

Obesity and Health-Care Costs in Switzerland: Dealing with Endogeneity in Non-Linear Regression Models

STEFAN MEYER^a

JEL-Classification: I11, I12, C26

Keywords: obesity, health expenditure, measurement errors, endogeneity, control functions

SUMMARY

We draw microdata from the Swiss Household Panel to estimate the causal effect of obesity on the number of physician visits, the amount of hospital days, and the respective costs incurred. We do so by simultaneously coping with three endogeneity issues, comprising reporting errors, omitted variables, and simultaneity. Using the conditional expectation approach, we first account for the reporting errors in weight and height. Second, we address endogeneity in the body mass index (BMI) by applying a control function approach. In contrast to the method of two-stage least squares, this technique is consistent in non-linear regression settings. Using the mean BMI of relatives as an instrument for the respondent's BMI, we show that naïve regression methods considerably underestimate the impact of weight on the use of inpatient care, outpatient care, and costs. Accordingly, an additional unit of BMI raises annual health-care costs by CHF 253 or 11.5%, while the non-IV estimate amounts to only CHF 34 or 1.5%. Several robustness checks suggest the average marginal effect to be in the range of between CHF 220 and CHF 294. The model also predicts that if the overweight and obese people in the sample lost weight to the threshold of being of normal weight (BMI = 25), health-care costs could be reduced by about -4.7%. We conclude that the negative external effects caused by overweight and obesity are considerably larger than previously thought.

a Department of Business and Economics (WWZ), University of Basel, Peter Merian-Weg 6, P.O. Box, CH-4002 Basel, Switzerland. Email: stefan.meyer@unibas.ch.

1. Introduction

In Switzerland, the prevalence of obesity, defined as a body mass index (BMI) of at least 30 kg/m^2 , has increased significantly over the last decades.¹ Between 1992 and 2012, the number of obese adults in Switzerland has more than doubled, reaching 10.3% of the total population in 2012. Over the same period, the share of people who were classified as overweight ($25 \leq \text{BMI} < 30$) rose from 24.9% to 30.9% (FSO, 2012).²

This trend is troubling, as overweight and obesity are linked to many severe health conditions. Among other diseases, obesity is associated with an increased risk of myocardial infarction, stroke, type 2 diabetes, cancer, hypertension, osteoarthritis, asthma, sleep apnoea and respiratory problems (CAWLEY and MEYER-HOEFER, 2012). Consequently, obesity is seen as a major driver of rising health-care expenditure (HCE) in the USA and other industrialised countries. Using a representative sample of the U.S. population, FINKELSTEIN, FIEBELKORN, and WANG (2003) computed aggregate overweight- and obesity-attributable medical spending for the United States. On the individual level, the study shows that in 1998 the average increase in annual medical spending associated with obesity was 37.4 per cent, amounting to USD 732 per person. For the U.S. population as a whole, the authors concluded that aggregate expenditure attributable to obesity was about 5.3% of total HCE.³ Using Swiss prevalence data and literature-based estimates of the relative risks, SCHMID et al. (2005) estimated the overweight-attributable cost fraction of 18 related diseases. Their findings indicate that in 2001, direct and indirect health-care costs of overweight and obesity amounted to CHF 2,691 million, representing 5.9% of total HCE in Switzerland. Taking into account uncertainty in parameter assumptions, the authors estimated the effective costs to be in a range between 4.7 and 7.1% of total HCE.

Obesity itself hardly causes economic distortions as long as its costs are borne by the individual concerned. However, health care in Switzerland is mainly financed through mandatory health insurance (premiums) and public spending (taxes). Accordingly, people only bear a minor part of the health-care costs incurred.

1 BMI is defined as the individual's body weight (kg) divided by the square of height (m^2).

2 According to the Federal Statistical Office (FSO), this trend has been most pronounced among men and elderly people. In 2012, for instance, 50.5 per cent of all Swiss men above the age of 15 were either overweight or obese, amounting to a BMI above 25 kg/m^2 (FSO, 2012).

3 In a follow-up study with 2006 data, FINKELSTEIN et al. (2009) report even more pronounced outcomes. Across all payers, obese people faced per capita spending that was up by 42 per cent (USD 1,429) compared to normal-weight individuals. In aggregate, the annual medical burden of obesity was found to be 9.1% of total medical spending.

Inevitably, this system gives rise to moral hazard, as part of the costs are external to the individual. For instance, people freely decide upon their level of prevention effort. With regard to overweight, they implicitly control their body weight by choosing the amount of caloric intake and the level of physical activity. When deciding upon their weight, they consider the private costs of being overweight. If people are insured against health risks, these costs are considerably smaller than the total social costs of an unhealthy body weight. As a result, this optimisation problem leads to a weight that does not necessarily coincide with the socially optimal weight. A higher equilibrium weight leads to increased HCE, which in turn is mainly borne by the public sector. We then face economic distortions, which arise from the disparity between social costs and private costs. Given these moral-hazard issues, government programmes aimed at reducing the prevalence of obesity are justified measures to increase social welfare (PARKS, ALSTON, and OKRENT, 2012).

To determine the extent of economic distortions, we aim to estimate the *causal* effect of overweight and obesity on health-care demand. Using survey data from the Swiss Household Panel (SHP), we analyse how overweight and obesity affect the utilisation of inpatient and outpatient care. To deal with reporting errors and endogeneity in weight, we first perform an error-correction approach. In a second step, the corrected BMI values are used to estimate an instrumental variable (IV) regression, which takes care of simultaneity and omitted variables. Applying this two-step method is crucial, since IV regression on its own is unable to eliminate biases arising from non-classical measurement errors (see O'NEILL and SWEETMAN, 2013).

We contribute to economic research in multiple ways. First, we can separately analyse the effects of obesity on two health-care sectors in Switzerland. As obese people face an increased risk of becoming multi-morbid, the effects on health-care demand are likely to be non-trivial. In order to quantify the impact on HCE, we further merge the data with information on the prices of physician visits and inpatient days. Second, we tackle the problem of endogeneity in the data. By applying a combined error-correction and IV regression approach, we can predict the *causal* effects of body weight on health outcomes, rather than the simple correlation between the two. Moreover, we take advantage of the household structure of the data by using the mean BMI of relatives as an instrument for the respondent's BMI. Finally, in order to test structural validity, we perform several robustness checks and a specific form of validity tests that are known as falsification testing.⁴

4 On the one hand, the robustness checks help us determine whether our findings are robust to changes in the underlying model structure. Even more importantly, the falsification tests will provide evidence with regard to the specificity of our instruments.

The remainder of this paper is organised as follows. Section 2 discusses the potential sources of endogeneity in self-reported body data from health surveys. In Section 3, we introduce the SHP data and outline our estimation strategy. In Section 4, we present the empirical results and perform robustness checks and falsification tests. Section 5 discusses the major findings and draws a conclusion.

2. Reporting Errors and Endogeneity in BMI

We are interested in measuring the causal effect of obesity on health-care demand. When using survey data on self-reported BMI, however, there are at least three limitations that may bias the results.

Numerous studies have shown that standard regression techniques lead to biased estimates due to systematic *misreporting* of self-reported weight and height (PLANKEY et al., 1997; VILLANUEVA, 2001; BURKHAUSER and CAWLEY, 2008; O'NEILL and SWEETMAN, 2013). There is a large body of evidence that suggests that self-reported BMI tends to underestimate true BMI. O'NEILL and SWEETMAN (2013) emphasise the importance of controlling for misreportings in weight and height, demonstrating that the reporting errors are likely to be non-classical errors.⁵ Using a representative sample of U.S. residents, CAWLEY and BURKHAUSER (2006) argue that (young) women are more likely to underestimate their body weight, while men are prone to overestimate their actual weight. It is well-known that measurement errors in survey data are associated with the so-called attenuation bias in OLS. The respective coefficients are dwarfed towards zero, while the effect is strongly related to the significance of the measurement error.⁶ Consequently, researchers may draw incorrect conclusions if they fail to account for the fact that their variable of interest is measured with error.

Second, as the optimisation problem stated in the introduction suggests, the individual's weight is – aside from genetic predisposition – a choice variable. So there might be several unobservable factors or *omitted variables*, which not only affect health-care demand, but also impact the amount of caloric intake, the level of physical activity, and therefore the BMI. An often-cited example of an omitted variable is the economic term of *time preference*. Empirical studies have shown that the rate of time preference (i.e., the discount rate) varies widely

5 Measurement errors are said to be non-classical if they are correlated with the true unobserved value.

6 Moreover, this bias does not vanish as the sample size increases.

among individuals.⁷ People who place a higher weight on future benefit may in turn be willing to take preventive measures now to maintain their health (e.g., stop smoking, drink responsibly, maintain a healthy body weight). As a result, this (unobserved) rate of time preference is likely to be correlated with health-care demand and BMI. With regard to our econometric model, treating the BMI as an exogenous variable will thus lead to biased coefficients due to its correlation with the error term. However, the direction of the bias cannot be predicted *a priori*. People with a low personal discount rate will be less prone to overweight and obesity. Still, the effect of patience on health-care use is ambiguous: A healthy lifestyle *per se* may lower the use of health-care services due to an improved state of health. Nevertheless, health-conscious individuals also undergo more medical check-ups than their counterparts, which, for instance, will be reflected in a higher number of outpatient visits.

Last but not least, we expect to face *simultaneity* due to the data collecting strategy of the SHP. As in many other surveys, the quantity of health care consumed during the last year and the body weight are reported at the same time. Thus, the flow of causality is ambiguous and, in general, non-trivial: As a matter of fact, the amount of health-care services consumed may impact body weight as well. On the one hand, general practitioners (GPs) will advise overweight patients to stick to a low-calorie diet while undertaking regular physical activity. These recommendations are supposed to support weight loss. Likewise, each year a substantial number of morbidly obese people are admitted to hospital for obesity surgery (i.e. sleeve gastrectomy, gastric bypass surgery). Assuming that medical and surgical treatment is effective to some extent, causality is reversed, suggesting that there is a partial negative correlation between health-care demand and weight. In addition, a long inpatient stay can have a significant impact on body weight due to a change in diet in combination with a different level of physical activity. However, the size and direction of this effect remains unclear. Aggregating the two arguments, health-care use is likely to have a negative effect on body weight and hence on BMI. If this negative relationship is unaccounted for, standard econometric methods will underestimate the causal weight effect on health-care demand.⁸

7 See FREDERICK, LOEWENSTEIN, and O'DONOGHUE (2002) for an extensive review of theoretical formulations and empirical research on intertemporal choice and time discounting.

8 This is a well-known and basic result from the simultaneous equations literature, often referred to as the *simultaneous equations bias* (CAMERON and TRIVEDI, 2005).

As there is no *a priori* justification for how the three sources of endogeneity add up, the direction of the bias is ambiguous. Put differently, to study the causal health effect of obesity, we have to address all three issues simultaneously.

3. Methods

3.1 Estimation Strategy

We aim to estimate the impact of obesity on health-care demand and costs. As we are interested in the overall effect of weight on health outcomes, we use a one-stage approach.⁹ We first present the basic nonlinear model that is used to conduct our analysis. We then show how reporting errors and endogeneity can be dealt with within this framework.

3.1.1 The Basic Model

Econometrically, we aim to estimate the following regression model

$$y_{it} = f(\beta_w BMI_{it} + \mathbf{x}'_{it}\boldsymbol{\beta} + \gamma u_{it}) + \varepsilon_{it}, \quad (1)$$

where $f(\cdot)$ is a known nonlinear function, BMI_{it} denotes the measured (over-) weight variable of individual i at time t , $\mathbf{x}_{it} = [1, x^1_{it}, x^2_{it}, \dots, x^K_{it}]$ is a vector of covariates, including individual characteristics, canton and time fixed effects, and the constant, b_w and $\boldsymbol{\beta}$, are the coefficients to be estimated, and e_{it} is the error term. The variable u_{it} is an unobservable confounder that is both correlated with the outcome y_{it} and the endogenous variable, BMI_{it} . For instance, u_{it} captures the time preference and the lifestyle choices of the individual.

In fact, though, there is no straightforward way to estimate the model in (1). First, we only observe \widetilde{BMI}_{it} , which is the reported value of BMI_{it} , consisting of

9 It is often argued that the number of outpatient visits are best described by a hurdle model. However, we are mainly interested in the causal effect of obesity on the *total* demand for outpatient care. Furthermore, the principal-agent relationship is largely irrelevant when studying the impact of weight on physician visits. Consequently, using a one-stage approach is the more efficient approach. Moreover, with regard to the inpatient sector, a two-stage model tends to be inadequate in any case. Usually, people are either admitted to hospital by their referring physician or they may have to go there in an emergency. Thus the inpatient admission is hardly the outcome of a decision-making process by the individual. As regards the health-care costs, BUNTIN and ZASLAVSKY (2004) found the differences between one-part and two-part models to be very small and non-systematic. The authors therefore suggest using the one-stage formulation “if total costs are of interest” (BUNTIN and ZASLAVSKY, 2004, p. 540).

the real value and a reporting error ($BMI_{it} + e_{it}$). Second, we do not observe the confounding variable u_{it} . As a consequence, the causal effect of weight on our outcome variable y_{it} cannot be determined in this framework.

We follow two steps in order to cope with the issues of endogeneity and reporting errors. First, we control for the reporting errors in body weight and height. The application of the IV regression is sufficient in most studies that exhibit endogeneity and measurement errors in an explanatory variable. As O'NEILL and SWEETMAN (2013) show, this is true even in the case of non-classical errors (i.e., $Cov(BMI_{it}, e_{it}) \neq 0$). Still, we have to account for the measurement errors before employing the IV regression. The reason is simple enough: In the second stage, we aim to use the mean BMI of relatives (i.e., parents, children) to instrument for the respondent's BMI. Consequently, our instrument is also subject to errors in reporting. In general, the IV estimates may still be unbiased if the instrument is measured with error. However, if the reporting errors of the instrument and the endogenous variable are correlated, this argument no longer holds.

3.1.2 *Measurement Errors and Endogeneity*

As we cannot rule out correlation in reporting errors, we employ the conditional expectation (CE) approach as described in LYLES and KUPPER (1997). This popular method of error correction was first mentioned by ARMSTRONG and OAKES (1982), who applied CE to a cohort mortality study.

Using estimate correction equations, we predict measured weight and height on the basis of self-reported weight and height, sex, and age. To do so, one requires validation data from a survey that covers self-reported and measured BMI. To our knowledge, however, there is no Swiss survey that offers information on the extent to which people misreport their body data. Therefore, we have to employ the formulas provided by CAWLEY and BURKHAUSER (2006) for white males and females in the United States to calculate expected BMI from self-reported data.¹⁰ Several other studies have made use of the estimates of CAWLEY and BURKHAUSER (2006) to cope with measurement errors in BMI (e.g., AULD, 2011; GILA and MORAB, 2011). Their formulas are based on the National Health and Nutrition Examination Survey (NHANES), which contains both measured and self-reported heights and weights.

10 CAWLEY and BURKHAUSER (2006) provide separate regression estimates for men and women of different ethnic groups on the basis of measured data, self-reported data and age. As the authors show, sex-specific error correction is crucial, since the reporting errors also depend on the sex of the respondent: For example, the authors found that (young) women are more likely to underestimate their body weight, while men are prone to overestimate their actual weight.

Expressed in mathematical terms, we estimate the following equations:

$$E(l_{it}) = \alpha^k + \gamma_{1k} \text{age}_{it} + \gamma_{2k} \text{age}_{it}^2 + \gamma_{3k} \tilde{l}_{it} + \gamma_{4k} \tilde{l}_{it}^2,$$

where $l \in \{\text{weight}_{\text{measure}}, \text{height}_{\text{measure}}\}$, $\tilde{l} \in \{\text{weight}_{\text{report}}, \text{height}_{\text{report}}\}$, $k \in \{\text{male}, \text{female}\}$, and α and γ are the parameters suggested by the authors. On the basis of these corrected values for weight and height, it is straightforward to obtain error-adjusted BMI.¹¹

In a second step, we account for other sources of endogeneity; i.e., omitted variables and simultaneity. At a first glance, it might be tempting to apply the 2SLS method to the nonlinear case.¹² In fact, applications of 2SLS in nonlinear health econometric contexts are quite common (see, e.g., HOLMES and DEB, 1998; CAWLEY, 2000; FOX, 2003; MEER and ROSEN, 2004). However, WINDMEIJER and SANTOS SILVA (1997) and, more recently, WOOLDRIDGE (2014) have pointed out that the consistency property of 2SLS in the fully linear models does not extend to the use of the two-stage approach in nonlinear models. This claim is supported, for instance, by the findings of TERZA, BRADFORD, and DISMUKE (2008). Using simulation studies, the authors show that substantial bias in the estimation of causal effects can result from applying conventional IV regression in nonlinear settings.

We follow the control function (CF) approach suggested by CAMERON and TRIVEDI (2013) to account for endogeneity in our nonlinear model. Akin to 2SLS, the CF method comprises two estimation steps. First, we estimate the first stage auxiliary regression

$$BMI_{it} = g(\mathbf{x}'_{it}\boldsymbol{\delta} + \mathbf{z}'_{it}\boldsymbol{\theta}) + u_{it}, \quad (2)$$

where \mathbf{x}_{it} is the same vector of covariates included in (1), and \mathbf{z}_{it} is a set of instrumental variables which have the property that changes in \mathbf{z}_{it} are associated with changes in BMI_{it} but do not lead to change in the demand for health care (aside from the indirect route via BMI_{it}). $\boldsymbol{\delta}$ and $\boldsymbol{\theta}$ are parameters to be estimated, and u_{it}

11 Applying the formulas to our data, the average BMI for males increases from 23.83 to 23.90 (+0.3%) and from 22.45 to 23.08 (+2.8%) for females.

12 In linear models, the most popular approach to dealing with endogeneity is the two-stage least squares (2SLS) estimator. In the first stage of 2SLS, auxiliary regressions are estimated using adequate instrumental variables. The results are then used to generate predicted values for the endogenous variables. In the second stage, the outcome variable of interest is regressed on the covariates and the endogenous variables, which have been replaced by the predicted values from the first stage.

is the unobservable confounder that is correlated with BMI_{it} , and y_{it} via (1). In a next step, we can simply compute $\hat{u}_{it} = BMI_{it} - g(\mathbf{x}'_{it}\hat{\boldsymbol{\delta}} + \mathbf{z}'_{it}\hat{\boldsymbol{\theta}})$. The estimated residual term, \hat{u}_{it} , is then substituted for the unobservable confounders in the nonlinear regression model. Replacing u_{it} for its estimate, (1) becomes

$$y_{it} = f(\beta_w BMI_{it} + \mathbf{x}'_{it}\boldsymbol{\beta} + \gamma\hat{u}_{it}) + \varepsilon_{it}. \quad (3)$$

As we now explicitly control for the unobservable confounders, BMI_{it} can be estimated consistently (TERZA, BASU, and RATHOUZ, 2008). Moreover, the exogeneity of BMI_{it} can be tested directly in the model. Under the hypothesis of no endogeneity, y_{it} is uncorrelated with the unobservable confounder, i.e. $\gamma=0$ (see HAUSMAN, 1978). To correct for the inclusion of an estimated value in (3), we calculate our standard errors using 5,000 bootstrap iterations that are clustered at the individual level (CAMERON and TRIVEDI, 2013).

The reason that the CF approach works is simple: If we knew the parameters from the auxiliary regression (i.e., $\boldsymbol{\delta}$ and $\boldsymbol{\theta}$), we would implicitly know the values of u_{it} by (2). These variables then could be included in (1) among the observable controls. In other words, the endogeneity in BMI_{it} would cease to exist. Even though we do not know u_{it} , we can predict it by performing the auxiliary regression in (2) and thereby obtain consistent estimates of the true u_{it} .¹³

3.1.3 Instrumental Variables

As with other IV regression approaches, however, the CF method strongly relies on valid instruments, \mathbf{z}_{it} . In our case, we have to find instruments that are powerful predictors for the BMI of the respondent, without being correlated with the unobserved confounders in the main model (i.e., $\text{Cov}[\mathbf{z}, u] = 0$). We aim to use the weight of a biological relative as an instrument for the BMI of the individual. Specifically, we calculate the mean BMI of the parents as instruments for their children's BMI, and vice versa. To increase the precision of the estimates, we further include the BMI squared and the BMI cubed of the relatives in our set of instruments.¹⁴

As our arguments will make clear, there is broad consensus that the weight of relatives meets the two requirements of a valid instrument mentioned above.

13 A couple of studies in the field of health economics have made use of the CF approach to cope with endogeneity in nonlinear frameworks (see, e.g., LINDROOTH and WEISBROD, 2007; SHEA et al., 2007; SHIN and MOON, 2007).

14 We also estimate the model (1) without the use of further polynomials of the instrument (i.e., the just-identified model), and (2) with BMI and BMI squared (see Section 4.1).

First, it is a sound predictor for the weight of the respondent, capturing common genetic factors that impact metabolic rate and body weight. Data from adoption studies, for instance, are consistent with genetic factors accounting for 20% to 60% of the variation in individuals' BMI values (MAES, NEALE, and EAVES, 1997). Analysing stratified data from 37,000 twin pairs in 8 different countries, SCHOUSBOE et al. (2003) find even higher correlations of BMI. Except for the relative low value in young males from Norway (0.45), all other estimates of heritability ranged from 0.64 to 0.84. This correlation is confirmed by the F -statistic from the first stage regression, which amounts to 304 (Cluster-adjusted: 73). This value clearly exceeds the minimum threshold suggested by STAIGER and STOCK (1997). We thus conclude that the predictive power of our instruments is high enough. The output from the first stage regression is depicted in Table 6 in the Appendix. Even more importantly, valid instruments are required to be exogenous. This is to say that they are uncorrelated with the unobservable confounders. A possible threat to the validity of the instruments is the household environment shared by the individuals. There may be unobserved household characteristics that are not only correlated with the weight of the individuals, but which also affect the demand for medical care. Even though it is impossible to directly test for exogenous instruments, a growing body of evidence indicates that the effect of the common household is rather negligible.¹⁵ Recently, HABERSTICK et al. (2010), using data from a large panel of adolescents in the United States, found no evidence for shared household factors that affect the BMI of the individuals considered. Moreover, the test for overidentification by HANSEN (1982) fails to reject the hypothesis of valid instruments, providing a further assurance that our set of instruments have been compiled sensibly.¹⁶

3.2 Data

Our survey data is the Swiss Household Panel (SHP). The yearly panel study follows a random sample of households in Switzerland over time, interviewing all household members. The SHP was established in 1999, collecting data on 12,931 persons living in 5,074 households. In 2004, an additional sample of

15 See CAWLEY and MEYERHOEFER (2012) for a comprehensive review of genetic studies that deal with shared household environment effects on weight.

16 Tests of overidentifying restrictions actually test two different aspects simultaneously. The first one is whether the instruments are uncorrelated with the error term. The other is that the equation is misspecified and that one or more of the excluded exogenous variables (z_i) should in fact be included in the second-stage equation. Thus a significant test statistic could signify either an invalid instrument or an incorrectly specified structural equation.

2,538 households was added to the existing data set. As a consequence of our IV method, our sample is restricted to households that comprise at least one parent and one biological child with complete data on body weight and height. The mean age of the children amounts to 18.9 years, while the mean age of the parents in the sample amounts to 49.6 years.¹⁷

The survey contains information on self-reported body weight and height. These two variables, which were first collected in 2004, are used to construct individual BMI scores.¹⁸ With regard to health-care demand, the SHP offers annual data on the number of outpatient visits¹⁹ and the number of days spent in hospital. To obtain individual health-care expenditure, we merge the SHP data with information on health-care costs. We use 2012 data on prices for all years of observation to set aside price changes. According to the FSO, the average costs of an outpatient visit (p_{out}) amounts to CHF 315. Likewise, the price per hospital day (p_{in}) comes to CHF 1,763, on average. Therefore, expected health-care expenditure on physician visits and inpatient stays can be calculated as

$$HCE_{it} = p_{out}DOCVIS_{it} + p_{in}HOSP_DAYS_{it} \quad (4)$$

We have to make the assumption that the prices p_{out} and p_{in} do not vary systematically across population subgroups.²⁰ Given the available information, we still consider (4) to produce a good approximation of the effective HCE incurred.

We control for a selected set of observable confounders that may contribute to the amount of medical care consumed. Interaction terms between age and sex are constructed to allow for sex-specific demand functions for health care over age groups. The individual health state is captured by several indicators. The subjective state of health, *HSTATUS*, is accounted for by an ordinal scale that ranges from (1) *very well* to (5) *very bad*. We include, as more objective measures, several variables on recent health issues, covering accidents, chronic diseases, back problems, weakness, sleeping problems, and headaches. All these health issues are

17 Figure 2 and 3 in the Appendix provide separate histograms for age and *BMI* within the two subgroups (parents, biological children).

18 Even though this index has become a very popular measure of overweight and obesity, it has also met with criticism. BAGUST and WALLEY (2000) argue that *BMI* is not supported by a sound theoretical basis, nor is it a valid measure for all populations. Nevertheless, as recent literature does not provide sophisticated, easy-to-use alternatives, and as we only observe weight and height in our data, we follow past practices.

19 Visits to the dentist are not included in this variable.

20 We may slightly underestimate (overestimate) the effect of obesity on HCE if the average cost of a physician visit or hospital stay is positively (negatively) correlated with weight.

assumed to increase the amount of health care consumed. We consider *NEW-BORN* as a crucial confounder in our analysis. The dummy variable indicates whether a woman has given birth to a child within the last 12 months. Pregnancy is likely to affect both the body weight and the demand for health care.

Educational attainment (*EDUCATION*) is added as an explanatory variable, as it is likely to be correlated with health-care demand and BMI. We further account for the different levels of opportunity costs by including the employment status (*EMPLOY*). The civil status of the respondent is captured by *CIVSTA*. While the degree of urbanity (*URBAN*) is a good proxy for access to health care, the first language spoken by the person at home is used to control for cultural differences (*LANG*). In addition, we allow for different levels of demand for Swiss citizens and non-citizens (*SWISS*).

Moreover, we account for unobservable differences across health-care areas by including canton fixed effects. These area-specific effects are intended to capture unobserved heterogeneity in the supply of inpatient and outpatient services across Switzerland (i.e., physician density, supply of hospital beds). Finally, to account for time trends, we include time fixed effects for the years from 2004 to 2011.

We include eight waves of the SHP, covering the observations between 2004 and 2011. After excluding any missing data from our file, we end up analysing an unbalanced panel of 4,968 individuals, amounting to a total sample size of 16,317. On average, each person is therefore observed approximately 3.3 times.

Table 1 provides variable definitions and presents summary statistics of our main variables.²¹ On average, each individual has a record of 2.88 physician visits and 0.71 days spent in hospital. As a result, the estimated HCE on physician services and inpatient care in our sample amounts to CHF 2,108 at the mean. According to the FSO, the mean total HCE in 2012 amounted to approximately CHF 8,456 per capita, which is substantially higher than our estimated costs. We clearly underestimate the total costs of health care for two reasons. First, HCE only involves visits to GPs and specialists, and hospital stays. All other aspects of health care are not part of our analysis (e.g., drug expenditure, dentists, long-term care, nursing care, prevention activities, administration costs). Second, as our study relies on data of multi-person households with children, the sample is very specific. Accordingly, the mean age amounts to only 34.8 years (Overall population: 41.6 years). In fact, the sample comprises very few elderly people, who are likely to cause very high health-care costs. Only 0.6% of the people in our sample are aged 70 or older (Overall population: 12.2%).

21 Detailed description and descriptive statistics of all variables is provided in Table 5 in the Appendix.

Table 1: Descriptive Statistics of the Main Variables

Variable	Definition	Mean	SD	Min	Max
<i>HCE</i>	Estimated health expenditure on physician and inpatient care (in 2012 CHF)	2,108	9,645	0	373,932
<i>DOCVIS</i>	Number of physician visits over the past 12 months	2.88	6.10	0	200
<i>HOSPDAYS</i>	Number of inpatient days over the past 12 months	0.71	5.16	0	200
<i>BMI</i>	Reported body mass index ($BMI = kg/m^2$)	23.11	3.95	12.12	63.18
Controls					
<i>NEWBORN</i>	Woman has given birth to a child within 12 months	0.001	0.037	0	1
<i>ACCIDENT</i>	Person has suffered a accident	0.052	0.223	0	1
<i>CHRONIC</i>	Person suffers from a chronic health condition	0.269	0.443	0	1
<i>BACKPAIN</i>	Person has been suffering from back pain (past 4 weeks)	0.416	0.493	0	1
<i>WEAK</i>	Person has been suffering from weakness or weariness (past 4 weeks)	0.447	0.497	0	1
<i>SLEEPING</i>	Person has been suffering from sleeping problems (past 4 weeks)	0.301	0.459	0	1
<i>HEAD</i>	Person has been suffering from headaches (past 4 weeks)	0.379	0.485	0	1
<i>HSTATUS</i>	Subjective health status (Basic category: <i>very well</i>)				
<i>well</i>	Health status is "well"	0.656	0.475	0	1
<i>fair</i>	Health status is "fair"	0.092	0.289	0	1
<i>bad</i>	Health status is "bad"	0.013	0.111	0	1
<i>very bad</i>	Health status is "very bad"	0.001	0.030	0	1

$N = 16,317$.

3.3 Model Specification

We specify three different generalised linear models (GLM) with a log link to estimate the impact of weight on the number of visits to the physician, the number of inpatient days, and the respective health-care costs incurred.

Both outpatient visits and hospital days are skewed count outcomes. In general, these health-care counts do not follow the well-defined Poisson distribution, but are highly overdispersed. In empirical work, the most common approach is to assume that the counts are drawn from a negative binomial distribution. This hypothesis is confirmed by Pearson's goodness-of-fit statistics that we calculate after running a Poisson regression.²²

As we calculate health-care expenditure on the basis of the two count outcomes, *HCE* tends to be highly skewed as well. This finding is not surprising, though. In health care, a very small share of the overall population is responsible for a substantial portion of total HCE.²³ In health economic literature, there is broad consensus that non-normal HCE data is best described by the Gamma distribution. MANNING and MULLAHY (2001) propose using a PARK (1966) test to determine whether the data actually fits the variance function of the Gamma distribution. In fact, the augmented Park test backs our choice of the cost function.²⁴

As our endogenous variable is continuous, one is inclined to estimate the first stage of our model by using OLS. Based on Akaike's information criterion (AIC) and the Bayesian information criterion (BIC), however, we decided to estimate the first stage equation (2) as a GLM of the Gaussian family with a log link.²⁵ To obtain estimates for u_{it} , we then calculate raw residuals via $\hat{u}_{it} = BMI_{it} - g(\mathbf{x}'_{it}\hat{\boldsymbol{\delta}} + \mathbf{z}'_{it}\hat{\boldsymbol{\theta}})$.²⁶

22 Accordingly, we can reject the hypotheses that *DOCVIS* or *HOSPDAYS* are Poisson distributed ($p < 0.01$).

23 With regard to our data, we can see this skewness by comparing the mean expenditure in Table 1 to its median: While the mean annual HCE amounts to CHF 2,108 in our sample, 50 per cent of the people spent less than CHF 315 per year.

24 Following the approach of MANNING and MULLAHY (2001), we run the modified version of the Park test by estimating a GLM with log link where the dependent variable is $(HCE_{it} - \widehat{HCE}_{it})^2$ and the explanatory variable is \widehat{HCE}_{it} from the initial GLM of HCE_{it} on \mathbf{x}_{it} . The parameter estimate of \widehat{HCE}_{it} , $\hat{\lambda}$, is supposed to capture the true variance function of our data. The respective distribution functions are Gaussian ($\lambda = 0$), Poisson ($\lambda = 1$), Gamma ($\lambda = 2$), Wald or inverse Gaussian ($\lambda = 3$). The test indicates that the conditional variance is proportional to the square of the conditional mean. $\hat{\lambda}$ amounts to 1.95 [95% CI: 1.48, 2.43] and is not significantly different from 2 ($p = 0.85$).

25 AIC: 87,625 (OLS), 87,580 (GLM log link); BIC: 88,211 (OLS), 88,166 (GLM log link).

26 According to WOOLDRIDGE (2014), the use of standardised residuals may be adequate in some cases. We perform the regression with standardised values for u_{it} as a robustness check in Section 4.1.

4. Results

The first stage regression shows that, apart from our three instruments, sex, and age, the differences in BMI are also driven by other covariates (see Table 6 in the Appendix). The level of education and the employment status are correlated with body weight. Accordingly, being unemployed is associated with an increase in BMI by 1.25 points. Furthermore, people living in urban and suburban areas (Mean BMI: 23.3) tend to be slightly less prone to overweight and obesity than people from more rural parts of Switzerland (BMI: 23.8). Finally, French-speaking people (22.8) have a significantly lower BMI than their German-Speaking (23.7) and Italian-Speaking counterparts (24.0).

The main results from the GLM regressions are presented in Table 2.²⁷ For comparison, we show the coefficients of the CF regression alongside the results from the non-IV regression approach.

As a general finding, the effect of weight tends to be considerably higher in the CF approach. Across all three models, an increase in BMI is associated with a higher amount of health-care use. The effect is significant for outpatient care, inpatient care, and costs ($p < 0.05$).

Table 2: Main Results from the GLM Regression

	<i>HCE</i> (gamma)		<i>DOCVIS</i> (negbin)		<i>HOSPDAYS</i> (negbin)	
	Non-IV	CF	Non-IV	CF	Non-IV	CF
<i>BMI</i>	0.015* (0.009)	0.115*** (0.041)	0.013** (0.005)	0.051** (0.020)	0.000 (0.017)	0.165** (0.065)
γ		-0.104** (0.045)		-0.040* (0.022)		-0.171** (0.070)
<i>SWISS</i>	0.183 (0.115)	0.204* (0.115)	0.101 (0.064)	0.106* (0.063)	0.096 (0.205)	0.124 (0.205)
<i>CIVSTAT</i>						
<i>married</i>	0.297 (0.252)	0.262 (0.252)	0.230 (0.152)	0.228 (0.152)	0.166 (0.604)	0.115 (0.597)
<i>separated</i>	0.382 (0.269)	0.334 (0.267)	0.260 (0.162)	0.252 (0.161)	0.179 (0.629)	0.107 (0.618)
<i>NEWBORN</i>	1.409*** (0.363)	1.555*** (0.348)	0.570 (0.357)	0.619* (0.346)	1.783*** (0.384)	2.055*** (0.385)
<i>ACCIDENT</i>	1.207*** (0.077)	1.203*** (0.078)	0.857*** (0.047)	0.853*** (0.047)	1.381*** (0.121)	1.373*** (0.122)

27 The complete outputs can be found in Tables 8, 9, and 10 in the Appendix.

	<i>HCE</i> (gamma)		<i>DOCVIS</i> (negbin)		<i>HOSPDAYS</i> (negbin)	
	Non-IV	CF	Non-IV	CF	Non-IV	CF
<i>CHRONIC</i>	0.874*** (0.068)	0.827*** (0.067)	0.765*** (0.037)	0.749*** (0.037)	0.862*** (0.111)	0.786*** (0.111)
<i>BACKPAIN</i>	0.081 (0.055)	0.054 (0.053)	0.064** (0.030)	0.057* (0.030)	0.042 (0.098)	-0.001 (0.095)
<i>WEAK</i>	0.118** (0.058)	0.116** (0.058)	0.153*** (0.030)	0.151*** (0.030)	0.157 (0.103)	0.153 (0.103)
<i>SLEEPING</i>	0.226*** (0.076)	0.228*** (0.074)	0.114*** (0.034)	0.115*** (0.034)	0.315** (0.126)	0.324*** (0.124)
<i>HEAD</i>	-0.007 (0.058)	0.004 (0.057)	0.082*** (0.030)	0.086*** (0.031)	-0.177* (0.107)	-0.163 (0.107)
<i>HSTATUS</i>						
<i>well</i>	0.227*** (0.068)	0.197*** (0.067)	0.209*** (0.043)	0.201*** (0.042)	0.346*** (0.118)	0.292** (0.118)
<i>fair</i>	0.942*** (0.089)	0.860*** (0.094)	0.774*** (0.061)	0.747*** (0.058)	1.223*** (0.148)	1.081*** (0.155)
<i>bad</i>	1.775*** (0.204)	1.664*** (0.214)	1.370*** (0.102)	1.330*** (0.106)	2.018*** (0.309)	1.823*** (0.320)
<i>very bad</i>	2.276*** (0.665)	2.293*** (0.668)	0.970*** (0.210)	0.970*** (0.207)	2.830** (1.141)	2.876** (1.308)
<i>EMPLOY</i>						
<i>nonworking</i>	0.055 (0.095)	0.041 (0.090)	0.029 (0.054)	0.027 (0.053)	0.166 (0.168)	0.147 (0.162)
<i>unemployed</i>	0.398* (0.241)	0.241 (0.246)	0.097 (0.132)	0.049 (0.134)	0.555 (0.410)	0.267 (0.422)
<i>URBAN</i>						
<i>urban</i>	0.008 (0.091)	0.074 (0.086)	0.059 (0.053)	0.081 (0.053)	-0.225 (0.163)	-0.099 (0.152)
<i>suburban</i>	0.057 (0.090)	0.115 (0.086)	0.037 (0.051)	0.057 (0.049)	-0.040 (0.157)	0.066 (0.148)
<i>LANG</i>						
<i>French</i>	0.135 (0.119)	0.228* (0.125)	0.195** (0.084)	0.226*** (0.085)	0.075 (0.179)	0.239 (0.192)
<i>Italian</i>	0.353 (0.230)	0.276 (0.223)	0.118 (0.128)	0.103 (0.129)	0.606* (0.367)	0.455 (0.357)

Notes: $N = 16,317$; coefficients of the canton and year fixed effects are not shown; standard errors are given in parentheses; standard errors are based on 5,000 bootstrap replications and adjusted for intra-household correlation; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Interestingly, we cannot reject the hypothesis of exogenous BMI in the regression with physician visits at the 5% level. Even though the impact of weight is stronger in the CF model, γ is significant only at the 10% level, indicating that endogenously determined BMI may not be causally important in modelling the demand for outpatient care. The marginal effects amount to 1.3% in the non-IV regression model, while in the CF model, a 1 point increase in *BMI* (e.g., from 28 to 29) is associated with a rise in *DOCVIS* of approximately 5.2%.

The situation, however, is considerably different in the analysis of hospital days. Our results indicate that we severely underestimate the effect of weight on the use of inpatient care. In the non-IV regression model, BMI tends to have no effect at all on the number of inpatient days. This result is reversed under the CF approach. Here, we not only observe a substantial causal effect of weight on the number of days, but we also find evidence that our variable of interest might be subject to a significant amount of endogeneity ($p < 0.05$). Taking account of the endogenously determined body weight, in turn, a 1 point rise in BMI may raise hospital days by no less than 18%.

As a consequence of these sectoral effects, we also find a correlation of weight and health outcome in the HCE model, which combines and prices the physician visits and inpatient days. On the one hand, BMI is positively correlated with HCE in the non-IV regression. The marginal effect of weight, however, is rather small and amounts to only 1.5%, on average. Given the results from the sectoral analysis, this moderate impact is not surprising, though. It mainly reflects the significant impact of BMI on the use of outpatient care that we obtained from the non-IV regression specification of *BMI* on *DOCVIS*. Turning to the CF estimation, we easily see that the causal effect of weight on costs seems to be much higher than the observed correlation in the naive regression model. First of all, the size and significance of γ suggest that the estimates from non-IV regression are subject to considerable bias due to the endogenous determination of BMI. In fact, confounding factors tend to dwarf the true effect of overweight and obesity on costs. That is, the output from the CF regression predicts a substantial effect of BMI that amounts to about +12.2% at the margin.

Table 3 shows the predicted means of the outcome variable at BMI = 30. The average marginal effects (AME) can be found in the bottom row of the table. For instance, the CF model predicts that, on average, an additional unit of BMI will raise HCE by about CHF 253. In contrast, the non-IV estimate amounts to only CHF 34 annually. These findings are slightly more pronounced than the estimates reported by CAWLEY and MEYERHOEFER (2012), who found the non-IV and CF effect to be USD 49 and USD 149, respectively. Moreover, CAWLEY and MEYERHOEFER (2012) noticed that the difference between the non-IV and

the CF estimates was larger for inpatient costs than for outpatient costs. This observation is actually confirmed by our analysis: At BMI 30, the CF estimate for doctor visits exceeds the non-IV prediction by about 28 per cent (3.163 vs. 4.046 visits), whereas the relative difference of *HOSPDAYS* amounts to +217% (0.738 vs. 2.337 days).

Table 3: Predicted Means at BMI=30 and Average Marginal Effects

	<i>HCE</i>		<i>DOCVIS</i>		<i>HOSPDAYS</i>	
	Non-IV	CF	Non-IV	CF	Non-IV	CF
<i>BMI</i> =30	2,415	4,743	3.163	4.046	0.738	2.337
<i>AME</i>	34	253	0.038	0.148	0.000 ⁺	0.122

⁺ Not significantly different from zero at the level of 10%.

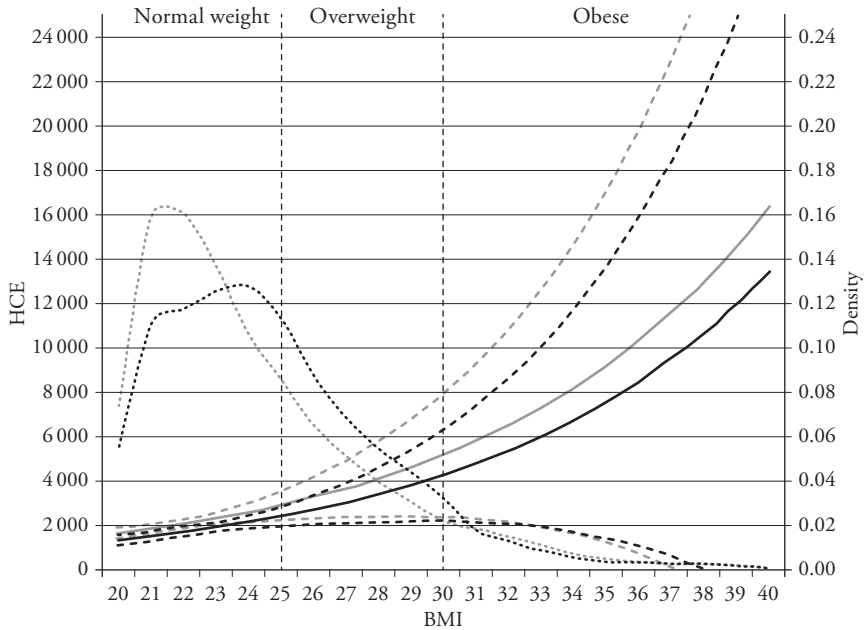
Reverse causation may be the major issue that drives the difference in *HOSPDAYS* between the non-IV and the CF specification. Bariatric surgery, which in general involves a hospital stay, is known to be an effective treatment for morbidly obese individuals. A recent Swiss study by PETERLI et al. (2012) reports 3-month weight loss to be in a range of about 16 kg, which amounts to -6 to -8 BMI points. As a consequence, the coefficient from the non-IV approach captures this negative effect of *HOSPDAYS* on *BMI* and thus offsets the positive impact of weight on inpatient days.

To a lesser extent, this argument also applies to the number of outpatient contacts; GPs can assist overweight patients in losing weight by advising them on weight-loss strategies or by prescribing medication.

The sex-specific effects of weight on health-care costs tend to be rather small (see Figure 1). Even though the predicted HCE are somewhat higher for women across all weight categories, the difference seems to be unrelated to the actual BMI level. More interestingly, the figure suggests that the increase in costs is relatively moderate within the normal weight category ($19 \leq BMI < 25$). A further increase in weight, though, is associated with a significant rise in HCE ($25 \leq BMI < 30$). The predicted cost almost doubles as individuals reach the threshold of being obese (i.e., $BMI = 30$).²⁸ As the distribution of BMI in the sample shows, these high costs are caused by a very small share of the overall population. Conversely,

28 We do not plot this graph for higher levels of obesity, as the predicted HCE and the respective confidence intervals become unreliable. The large uncertainty in these predictions is mainly caused by the small subsample of people who are obese or morbidly obese.

Figure 1: Predicted Relationship between BMI and HCE for Men and Women



Note: The figure depicts the estimates from the CF model; the black (grey) solid line indicates predicted HCE for men (women); the respective 90% confidence intervals are represented by the dashed lines; dotted lines denote the distribution of BMI in the sample

the model predicts that a reduction in weight may help lower costs significantly. To see that, let us assume that all the people in the sample who are classified as overweight (23.2%) or obese (6.5%) manage to reduce their body weight to the threshold value for normal weight ($BMI=25$). For our sample, the CF model finds that this *ceteris paribus* reduction in weight would be followed by a decline in total expenditure of about -4.7% . In turn, the respective figure reported by the non-IV regression model amounts to only -2.2% , which is less than half of the prediction obtained from the CF approach.

Most coefficients of the regression covariates in Table 2 have the expected sign. The estimates for the recent health issues (*BACKPAIN*, *WEAK*, *SLEEPING*, *HEAD*) suggest that, in the first place, most people seek care from a GP or specialist. Accordingly, all four indicator variables are positively correlated with the number of outpatient contacts, while only sleeping problems tend to affect the inpatient sector. One reason for this finding might be that people who experience

severe insomnia are admitted to hospitals for further tests (e.g., for overnight sleep studies). However, the direction of causality is not very clear: Inpatients may find it hard to fall asleep while staying in hospital. Thus, they are more likely to report having suffered from recent insomnia.

Subjective health status (*HSTATUS*) is an excellent predictor for the level of health-care costs. The relationships between perceived status of health and the monthly health costs per individual for the five health categories are as follows: *very well* is associated with CHF 124, *well* with CHF 151, *fair* with CHF 293, and *bad* with CHF 654. Individuals who are in a *very bad* state of health face, on average, health-related costs of CHF 1,219.

Unemployed persons experience higher HCE than employed, even though the significance of the effects tends to be rather low. According to DOOLEY, CATALANO, and HOUGH (1992), there is sound evidence that unemployment negatively affects physical and mental health and well-being. This finding has been linked to risk-taking behaviour and a decreased level of prevention effort among those individuals. Several authors have demonstrated how unemployment leads to a rise in unhealthy habits such as alcohol abuse, smoking, poor diet, low levels of physical activity, and other health-related behaviours (see HALFORD and LEARNER, 1984; LEE et al., 1991; DOOLEY, CATALANO, and HOUGH, 1992; DEB et al., 2011).

As we control for the canton of residence, it is not surprising that the effect of urbanity is not significantly different from zero. The canton fixed effects are likely to absorb most of the potential variation in access across urban and rural regions. Conversely, though, we find a positive effect of *French* on the number of outpatient visits. Apparently, there seem to be regional or cultural differences that shape the demand for outpatient care and which cannot be explained by cantonal supply-side factors. The absolute disparity is actually substantial. The predicted number of visits amounts to 3.47 for French-speaking individuals while the number for German-speakers is only 2.75. As regards the inpatient sector, we find little variance between language groups. The mean *HOSP_DAYS* of German-speakers (0.693) is not significantly different from the mean number of inpatient days of French-speakers (0.886).²⁹

29 This finding can be explained by different processes of decision-making in health care: In general, a person can decide whether to consult a GP or not, and, to a certain extent, on the number of follow-up visits. On the other hand, however, people are referred to the hospital by their doctor or they may have to go to A&E in an emergency. Thus utilisation of inpatient care is mainly driven by exogenous factors that cannot be influenced by the patient in question.

4.1 Robustness Checks

4.1.1 Alternative Model Specifications

For reasons of brevity, we limit our robustness analysis to the health-care costs. If the estimates from the cost regression are robust to changes in specification, this in turn indicates that the underlying models (*DOCVIS* and *HOSPDAYS*) are robust as well. The average marginal effects from these robustness checks can be found in Table 11 in the Appendix.

As a first robustness check, we run two more flexible specifications of the cost regression (HCE). To do so, we not only include *BMI* as an explanatory variable, but also BMI squared (Model BMI2), and BMI squared and BMI cubed (Model BMI3), respectively. As we have constructed three instruments, these models are still identified. The marginal effect of BMI on health-care costs does not change significantly. Compared to the standard model (AME: CHF 253), the AMEs amount to CHF 241 in the BMI2 specification and CHF 250 under BMI3.

Second, we change the set of instruments to see whether our results are driven by the choice of polynomials used in the IV regression. Using only BMI_{IV} and BMI_{IV}^2 , the marginal effect of a one-unit change in BMI becomes CHF 247 (Model BMI_IV2). The respective AME amounts to CHF 258 in the just-identified model, where BMI_{IV} is included as the only instrument (Model BMI_IV1). Both estimates are very close to those of the standard model.

To test whether our findings are sensitive to the set of covariates used in the regression, we drop all covariates except for age, sex and its interaction terms and re-estimate the HCE model. The marginal effect of BMI now amounts to CHF 297 (Model FC). This coefficient is somewhat larger than in the model with all control variables, but we cannot reject the hypothesis that the two coefficients are identical.

Fourth, instead of a GLM, we perform a Poisson regression using a generalised methods of moments (GMM) estimator. As a main advantage over the CF approach, the Poisson GMM estimator relies on fewer distributional assumptions and is therefore likely to be the more efficient approach. The respective AME is CHF 222, which is slightly lower than the CF estimate, but not significantly.³⁰

Fifth, we calculate standardised values of the residual u_{it} , which are then included in the second stage regression (Model SR). According to Wooldridge (2014), we can obtain standardised values of the residual term (u'_{it}) by calculating

30 We have demonstrated that HCE follows a gamma distribution. Thus the estimates from the Poisson regression are likely biased as the model is based on strict distribution assumptions.

$$\hat{u}_{ii}^s = \frac{[BMI_{ii} - \hat{g}(\mathbf{x}_{ii}, \mathbf{z}_{ii})]}{[\hat{h}(\mathbf{x}_{ii}, \mathbf{z}_{ii})]^{1/2}} = \frac{\hat{u}_{ii}}{[\hat{h}(\mathbf{x}_{ii}, \mathbf{z}_{ii})]^{1/2}},$$

where $\hat{h}(\cdot)$ is the estimated variance function of *BMI*. WOOLDRIDGE (2014) suggests using $\exp(\mathbf{x}_{ii}, \mathbf{z}_{ii})$ as the variance function of the left-hand variable.³¹ With the inclusion of the standardised value of u_{ii} , the predicted marginal effect of weight is reduced to CHF 220. As with the robustness checks above, though, the effect is not significantly different from the standard specification.

In addition, we take advantage of the longitudinal structure of our dataset to estimate a panel model first with random (RE) and then with fixed effects (FE). We use the zero-skewed log of HCE as the dependent variable.³² To estimate the FE model, we also have to drop all variables that are constant over time (e.g., sex, canton of residence). The marginal effect of BMI in the RE model is approximately +1.72% and amounts to CHF 36 at the sample means. This figure is relatively close to the non-IV estimate of CHF 34 in the GLM regression. The FE estimator implicitly controls for the bias caused by omitted variables that are constant over time. It thus only captures a very specific form of endogeneity. Hence, not surprisingly, the marginal effect reported by the FE model lies between the non-IV regression estimates and the prediction from the CF model. Accordingly, a 1-unit increase in BMI is followed by a rise in costs of about CHF 48.

Finally, to test for the sensitivity of our instrument, we use a different set of IVs to instrument the children's BMI. Instead of using the mean of the parents' BMI (AME: CHF 281), we alternatively take the BMI of the mother (AME: CHF 319), the father's BMI (AME: CHF 220), and the separate BMI scores of both parents (AME: CHF 281). Even though there is some variation in the point estimates, they are not significantly different from each other.

We conclude that the marginal effect of weight on health-care expenditure tends to be robust against changes in the model specification. Across all alternative specifications with IVs, the effects of BMI on costs are very much the same (CHF 220–CHF 297).

31 As a main advantage over the non-standardised residuals, \hat{u}_{ii}^s accounts for potential heteroscedasticity in the error term (WOOLDRIDGE, 2014).

32 We transform the dependent variable, $\ln HCE = \ln(HCE + k)$, choosing k so that the skewness of $\ln HCE$ is zero (Royston and Lambert, 2011).

4.1.2 Falsification Tests

In order to test the validity of our instruments, we also perform a set of IV tests that are often referred to as falsification tests. If there are unobserved household characteristics that affect the weight of our household members, the observed correlation between BMI of parents and children may partially be caused by non-genetic factors (e.g., due to a similar diet, exercise habits). As a result, the validity of our instruments would become questionable. To cope with this issue, we take advantage of the fact that there are non-relatives in some households (e.g., foster children, friends, flatmates etc.). If we found a correlation of BMI between non-relatives, this would suggest that the shared household environment may play a role. In order to test for shared environmental factors, we instrument the BMI of the parents with the mean BMI of the non-relatives in the household and *vice versa*. We then re-estimate the auxiliary regression (2) for this subsample ($N=793$) (see Table 7 in the Appendix for the complete regression output). As it can be seen from the F -statistic, the mean BMI of non-relatives tends to be a rather weak instrument ($F=1.361$). In fact, we cannot reject the joint hypothesis that the coefficients of all three instruments are zero ($p=0.254$). This finding suggests that the shared home has a negligible effect on the weight of the household members.

As a second check, we analyse how BMI is correlated with certain health conditions. If we noticed that the BMI was strongly correlated with a condition it is supposed to be unrelated to, our instruments would be subject to severe validity problems. In other words, if our falsification tests yield implausible results, this indicates that our IV method is not sufficiently specific: It fails to separate true obesity effects from spurious correlations, resulting from a lack of specifications in the model (e.g., insufficient control variables for the local environment and individual characteristics) (CAWLEY and RUHM, 2011).

The results from the four falsification tests are reported in Table 4.³³ First, we would expect heavier people to be more prone to back problems (see SHIRI et al., 2009). This link is actually confirmed by the logit regression: Heavier people are more likely to report back problems than people with lower BMI scores. Second, the same is true for having experienced weakness or weariness over the last 4 weeks. BMI is associated with an increase in the odds of reporting a positive outcome. In both regressions, the effects are more pronounced within the CF model. Third, as regards sleeping problems, several studies have found that inadequate sleep (i.e., insufficient length or quality of sleep) can result in weight

33 As the dependent variables are indicators, we use the logit specification.

gain (see, e.g., GANGWISCH et al., 2005). Conversely, obesity may also have an impact on the perceived quality of sleep. However, no such direct link has been established in the literature. Of course, this does not rule out the possibility that overweight actually has an immediate impact on sleep. In fact, however, we do not find a causal effect of BMI on the risk of suffering from insomnia. The non-IV regression approach, though, reports a significantly positive correlation between weight and *SLEEPING*. This clearly points to the aforementioned simultaneity issue: Inadequate sleep may put people at risk of becoming overweight or obese. Therefore, the naive estimator predicts a positive relationship between sleeping issues and body weight. Fourth, BMI is a potential risk factor for headaches. In a study that questioned more than 30,000 participants, BIGAL, LIBERMAN, and LIPTON (2006) have demonstrated that obesity is a significant risk factor for transformed migraine and a moderate risk factor for tension-type headache. Using the CF approach, our model indicates that BMI, as supposed, is a strong predictor for reporting recent headaches. In turn, the naive approach fails to identify a direct link between weight and headaches, which contradicts the findings described by BIGAL, LIBERMAN, and LIPTON (2006).

Table 4: Logit Regression of BMI on Reported Health Issues

	<i>BACKPAIN</i>	<i>WEAK</i>	<i>SLEEPING</i>	<i>HEAD</i>
BMI	0.097* (0.051)	0.099** (0.043)	0.049 (0.051)	0.097** (0.045)
γ	-0.082 (0.051)	-0.086** (0.044)	-0.022 (0.052)	-0.100** (0.047)
<i>BMI</i> (non-IV)	0.019* (0.010)	0.018* (0.010)	0.028*** (0.010)	0.003 (0.010)

Notes: Cluster robust standard errors are given in parentheses; The models have been estimated with the same covariates as in the basic models, except for *BACKPAIN*, *WEAK*, *SLEEPING*, and *HEAD*; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Even though falsification tests are unable to give a final answer, they can provide reassurance with regard to the adequacy of the instruments. Taken together, the performed falsification tests indicate that the BMI of relatives (and its polynomials) is specific enough to separate obesity-related effects from unobserved heterogeneity.

5. Conclusion

In this paper, we have provided new insights into the relationship between obesity and the use of inpatient and outpatient care in Switzerland. Individual data were drawn from the SHP, providing a total sample size of 4,968 individuals over the period 2004–2011 (16,317 observations). Using a combined two-stage approach, which adjusts for measurement errors and endogeneity, we estimated the impact of BMI on the number of outpatient visits, the number of hospital days, and health-care costs. In a first step, we calculated error-adjusted height and weight for each individual on the basis of self-reported data using the CE approach suggested by CAWLEY and BURKHAUSER (2006). To account for omitted variables and simultaneity, we then applied IV regression to calculate the causal effect of overweight on health-care use. To obtain unbiased results, we employed the CF method suggested by TERZA, BASU, and RATHOUZ (2008). The mean BMI of the parents thereby served as an instrument for the BMI of the biological children, and vice versa.

Our results suggest that unadjusted models highly underestimate the impact of obesity on the demand for inpatient and outpatient treatment. Controlling for individual characteristics, the state of health, and regional differences, the CF model predicts that a 1-point increase in BMI raises HCE by CHF 253. In contrast, the naive non-IV regression suggests an increase of only CHF 34. Several robustness checks confirm that the marginal effect of BMI is likely to range between CHF 220 and CHF 294. Our model also predicts that total HCE could be reduced by approximately 4.7% if all people in the sample who are classified as overweight or obese managed to reduce their body weight to the threshold value for normal weight.

The sector-specific results further show that the considerable difference between the CF and the non-IV regression model are mainly driven by the number of hospital days. Accordingly, we severely underestimate the effect of weight on the use of inpatient care. In non-IV regression, BMI tends to have no effect at all on the number of inpatient days. This result is reversed under the CF approach. Here, we not only observe a substantial causal effect of weight on the number of days (AME: +0.12 days), but we also find evidence that our variable of interest might be subject to a significant amount of endogeneity ($p < 0.05$). We argue that reverse causation probably causes the difference in the marginal effects between the two models.

We further find some remarkable effects from our set of covariates. Subjective health status, on the one hand, is an excellent predictor for the level of health-care costs. People who perceived their health status as being *very well* exhibit monthly

costs of CHF 124. While *fair* is associated with monthly costs of CHF 293, individuals who are in a *very bad* state of health face, on average, health-related costs of CHF 1,219.

This paper has several limitations to consider. First, IV regression always raises concern about the validity of the instruments, as it is impossible to directly test whether the instrument is exogenous. The shared household environment is a potential threat to using a relative's BMI as an instrument. Still, the testable hypotheses with regard to our instrumental variables do not find any severe misspecification. Moreover, other studies that rely on similar IVs for BMI report no validity problems (CAWLEY and MEYERHOEFER, 2012). A more latent threat to the validity of our findings is the potential correlation of reporting errors in BMI among household members (i.e., reporting errors associated with the BMI of parents and children). Even though we applied the CE approach, this method is unlikely to fully account for this unobserved correlation. Consequently, the CF estimates may still suffer from minor biases due to this.³⁴

As a second shortcoming, we only observe BMI as a measure of overweight. This measure of overweight has been criticised due to its inability to distinguish between muscle and fat (see, e.g., DALTON et al., 2003). Thus, for instance, people who have a large muscle mass, such as sports persons and athletes may be categorised as overweight or obese, while inactive individuals with a high percentage of body fat may exhibit a desirable BMI. However, recent literature still lacks more accurate predictors of obesity.

Third, our analysis is limited to the use of outpatient care (GPs, specialists) and hospital services. Therefore, we can only estimate the extent to which obesity raises the demand for care in these sectors. We cannot predict the impact of excess weight on other areas of health care (e.g., drug expenditure, dentists, long-term care, nursing care).

Finally, limited by our estimation strategy, we can only report estimates for multi-person households that consist of at least one parent and one biological child. Thus, it remains questionable whether our findings also apply to the entire Swiss population. Nonetheless, since the study population is relatively young and many obesity-related diseases show up later in life, we expect the actual size of the effect to be far more pronounced.

34 Nevertheless, we consider the issue of proxy-error reporting to be rather small and irrelevant, since, unlike other studies, the BMI of the children in our study is reported by themselves. Thus the intra-household dependence of reporting errors of weight and height will be mostly random.

In conclusion, our findings suggest that the external costs of obesity that operate through social insurance are likely to be greater than previously estimated (see SCHMID et al., 2005; HÄUSSLER, 2014). Moreover, the first stage regression makes clear that only a minor part of the variance in BMI can be explained by genetics. The partial R^2 of the instruments amounts to 0.052, which predicts that only about 5% of the variance in BMI can be attributed to the variance of our instruments. What is more, the R^2 of the complete first stage comes to 0.289. Put differently, more than 70% of the observed variance in BMI across individuals remains unexplained by genetics, age, sex, educational attainment, the health status, etc. While the differences in weight that are linked to genetic predisposition do not call for public action, the presence of moral hazard leads to economic inefficiencies. We expect the inefficiencies that are caused by *ex-ante* moral hazard to be substantial, as our findings suggest that the unobserved behaviour of the individual plays a major role in the determination of his body weight. Thus, from an economic perspective, the situation could be alleviated by offering the individuals proper incentives to internalise the social costs of overweight.

As Swiss insurers are not allowed to charge risk-rated premiums, other measures can be implemented to internalise these externalities. BATTLE and BROWNELL (1996), for instance, have suggested subsidising healthy foods, while imposing a *fat tax* on energy-dense foods that are low in nutritional value. In addition, the tax revenue could then be used to promote a healthier lifestyle for the population by providing more opportunities for physical activity. However, empirical research has provided only little evidence that food taxes affect weight (see, e.g., SCHROETER, LUSK, and TYNER, 2008; FLETCHER, FRISVOLD, and TEFFT, 2010). It is widely believed that the efficacy of taxation mainly depends on how regulators can anticipate the behaviour of the target audience. For instance, imposing an arbitrary taxation on some energy-dense foods may not help reduce net calorie consumption, as consumers can substitute these foods with (untaxed) alternative products that are high in calories. Moreover, as genetic factors may be an important co-driver of obesity, the potential effect of taxation and subsidies is rather limited.

Appendix

Table 5: Descriptive Statistics of all Variables

Variable	Definition	Mean	SD	Min	Max
<i>HCE</i>	Estimated health expenditure on physician and inpatient care (past 12 months, in 2012 CHF)	2,108	9,645	0	373,932
<i>DOCVIS</i>	Number of physician visits over the past 12 months	2.88	6.10	0	200
<i>HOSPDAYS</i>	Number of inpatient days over the past 12 months	0.71	5.16	0	200
<i>BMI</i>	Reported body mass index ($BMI = kg/m^2$)	23.11	3.95	12.12	63.18
Controls					
<i>FEMALE</i>	Person is female	0.523	0.499	0	1
<i>AGE</i>	Age of the person (Basic category: 15–20 years)				
<i>20_30</i>	Age between 20 and 29	0.159	0.365	0	1
<i>30_50</i>	Age between 30 and 49	0.287	0.453	0	1
<i>50_70</i>	Age between 50 and 69	0.234	0.423	0	1
<i>70up</i>	Person is older than 69	0.006	0.074	0	1
<i>20_30_female</i>	Person is female, age 20–29	0.070	0.254	0	1
<i>30_50_female</i>	Person is female, age 30–49	0.179	0.383	0	1
<i>50_70_female</i>	Person is female, age 50–69	0.115	0.319	0	1
<i>70up_female</i>	Person is female, older than 69	0.002	0.049	0	1
<i>SWISS</i>	Person is Swiss	0.946	0.225	0	1
<i>CIVSTAT</i>	Civil status (Basic category: <i>single</i>)				
<i>married</i>	Person is married	0.441	0.497	0	1
<i>separated</i>	Person is separated or divorced	0.069	0.253	0	1
<i>NEWBORN</i>	Woman has given birth to a child within 12 months	0.001	0.037	0	1
<i>ACCIDENT</i>	Person has suffered an accident	0.052	0.223	0	1
<i>CHRONIC</i>	Person suffers from a chronic health condition	0.269	0.443	0	1
<i>BACKPAIN</i>	Person has been suffering from back pain (past 4 weeks)	0.416	0.493	0	1

Variable	Definition	Mean	SD	Min	Max
<i>WEAK</i>	Person has been suffering from weakness or weariness (past 4 weeks)	0.447	0.497	0	1
<i>SLEEPING</i>	Person has been suffering from sleeping problems (past 4 weeks)	0.301	0.459	0	1
<i>HEAD</i>	Person has been suffering from headaches (past 4 weeks)	0.379	0.485	0	1
<i>HSTATUS</i>	Subjective health status (Basic category: <i>very well</i>)				
<i>well</i>	Health status is "well"	0.656	0.475	0	1
<i>fair</i>	Health status is "fair"	0.092	0.289	0	1
<i>bad</i>	Health status is "bad"	0.013	0.111	0	1
<i>very bad</i>	Health status is "very bad"	0.001	0.030	0	1
<i>EDUCATION</i>	Educational attainment (Basic category: (1) <i>incomplete schooling</i>)				
2	Compulsory school	0.177	0.381	0	1
3	Domestic science course	0.023	0.150	0	1
4	General training school	0.013	0.115	0	1
5	Apprenticeship	0.246	0.430	0	1
6	Full-time vocational school	0.034	0.182	0	1
7	Bachelor, maturity	0.128	0.334	0	1
8	Vocational high school, master certificate	0.049	0.215	0	1
9	Technical or vocational school	0.024	0.154	0	1
10	Vocational high school ETS, HTL	0.041	0.198	0	1
11	University	0.097	0.296	0	1
<i>EMPLOY</i>	Employment status (Basic category: <i>employed</i>)				
<i>nonworking</i>	Not in labour force (e.g., student, pensioner)	0.492	0.500	0	1
<i>unemployed</i>	Person is unemployed	0.011	0.102	0	1
<i>URBAN</i>	Degree of urbanisation (Basic category: <i>rural</i>)				
<i>urban</i>	Urban community, city	0.203	0.402	0	1
<i>suburban</i>	Suburban community	0.480	0.500	0	1
<i>LANG</i>	First language (Basic category: <i>German</i>)				
<i>French</i>	Person speaks French	0.249	0.450	0	1
<i>Italian</i>	Person speaks Italian	0.032	0.177	0	1

N = 16,317

Table 6: Output from the First Stage Regression

<i>BMI</i> (Gaussian, log link)	Coefficient	SE
Instruments		
<i>BMI_{IV}</i>	0.011***	(0.000)
<i>BMI_{IV}²</i>	-0.007	(0.007)
<i>BMI_{IV}³</i>	-0.074*	(0.041)
Covariates		
<i>FEMALE</i>	0.014***	(0.005)
<i>AGE</i>		
<i>20_30</i>	0.073***	(0.006)
<i>30_50</i>	0.158***	(0.012)
<i>50_70</i>	0.135***	(0.013)
<i>70up</i>	0.117***	(0.022)
<i>20_30_female</i>	-0.033***	(0.008)
<i>30_50_female</i>	-0.115***	(0.006)
<i>50_70_female</i>	-0.082***	(0.006)
<i>70up_female</i>	-0.031	(0.028)
<i>SWISS</i>	-0.004	(0.005)
<i>CIVSTAT</i>		
<i>married</i>	0.018	(0.011)
<i>separated</i>	0.008	(0.012)
<i>NEWBORN</i>	-0.053	(0.034)
<i>ACCIDENT</i>	0.006	(0.005)
<i>CHRONIC</i>	0.014***	(0.003)
<i>BACKPAIN</i>	0.010***	(0.002)
<i>WEAK</i>	0.002	(0.002)
<i>SLEEPING</i>	0.000	(0.003)
<i>HEAD</i>	-0.006**	(0.002)
<i>HSTATUS</i>		
<i>well</i>	0.013***	(0.003)
<i>fair</i>	0.031***	(0.005)
<i>bad</i>	0.047***	(0.010)
<i>very bad</i>	-0.001	(0.034)
<i>EDUCATION</i>		
2	0.038***	(0.004)
3	0.036***	(0.009)
4	0.014	(0.011)

<i>BMI</i> (Gaussian, log link)	Coefficient	SE
5	0.047***	(0.006)
6	0.000	(0.008)
7	0.009	(0.006)
8	0.047***	(0.007)
9	0.007	(0.009)
10	0.024***	(0.008)
11	0.024***	(0.006)
<i>EMPLOY</i>		
<i>nonworking</i>	0.002	(0.004)
<i>unemployed</i>	0.052***	(0.011)
<i>URBAN</i>		
<i>urban</i>	-0.022***	(0.004)
<i>suburban</i>	-0.019***	(0.003)
<i>LANG</i>		
<i>French</i>	-0.037***	(0.007)
<i>Italian</i>	0.014	(0.012)
Year fixed effects		yes
Canton fixed effects		yes
R^2		0.289
Shea's partial R^2		0.052
$F(3,16241)$ (<i>p-value</i>)		303.693 (0.000)
$F(3,4967)$ with individual clusters (<i>p-value</i>)		73.050 (0.000)
$F(3,1540)$ with household clusters (<i>p-value</i>)		41.431 (0.000)

Table 7: Explaining the BMI of Non-Relatives

<i>BMI</i> (Gaussian, log link)	Coefficient	SE
Instruments		
BMI_{IV}	-0.050	(0.038)
BMI_{IV}^2	0.002	(0.001)
BMI_{IV}^3	0.000	(0.000)
Covariates		

<i>BMI</i> (Gaussian, log link)	Coefficient	SE
<i>FEMALE</i>	0.034	(0.033)
<i>AGE</i>		
20_30	0.126***	(0.041)
30_50	0.187***	(0.047)
50_70	0.210***	(0.056)
70up	0.155**	(0.066)
20_30_female	-0.077*	(0.040)
30_50_female	-0.151***	(0.042)
50_70_female	-0.141***	(0.054)
70up_female	0.017	(0.080)
<i>SWISS</i>	0.011	(0.032)
<i>CIV STAT</i>		
<i>married</i>	0.047	(0.031)
<i>separated</i>	-0.002	(0.032)
<i>NEWBORN</i>	-0.200**	(0.087)
<i>ACCIDENT</i>	0.018	(0.029)
<i>CHRONIC</i>	0.008	(0.015)
<i>BACKPAIN</i>	0.036**	(0.015)
<i>WEAK</i>	0.008	(0.013)
<i>SLEEPING</i>	0.007	(0.016)
<i>HEAD</i>	0.037***	(0.014)
<i>HSTATUS</i>		
<i>well</i>	0.012	(0.013)
<i>fair</i>	0.024	(0.023)
<i>bad</i>	-0.005	(0.042)
<i>very bad</i>	-0.086*	(0.047)
<i>EDUCATION</i>		
2	0.088***	(0.032)
3	0.056	(0.097)
4	-0.018	(0.059)
5	0.108**	(0.043)
6	0.097*	(0.056)
7	-0.008	(0.044)
8	0.081	(0.051)
9	0.013	(0.064)
10	0.048	(0.056)
11	0.049	(0.043)

<i>BMI</i> (Gaussian, log link)	Coefficient	SE
<i>EMPLOY</i>		
<i>nonworking</i>	0.046*	(0.024)
<i>unemployed</i>	0.116	(0.074)
<i>URBAN</i>		
<i>urban</i>	-0.013	(0.029)
<i>suburban</i>	-0.006	(0.033)
<i>LANG</i>		
<i>French</i>	-0.048	(0.036)
<i>Italian</i>	0.266***	(0.084)
Year fixed effects		yes
Canton fixed effects		yes
R^2		0.323
Shea's partial R^2		0.013
$F(3,432)$		1.361
(<i>p-value</i>)		(0.254)

Table 8: Complete Output from the Cost Regression

<i>HCE</i> (Gamma, log link)	Non-IV		CF	
<i>BMI</i>	0.015*	(0.009)	0.115***	(0.041)
γ			-0.104**	(0.045)
<i>FEMALE</i>	0.135	(0.130)	0.109	(0.125)
<i>AGE</i>				
<i>20_30</i>	-0.089	(0.136)	-0.240	(0.166)
<i>30_50</i>	-0.016	(0.314)	-0.411	(0.351)
<i>50_7</i>	0.375	(0.355)	0.008	(0.362)
<i>70up</i>	0.850*	(0.445)	0.507	(0.484)
<i>20_30_female</i>	0.093	(0.192)	0.177	(0.185)
<i>30_50_female</i>	0.031	(0.170)	0.322*	(0.183)
<i>50_70_female</i>	-0.550***	(0.209)	-0.330*	(0.181)
<i>70up_female</i>	0.328	(0.597)	-0.320	(0.589)
<i>SWISS</i>	0.183	(0.115)	0.204*	(0.115)
<i>CIVSTAT</i>				
<i>married</i>	0.297	(0.252)	0.262	(0.252)
<i>separated</i>	0.382	(0.269)	0.334	(0.267)
<i>NEWBORN</i>	0.179***	(0.330)	0.229***	(0.323)

<i>HCE</i> (Gamma, log link)	Non-IV		CF	
<i>ACCIDENT</i>	1.207***	(0.077)	1.203***	(0.078)
<i>CHRONIC</i>	0.874***	(0.068)	0.827***	(0.067)
<i>BACKPAIN</i>	0.081	(0.055)	0.054	(0.053)
<i>WEAK</i>	0.118**	(0.058)	0.116**	(0.058)
<i>SLEEPING</i>	0.226***	(0.076)	0.228***	(0.074)
<i>HEAD</i>	-0.007	(0.058)	0.004	(0.057)
<i>HSTATUS</i>				
<i>well</i>	0.227***	(0.068)	0.197***	(0.067)
<i>fair</i>	0.942***	(0.089)	0.860***	(0.094)
<i>bad</i>	1.775***	(0.204)	1.664***	(0.214)
<i>very bad</i>	2.276***	(0.665)	2.293***	(0.668)
<i>EDUCATION</i>				
2	0.041	(0.138)	-0.060	(0.121)
3	-0.403**	(0.176)	-0.505***	(0.176)
4	-0.673***	(0.237)	-0.726***	(0.235)
5	-0.177	(0.164)	-0.302*	(0.160)
6	-0.338*	(0.197)	-0.350*	(0.201)
7	-0.152	(0.191)	-0.188	(0.186)
8	-0.264	(0.201)	-0.385*	(0.197)
9	-0.126	(0.248)	-0.147	(0.251)
10	-0.281	(0.213)	-0.337	(0.213)
11	-0.176	(0.194)	-0.231	(0.191)
<i>EMPLOY</i>				
<i>nonworking</i>	0.055	(0.095)	0.041	(0.090)
<i>unemployed</i>	0.398*	(0.241)	0.241	(0.246)
<i>URBAN</i>				
<i>urban</i>	0.008	(0.091)	0.074	(0.086)
<i>suburban</i>	0.057	(0.090)	0.115	(0.086)
<i>LANG</i>				
<i>French</i>	0.135	(0.119)	0.228*	(0.125)
<i>Italian</i>	0.353	(0.230)	0.276	(0.223)
Year fixed effects	<i>yes</i>		<i>yes</i>	
Canton fixed effects	<i>yes</i>		<i>yes</i>	

Table 9: Complete Output from the Outpatient Regression

<i>DOCVIS</i> (Negbin, log link)	Non-IV		CF	
<i>BMI</i>	0.013**	(0.005)	0.051**	(0.020)
γ			-0.040*	(0.022)
<i>FEMALE</i>	0.173***	(0.066)	0.166**	(0.065)
<i>AGE</i>				
<i>20_30</i>	-0.130*	(0.075)	-0.184**	(0.083)
<i>30_50</i>	-0.298*	(0.167)	-0.434**	(0.177)
<i>50_70</i>	-0.036	(0.182)	-0.161	(0.192)
<i>70up</i>	0.301	(0.240)	0.172	(0.260)
<i>20_30_female</i>	0.102	(0.091)	0.124	(0.092)
<i>30_50_female</i>	0.156	(0.099)	0.250**	(0.106)
<i>50_70_female</i>	-0.123	(0.103)	-0.057	(0.106)
<i>70up_female</i>	-0.483	(0.321)	-0.461	(0.325)
<i>SWISS</i>	0.101	(0.064)	0.106*	(0.063)
<i>CIVSTAT</i>				
<i>married</i>	0.230	(0.152)	0.228	(0.152)
<i>separated</i>	0.260	(0.162)	0.252	(0.161)
<i>NEWBORN</i>	0.570	(0.357)	0.619*	(0.346)
<i>ACCIDENT</i>	0.857***	(0.047)	0.853***	(0.047)
<i>CHRONIC</i>	0.765***	(0.037)	0.749***	(0.037)
<i>BACKPAIN</i>	0.064**	(0.030)	0.057*	(0.030)
<i>WEAK</i>	0.153***	(0.030)	0.151***	(0.030)
<i>SLEEPING</i>	0.114***	(0.034)	0.115***	(0.034)
<i>HEAD</i>	0.082***	(0.030)	0.086***	(0.031)
<i>HSTATUS</i>				
<i>well</i>	0.209***	(0.043)	0.201***	(0.042)
<i>fair</i>	0.774***	(0.061)	0.747***	(0.058)
<i>bad</i>	1.370***	(0.102)	1.330***	(0.106)
<i>very bad</i>	0.970***	(0.210)	0.970***	(0.207)
<i>EDUCATION</i>				
2	0.012	(0.060)	-0.015	(0.061)
3	-0.046	(0.113)	-0.075	(0.114)
4	-0.265*	(0.150)	-0.277*	(0.147)
5	-0.039	(0.079)	-0.076	(0.081)
6	-0.160	(0.104)	-0.157	(0.105)
7	0.024	(0.087)	0.020	(0.087)
8	-0.061	(0.103)	-0.097	(0.104)
9	0.107	(0.157)	0.112	(0.156)

<i>DOCVIS</i> (Negbin, log link)	Non-IV		CF	
<i>I0</i>	-0.071	(0.113)	-0.084	(0.116)
<i>I1</i>	-0.027	(0.096)	-0.038	(0.098)
<i>EMPLOY</i>				
<i>nonworking</i>	0.029	(0.054)	0.027	(0.053)
<i>unemployed</i>	0.097	(0.132)	0.049	(0.134)
<i>URBAN</i>				
<i>urban</i>	0.059	(0.053)	0.081	(0.053)
<i>suburban</i>	0.037	(0.051)	0.057	(0.049)
<i>LANG</i>				
<i>French</i>	0.195**	(0.084)	0.226***	(0.085)
<i>Italian</i>	0.118	(0.128)	0.103	(0.129)
Year fixed effects		<i>yes</i>		<i>yes</i>
Canton fixed effects		<i>yes</i>		<i>yes</i>

Table 10: Complete Output from the Inpatient Regression

<i>HOSPDAYS</i> (Negbin, log link)	Non-IV		CF	
<i>BMI</i>	0.000	(0.017)	0.165**	(0.065)
γ			-0.171**	(0.070)
<i>FEMALE</i>	-0.019	(0.232)	-0.056	(0.221)
<i>AGE</i>				
<i>20_30</i>	0.012	(0.249)	-0.248	(0.291)
<i>30_50</i>	0.460	(0.544)	-0.278	(0.601)
<i>50_70</i>	1.010*	(0.580)	0.312	(0.602)
<i>70up</i>	1.740**	(0.722)	1.074	(0.769)
<i>20_30_female</i>	-0.015	(0.392)	0.128	(0.372)
<i>30_50_female</i>	-0.004	(0.295)	0.487*	(0.293)
<i>50_70_female</i>	-0.874**	(0.344)	-0.503*	(0.298)
<i>70up_female</i>	0.065	(1.014)	0.079	(0.950)
<i>SWISS</i>	0.096	(0.205)	0.124	(0.205)
<i>CIV_STAT</i>				
<i>married</i>	0.166	(0.604)	0.115	(0.597)
<i>separated</i>	0.179	(0.629)	0.107	(0.618)
<i>NEWBORN</i>	1.783***	(0.384)	2.055***	(0.385)
<i>ACCIDENT</i>	1.381***	(0.121)	1.373***	(0.122)
<i>CHRONIC</i>	0.862***	(0.111)	0.786***	(0.111)

<i>HOSPDAYS</i> (Negbin, log link)	Non-IV		CF	
<i>BACKPAIN</i>	0.042	(0.098)	-0.001	(0.095)
<i>WEAK</i>	0.157	(0.103)	0.153	(0.103)
<i>SLEEPING</i>	0.315**	(0.126)	0.324***	(0.124)
<i>HEAD</i>	-0.177*	(0.107)	-0.163	(0.107)
<i>HSTATUS</i>				
<i>well</i>	0.346***	(0.118)	0.292**	(0.118)
<i>fair</i>	1.223***	(0.148)	1.081***	(0.155)
<i>bad</i>	2.018***	(0.309)	1.823***	(0.320)
<i>very bad</i>	2.830**	(1.141)	2.876**	(1.308)
<i>EDUCATION</i>				
2	0.106	(0.232)	-0.060	(0.203)
3	-0.796**	(0.373)	-0.967***	(0.366)
4	-0.932*	(0.550)	-1.044*	(0.540)
5	-0.223	(0.289)	-0.434	(0.278)
6	-0.475	(0.386)	-0.472	(0.392)
7	-0.201	(0.348)	-0.254	(0.338)
8	-0.363	(0.344)	-0.569*	(0.336)
9	-0.229	(0.433)	-0.222	(0.435)
10	-0.428	(0.366)	-0.508	(0.371)
11	-0.275	(0.347)	-0.350	(0.340)
<i>EMPLOY</i>				
<i>nonworking</i>	0.166	(0.168)	0.147	(0.162)
<i>unemployed</i>	0.555	(0.410)	0.267	(0.422)
<i>URBAN</i>				
<i>urban</i>	-0.225	(0.163)	-0.099	(0.152)
<i>suburban</i>	-0.040	(0.157)	0.066	(0.148)
<i>LANG</i>				
<i>French</i>	0.075	(0.179)	0.239	(0.192)
<i>Italian</i>	0.606*	(0.367)	0.455	(0.357)
Year fixed effects		<i>yes</i>		<i>yes</i>
Canton fixed effects		<i>yes</i>		<i>yes</i>

Table 11: Estimated Marginal Effects from the Robustness Checks

	Average marginal effect	95% confidence interval	
Alternative specifications			
BMI2	241.274	51.282	431.266
BMI2 (non-IV)	24.586	-16.518	65.690
BMI3	249.981	44.304	455.657
BMI3 (non-IV)	12.538	-36.736	61.812
BMI IV1	258.231	73.258	443.205
BMI IV2	246.873	70.600	423.147
FC	296.855	108.413	485.296
FC (non-IV)	34.234	-17.518	85.986
GMM	222.367	34.223	410.512
SR	219.807	55.004	384.610
RE	36.320	15.153	57.487
FE	48.111	-3.485	99.706
Different instruments for the BMI of children			
Non-IV ($N = 7882$)	63.561	14.469	112.653
Mean BMI of parents (and its polynomials) ($N = 7882$)	281.275	80.749	481.802
BMI of the biological mother ($N = 7791$)	319.482	77.911	561.053
BMI of the biological father ($N = 4833$)	219.711	-42.221	481.643
BMI of mother and father ($N = 4742$)	281.218	66.025	496.411

Figure 2: Age Distribution of Parents and their Biological Children

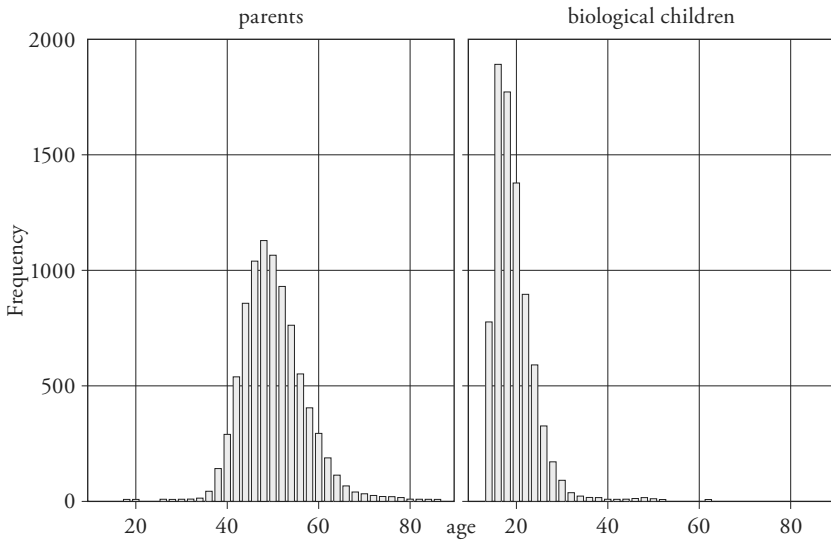
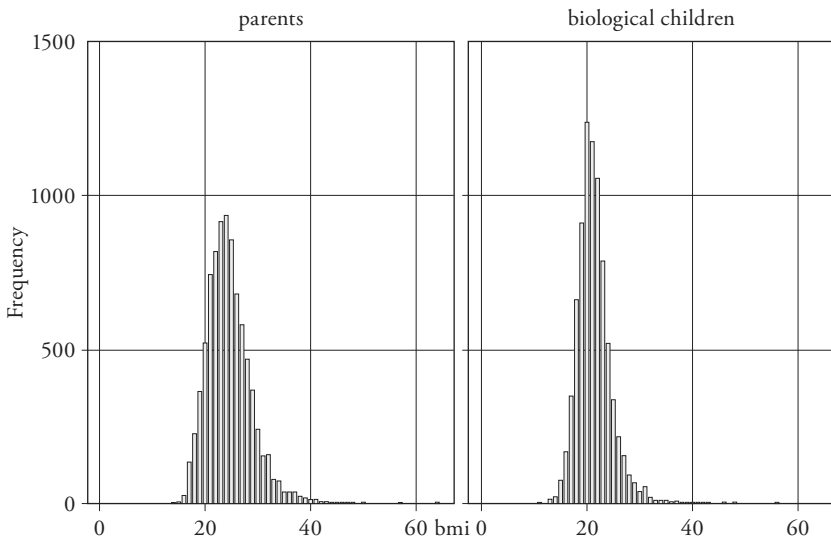


Figure 3: BMI Distribution of Parents and their Biological Children



References

- ARMSTRONG, BENEDICT G., and DAVID OAKES (1982), "Effects of Approximation in Exposure Assessments on Estimates of Exposure-Response Relationships", *Scandinavian Journal of Work, Environment and Health*, 8(suppl. 1), pp. 20–23.
- AULD, M. CHRISTOPHER (2011), "Effect of Large-Scale Social Interactions on Body Weight", *Journal of Health Economics*, 30(2), pp. 303–316.
- BAGUST, ADRIAN, and TOM WALLEY (2000), "An Alternative to Body Mass Index for Standardizing Body Weight for Stature", *QJM*, 93(9), pp. 589–596.
- BATTLE, E. KATHERINE, and KELLY D. BROWNELL (1996), "Confronting a Rising Tide of Eating Disorders and Obesity: Treatment vs. Prevention and Policy", *Addictive Behaviors*, 21(6), pp. 755–765.
- BIGAL, MARCELO E., JOSHUA N. LIBERMAN, and RICHARD B. LIPTON (2006), "Obesity and Migraine: A Population Study", *Neurology*, 66(4), pp. 545–550.
- BUNTIN, MELINDA BEEUWKES, and ALAN M. ZASLAVSKY (2004), "Too Much Ado about Two-Part Models and Transformation? Comparing Methods of Modeling Medicare Expenditures", *Journal of Health Economics*, 23(3), pp. 525–542.
- BURKHAUSER, RICHARD V., and JOHN CAWLEY (2008), "Beyond BMI: The Value of More Accurate Measures of Fatness and Obesity in Social Science Research", *Journal of Health Economics*, 27(2), pp. 519–529.
- CAMERON, A. COLIN, and PRAVIN K. TRIVEDI (2005), *Microeconometrics: Methods and Applications*, Cambridge, UK: Cambridge University Press.
- CAMERON, A. COLIN, and PRAVIN K. TRIVEDI (2013), *Regression Analysis of Count Data*, 2nd edn., Cambridge, UK: Cambridge University Press.
- CAWLEY, JOHN (2000), "An Instrumental Variables Approach to Measuring the Effect of Body Weight on Employment Disability", *Health Serv. Res.*, 35(5 Pt 2), pp. 1159–1179.
- CAWLEY, JOHN, and RICHARD V. BURKHAUSER (2006), "Beyond BMI: The Value of More Accurate Measures of Fatness and Obesity in Social Science Research", Working Paper 12291, National Bureau of Economic Research.
- CAWLEY, JOHN, and CHAD MEYERHOFER (2012), "The Medical Care Costs of Obesity: An Instrumental Variables Approach", *Journal of Health Economics*, 31(1), pp. 219–230.
- CAWLEY, JOHN, and CHRISTOPHER J. RUHM (2011), "The Economics of Risky Health Behaviors", in *Handbook of Health Economics*, Mark V. Pauly, Thomas G. McGuire and Pedro P. Barros, eds., vol. 2, chap. 3, pp. 95–199, Philadelphia, PA: Elsevier.

- DALTON, M., A.J. CAMERON, P.Z. ZIMMET, J.E. SHAW, D. JOLLEY, D.W. DUNSTAN, and T.A. WELBORN (2003), "Waist Circumference, Waist-Hip Ratio and Body Mass Index and Their Correlation with Cardiovascular Disease Risk Factors in Australian Adults", *Journal of Internal Medicine*, 254(6), pp. 555–563.
- DEB, PARTHA, WILLIAM T. GALLO, PADMAJA AYYAGARI, JASON M. FLETCHER, and JODY L. SINDELAR (2011), "The Effect of Job Loss on Overweight and Drinking", *Journal of Health Economics*, 30(2), pp. 317–327.
- DOOLEY, DAVID, RALPH CATALANO, and RICHARD HOUGH (1992), "Unemployment and Alcohol Disorder in 1910 and 1990: Drift Versus Social Causation", *Journal of Occupational and Organizational Psychology*, 65(4), pp. 277–290.
- FINKELSTEIN, ERIC A., IAN C. FIEBELKORN, and GUIJING WANG (2003), "National Medical Spending Attributable to Overweight and Obesity: How Much, and Who's Paying?", *Health Affairs*, W3 (suppl.), pp. 219–226.
- FINKELSTEIN, ERIC A., JUSTIN G. TROGDON, JOEL W. COHEN, and WILLIAM DIETZ (2009), "Annual Medical Spending Attributable to Obesity: Payer- and Service-Specific Estimates", *Health Affairs*, 28(5), pp. w822–w831.
- FLETCHER, JASON M., DAVID FRISVOLD, and NATHAN TEFFT (2010), "Can Soft Drink Taxes Reduce Population Weight?", *Contemporary Economic Policy*, 28(1), pp. 23–35.
- FOX, MARC (2003), "Medical Student Indebtedness and the Propensity to Enter Academic Medicine", *Health Economics*, 12(2), pp. 101–112.
- FREDERICK, SHANE, GEORGE LOEWENSTEIN, and TED O'DONOGHUE (2002), "Time Discounting and Time Preference: A Critical Review", *Journal of Economic Literature*, 40(2), pp. 351–401.
- FSO (2012), "Schweizerische Gesundheitsbefragung (SGB): CD-Rom Standardtabellen SGB12 (Auszug)", Available online <http://www.portal-stat.admin.ch/sgb2012/> accessed 03.03.2014.
- GANGWISCH, J.E., D. MALASPINA, B. BODEN-ALBALA, and S.B. HEYMSFIELD (2005), "Inadequate Sleep as a Risk Factor for Obesity: Analyses of the NHANES I", *Sleep*, 28(10), pp. 1289–1296.
- GILA, JOAN, and TONI MORAB (2011), "The Determinants of Misreporting Weight and Height: The Role of Social Norms", *Economics and Human Biology*, 9(1), pp. 78–91.
- HABERSTICK, BRETT C., JEFFERY M. LESSEM, MATTHEW B. MCQUEEN, JASON D. BOARDMAN, CHRISTIAN J. HOPFER, ANDREW SMOLEN, and JOHN K. HEWITT (2010), "Stable Genes and Changing Environments: Body Mass Index across Adolescence and Young Adulthood", *Behavior Genetics*, 40(4), pp. 495–504.

- HALFORD, W. KIM, and E. LEARNER (1984), "Correlates of Coping with Unemployment in Young Australians", *Australian Psychologist*, 19(3), pp. 333–344.
- HANSEN, LARS PETER (1982), "Large Sample Properties of Generalized Method of Moments Estimators", *Econometrica*, 50(4), pp. 1029–1054.
- HAUSMAN, JERRY A (1978), "Specification Tests in Econometrics", *Econometrica*, 46(6), 1251–1271.
- HÄUSSLER, JAN (2014), "Effects of Obesity and Physical Activity on Health Care Utilization and Costs", Working Paper Series 2014-07, University of Konstanz.
- HOLMES, ANN M., and PARTHA DEB (1998), "Provider Choice and Use of Mental Health Care: Implications for Gatekeeper Models", *Health Serv. Res.*, 33(5 Pt 1), pp. 1263–1284.
- LEE, AMANDA J., IAIN K. CROMBIE, WILLIAM C. S. SMITH, and HUGH D. TUNSTALL-PEDOE (1991), "Cigarette Smoking and Employment Status", *Soc. Sci. Med.*, 33(11), pp. 1309–1312.
- LINDROOTH, RICHARD C. and BURTON A. WEISBROD (2007), "Do Religious Nonprofit and For-Profit Organizations Respond Differently to Financial Incentives? The Hospice Industry", *Journal of Health Economics*, 26(2), pp. 342–357.
- LYLES, ROBERT H., and LAWRENCE L. KUPPER (1997), "A Detailed Evaluation of Adjustment Methods for Multiplicative Measurement Error in Linear Regression with Applications in Occupational Epidemiology", *Biometrics*, 53(3), pp. 1008–1025.
- MAES, HERMINE H.M., MICHAEL C. NEALE, and LINDON J. EAVES (1997), "Genetic and Environmental Factors in Relative Body Weight and Human Adiposity", *Behavior Genetics*, 27(4), pp. 325–351.
- MANNING, WILLARD G., and JOHN MULLAHEY (2001), "Estimating Log Models: To Transform or not to Transform?", *Journal of Health Economics*, 20(4), pp. 461–494.
- MEER, JONATHAN, and HARVEY S. ROSEN (2004), "Insurance and the Utilization of Medical Services", *Soc. Sci. Med.*, 58(9), pp. 1623–1632.
- O'NEILL, DONAL, and OLIVE SWEETMAN (2013), "The Consequences of Measurement Error when Estimating the Impact of Obesity on Income", *IZA Journal of Labor Economics*, 2(3), pp. 1–20.
- PARK, ROLLA E. (1966), "Estimation with Heteroscedastic Error Terms", *Econometrica*, 34, p. 888.
- PARKS, JOANNA C., JULIAN M. ALSTON, and ABIGAIL M. OKRENT (2012), "The Marginal External Cost of Obesity in the United States", RMICWE Working Paper 1201, Robert Mondavi Institute Center for Wine Economics.

- PETERLI, RALPH, ROBERT E. STEINERT, BETTINA WOELNERHANSEN, THOMAS PETERS, CAROLINE CHRISTOFFEL-COURTIN, MARKUS GASS, BEATRICE KERN, MARKUS VON FLUEE, and CHRISTOPH BEGLINGER (2012), "Metabolic and Hormonal Changes after Laparoscopic Roux-en-Y Gastric Bypass and Sleeve Gastrectomy: A Randomized, Prospective Trial", *Obesity Surgery*, 22(5), pp. 740–748.
- PLANKEY, MICHAEL W., JUNE STEVENS, KATHERINE M. FLEGAL, and PHILIP F. RUST (1997), "Prediction Equations Do not Eliminate Systematic Error in Self-Reported Body Mass Index", *Obesity Research*, 5(4), pp. 308–314.
- ROYSTON, PATRICK, and PAUL C. LAMBERT (2011), *Flexible Parametric Survival Analysis Using Stata: Beyond the Cox Model*, College Station, TX: Stata Press.
- SCHMID, ALEXANDRA, HEINZ SCHNEIDER, ALAIN GOLAY, and ULRICH KELLER (2005), "Economic Burden of Obesity and its Comorbidities in Switzerland", *Sozial- und Praeventivmedizin*, 50(2), pp. 87–94.
- SCHOUSBOE, KAROLINE, GONNEKE WILLEMSSEN, KIRSTEN O. KYVIK, JAKOB MORTENSEN, DORRET I. BOOMSMA, BELINDA K. CORNES, CHAYNA J. DAVIS, CORRADO FAGNANI, JACOB HJELMBORG, JAAKKO KAPRIO, MARLIES DE LANGE, MICHELLE LUCIANO, NICHOLAS G. MARTIN, NANCY PEDERSEN, KIRSI H. PIETILAINEN", AILA RISSANEN, SUOMA SAARNI, THORKILD I.A. SØRENSEN, G. CAROLINE M. VAN BAAL, and JENNIFER R. HARRIS (2003), "Sex Differences in Heritability of BMI: A Comparative Study of Results from Twin Studies in Eight Countries", *Twin Research and Human Genetics*, 6, pp. 409–421.
- SCHROETER, CHRISTIANE, JAYSON LUSK, and WALLACE TYNER (2008), "Determining the Impact of Food Price and Income Changes on Body Weight", *Journal of Health Economics*, 27(1), pp. 45–68.
- SHEA, DENNIS G., JOSEPH V TERZA, BRUCE C STUART, and BECKY BRIESACHER (2007), "Estimating the Effects of Prescription Drug Coverage for Medicare Beneficiaries", *Health Services Research*, 42(3p1), pp. 933–949.
- SHIN, JAEUN, and SANGHO MOON (2007), "Do HMO Plans Reduce Health Care Expenditure in the Private Sector?", *Economic Inquiry*, 45(1), pp. 82–99.
- SHIRI, RAHMAN, JARO KARPPINEN, PAIVI LEINO-ARJAS, SVETLANA SOLOVIEVA, and EIRA VIHKARI-JUNTURA (2009), "The Association between Obesity and Low Back Pain: A Meta-Analysis", *American Journal of Epidemiology*, 171(2), pp. 135–154.
- STAIGER, DOUGLAS, and JAMES H. STOCK (1997), "Instrumental Variables Regression with Weak Instruments", *Econometrica*, 65(3), pp. 557–586.

- TERZA, JOSEPH V., ANIRBAN BASU, and PAUL J. RATHOUZ (2008), “Two-Stage Residual Inclusion Estimation: Addressing Endogeneity in Health Econometric Modeling”, *Journal of Health Economics*, 27(3), pp. 531–543.
- TERZA, JOSEPH V., W. DAVID BRADFORD, and CLARA E. DISMUKE (2008), “The Use of Linear Instrumental Variables Methods in Health Services Research and Health Economics: A Cautionary Note”, *Health Services Research*, 43(3), pp. 1102–1120.
- VILLANUEVA, ELMER V. (2001), “The Validity of Self-Reported Weight in US Adults: A Population Based Cross-Sectional Study”, *BMC Public Health*, 1, p. 11.
- WINDMEIJER, FRANK A.G., and JOAO M. C. SANTOS SILVA (1997), “Endogeneity in Count Data Models: An Application to Demand for Health Care”, *Journal of Applied Econometrics*, 12(3), pp. 281–294.
- WOOLDRIDGE, JEFFREY M. (2014), “Quasi-Maximum Likelihood Estimation and Testing for Nonlinear Models with Endogenous Explanatory Variables”, *Journal of Econometrics*, 182(1), pp. 226–234.