through a written questionnaire. This renders item non-response an issue, which further reduced the size of the estimation sample to 485 individuals in the self-rated health estimation and 468 in the physical well-being estimation, for which weight and health information is available for either time of measurement.

2.2 Variables used in the Empirical Analysis

We employ two variables to measure the outcome subjective health perception: (i) self-rated health (*SRH*) and (ii) physical well-being (*PWB*). Self-rated health is measured by asking the respondents "how would you describe your current health status?" and allowing for five possible answers: "very good", "good", "satisfactory", "poor" and "bad".[6] Physical well-being is measured by asking the respondents "how would you describe your current physical well-being?", allowing for the same five possible answers.

While either variable measures subjective health perception, they potentially capture different aspects of it. *PWB* emphasizes subjectiveness in health perception even stronger, while *SRH* leaves more room for objectifying the reported health status. For

[6] A wide variety of methods to assess subjective health perception have been suggested in the literature. These methods include multi-item measures as well as single item-measures. An example of a multi-item measure is the often used Medical Outcomes Study Short Form 36 (SF-36) (Ware et al., 1993). Most studies using the SF-36 find obesity to be associated with poor subjective health perception (see Kroes et al., 2016; Ul-Haq et al., 2013a; Kolotkin et al., 2001; Fontaine and Barofsky, 2001, for reviews).

Table 1: Joint and Marginal Distribution of SRH_0 and SRH_1

		SRH_1 bad	poor	satisfactory	good	very good	marginal distribution
SRH_0	bad	7	5	6	0	0	18
	poor	14	37	50	18	2	121
	satisfactory	5	28	106	67	4	210
	good	2	8	38	69	8	125
	very good	0	2	3	4	2	11
marginal distribution		28	80	203	158	16	485

instance an obese individual without any health impairments might rate her physical well-being as very good. At the same time, she is probably aware, that her excess weight is a risk for her health. Although feeling healthy she might therefore report a relatively poor *SRH*, to account for potential health risks.

While any questionnaire the participants were asked to fill in included questions about *SRH* and *PWB*, the present empirical analysis focusses on *SRH* and *PWB* that was reported by the end of the four-month weight reduction phase. These variables, denoted as SRH_1 and PWB_1 enter the econometric model at the left-hand side.[7] The analysis also makes use of self-rated health and physical well-being reported at rehab discharge, i.e. at the outset of the weight reduction phase. As single item measures that do not refer to any objective health indicator but are purely subjective in nature, *SRH* and *PWB* are well suited for analyzing self-perceived rather than objectively measured health effects.

Table 1 displays the (joint and marginal) sample distribution of *SRH* for both considered times of measurement. Not surprisingly – all respondents underwent medical rehabilitation for some reason – the share of individuals who regarded themselves in very good or good health is smaller than in general population surveys such as the German Socioeconomic Panel (SOEP). Nevertheless, *SRH* exhibits substantial heterogeneity between individuals. From Table 1 it also becomes obvious that self-rated health considerably varies at the individual level over the observation period.[8] For 54% of the participants we observe a change in *SRH* (off-diagonal elements in Table 1), while 46% report the same category of *SRH* at the beginning and by the end of the weight reduction phase (cells highlighted grey in Table 1). 60% of all changes are improvements in *SRH* (cells above the principal diagonal). Among the participants who reported *SRH* changes, 81% report a change to an adjacent category. Yet, some rather drastic shifts in *SRH*, e.g. from ‘very good’ to ‘poor’ or the other way round, are observed.

The corresponding (joint and marginal) sample distribution of *PWB* at rehab dis-

[7] The subscript 1 is a time index that refers to the information gathered by the end of the weight loss phase (period 1). The subscript 0 indicates pre-intervention values that is SRH_0 (PWB_0) denotes self-rated heath (physical well-being) at rehab (period 0) discharge. This notation analogously applies to all variables that are measured at different points in time such as the body mass index BMI_1 and BMI_0.

[8] If no within-individual variation of *SRH* was observed, linking changes in *SRH* to weight change would arguably make little sense.

Table 2: Joint and Marginal Distribution of PWB_0 and PWB_1

		PWB_1					
		bad	poor	satisfactory	good	very good	marginal distribution
PWB_0	bad	9	9	5	5	1	29
	poor	22	48	58	20	3	151
	satisfactory	10	37	82	53	5	187
	good	0	6	29	48	8	91
	very good	0	1	2	5	2	10
marginal distribution		41	101	176	131	19	468

Table 3: Joint and Marginal Distribution of PWB_1 and SRH_1

		SRH_1					
		bad	poor	satisfactory	good	very good	marginal distribution
PWB_1	bad	19	16	7	1	0	43
	poor	9	40	43	9	0	101
	satisfactory	1	18	122	40	3	184
	good	0	2	29	102	2	135
	very good	1	0	2	6	10	19
marginal distribution		30	76	203	158	15	482

charge (PWB_0) and at the end of the weight loss phase (PWB_1) is displayed in Table 2.[9] Comparable to *SRH*, physical well-being exhibits substantial heterogeneity between individuals and varies at the individual level over time. For 60% of the participants we observe a change in *PWB* (off-diagonal elements in Table 2), while 40% report the same category of *PWB* at the beginning and by the end of the weight reduction phase (cells highlighted grey in Table 2). 60% of all changes are improvements in *PWB* (cells above the principal diagonal). Among the participants for which reported *PWB* changes, 79% report a change to an adjacent category.

Self-rated health and physical well-being are obviously closely related measures and are strongly correlated in the sample. However, as their correlation is far from perfect the two variables seem to capture different aspects of subjective health perception. Table 3 displays the (joint and marginal) sample distribution of *SRH* and *PWB* at the end of the weight loss phase. Most respondents report the identical answer category for both variables (61%). However, 25% of the respondents reported better *SRH*, while 14% of the respondents reported better *PWB*.[10] Only 1% of the respondents deviated by more than two answer categories (bad *SRH* and very good *PWB*).[11]

[9] The number of observations for the variables measuring subjective health perception differ - individuals reported their self-rated health status slightly more often.

[10] This pattern is similar when we look at the relationship of self-rated health and physical well-being at the end of the rehab-phase (correlation coefficient of 0.64). Here 58% of respondents reported the same answer category for both variables, while 28% of respondents reported better *SRH* and 14% reorted better *PWB*. See Table A1 in the Appendix.

[11] Excluding these individuals from the analysis does not change our results in qualitative terms.

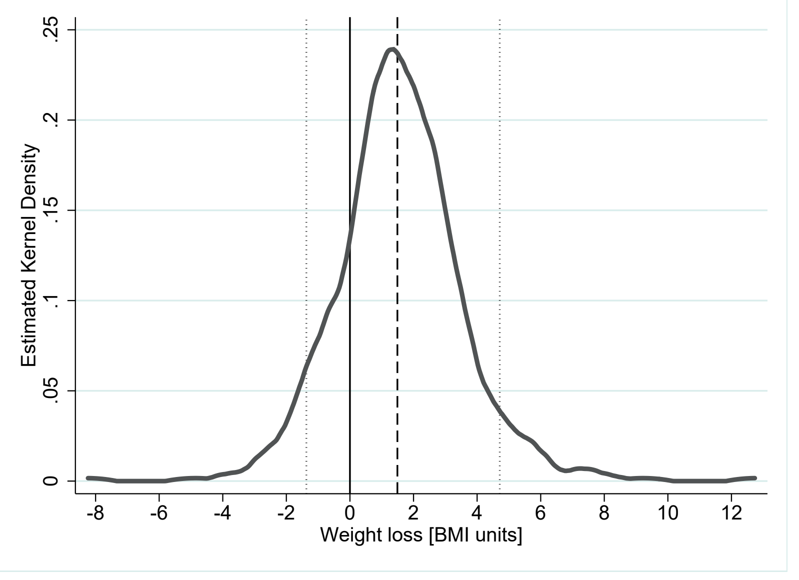

Figure 1: Distribution of *Weight loss* in the sample.

Notes: Estimated kernel density; dashed line marks the median (1.51); dotted lines mark the 95th (4.71) and the 5th (−1.37) percentile.

Body weight, which is the key explanatory variable in the present analysis, is measured in terms of the body mass index.[12] Rather than its level, we consider the absolut change (*Weight loss* $\equiv BMI_0 - BMI_1$) between rehab discharge and the end of the weight-reduction phase as regressor. By this choice, we emphasize that the focus of the analysis is on the effects of within-individual weight loss rather than between-individual heterogeneity in the level of *BMI*.[13]

The variation of weight change in the sample is quite substantial. 81% of the participants lost weight. Mean weight change is 1.56 *BMI* units. The median of the weight loss distribution (1.51) is close to the mean. The 95 percent quantile is 4.71, indicating that a substantial share of participants managed to materially reduce body

[12]For decades, *BMI* and its commonly used threshold value of $30 kg/m^2$ (WHO, 2016) have been criticized as an, at least in certain circumstances, inappropriate measure of clinical obesity (Garn et al., 1986). We nevertheless stick to this frequently used measure. Since we consider changes of *BMI* over a relatively short period of time, rather than comparing the level of *BMI* between individuals, several shortcomings (age dependence, indifference regarding lean and fat tissue, etc.) of the *BMI* are arguably of little importance. Using percentage change in body weight instead of absolute change in *BMI* as weight change measure yields largely equivalent results in our empirical analysis. Moreover, the problem of misreported height and weight (cf. Gorber et al., 2007) is of little relevance to our study since body weight is not self-reported but measured by staff clinic or pharmacy staff.

[13]We include the pre-intervention level BMI_0 as control. Hence technically, our preferred specification is equivalent to including both the pre- and post-intervention level BMI_0 and BMI_1 at the right-hand-side of the regression model.

Table 4: Self-Rated Health and Physical Well-Being by weight loss

		bad	poor	satisfactory	good	very good	marginal distribution*
SRH_1	No Weight Loss	16.48%	21.98%	42.86%	15.38%	3.30%	91
	Weight loss	3.30%	15.23%	41.62%	36.55%	3.30%	394
PWB_1	No Weight Loss	13.95%	32.56%	38.37%	11.63%	3.49%	86
	Weight loss	7.59%	19.11%	37.43%	31.68%	4.19%	382
SRH_0	No Weight Loss	8.79%	30.77%	34.07%	25.27%	1.10%	91
	Weight loss	2.54%	23.60%	45.43%	25.89%	2.54%	394
PWB_0	No Weight Loss	6.98%	38.37%	36.05%	18.60%	0.00%	86
	Weight loss	6.02%	30.89%	40.84%	19.63%	2.62%	382

Notes: Shares and *absolut numbers of individuals

weight over the four month weight loss phase. Yet, the 5 percent quantile is -1.37, pointing to substantial weight gain being not a rare phenomenon in the sample; see Figure 1 for sample distribution of *Weight loss*.

From the first panel of Table 4 one can see that – in a descriptive sense – participants who lost weight are more likely ro report good health. Among this group of the participants around 40% reported good or very good health, while this only holds for around 19% of the participants who gained weight. 38% of the latter reported poor or bad health. In contrast, the corresponding share of participants who lost weight is only 19%. According to a Wilcoxon rank-sum test, the distribution of SRH_1 clearly differs (p-value 0.000) between individuals who lost weight and individuals who did not. The estimated probability for an individual from the former group to be in better health than an individual from the latter is 0.65. These descriptive findings line up with the general pattern of results found in the literature that less body weight is associated with better self-ratings of health. Considering physical well-being instead of self-rated health yields a very similar picture. Again, according to a Wilcoxon rank-sum test, the distribution of PWB_1 clearly differs (p-value 0.000) between individuals who lost weight and individuals who did not.

If the same descriptive analysis is applied to self-rated health measured at the beginning of the weight loss phase, i.e. to SRH_0 instead of SRH_1, we still find a significant (p-value 0.044), though less distinct, deviation in the distribution of self-rated health. At the one hand, this suggests that weight loss might be endogenous and, in turn, calls for an empirical approach that does not interpret the mere correlation as causal effect. On the other hand, this pattern suggests analyzing the effect of weight loss on SRH conditionally on its initial level SRH_0 in order to account for persistent unobserved heterogeneity and to eliminate variation in the dependent variable that cannot be explained by a change in BMI. For this reason, SRH_0 enters the econometric analysis as control variable.

If the same analysis is applied to physical well-being measured at the beginning of the weight loss phase (PWB_0) we do not find a clearly significant difference (p-value 0.159) in the distribution of physical well-being. Yet, the share of respondents who rate their physical well-being at least satisfactory is still higher for respondents who lost

Table 5: Descriptive Statistics for Estimation Sample

	Obs.	Mean	S.D.	Min.	Max.	Median
dependent variables:						
SRH_1	485	3.11	0.92	1	5	3
$PWB_1{}^{\dagger}$	468	2.97	1.00	1	5	3
explanatory variables:						
Weight loss	497	1.56	1.98	-8.28	12.78	1.51
BMI_0	497	37.26	6.16	28.03	60.22	36.03
SRH_0	497	2.97	0.87	1	5	3
female	497	0.33	0.47	0	1	0
age	497	49.16	8.56	20	68	50
instrumental variable:						
incentive	497	0.69	0.46	0	1	1

Notes: The number of observations differs between the main (485, 468) and the auxiliary (497, 487) equation of the econometric model. Descriptive statistics for the explanatory variables are for the auxiliary equation in the model explaining SRH_1. †See Appendix, Table A2 for comprehensive descriptive statistics of the estimation sample with PWB_1 serving as dependent variable.

weight. Hence, analogously to the regression explaining self-rated health we control for PWB_0 when our outcome variable is PWB_1 in order to account for persistent unobserved heterogeneity.[14]

As another approach to account for unobserved heterogeneity, we also control for initial body mass index BMI_0. Though all participants were obese at the time of recruitment, BMI_0 exhibits pronounced heterogeneity ranging from 28 up to 60.[15] The average of the initial *BMI* is 37.26 while the median value of 36.03 is somewhat smaller, indicating that distribution of initial *BMI* is skewed to the right.

Due to the relatively small estimation sample, we abstain from specifying a rich regression model with a large number of controls. As basic socioeconomic characteristics we only control for age and gender.[16] Roughly two-thirds of the individuals in the sample are men. Since earlier studies found gender differences in the relationship between *BMI* and *SRH* (Imai et al., 2008), besides the pooled model, we also conduct separate regression analyses for males and females.

As discussed above, we use exposition to monetary weight loss incentives as instrument for weight change. Though the experiment involved two treatment groups which were offered incentives of different size, in the regression analysis we use a simple dummy that indicates random assignment to one of the treatment groups. Pooling the treatment groups is in line with the finding of Augurzky et al. (2012) that the size of offered monetary reward proved to be immaterial for realized weight loss. Descriptive statistics for all variables that enter the preferred regression model are provided in Table 5.

[14] Both SRH_0 and PWB_0 enter the model in a linear way. Estimating the model with dummy variables for the different categories of *SRH* and *PWB* does not alter our results.

[15] Many individuals already lost weight over the rehab stay. This is the reason for some participants entering the weight loss phase with a *BMI* smaller than 30.

[16] We also estimated models with more explanatory variables (controlling for education, income and employment), however, the results of those models (reported in Table A4 and A5 in the Appendix) are similar to the results of our preferred specification, where the number of observations is higher.

3 Estimation Procedure

In order to take the ordered categorical nature of our dependent variables SRH_1 and PWB_1 into account, the econometric analysis rests on ordered probit models. We start with estimating a conventional specification of this model that regards all regressors as exogenous. Besides the key explanatory variable *Weight loss*, pre-intervention body weight BMI_0, age and gender enter the models at the right-hand-side. Additionally, we control for pre-intervention self-rated health (SRH_0) or pre-intervention physical well-being (PWB_0), depending on the dependent variable that is used. This basic model specification serves as reference.

Yet, as discussed above, results from conventional ordered probit estimation are most likely biased, due to unobserved confounders affecting both subjective health perception and *Weight loss*, as well as reverse causality. To tackle possible endogeneity bias, and to allow for identifying a causal effect of *Weight loss* on subjective health perception, in our preferred empirical model we do not rely on naïve ordered probit estimation, but tap an exogenous source of variation in body weight for identifying the effect under scrutiny. Random assignment to either the control or the treatment arm of the experiment generates weight variation, which by the experimental design is exogenous. Moreover, as shown elsewhere (Augurzky et al., 2012), the incentive treatment was clearly effective and hence induced exogenous variation in *Weight loss*. Technically, the binary indicator *incentive*, which indicates assignment to one of the two incentive groups, serves as instrument for *Weight loss*.[17]

If health was measured on a continuous scale, two-stage least squares would be an obvious choice for the estimation procedure. However, this choice would conflict with the ordered categorical nature of SRH_1 and PWB_1.[18] We, hence, opt for a more parametric approach to instrumental variables estimation. That is, we augment the equation of prime interest by a second equation that specifies the endogenous regressor *Weight loss* as a function of the instrument *incentive* and the covariates that enter the main equation, and assume joint normality of the two error terms. The cross-equation error-correlation, hence, captures possible endogeneity of *Weight loss*. Joint estimation by full-information maximum likelihood (ML) is straightforward for this model.[19]

[17] As an alternative model specification we used two indicators, each indicating membership in one of the two incentive groups, as instrument. This affects the results just marginally.

[18] We also estimated the model by two stage least squares. In this robustness check we reduced the number of categories to just two. In qualitative terms, the results from this less parametric model are largely equivalent to our preferred specification; see Table A3 in the Appendix for detailed results.

[19] We used the user-written Stata command *cmp* (Roodman, 2009) for estimation. It generalizes the familiar full information ML approach to estimating binary probit models with endogenous explanatory variables (cf. Wooldridge, 2002, 472-477). Although, referring to joint ML estimation with exclusion restrictions as 'instrumental variables estimation' is a questionable choice of terminology, we stick to this nomenclature, which is common in the applied empirical literature.

4 Estimation Results

4.1 Results for the Basic Model

In this section, we present and discuss results for the regression models we introduced in the previous section. Table 6 displays the estimation results of the naïve model that does not take possible endogeneity into account. Results for the whole estimation sample (columns 1 and 2) and results for stratified samples of males and females (columns 3 to 6) are displayed.

The results are in line with those of the majority of the related literature. In all specifications we find a statistically significant association between weight change and subjective health perception, where weight loss is positively associated with the inclination to report a better status of subjective health. In terms of magnitude, the estimated coefficient does not differ much between the different specifications.

In the case of self-rated health the point estimate of *Weight loss* for the female sample is slightly higher than the point estimate in the male sample. Yet, this difference is statistically insignificant.[20] The reverse pattern is found for the impact of *Weight loss* on physical well-being. Here in the male sample the point estimate is slightly higher than the point estimate in the female sample. Again, the difference is not statistically significant. In quantitative terms the point estimate of 0.15 (full sample, self-rated health) translates into an increase in the probability of rating one's health 'satisfactory' or better of 3.7 percentage points if one reduces her or his body weight by one *BMI* unit.

Turning to the coefficients of the control variables, BMI_0 is not significantly associated with subjective health perception. Yet, not surprisingly, the coefficients estimated for initial subjective health perception (SRH_0 and PWB_0) are positive and highly significant, revealing pronounced persistence in subjective health perception, which has already been observed in Table 1 and 2. The simple ordered probit regressions indicate a gender differential in subjective health perception, with women exhibiting a less favorable subjective health rating both in terms of *SRH* and *PWB*. The role age plays as determinant of subjective health perception is less clear. While the regression analysis does not yield a significant association between age and *SRH*[21] for the pooled and the women's sample, a marginally significant and inverse association is found for males. We find no significant influence of age on physical well-being.

4.2 Results from IV Estimation

As discussed above, the results presented in Table 6 might suffer from endogeneity bias regarding the coefficient attached to *Weight loss*. In this subsection, we discuss estimates

[20] Estimating a model that includes the interaction of gender and weight-change either does not reveal a distinct gender-pattern in the association of weight-change and subjective health perception.

[21] Including age^2 as additional regressor does not point to a non-linear relationships between *SRH* and age.

Table 6: Naïve Ordered Probit Estimation (coefficient estimates)

	All		Males		Females	
dependent variable	SRH_1	PWB_1	SRH_1	PWB_1	SRH_1	PWB_1
Weight loss	0.150***	0.160***	0.135***	0.174***	0.192***	0.131***
	(0.026)	(0.026)	(0.031)	(0.032)	(0.047)	(0.047)
BMI_0	-0.005	-0.004	-0.016	-0.011	0.009	0.008
	(0.008)	(0.008)	(0.011)	(0.011)	(0.013)	(0.014)
SRH_0	0.598***	–	0.611***	–	0.586***	–
	(0.062)		(0.076)		(0.106)	
PWB_0	–	0.542***	–	0.574***	–	0.467***
		(0.059)		(0.073)		(0.103)
female	-0.362***	-0.339***	–	–	–	–
	(0.107)	(0.107)				
age	-0.009	-0.005	-0.020***	-0.009	0.010	0.001
	(0.006)	(0.006)	(0.007)	(0.007)	(0.010)	(0.010)
No. of observations	485	468	329	319	156	149

Notes: Standard errors in parentheses; * $p \leq 0.1$, ** $p \leq 0.05$, *** $p \leq 0.01$; cut points of the ordered probit estimation are displayed in Table A6 in the Appendix.

that address this issue by the use of an instrumental variable.[22] Table 7 displays coefficients for the model that relies on weight variation induced by randomly assigned weight loss incentives for identifying the coefficient of prime interest. Besides the coefficients of the equation of prime interest (upper panel), estimates for the auxiliary equation explaining *Weight loss* (lower panel) are also displayed. Analogous to what is reported for naïve ordered probit estimation, results for the pooled sample and gender specific estimates are reported.

Starting with the instrumental equation, the second panel in Table 7 indicates that cash incentives have a substantial effect on achieved weight loss. This result, which has already been established in the literature (e.g. Augurzky et al., 2012; Volpp et al., 2008; John et al., 2011; Cawley and Price, 2013; Paloyo et al., 2014), is important for the present analysis, as it points to the experiment generating exogenous variation in *BMI* that can be used for identification. Indeed, the indicator *incentive* proved to be a rather strong instrument for *Weight loss*. For the whole sample, the relevant F-statistic[23] is 27.72 and 28.54, respectively,[24] which clearly exceed the conventional threshold value of 10 (Stock et al., 2002). For the stratified samples, the corresponding F-statistics likewise exceed this threshold. Nevertheless, the strength of the instrument is substantially reduced in the male sample. Due to the small sample size in the female sample and the relative small F-statistic for instrument relevance for males, the IV-results have to be interpreted with some caution in the stratified samples. Besides this key result

[22]From a purely technical perspective, one could argue that instrumental variables are not required for identification that could solely rest on the non-linearity of the model. Yet, this will rarely work in practice. Indeed, in the present application the optimization procedure runs into serious convergence problems if *incentive* is not included as instrument.

[23]It is calculated from estimating the auxiliary equation separately by least squares.

[24]The first stage regressions differ slightly between the models explaining SRH_1 and PWB_1, since either SRH_0 or PWB_0 enters the model at the right-hand-side.

Table 7: Instrumental Variable Ordered Probit Estimation (coefficient estimates)

	All		Males		Females	
dependent variable	SRH_1	PWB_1	SRH_1	PWB_1	SRH_1	PWB_1
Weight loss	0.126	0.072	0.146	0.018	0.038	0.130
	(0.111)	(0.110)	(0.155)	(0.151)	(0.147)	(0.144)
BMI_0	-0.004	0.001	-0.017	-0.001	0.015	0.008
	(0.010)	(0.011)	(0.014)	(0.015)	(0.014)	(0.016)
SRH_0	0.601***	–	0.608***	–	0.569***	–
	(0.062)		(0.085)		(0.111)	
PWB_0	–	0.550***	–	0.580***	–	0.467***
		(0.059)		(0.074)		(0.105)
female	-0.367***	-0.357***	–	–	–	–
	(0.109)	(0.108)				
age	-0.009	-0.007	-0.020***	-0.011	0.007	0.001
	(0.006)	(0.006)	(0.008)	(0.007)	(0.010)	(0.010)
auxiliary equation (dependent variable: *Weight loss*)						
incentive	0.978***	1.004***	0.821***	0.857***	1.402***	1.412***
	(0.185)	(0.187)	(0.226)	(0.226)	(0.318)	(0.332)
BMI_0	0.051***	0.056***	0.061***	0.066***	0.033	0.040*
	(0.014)	(0.014)	(0.018)	(0.019)	(0.021)	(0.022)
SRH_0	0.167*	–	0.232*	–	0.047	–
	(0.099)		(0.126)		(0.154)	
PWB_0	–	0.196**	–	0.221*	–	0.164
		(0.097)		(0.121)		(0.160)
female	-0.299	-0.318*	–	–	–	–
	(0.182)	(0.184)				
age	-0.014	-0.017*	-0.010	-0.013	-0.023	-0.027
	(0.010)	(0.010)	(0.013)	(0.013)	(0.016)	(0.017)
constant	-0.721	-0.820	-1.347	-1.360	0.114	-0.257
	(0.871)	(0.854)	(1.152)	(1.136)	(1.308)	(1.278)
cross-eq. error corr.	0.048	0.173	-0.023	0.306	0.295	0.002
	(0.214)	(0.208)	(0.309)	(0.292)	(0.267)	(0.272)
Instrument relevance (F-statistic)	27.72	28.54	12.96	14.16	18.80	17.47
No. of observations (over all)	497	487	335	329	162	158
No. of observations (main equation)	485	468	329	319	156	149

Notes: Standard errors in parentheses; * $p \leq 0.1$, ** $p \leq 0.05$, *** $p \leq 0.01$; number of observations differ. Due to joint estimation, participants for which SRH_1 and PWB_1, respectively, is not observed but the endogenous regressor *Weight loss* is observed contribute to the log-likelihood function and enter the estimation sample. Cut points of the ordered probit estimation are displayed in Table A6 in the Appendix.

regarding instrument relevance, the estimates for the auxiliary equation indicate that obese women seem to be more responsive to monetary weight loss incentives than men[25] and that those who start with a high initial *BMI* are more likely to lose weight.

Turning to the equation of main interest, we see almost no change in the coefficients of the control variables as compared to simple ordered probit. Yet, with respect to the effect of weight change on subjective health perception, we find a pattern of results that in some respect deviates from its counterpart from naïve estimation. For all

[25]Not surprisingly, this coincides with the result of Augurzky et al. (2012) who use the same data as the present analysis.

six considered variants of the model, the coefficient of *Weight loss* turns statistically insignificant. Moreover the point estimates, except for the regression explaining SRH_1 in the male sample, are smaller compared to their counterparts in Table 6. However, all estimated weight loss coefficients are still positive. Moreover, for several model variants they are of similar magnitude as their counterparts from the naïve model. The lack of statistical significance seems first of all to be a standard error issue, and can hardly be interpreted as evidence for the absence of a weight loss effect on subjective health. Evidently – although the instrument is not weak – augmenting the naïve model by the instrumental equation inflates the noisiness of the estimates substantially.

The finding that instrumenting *Weight loss* does not fundamentally change the estimated key coefficients is mirrored by the estimates of the cross-equation error correlation. With a single exception, the estimate is positive but of moderate or even negligible magnitude and statistically insignificant. Though the positive sign argues in favour of unobserved confounders may play some role in the correlation of *SRH* and *BMI*, the estimates still provide little evidence for endogeneity of *Weight loss* being a major issue. Moreover, based on linear specifications (see next subsection) that employ OLS and 2SLS rather than ordered probit and IV ordered probit, Hausman-Wu tests do not yield evidence for any systematic deviation between instrumental variables and naïve estimation.[26] One possible reason for this pattern is that relying on short-term, within-individual variation cuts or at least weakens already several channels – for instance, health-conscious attitudes, educational and family background, and certain genetic endowments – that are likely to be major sources of endogeneity in cross-sectional data based analyses.

To sum up this discussion, though the instrumental variables estimation does not yield a clear cut result regarding the effect of weight change on subjective health perception, the pattern of results is still telling. While, the point estimates argue for the naïve empirical approach suffering from some upward bias, the rather noisy IV estimates do not put the general finding of weight-gain in obese individuals affecting subjective health perception detrimentally into question. In qualitative terms our results, hence, do not conflict with the bulk of the literature. They are also largely in line with those of Cullinan and Gillespie (2015)[27], who carefully designed their analysis to allow for a causal interpretation of the link between body weight and self-rated health. This appears to be an interesting finding, given that the present analysis exploits a source of variation for identification that is very different from what is used in Cullinan and Gillespie (2015) and that, in consequence, the nature of the estimated effect also differs. While Cullinan and Gillespie (2015) rest on genetics as a persistent and long term determinant of body weight, the present analysis uses short-term extrinsic incentives. Thus, the result of the former can be interpreted such that permanently reducing the *BMI* of an obese individual to a normal level will improve her *SRH* substantially. In

[26] For the majority of specifications the p-value exceeds 0.9.

[27] They identify a strong link between *BMI* and *SRH* in obese individuals, while their instrumental variables estimation results are inconclusive for moderately overweight individuals.

contrast, our results – at least in terms of the point estimates – suggest that also a small reduction in body weight will make an obese individual instantaneously feel healthier. This distinction is important for the question of how to motivate obese individuals to lose weight. Even if substantial weight loss is known to pay-off in the long-run for sure, obese individuals may still need some instantaneous improvement in subjectively perceived health in order to keep the discipline to continue their weight loss efforts.

4.3 Robustness Checks

We ran several robustness checks in order to test how sensitive the results are to changes of the model specification. (i) We estimated all discussed model specifications with additional controls for education, employment, and income measured at rehab discharge; see Tables A4 and A5 in the Appendix. This does not change the overall pattern of results. The coefficients of the naïve model are hardly affected. Their counterparts from instrumental variables estimation remain inconclusive, as they do not change consistently in the same direction. For some specifications (whole sample *SRH*, male sample *SRH*, male sample *PWB*) the point estimate of the weight loss coefficient gets smaller, while it gets bigger for others (whole sample *PWB*, female sample *SRH*). The coefficient of weight loss for the female sample in the physical well-being case even turns significant. (ii) We reduced the number of health categories to just three merging 'good' with 'very good' and 'bad' with 'poor'. This just marginally affects the coefficient estimates; see Table A3 first panel in the Appendix. As another robustness check, (iii) we excluded individuals with extreme changes in *BMI*, in order to check for few extraordinary cases possibly driving the empirical results. We considered two definitions of extreme, (*Weight loss* > 5) and (*Weight loss* < -2 | *Weight loss* > 5). For both, the pattern of results remains largely unchanged; see Table A9 and A10 in the Appendix.

In order not to rely exclusively on fully parametric model, (iv) we also ran two-stage least squares (2SLS) regression. To avoid interpreting *SRH* as being measured on an interval scale, we recoded the dependent variable to have just two categories and, in consequence, we estimated linear probability models by 2SLS and as reference also by OLS. Since transforming five-category *SRH* to a binary health indicator involves the somewhat arbitrary choice of a cutoff category, we tried all four possible variants; see Table A3 second to fifth panel and Figures A1 to A6 in the Appendix. If the left-hand side variable is specified to indicate one of the two extreme categories ('better than good' or 'bad') the effect of *Weight loss* gets very small or even vanishes. Interestingly, this holds for both OLS and 2SLS. One possible explanation is that these categories are rather rare in the sample, hampering the identification of effects on these extreme categories; see Table 1. An alternative, less technical, explanation is that relatively small changes in body weight will rarely be the reason for a shift in subjective health perception to an extreme. If an interior category is chosen as cutoff, the linear model insofar mirrors the results of ordered probit estimation as the naïve estimator yields a significant and favorable effect of weight loss on *SRH*, while IV does not. For the variant with the left-hand side variable indicating poor or bad *SRH*, the 2SLS coefficient even turns negative. This does not apply to the variant with an indicator for '*SRH* neither good nor very good' serving as dependent variable. There, the 2SLS coefficients are similar or larger than their OLS counterparts.

Since due to drop-out the estimation sample is considerably smaller than the initial 695 recruited participants, non-random sample attrition might bias our results. (v) To address the concern, we estimated model specifications that include a third equation explaining experiment attrition that is jointly estimated with the remaining two. As an 'instrument' for attrition this additional equation includes a dummy variable 'pharmacy in town' which indicates whether the assigned pharmacy for the weigh-in lies in the same zip-code as the respondent's place of residence. The results from these specifications suggest, that non-random sample attrition is likely to be no issue; see Table A7 and A8 in the Appendix.[28]

As displayed in Table 1, initial body weight varies substantially in our sample. To check whether this heterogeneity is mirrored by heterogeneity in the effect of weight loss on self-perceived health, (vi) we stratified the analysis by initial *Weight loss*. Figure A7 in the Appendix displays the respective estimated coefficients[29] of *Weight loss*, which do not reveal a striking pattern of heterogeneity with respect to BMI_0.

5 Discussion and Conclusion

This paper analyzes the relationship between moderate weight loss and subjective health perception in obese individuals. We confirm the results of the related literature, which find a significant association of body weight and subjective health perception. Unlike the bulk of the existing literature, the present analysis is not only concerned with the association of body weight and self-rated health, but employs instrumental variables estimation to establish a causal link. In doing this, it follows Cullinan and Gillespie (2015) who also use instrumental variables estimation to identify a causal effect of *BMI* on *SRH*. Our analysis differs from this key reference, by tapping a completely different source of exogenous weight variation. While Cullinan and Gillespie (2015) use *BMI* of biological relatives as instrument and rest on exogenous genetically determined, long-term, between-individuals weight variation for identification, we use cash incentives of a weight loss intervention as instrument and, hence, rely on short-term, within-individual variation. Though, our instrumental variable approach does not establish statistical significance of the effect under scrutiny, the pattern of results suggests that the positive association of subjective health perception and weight loss is not primarily due to unobserved confounders. Our results are hence largely in line with the finding of Cullinan and Gillespie (2015). It, nevertheless, adds a relevant aspect to the insights into how weight loss affects subjective health perception in obese individuals. While

[28] Augurzky et al. (2012) who use the same data as the present analysis also found, that selection-bias does not drive their results.

[29] We use our preferred set of right-hand-side variables and simple OLS as estimation method.

Cullinan and Gillespie (2015) establish that obese individuals' health perception will improve if they manage to become normal-weight, our analysis yields some evidence for subjective health improvements accompanying even small initial weight reduction. This finding may encourage obese individuals in their weight loss attempts, since they can expect to be immediately rewarded for their efforts by subjective health improvements.

References

Ali, M. M., Rizzo, J. A., Amialchuk, A., and Heiland, F. (2014). Racial differences in the influence of female adolescents' body size on dating and sex. *Economics and Human Biology*, 12:140–152.

Augurzky, B., Bauer, T. K., Reichert, A. R., Schmidt, C. M., and Tauchmann, H. (2012). Does money burn fat? Evidence from a randomized experiment. *Ruhr Economic Papers*, 368.

Augurzky, B., Bauer, T. K., Reichert, A. R., Schmidt, C. M., and Tauchmann, H. (2014). Small cash rewards for big losers - Experimental insights into the fight against the obesity epidemic. *Ruhr Economic Papers*, 530.

Baruth, M., Becofsky, K., Wilcox, S., and Goodrich, K. (2014). Health characteristics and health behaviors of African American adults according to self-rated health status. *Ethnicity & Disease*, 24(1):97–103.

Blackburn, G. (1995). Effect of degree of weight loss on health benefits. *Obesity Research*, 3(S2):211s–216s.

Calle, E. E., Rodriguez, C., Walker-Thurmond, K., and Thun, M. J. (2003). Overweight, obesity, and mortality from cancer in a prospectively studied cohort of us adults. *New England Journal of Medicine*, 348(17):1625–1638.

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

Cawley, J. and Price, J. A. (2013). A case study of a workplace wellness program that offers financial incentives for weight loss. *Journal of Health Economics*, 32(5):794–803.

Cullinan, J. and Gillespie, P. (2015). Does overweight and obesity impact on self-rated health? Evidence using instrumental variables ordered probit models. *Health Economics*, 25(10):1341–1348.

Darviri, C., Fouka, G., Gnardellis, C., Artemiadis, A. K., Tigani, X., and Alexopoulos, E. C. (2012). Determinants of self-rated health in a representative sample of a rural population: a cross-sectional study in Greece. *International Journal of Environmental Research and Public Health*, 9(3):943–954.

Di Angelantonio, E. (2016). Body-mass index and all-cause mortality: Individual-participant-data meta-analysis of 239 prospective studies in four continents. *The Lancet*, 388:776–786.

Elks, C., Den Hoed, M., Zhao, J. H., Sharp, S., Wareham, N., Loos, R., and Ong, K. (2012). Variability in the heritability of body mass index: A systematic review and meta-regression. *Frontiers in Endocrinology*, 3:29.

Ferraro, K. F. and Yu, Y. (1995). Body weight and self-ratings of health. *Journal of Health and Social Behavior*, 36(3):274–284.

Fletcher, J. M. and Lehrer, S. F. (2011). Genetic lotteries within families. *Journal of Health Economics*, 30(4):647–659.

Fontaine, K. and Barofsky, I. (2001). Obesity and health-related quality of life. *Obesity Reviews*, 2(3):173–182.

Frange, C., de Queiroz, S. S., da Silva Prado, J. M., Tufik, S., and de Mello, M. T. (2014). The impact of sleep duration on self-rated health. *Sleep Science*, 7(2):107–113.

Garn, S., Leonard, W., and Hawthorne, V. (1986). Three limitations of the body mass index. *American Journal of Clinical Nutrition*, 44(6):996–997.

Gorber, S. C., Tremblay, M., Moher, D., and Gorber, B. (2007). A comparison of direct vs. self-report measures for assessing height, weight and body mass index: a systematic review. *Obesity Reviews*, 8(4):307–326.

Guallar-Castillón, P., López, G. E., Lozano, P. L., Gutiérrez-Fisac, J., Banegas, B. J., Lafuente, U. P., and Rodriguez, A. F. (2002). The relationship of overweight and obesity with subjective health and use of health-care services among Spanish women. *International Journal of Obesity and related Metabolic Disorders: Journal of the International Association for the Study of Obesity*, 26(2):247–252.

Hubert, H. B., Feinleib, M., McNamara, P. M., and Castelli, W. P. (1983). Obesity as an independent risk factor for cardiovascular disease: a 26-year follow-up of participants in the framingham heart study. *Circulation*, 67(5):968–977.

Imai, K., Gregg, E. W., Chen, Y. J., Zhang, P., Rekeneire, N., and Williamson, D. F. (2008). The association of BMI with functional status and self-rated health in US adults. *Obesity*, 16(2):402–408.

John, L., Loewenstein, G., Troxel, A., Norton, L., Fassbender, J., and Volpp, K. (2011). Financial incentives for extended weight loss: A randomized, controlled trial. *Journal of General Internal Medicine*, 26:621–626.

Kepka, D., Ayala, G. X., and Cherrington, A. (2007). Do Latino immigrants link self-rated health with BMI and health behaviors? *American Journal of Health Behavior*, 31(5):535–544.

Kline, B. and Tobias, J. L. (2008). The wages of BMI: Bayesian analysis of a skewed treatment–response model with nonparametric endogeneity. *Journal of Applied Econometrics*, 23(6):767–793.

Kolotkin, R., Meter, K., and Williams, G. (2001). Quality of life and obesity. *Obesity Reviews*, 2(4):219–229.

Kroes, M., Osei-Assibey, G., Baker-Searle, R., and Huang, J. (2016). Impact of weight change on quality of life in adults with overweight/obesity in the United States: a systematic review. *Current Medical Research and Opinion*, 32(3):485–508.

Loveman, E., Frampton, G., Shepherd, J., Picot, J., Cooper, K., Bryant, J., Welch, K., and Clegg, A. (2011). The clinical effectiveness and cost-effectiveness of long-term weight management schemes for adults: a systematic review. *Health Technology Assessment*, 15(2).

Macmillan, R., Duke, N., Oakes, J. M., and Liao, W. (2011). Trends in the association of obesity and self-reported overall health in 30 years of the integrated health interview series. *Obesity*, 19(5):1103–1105.

Maes, H. H. M., Neale, M. C., and Eaves, L. J. (1997). Genetic and environmental factors in relative body weight and human adiposity. *Behavior Genetics*, 27(4):325–351.

Marques-Vidal, P., Ravasco, P., and Paccaud, F. (2012). Differing trends in the association between obesity and self-reported health in Portugal and Switzerland. Data from national health surveys 1992–2007. *BMC Public Health*, 12(588):1–8.

Mokdad, A. H., Ford, E. S., Bowman, B. A., Dietz, W. H., Vinicor, F., Bales, V. S., and Marks, J. S. (2003). Prevalence of obesity, diabetes, and obesity-related health risk factors, 2001. *JAMA*, 289(1):76–79.

Molarius, A., Berglund, K., Eriksson, C., Lambe, M., Nordström, E., Eriksson, H. G., and Feldman, I. (2007). Socioeconomic conditions, lifestyle factors, and self-rated health among men and women in Sweden. *The European Journal of Public Health*, 17(2):125–133.

Ng, M., Fleming, T., Robinson, M., Thomson, B., Graetz, N., Margono, C., Mullany, E. C., Biryukov, S., Abbafati, C., Abera, S. F., et al. (2014). Global, regional, and national prevalence of overweight and obesity in children and adults during 1980–2013: a systematic analysis for the global burden of disease study 2013. *The Lancet*, 384(9945):766–781.

Norton, E. C. and Han, E. (2008). Genetic information, obesity, and labor market outcomes. *Health Economics*, 17(9):1089–1104.

Okosun, I. S., Choi, S., Matamoros, T., and Dever, G. A. (2001). Obesity is associated with reduced self-rated general health status: evidence from a representative sample of white, black, and Hispanic Americans. *Preventive Medicine*, 32(5):429–436.

Paloyo, A., Reichert, A. R., Reinermann, H., and Tauchmann, H. (2014). The causal link between financial incentives and weight loss: An evidence-based survey of the literature. *Journal of Economic Surveys*, 28(3):401–420.

Patel, S. R. and Hu, F. B. (2008). Short sleep duration and weight gain: a systematic review. *Obesity*, 16(3):643–653.

Phillips, L. J., Hammock, R. L., and Blanton, J. M. (2005). Predictors of self-rated health status among Texas residents. *Preventing Chronic Disease*, 2(4).

Prosper, M.-H., Moczulski, V. L., and Qureshi, A. (2009). Obesity as a predictor of self-rated health. *American Journal of Health Behavior*, 33(3):319–329.

Reichert, A. R. (2015). Obesity, weight loss, and employment prospects – evidence from a randomized trial. *Journal of Human Resources*, 50(3):759–810.

Reichert, A. R., Tauchmann, H., and Wübker, A. (2015). Weight loss and sexual activity in adult obese individuals: Establishing a causal link. *Ruhr Economic Papers*, 561.

Roodman, D. (2009). Estimating fully observed recursive mixed-process models with cmp. *Stata Journal*, 11:159–206.

Sabia, J. J. and Rees, D. I. (2011). The effect of body weight on adolescent sexual activity. *Health Economics*, 20(11):1330–1348.

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.

Stock, J. H., Wright, J. H., and Yogo, M. (2002). A survey of weak instruments and weak identification in generalized method of moments. *Journal of Business & Economic Statistics*, 20(4):518–529.

Tompson, T., Benz, J., Agiesta, J., Brewer, K., Bye, L., Reimer, R., and Junius, D. (2012). Obesity in the United States: public perceptions. *The Food Industry*, 53(26):21.

Ul-Haq, Z., Mackay, D. F., Fenwick, E., and Pell, J. P. (2013a). Meta-analysis of the association between body mass index and health-related quality of life among adults, assessed by the SF-36. *Obesity*, 21(3):E322–E327.

Ul-Haq, Z., Mackay, D. F., Martin, D., Smith, D. J., Gill, J. M., Nicholl, B. I., Cullen, B., Evans, J., Roberts, B., Deary, I. J., et al. (2013b). Heaviness, health and happiness: a cross-sectional study of 163 066 UK biobank participants. *Journal of Epidemiology and Community Health*, 68:340–348.

Vogler, G., Sørensen, T., Stunkard, A., Srinivasan, M., and Rao, D. (1995). Influences of genes and shared family environment on adult body mass index assessed in an adoption study by a comprehensive path model. *International Jornal of Obesesity and Related Metablic Disorders*, 19(1):193–197.

Volpp, K. G., John, L. K., Troxel, A. B., Norton, L., Fassbender, J., and Loewenstein, G. (2008). Financial incentive-based approaches for weight loss. *JAMA: The Journal of the American Medical Association*, 300(22):2631–2637.

von Hinke, S., Smith, G. D., Lawlor, D. A., Propper, C., and Windmeijer, F. (2016). Genetic markers as instrumental variables. *Journal of Health Economics*, 45:131–148.

Ware, J. E., Snow, K. K., Kosinski, M., and Gandek, B. (1993). *SF-36 Health Survey Manual and Interpretation Guide*. The Health Institute, New England Medical Center, Boston Massachusetts.

WHO (2000). Obesity: preventing and managing the global epidemic. Technical Report 894.

WHO (2016). Obesity and overweight, Fact Sheet 311. http://www.who.int/mediacentre/factsheets/fs311/en/.

Willage, B. (2017). The effect of weight on mental health: New evidence using genetic IVs. unpublished, https://docs.wixstatic.com/ugd/56beed_f3527ba699e64784a2c32d4a09d08dcd.pdf.

Wing, R. R., Lang, W., Wadden, T. A., Safford, M., Knowler, W. C., Bertoni, A. G., Hill, J. O., Brancati, F. L., Peters, A., Wagenknecht, L., et al. (2011). Benefits of modest weight loss in improving cardiovascular risk factors in overweight and obese individuals with type 2 diabetes. *Diabetes Care*, 34(7):1481–1486.

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

Zellner, D. A., Loaiza, S., Gonzalez, Z., Pita, J., Morales, J., Pecora, D., and Wolf, A. (2006). Food selection changes under stress. *Physiology & Behavior*, 87(4):789–793.

A Appendix

A.1 Tables

Table A1: Joint and Marginal Distribution of PWB_0 and SRH_0

		SRH_0					
		bad	poor	satisfactory	good	very good	marginal distribution
PWB_0	bad	12	14	3	1	0	30
	poor	6	73	55	14	1	149
	satisfactory	0	27	127	42	1	197
	good	0	3	25	63	3	94
	very good	0	0	1	3	6	10
marginal distribution		18	117	211	123	11	480

Table A2: Descriptive Statistics for Estimation Sample (PWB_1 as dep. var.)

	Obs.	**Mean**	**S.D.**	**Min.**	**Max.**	**Median**
dependent variable:						
PWB_1	468	2.97	1.00	1	5	3
explanatory variables:						
Weight loss	487	1.57	1.99	-8.28	12.78	1.52
BMI_0	487	37.26	6.15	28.03	60.22	36.05
PWB_0	487	2.79	0.89	1	5	3
female	487	0.32	0.47	0	1	0
age	487	49.05	8.50	20	68	50
instrumental variable:						
incentive	487	0.70	0.46	0	1	1

Notes: The number of observations differs between the main (468) and the auxiliary (487) equation of the econometric model. Descriptive statistics for the explanatory variables are for the auxiliary equation.

Table A3: Estimated coefficients of *Weight loss* for different specifications of the dependent variable

	All		Males		Females	
dependent variable	SRH_1	PWB_1	SRH_1	PWB_1	SRH_1	PWB_1
Trinary						
Ordered Probit	0.172***	0.174***	0.162***	0.187***	0.200***	0.151***
	(0.028)	(0.029)	(0.034)	(0.035)	(0.050)	(0.050)
IV-Ordered Probit	0.110	0.047	0.166	0.007	-0.024	0.086
	(0.120)	(0.119)	(0.167)	(0.162)	(0.148)	(0.156)
instrument relevance	27.38	28.05	12.80	13.84	18.69	17.24
No. of observations (over all)	497	487	335	329	162	158
No. of observations (main equation)	485	468	329	319	156	149
Binary$_{\text{bad}}$						
OLS	0.014***	0.015**	0.011**	0.015**	0.022*	0.013
	(0.005)	(0.007)	(0.005)	(0.006)	(0.013)	(0.017)
2 SLS	0.030	0.014	0.015	-0.001	0.045	0.032
	(0.024)	(0.027)	(0.028)	(0.034)	(0.038)	(0.045)
instrument relevance	28.7	28.82	12.96	14.25	20.05	17.23
Binary$_{\text{poor}}$						
OLS	0.040***	0.045***	0.038***	0.047***	0.050**	0.042*
	(0.009)	(0.010)	(0.009)	(0.010)	(0.020)	(0.023)
2 SLS	-0.026	-0.015	-0.012	-0.007	-0.059	-0.031
	(0.039)	(0.043)	(0.050)	(0.059)	(0.056)	(0.062)
instrument relevance	27.42	28.18	13.09	13.81	18.73	17.24
Binary$_{\text{satisfactory}}$						
OLS	0.055***	0.059***	0.048***	0.061***	0.071***	0.055***
	(0.010)	(0.010)	(0.012)	(0.013)	(0.017)	(0.016)
2 SLS	0.083*	0.042	0.084	0.002	0.054	0.091*
	(0.044)	(0.043)	(0.064)	(0.064)	(0.052)	(0.049)
instrument relevance	27.10	27.65	12.38	13.55	18.80	17.19
Binary$_{\text{good}}$						
OLS	0.004	0.014**	0.004	0.020**	0.003	0.002
	(0.004)	(0.007)	(0.006)	(0.009)	(0.003)	(0.002)
2 SLS	-0.008	0.001	-0.008	0.003	-0.012	-0.002
	(0.019)	(0.020)	(0.029)	(0.031)	(0.019)	(0.021)
instrument relevance	26.91	27.92	12.32	13.49	18.59	17.40
No. of observations	485	468	329	319	156	149

Notes: Trinary specification of *SRH* and *PWB*: standard errors in parentheses; *SRH* / *PWB* (trinary) equals 1, for poor and bad *SRH* / *PWB*, 2 for satisfactory *SRH* / *PWB* and 3 for good and very good *SRH* / *PWB*; Binary specifications of *SRH* / *PWB*: Robust standard errors in parantheses; binary$_{\text{bad}}$ equals 0, for bad *SRH* / *PWB*, 1 for poor, satifactory, good and very good *SRH* / *PWB*; binary$_{\text{poor}}$ equals 0, for bad and poor *SRH* / *PWB*, 1 for satisfactory, good and very good *SRH* / *PWB*. binary$_{\text{satisfactory}}$ equals 0, for bad, poor and satisfactory *SRH* / *PWB*, 1 for good and very good *SRH* / *PWB*. binary$_{\text{good}}$ equals 0, for bad, poor, satisfactory and good *SRH* / *PWB*, 1 for very good *SRH* / *PWB*. OLS stands for ordinary least squares, 2 SLS stands for two-stage least squares. * $p \leq 0.1$, ** $p \leq 0.05$, *** $p \leq 0.01$;

Table A4: Estimated Coefficients of Ordered Probit Model with additional controls

	All		Males		Females	
dependent variable	SRH_1	PWB_1	SRH_1	PWB_1	SRH_1	PWB_1
Weight loss	0.166***	0.163***	0.155***	0.178***	0.197***	0.118**
	(0.030)	(0.030)	(0.036)	(0.036)	(0.055)	(0.054)
BMI_0	-0.010	-0.009	-0.024**	-0.014	0.014	0.003
	(0.009)	(0.010)	(0.012)	(0.012)	(0.016)	(0.017)
SRH_0	0.636***	–	0.670***	–	0.602***	–
	(0.069)		(0.087)		(0.119)	
PWB_0	–	0.526***	–	0.562***	–	0.483***
		(0.065)		(0.080)		(0.116)
female	-0.345***	-0.304**	–	–	–	–
	(0.120)	(0.120)				
age	-0.010	-0.003	-0.018**	-0.005	0.006	0.002
	(0.007)	(0.007)	(0.008)	(0.008)	(0.012)	(0.012)
no secondary	-1.044**	-0.494	-1.366**	-0.491	-0.660	-0.532
school education	(0.485)	(0.462)	(0.590)	(0.577)	(0.929)	(0.824)
lower secondary	-0.573*	-0.368	-0.901**	-0.392	0.137	-0.222
education	(0.300)	(0.292)	(0.379)	(0.364)	(0.510)	(0.500)
medium secondary	-0.556*	-0.287	-0.832**	-0.314	0.065	-0.211
education	(0.311)	(0.304)	(0.392)	(0.378)	(0.539)	(0.536)
higher secondary	-0.779**	-0.247	-1.039**	-0.327	-0.407	0.040
education	(0.351)	(0.344)	(0.432)	(0.417)	(0.627)	(0.641)
$employment_0$	0.071	-0.001	0.126	-0.135	-0.014	0.269
	(0.160)	(0.162)	(0.210)	(0.206)	(0.251)	(0.270)
low $income_0$	-0.151	-0.299*	-0.101	-0.171	-0.239	-0.678**
	(0.172)	(0.171)	(0.209)	(0.207)	(0.345)	(0.342)
medium $income_0$	-0.037	-0.113	-0.020	-0.103	0.075	-0.305
	(0.143)	(0.141)	(0.163)	(0.161)	(0.322)	(0.318)
No. of observations	412	397	281	273	131	124

Notes: Standard errors in parentheses; * $p \leq 0.1$, ** $p \leq 0.05$, *** $p \leq 0.01$; for education/income, individuals with university degree/high income are the reference category respectively.

Table A5: Estimated Coefficients of Instrumental Variable Ordered Probit Model with additional controls

	All		Males		Females	
dependent variable	SRH_1	PWB_1	SRH_1	PWB_1	SRH_1	PWB_1
Weight loss	0.082	0.114	0.035	0.017	0.078	0.250*
	(0.124)	(0.119)	(0.177)	(0.167)	(0.174)	(0.150)
BMI_0	-0.006	-0.006	-0.018	-0.006	0.020	-0.005
	(0.011)	(0.011)	(0.015)	(0.014)	(0.017)	(0.019)
SRH_0	0.640***	–	0.671***	–	0.603***	–
	(0.069)		(0.090)		(0.120)	
PWB_0	–	0.530***	–	0.559***	–	0.443***
		(0.065)		(0.084)		(0.132)
female	-0.350***	-0.308**	–	–	–	–
	(0.120)	(0.120)				
age	-0.012	-0.005	-0.020**	-0.008	0.003	0.007
	(0.007)	(0.007)	(0.009)	(0.009)	(0.013)	(0.013)
no secondary school education	-1.118**	-0.536	-1.641**	-0.877	-0.337	-0.837
	(0.491)	(0.471)	(0.676)	(0.673)	(1.031)	(0.877)
lower secondary education	-0.551*	-0.357	-0.917**	-0.424	0.247	-0.368
	(0.302)	(0.293)	(0.376)	(0.361)	(0.527)	(0.521)
medium secondary education	-0.535*	-0.277	-0.872**	-0.379	0.201	-0.398
	(0.312)	(0.305)	(0.390)	(0.378)	(0.565)	(0.570)
higher secondary education	-0.808**	-0.267	-1.149***	-0.489	-0.292	-0.122
	(0.351)	(0.347)	(0.445)	(0.439)	(0.646)	(0.663)
$employment_0$	0.074	0.004	0.105	-0.123	0.018	0.253
	(0.159)	(0.163)	(0.212)	(0.206)	(0.254)	(0.270)
low $income_0$	-0.146	-0.297*	-0.076	-0.117	-0.250	-0.604*
	(0.171)	(0.171)	(0.211)	(0.214)	(0.344)	(0.357)
medium $income_0$	-0.026	-0.105	0.010	-0.058	0.065	-0.289
	(0.143)	(0.142)	(0.168)	(0.167)	(0.321)	(0.318)
auxiliary equation (dependent variable: *Weight loss*)						
incentive	0.982***	1.017***	0.808***	0.848***	1.358***	1.368***
	(0.193)	(0.197)	(0.232)	(0.234)	(0.343)	(0.365)
BMI_0	0.047***	0.051***	0.047**	0.050**	0.043*	0.052**
	(0.015)	(0.016)	(0.019)	(0.020)	(0.023)	(0.025)
SRH_0	0.132	–	0.139	–	0.101	–
	(0.106)		(0.133)		(0.168)	
PWB_0	–	0.149	–	0.148	–	0.203
		(0.103)		(0.126)		(0.174)
female	-0.167	-0.152	–	–	–	–
	(0.195)	(0.198)				
age	-0.019*	-0.021*	-0.019	-0.018	-0.025	-0.030
	(0.011)	(0.011)	(0.014)	(0.014)	(0.018)	(0.019)
no secondary school education	-0.860	-0.854	-2.405**	-2.424**	2.633*	1.850
	(0.787)	(0.771)	(0.955)	(0.949)	(1.399)	(1.295)
lower secondary education	0.300	0.226	-0.150	-0.122	0.794	0.615
	(0.470)	(0.489)	(0.602)	(0.603)	(0.723)	(0.817)
medium secondary education	0.257	0.209	-0.318	-0.293	0.828	0.737
	(0.489)	(0.510)	(0.625)	(0.627)	(0.778)	(0.879)
higher secondary education	-0.355	-0.397	-0.957	-0.925	0.701	0.641
	(0.559)	(0.576)	(0.695)	(0.693)	(0.923)	(1.020)
$employment_0$	0.132	0.194	-0.046	0.121	0.350	0.222
	(0.263)	(0.270)	(0.347)	(0.347)	(0.381)	(0.417)
low $income_0$	-0.113	-0.128	0.094	0.182	-0.525	-0.761
	(0.283)	(0.287)	(0.344)	(0.346)	(0.535)	(0.551)
medium $income_0$	0.054	0.084	0.223	0.235	-0.428	-0.398
	(0.235)	(0.237)	(0.269)	(0.269)	(0.500)	(0.518)
constant	-0.586	-0.703	0.035	-0.331	-0.816	-0.871
	(1.053)	(1.056)	(1.324)	(1.343)	(1.741)	(1.732)
cross-eq. error corr.	0.159	0.093	0.226	0.299	0.218	-0.270
	(0.225)	(0.218)	(0.324)	(0.306)	(0.300)	(0.310)
Instrument relevance (F-statistic)	24.96	25.72	11.59	12.6	14.24	12.74
No. of observations (over all)	421	411	286	281	135	130

Notes: Standard errors in parentheses; * $p \leq 0.1$, ** $p \leq 0.05$, *** $p \leq 0.01$; for education/income, individuals with university degree/high income are the reference category respectively.

Table A6: Cut Points of Naïve Ordered Probit and Instrumental Variable Ordered Probit Estimation (coefficient estimates)

	All		Males		Females	
dependent variable	SRH_1	PWB_1	SRH_1	PWB_1	SRH_1	PWB_1
	Naive Ordered Probit Estimation					
cut 1: bad to poor	-0.626	-0.363	-1.517**	-0.719	1.136	0.513
	(0.501)	(0.491)	(0.647)	(0.635)	(0.818)	(0.782)
cut 2: poor to satisfactory	0.360	0.625	-0.619	0.299	2.253***	1.466*
	(0.500)	(0.491)	(0.644)	(0.634)	(0.824)	(0.784)
cut 3: satisfactory to good	1.715***	1.797***	0.806	1.469**	3.515***	2.654***
	(0.504)	(0.494)	(0.644)	(0.636)	(0.844)	(0.799)
cut 4: good to very good	3.435***	3.304***	2.505***	2.988***	5.420***	4.161***
	(0.518)	(0.508)	(0.653)	(0.649)	(0.911)	(0.852)
No. of observations	485	468	329	319	156	149
	Instrumental Variable Ordered Probit Estimation					
cut 1: bad to poor	-0.621	-0.335	-1.525**	-0.588	1.041	0.513
	(0.501)	(0.491)	(0.653)	(0.648)	(0.822)	(0.782)
cut 2: poor to satisfactory	0.365	0.640	-0.627	0.386	2.116**	1.466*
	(0.500)	(0.490)	(0.652)	(0.633)	(0.844)	(0.785)
cut 3: satisfactory to good	1.718***	1.796***	0.797	1.506**	3.333***	2.654***
	(0.504)	(0.494)	(0.653)	(0.632)	(0.886)	(0.799)
cut 4: good to very good	3.437***	3.282***	2.496***	2.960***	5.171***	4.161***
	(0.517)	(0.514)	(0.664)	(0.661)	(0.992)	(0.852)
No. of observations (over all)	497	487	335	329	162	158
No. of observations (main equation)	485	468	329	319	156	149

Notes: Standard errors in parentheses; * $p \leq 0.1$, ** $p \leq 0.05$, *** $p \leq 0.01$

Table A7: Naïve Ordered Probit Estimation, controlling for selection (coefficient estimates)

	All		Males		Females	
dependent variable	SRH_1	PWB_1	SRH_1	PWB_1	SRH_1	PWB_1
Weight loss	0.145***	0.142***	0.122***	0.172***	0.192***	0.122***
	(0.029)	(0.027)	(0.033)	(0.032)	(0.047)	(0.046)
BMI_0	-0.002	-0.010	-0.009	-0.010	0.009	0.002
	(0.009)	(0.008)	(0.011)	(0.011)	(0.015)	(0.014)
SRH_0	0.563***	–	0.543***	–	0.587***	–
	(0.095)		(0.104)		(0.119)	
PWB_0	–	0.513***	–	0.561***	–	0.475***
		(0.065)		(0.079)		(0.101)
female	-0.359***	-0.253**	–	–	–	–
	(0.107)	(0.111)				
age	-0.012*	0.002	-0.024***	-0.011	0.010	0.007
	(0.007)	(0.006)	(0.007)	(0.008)	(0.014)	(0.010)
No. of observations	665	650	453	447	212	203

Notes: Standard errors in parentheses; * $p \leq 0.1$, ** $p \leq 0.05$, *** $p \leq 0.01$

Table A8: Instrumental Variable Ordered Probit Estimation, controlling for selection (coefficient estimates)

	All		Males		Females	
dependent variable	SRH_1	PWB_1	SRH_1	PWB_1	SRH_1	PWB_1
Weight loss	0.061	0.070	0.043	-0.049	0.018	0.209
	(0.126)	(0.131)	(0.139)	(0.134)	(0.198)	(0.131)
BMI_0	0.002	-0.001	-0.003	0.005	0.017	-0.003
	(0.012)	(0.015)	(0.014)	(0.014)	(0.016)	(0.016)
SRH_0	0.575***	–	0.527***	–	0.556***	–
	(0.088)		(0.094)		(0.151)	
PWB_0	–	0.552***	–	0.557***	–	0.456***
		(0.061)		(0.085)		(0.109)
female	-0.377***	-0.331**	–	–	–	–
	(0.106)	(0.142)				
age	-0.013*	-0.004	-0.025***	-0.013	0.006	0.009
	(0.007)	(0.010)	(0.007)	(0.008)	(0.014)	(0.011)
No. of observations	665	650	453	447	212	203

Notes: Standard errors in parentheses; * $p \leq 0.1$, ** $p \leq 0.05$, *** $p \leq 0.01$

Table A9: Naïve Ordered Probit Estimation, excluding extreme cases (coefficient estimates)

	Whole Sample		*Weight loss* ≤ 5		$-2 \leq$ *Weight loss* ≤ 5	
dependent variable	SRH_1	PWB_1	SRH_1	PWB_1	SRH_1	PWB_1
Weight loss	0.150***	0.160***	0.179***	0.166***	0.204***	0.215***
	(0.026)	(0.026)	(0.031)	(0.031)	(0.036)	(0.037)
BMI_0	-0.005	-0.004	-0.001	-0.003	-0.003	-0.005
	(0.008)	(0.008)	(0.009)	(0.009)	(0.009)	(0.009)
SRH_0	0.598***	–	0.594***	–	0.590***	–
	(0.062)		(0.063)		(0.064)	
PWB_0	–	0.542***	–	0.558***	–	0.554***
		(0.059)		(0.060)		(0.062)
female	-0.362***	-0.339***	-0.380***	-0.317***	-0.375***	-0.327***
	(0.107)	(0.107)	(0.109)	(0.109)	(0.110)	(0.111)
age	-0.009	-0.005	-0.008	-0.005	-0.006	-0.003
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
No. of observations	485	468	464	448	449	433

Notes: Standard errors in parentheses; * $p \leq 0.1$, ** $p \leq 0.05$, *** $p \leq 0.01$

Table A10: Instrumental Variable Ordered Probit Estimation, excluding extreme cases (coefficient estimates)

	Whole Sample		*Weight loss* ≤ 5		$-2 \leq$ *Weight loss* ≤ 5	
dependent variable	SRH_1	PWB_1	SRH_1	PWB_1	SRH_1	PWB_1
Weight loss	0.126	0.072	0.117	0.037	0.143	0.073
	(0.111)	(0.110)	(0.132)	(0.125)	(0.169)	(0.161)
BMI_0	-0.004	0.001	0.000	0.001	-0.002	-0.001
	(0.010)	(0.011)	(0.009)	(0.009)	(0.010)	(0.010)
SRH_0	0.601***	–	0.600***	–	0.597***	–
	(0.062)		(0.063)		(0.065)	
PWB_0	–	0.550***	–	0.569***	–	0.570***
		(0.059)		(0.060)		(0.062)
female	-0.367***	-0.357***	-0.376***	-0.314***	-0.372***	-0.318***
	(0.109)	(0.108)	(0.109)	(0.109)	(0.111)	(0.112)
age	-0.009	-0.007	-0.008	-0.006	-0.007	-0.005
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
No. of obs. (over all)	497	487	476	467	461	452

Notes: Standard errors in parentheses; * $p \leq 0.1$, ** $p \leq 0.05$, *** $p \leq 0.01$

A.2 Figures

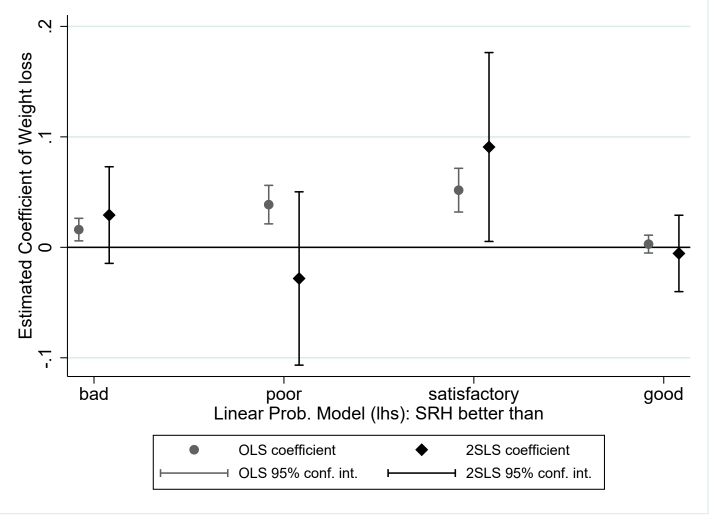

Figure A1: Est. Coef. for Linear Prob. Model (**Self-Rated Health**, **full sample**)

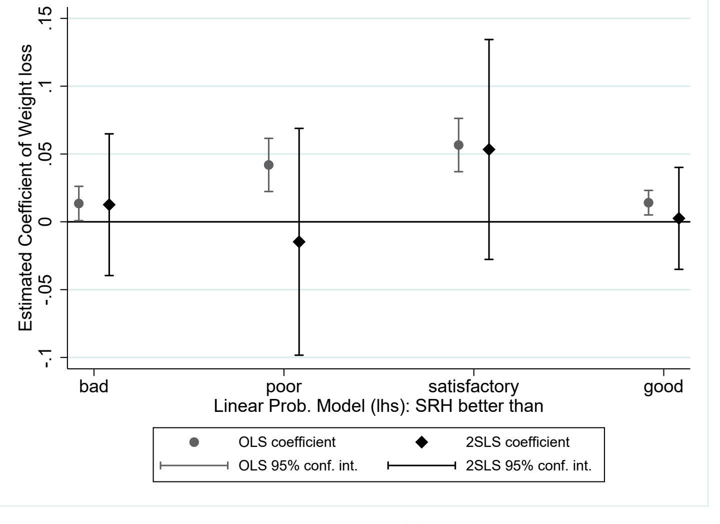

Figure A2: Est. Coef. for Linear Prob. Model (**Physical Well-Being**, **full sample**)

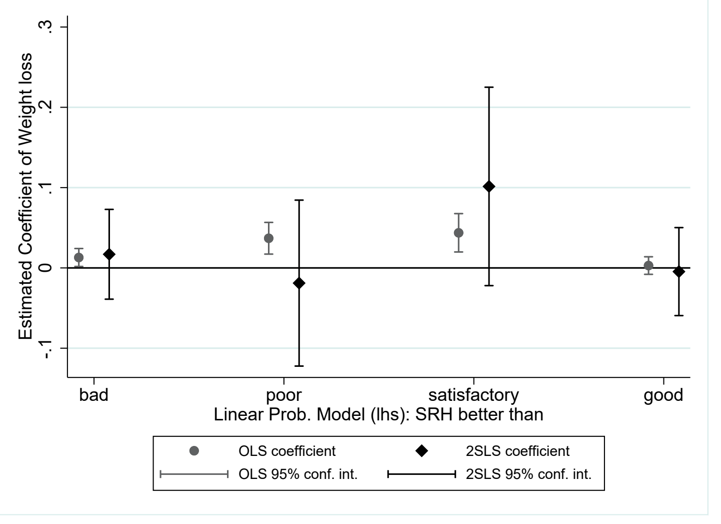

Figure A3: Est. Coef. for Linear Prob. Model (**Self-Rated Health**, **men**)

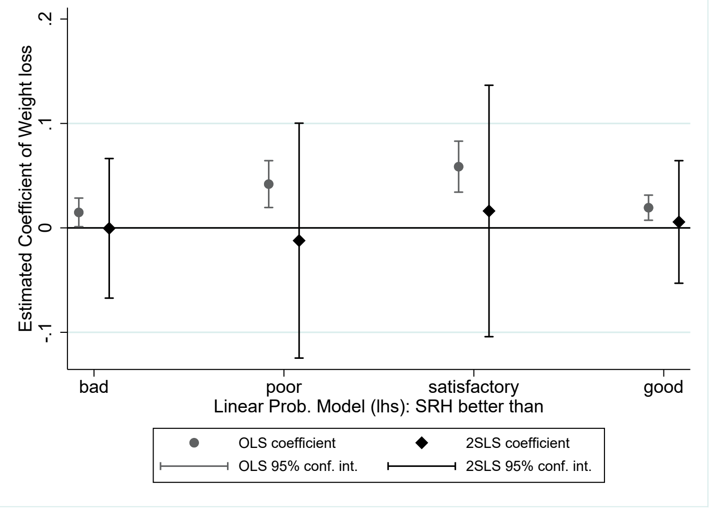

Figure A4: Est. Coef. for Linear Prob. Model (**Physical Well-Being**, **men**)

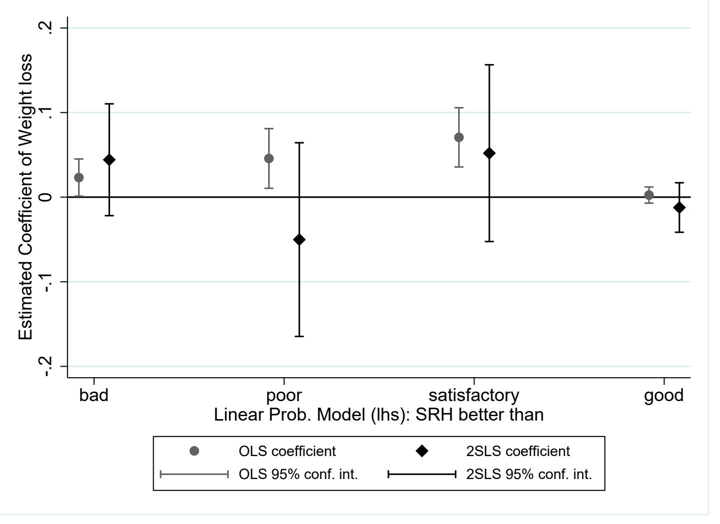

Figure A5: Est. Coef. for Linear Prob. Model (**Self-Rated Health**, **women**)

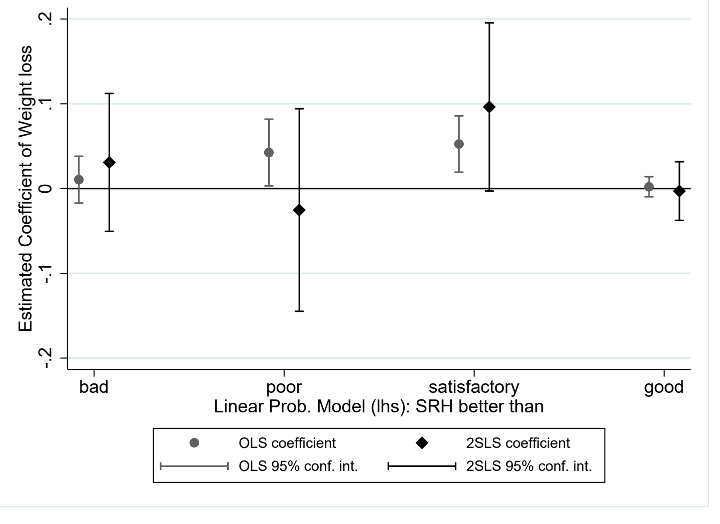

Figure A6: Est. Coef. for Linear Prob. Model (**Physical Well-Being**, **women**)

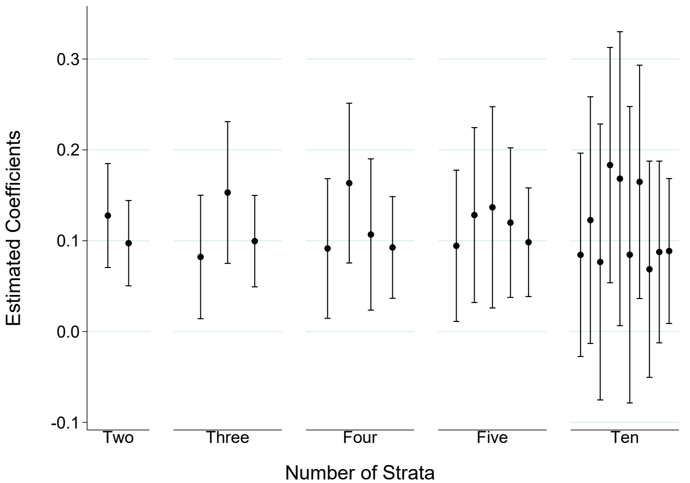

Figure A7: Est. Coef. (with 95% conf. int.) from OLS estimation stratified by different quantiles of the initial weight distribution (dependent variable: **Self-Rated Health**)