Estimating Obesity Rates in Europe in the Presence of Self-Reporting Errors

Donal O'Neill [**]
Dept. of Economics, National University of Ireland Maynooth, Maynooth, Co. Kildare, Ireland
donal.oneill@nuim.ie; tel 353-1-7083555.

and

Olive Sweetman

Dept. of Economics, National University of Ireland Maynooth, Maynooth, Co. Kildare, Ireland
olive.sweetman@nuim.ie

Abstract: Reliable measures of obesity are essential in order to develop effective policies to tackle the costs of obesity. We examine what, if anything, we can learn about obesity rates using self-reported BMI once we allow for possible measurement error. Existing approaches that correct for self-reporting errors often require strong assumptions. In this paper we combine self-reported data on BMI with estimated misclassification rates obtained from auxiliary data to derive upper and lower bounds for the population obesity rate for ten European countries using minimal assumptions on the error process. For men it is possible to obtain meaningful comparisons across countries even after accounting for measurement error. In particular the self-reported data identifies a set of low obesity countries consisting of Denmark, Ireland, Italy, Greece and Portugal and a set of high obesity countries consisting of Spain and Finland. However, it is more difficult to rank countries by female obesity rates. Meaningful rankings only emerge when the misclassification rate is bounded at a level that is much lower than that observed in auxiliary data. A similar limit on misclassification rates is also needed before we can begin to observe meaningful gender differences in obesity rates within countries.

Key Words: Obesity, Self-Reporting Errors, Bounds

JEL Classification: I18, C38

[**] Corresponding author

Introduction

Obesity is an important cause of morbidity, disability and premature death and increases the risk for a wide range of chronic diseases [1-3]. Reliable measures of obesity are essential in order to develop effective policies aimed at reducing the substantial costs associated with obesity. Using self-reported data from the European Community Household Panel (ECHP) Brunello and D'Hombres [4] find substantial differences in the estimated obesity rates across nine European countries. For men during the period 1998-2001 the obesity rate ranges from a low of 5% in Ireland to a high of 10% in Finland. Denmark has the highest percentage of obese females (9%), while Italy has the lowest (3%). However, there is a large body of evidence that suggests that individuals underreport their weight and overstate their height when surveyed. Reporting error in height and weight can lead to estimates of the prevalence of obesity which are biased downwards [5-6]. In addition there is some evidence that misclassification errors are increasing over time [7]. Errors in self-reported BMI may have serious consequences for policy making since these data are often used to generate national estimates of obesity and are in turn used by policy makers when setting priorities in health policy.[1] Because of the limitations associated with self-reported measures, objective or direct measures of obesity have been recommended. However, the costs of obtaining these direct measures can sometimes be prohibitively high, and their intrusive nature may also impact on response rates. As a result reliance on self-reported BMI remains high. The WHO global infobase[2], for example, is a data warehouse that collects, stores and displays information on chronic diseases and their risk factors for all WHO member states. This is a key source for international comparable statistics on a range of health indicators, including obesity rates. However, an examination of the underlying data sources reveals that for many countries in the database the information on obesity is based on self-reported measures of weight and height.

A number of correction strategies have been proposed to deal with the problem of measurement error in self-reported BMI when estimating true underlying prevalence

[1] As well as making it difficult to determine true underlying prevalence rates measurement error also causes problems when trying to relationships between the mismeasured variable and other correctly measured outcomes. Carroll et al. [8] provide an excellent general summary of work in this area, while O'Neill and Sweetman [9] provide a recent discussion and analysis in the specific context of mismeasured obesity. However, this is not the focus of the current paper.

[2] For a detailed description see https://apps.who.int/infobase/.

rates. These include adjusting the self-reported threshold for obesity [10-11] and adjusting self-reported height and weight using prediction equations derived from auxiliary data [12-20]. In general these studies find that while such corrections tend to narrow misclassification of obese people, they do not eliminate the misclassification bias. For instance, Faeh et al. [16] report an odds ratios for female obesity based on adjusted measures relative to measured BMI of approximately .70 for both the adjusted threshold and the equation corrected measures of BMI.[3] The estimated odds ratio was significantly less than one, implying that the prevalence of obesity with the self-reported data was still statistically significantly lower than the true obesity rate even after adjustments were made to the self-reported data (see also [19, 21]). The results of these adjustments are also sensitive to the specific correction algorithm adopted. For instance Nyholm et al. [15] find that the corrected prevalence rate may underestimate or overestimate the true prevalence rate depending on which correction algorithm is used.

Latent class analysis (LCA) has also been proposed as a method to correct for measurement error in general [22]. This approach has been applied to correct for measurement error in studies of drug-use [23], unemployment status [24] and poverty dynamics [25], however we are not aware of any applications of LCA to correct for self-reported measurement error in studies of obesity. In addition, LCA often requires strong assumptions on the data generating process in order to identify key parameters. One such assumption is that of local independence [26]. When multiple manifest (or observed) variables are used in the model, local independence requires that the errors from the observed measures are independent conditional on the true unobserved obesity status. Albert and Dodd [27] show how misspecification of the conditional dependence between measures can bias estimates of misclassification error as well as the estimates of the true underlying prevalence rates. An alternative identification strategy relies on partitioning the overall population into distinct groups (e.g. men and women) all having different prevalence rates [28]. Identification using this approach requires strong assumptions on the nature of the measurement error process across groups. In addition when only two observed measures are available (as is often the case in obesity studies) this model is exactly identified so that none of the identifying restrictions can be tested. Hausman et al. [29] propose a latent variable maximum

[3] The estimated odds ratio based on self-reported BMI relative to measured BMI was .52 [16].

likelihood estimator that can be used to estimate misclassification rates when only one observed variable is variable. Identification in this approach relies on group differences in obesity rates and also non-linearities in the underlying parametric model. Bergin [30] explores the identification problems that arise with this approach.

In this paper we adopt a different strategy. Rather than trying to correct self-reported BMI we examine, what, if anything one can learn about obesity rates based on self-reported BMI using only a minimal set of assumptions on the likely rates of misclassification. In particular we use self-reported data on height and weight, along with estimates of the misclassification rates obtained from auxiliary data, to derive upper and lower bounds for the population obesity rate in ten European countries. These bounds are sharp under the maintained assumptions, in that they exhaust all the information available in the self-reported data. We show that although the presence of measurement error reduces the information in the self-reported data, these data are still capable of producing meaningful comparisons across countries for men. When comparing male obesity rates we can still identify a set of low obesity countries consisting of Denmark, Ireland, Italy, Greece and Portugal and a set of high obesity countries consisting of Spain and Finland. It is more difficult however to rank countries by female obesity rates. For women meaningful rankings only emerge when the misclassification rate is bounded at a level that is substantially lower than the rate observed in practice.

Methods

Obesity is typically measured using an individual's Body Mass Index (BMI), where BMI = weight in kg/height in m^2. Individuals are classified as overweight if their BMI is between 25 and 30 and are classified as obese if their BMI exceeds 30. However, a number of studies have shown that self-reported height and weight suffer from serious measurement error problems.[4] For example O'Neill and Sweetman [9] report that while 14% of a sample of Irish mothers were estimated to be obese on the basis of self-reported data, the estimated obesity rate based on recorded data was 17%. In a U.S. sample of women they found that the estimated obesity rate was 18%

[4] For a systematic review of the literature on measurement error in self-reported BMI see [5].

when based on self-reported data and 23% when using recorded data. The estimated obesity rate based on self-reported data tends to be too low both because respondents over estimate their height and underestimate their weight.

Clearly measurement error in self-reported BMI can have a significant effect on measured obesity rates. In this paper we examine the informational content of self-reported BMI for true prevalence rates using bounds developed in [31-32]. These bounds are sharp in the sense that they exhaust all the available information given the sampling process and the maintained assumptions. In this paper we apply these techniques to estimate bounds for obesity rates across ten European countries. To understand the bounds let X^*, denote the true measure of BMI. Let $D_X{}^*$ be a true obesity indicator equal to one if $X^*>30$ and zero otherwise. The true obesity rate is given by $\Pr(D_X{}^*=1)=\Pr(X^*>30)$.[5] However, in survey data we typically do not have access to X^* but instead must rely on a self-reported (possibly mismeasured) measure X. The observed obesity indicator D_X is equal to one if $X>30$ and is equal to zero otherwise and the observed obesity rate is $\Pr(D_X=1)=\Pr(X>30)$. When $X^*\neq X$ the observed BMI level is measured with error and ignoring this problem may lead to biased estimates of the population obesity rate. Molinari [31] provides direct bounds for the true obesity rate by exploiting the following identity:

$$\begin{aligned}\Pr(D_X = 1) = {} & \Pr(D_X = 1|D_{X^*} = 1)\Pr(D_{X^*} = 1) \\ & + \Pr(D_X = 1|D_{X^*} = 0)\Pr(D_{X^*} = 0)\end{aligned}$$

This is simply a statement of the law of total probability and places no restriction on the relationship between the true recorded measure of BMI and the self-reported measure. By imposing restrictions on the misclassification rates one can determine upper and lower bounds for the true obesity rate. The simplest bounds are obtained under the assumption that

Assumption 1: $\Pr(D_{X^*} \neq D_X) \leq \lambda_1 < 1$.

---

[5] The outcome of interest in our study (obesity status) is derived by dichotomising a continuous variable (BMI). Gustafson and Le [33] consider biases that arise in regression analysis when a dichotomised measure derived from a mismeasured continuous predictor is used as an explanatory variable. They show that differential misclassification can arise even when the error in the original continuous variable measured is non-differential. While this finding is important for regression analysis it is not relevant for our study which focuses on estimating prevelance rates.

Under this assumption Molinari [31] shows that tight bounds on $\Pr(D_{X^*} = 1)$ are given by

$$\begin{aligned} UB_1 &= \min\{\Pr(D_X = 1) + \lambda_1, 1\} \\ LB_1 &= \max\{\Pr(D_X = 1) - \lambda_1, 0\} \end{aligned} \qquad (1)$$

Alternative bounds follow from the imposition of alternative restrictions on the misclassification probabilities. In particular if we assume

Assumption 2: $\Pr(D_X = 1|D_{X^*} = 0) \leq \Pr(D_X = 0|D_{X^*} = 1) \leq \lambda_2 < 1$,

then following Proposition 8 of Molinari [31] we can establish the following set of bounds on the population obesity rate[6]

$$\begin{aligned} \mathrm{UB}_2 &= \min\left\{\frac{\Pr(D_X=1)}{1-\lambda_2}, 1\right\} \\ LB_2 &= \max\left\{\frac{\Pr(D_X=1)-\lambda_2}{1-2\,\lambda_2}, 0\right\} \end{aligned} \qquad (2)$$

Assumption 2 states that it is more likely for obese people to report a BMI below the obesity threshold than it is for non-obese people to report a BMI above the threshold. This condition seems plausible though we will check its validity in the next section.

In addition, Nicoletti et al. [32] derive alternative bounds by considering restrictions on the indirect misclassification probabilities, $\Pr(D_{X^*} = x^*|D_X = x)$. They consider the following monotonicity assumption:

Assumption 3: $\Pr(D_{X^*} = 0|D_X = 1) \leq \Pr(D_{X^*} = 1|D_X = 0) \leq \lambda_3 < 1$.

Under this assumption they derive the following bounds:[7]

$$\begin{aligned} \mathrm{UB}_3 &= \Pr(D_X = 1)\,(1 - \lambda_3) + \lambda_3 \\ LB_3 &= \Pr(D_X = 1) \end{aligned} \qquad (3)$$

[6] These bounds hold provided $\Pr(D_X = 1) < .5$ and $\lambda < .5$. The summary statistics show that the first condition is true for each of the countries in our sample, while analysis of the auxiliary data in the next section will also verify the second condition.

[7] These bounds hold provided $\Pr(D_X = 1) < .5$, which is true for all countries in our analysis.

If we assume that Assumptions 1-3 hold at the same time then we can obtain narrower bounds by combing the information from the three individual bounds. The resulting identification interval is given by $\{LB^*,UB^*\}$ where LB^* is the maximum between $\{LB_1,LB_2,LB_3\}$ and UB^* is the minimum between $\{UB_1,UB_2,UB_3\}$. In the remainder of the paper we combine auxiliary data, which provides estimates of the λs, with the self-reported data on BMI from the ECHP in order to estimate these obesity bounds for ten European countries.

Data

In order to estimate the bounds on the population obesity rate we need to be able to put limits on the rate of misclassification with self-reported BMI data. To establish these limits we use two data sets; the National Health and Nutrition Examination Survey (NHANES) in the U.S. and the Surveys of Lifestyle Attitudes and Nutrition (SLAN) for Ireland. The NHANES III is a nationally representative survey of 33,994 individuals in the U.S. aged two months of age and older. The interviews were carried out over the period from 1988-1994. The NHANES data have been used previously to examine the extent and nature of misclassification error in self-reported BMI [34-35], and also in studies that have sought to correct for misclassification error when examining the impact of obesity of labour market outcomes [36-38]. The SLAN data are interview based cross-sectional surveys of a nationally representative sample of Irish men and women in 1998, 2002 and 2007. The SLAN data have been used to examine trends in obesity in Ireland [7] and also provide key inputs into health policy making in Ireland [39].

A key feature of both the NHANES data and the SLAN data is that, in addition to self-reported measures of height and weight, both data sets also contain independent measures of the respondent's height and weight. We refer to the latter as recorded measures and treat them as the true height and weight of the respondents. In the NHANES data these recorded outcomes were obtained by a team of physicians, medical and health technicians in specially-designed and equipped mobile centres. In the SLAN data the physical examinations were carried out by nurses with specific training who followed documented procedures. Comparing obesity status on the basis of self-reported and recorded measures of BMI, allows us to derive bounds for the misclassification rates and also to examine the validity of the monotonicity assumptions presented in Section 2. Since the misclassification bounds are a key component in the construction of the obesity bounds the availability of two independent auxiliary data sources, is attractive in that it allows us check the robustness of our estimated misclassification rates. Both auxiliary data sets have advantages and disadvantages. The NHANES data has much larger samples than the SLAN data (the 2002 SLAN data used in this analysis only contains recorded measures for 147 men and 184 women). On the other hand the timing of the SLAN survey is more consistent with the timing of the ECHP data on which on our overall analysis is based and there is no guarantee that misclassification rates based on US data will necessarily apply to European countries. The availability of the SLAN data allows us to consider the extent to which misclassification bounds based on U.S. data may be applicable more generally.

In order to compare obesity rates across Europe we use data from the European Community Household Panel (ECHP). The ECHP is a dataset explicitly designed to facilitate international comparisons and has been used by Brunello and D'Hombres [4] to examine the impact of body weight on wages. The ECHP provides self-reported BMI for ten European countries for the periods 1998-2001.[8] We focus on data for the latest year and restrict attention to individuals aged between 18 and 65. Summary statistics for each of the ten countries are given in Table 1. The sample size ranges from 3109 in Denmark to 10866 in Italy. In general obesity rates are higher for men than for women. In keeping with [4] we find that obesity, based on self-reported height and weight, varies across countries.[9] The countries in Table 1 are ordered on the basis of overall obesity rates; Italy has the lowest obesity rate at 7.5%, while Finland has the largest reported obesity rate at 12.7%. These differences across countries are also apparent when we condition on gender. For example the female obesity rate is twice as high in Finland (13%) than in Italy (6.6%). In this paper we examine the extent to which these differences across countries remain after accounting for misclassification in self-reported BMI. To do this we combine the estimated misclassification rates based on the auxiliary data with the self-reported

---

[8] France, Germany, the Netherlands, the U.K. and Luxembourg also participated in the ECHP but the height and weight data needed to construct BMI was not available for these countries.

[9] Our obesity rates differ to those reported in [4] because we look at all respondents, whereas they focus on employees working at least 15 hours. They also trim the sample excluding people with BMI<15 or BMI>35. These cut-off points correspond approximately to the bottom .05% and top 2% of the sample respectively. We include all observations in our analysis.

measures of BMI in the ECHP to estimate the obesity bounds for each of the ECHP countries.

Results

Table 2 reports the estimated misclassification probabilities using the NHANES and SLAN data. The first two columns report the results for women, while the third and fourth columns provide the estimates for men. Looking at the first row we see that the estimated misclassification rate in the self-reported data was approximately 6% for both men and women in the NHANES data and 10-11% in the SLAN data. However, the Irish and U.S. misclassification rates estimates are not statistically significantly different from each other given the standard error on the SLAN estimate.

We next consider the empirical validity of the monotonicity assumptions discussed in Section 2. Both auxiliary data sets provide clear support for the direct monotonicity assumption (Assumption 2). This can be seen by comparing the probabilities in the second and third rows of Table 2. Very few people report BMI's above the obesity threshold when their true BMI is below 30. In contrast the proportion of the NHANES sample that report BMI's below 30 when their recorded measure exceeds the obesity threshold is 27% for women and 25% for men. The corresponding estimates based on the Irish data are 32% and 40% respectively. From this it is clear that the likelihood of misclassification is greater among those who are actually obese than among the non-obese. The auxiliary data also provide some support for the indirect monotonicity assumption (Assumption 3). The condition is only violated in one of the four samples we consider (women in the SLAN data).[10]

Although the misclassification rates in the Irish data are slightly higher than in the U.S. data, the estimates across the two data sets are consistent with each other. Given the larger sample sizes available in the NHANES data we use the point estimates from these data as the basis of our misclassification bounds. We follow Nicoletti et al. [32] and set the bounds on the misclassification probabilities equal to the estimated values plus twice their standard errors. Therefore we choose λ_1= .077, λ_2=.288 and λ_3=.085

[10] Although we use the 2002 SLAN data because its timing corresponds to the timing of the ECHP data we also checked misclassification rates using the SLAN 2007 data. Assumption 2 and Assumption 3 hold for both men and women in the later SLAN data.

for women and λ_1= .07, λ_2=.267 and λ_3=.071 for men. Later we conduct a sensitivity analysis to see how the results change as we vary the misclassification bounds.

Table 3 reports the upper and lower bounds $\{LB^*, UB^*\}$ on the female and male obesity rates for all ten of the countries. The first row for each country gives the point estimates for the lower and upper bounds, while the corresponding upper and lower limits of bootstrapped confidence intervals are given in the second row. We first compare the male and female obesity rates within countries. Despite the general tendency for male BMI to be higher than females we see that the identification bounds for men and women overlap in every country. As a result it is not possible to make any comparisons across gender once measurement error is accounted for.

By comparing the rows in table 3 we can determine the extent to which it is possible to make rankings across countries. Looking at the results for females we see that, once we account for likely misclassification in self-reported BMI, it becomes difficult to make strong statements regarding the ranking of obesity rates across countries. To distinguish between countries we require the upper bound for one country to be less than the lower bound for another country. When looking at females we see that, with our baseline estimates of the misclassification bounds, the data can only distinguish between Italy (a low obesity country) relative to Spain and Finland (high obesity countries). It is not possible to classify any of the other countries. However, more meaningful comparisons are possible when we consider the male obesity rates. For men the set of low obesity countries is expanded considerably to include Denmark, Ireland, Greece and Portugal along with Italy. For men it would appear that minimal assumptions on misclassification errors are sufficient to identify bounds that are narrow enough to be informative about the ranking of countries by obesity levels.

Sensitivity Analysis

Even though we derived our misclassification bounds from validation data, the choice of bounds is still to some extent arbitrary. One can examine the sensitivity of our findings to changes in the misclassification probabilities by altering λ_1 , λ_2 and λ_3 For instance, in the analysis in section 4, the misclassification bounds used for women were larger than those used for men. To examine whether this accounts for the gender differences noted in section 4 we repeat the analysis for females except this time we use the male bounds on the misclassification rates. Since these are lower we will observe tighter bounds on the true female obesity rate, which in turn may facilitate more meaningful ranking for women. The results in the first two columns of Table 4, show that using the lower male misclassification bounds when constructing bounds on the true female obesity adds Portugal and Austria to the set of countries which have substantially higher obesity rates than Italy, though it still is not possible to classify many of the countries.

Given this finding one might be interested in knowing the largest misclassification error that one could tolerate and still make meaningful obesity rankings across countries using our raw data for women. Since we know that male misclassification rates are too high to permit broad rankings we use λ_1= .07 as a starting point and then reduce the misclassification rate in increments of .005. We adjust λ_2 and λ_3 accordingly so as to keep the ratio between these parameters and λ_1 equal to the ratio implied by the estimates used in the previous section. We then recalculate $\{LB^*,UB^*\}$ for each new limit and examine the results. The key findings are reported in columns 4-7 of Table 4. The results in the fourth and fifth columns show that reducing λ_1 to .06 adds Belgium and Sweden to the set off low obesity countries (along with Italy) relative to Finland and Spain. However, even with this lower limit it is still difficult to rank most of the countries. The results in columns 6 and 7 show that an upper bound of λ_1 equal to .05 (approximately 75% of the point estimate obtained in the NHANES data) is required in order to substantially expand the set of low income countries. If one could bound the misclassification rate at this lower level then the raw data would identify a set of low obesity countries consisting of Denmark, Belgium, Ireland, Italy Greece and Sweden, a high obesity set consisting of Spain and Finland and an indeterminate group consisting of only Austria and Portugal. Comparing the male bounds in Table 3 with these latest female bounds in Table 4 also shows that this lower limit on misclassification also permits gender rankings within countries. In particular, with an upper bound of λ_1 equal to .05 there is no overlap between the male and female obesity bounds in Belgium, Italy or Spain. If we could accept this limit on measurement error then the raw data would identify the higher male obesity rates in these countries.

Discussion

We examine the robustness of obesity rankings across ten European countries taking account of potential measurement error in self-reported BMI. Our results for men are promising. Despite the presence of measurement error our analysis reveals that minimal assumptions on the rates of misclassification error are sufficient to construct bounds which are narrow enough to be informative about the ranking of countries by male obesity levels.

However, it is more difficult to obtain meaningful rankings by female obesity levels. With our baseline estimates it is only possible to rank three of the 10 countries on the basis of female obesity rates. Given the levels of measurement error observed in the data no other meaningful comparisons are possible. Further sensitivity analysis suggests that for women meaningful rankings only emerge when the misclassification rate is bounded at approximately 75% of the rate observed in auxiliary data. A similar limit on misclassification rates is also needed before we can begin to observe meaningful gender differences in obesity rates within countries.

Limitations

In order to bound the observed obesity rates it is necessary to first obtain bounds on the possible misclassification rates in self-reported BMI. Constructing these misclassification bounds requires auxiliary data containing both true and self-reported BMI. Ideally one would like country specific misclassification rates. However the required auxiliary data are typically not widely available and therefore one may be forced to use misclassification bounds derived using data from one country when calculating the obesity bounds for other countries. In this paper we use two independent auxiliary data sources, one for Ireland and one for the US to check the robustness of the analysis to the choice of estimated misclassification rates. In general the misclassification rate was higher in the Irish data than in the US data, however the estimates of the misclassification rates in the two countries were not statistically significantly different from each other. While this is reassuring it is not possible to determine if the misclassification rates differ among the broader set of ECHP countries considered in our analysis.

In the same way that misclassification rates may differ between countries they may also differ over time, which would require the bounds derived in our analysis to be updated over time. The evidence on the temporal trend in misclassification rates in self-reported BMI is mixed. For instance Gorber et al. [40] found that the difference between self-reported and measured obesity increased in Canada but remained stable in the US. Using more recent waves of the SLAN data, Shiely et al. [7] found that underreporting of BMI in Ireland increased between 1998 and 2007. To examine the possible implications of this for our analysis we compare misclassification bounds derived from the 2002 and 2007 data. In comparing these bounds it is important to remember that we follow [32] and set the misclassification bounds equal to their point estimates plus twice their standard errors. Therefore both the actual value of the point estimate and the precision with which it is estimated will be important in determining the bounds. For instance our estimate of λ_1 for women using the 2003 and SLAN data equals .147 for 2002 and .122 for 2007. Thus despite the fact that the estimated misclassification rate is higher in the 2007 data the bound is actually lower. This is because the larger samples sizes in 2007 result in much more precise point estimates. Therefore while it is not necessarily the case that higher point estimates translate into higher bounds, it is important to recognise that larger upper bounds on the misclassification rate will only exacerbate the identification problems already highlighted in our analysis.

The fact that the range of our estimated bounds are relatively large, especially for women, and overlap for a number of countries may be seen as a weakness of our approach. However, we do not see it this way. Our objective is to determine what, if anything can be learned about obesity rates using self-reported BMI, making only minimal assumptions on the nature of the misclassification error. Wide bounds imply that the raw informational content of self-reported BMI is limited. Tighter bounds can only be obtained by imposing additional assumptions on the data generating process. Our approach makes this explicit and puts the onus on researchers to substantiate any additional assumptions required to obtain more precise bounds.

Finally, throughout the paper we have focused on BMI as a measure of obesity throughout our analysis. The usefulness of BMI as a measure of fatness has been challenged in recent years [37]. While the relative merits of alternative measures of fatness raises interesting policy issues we still believe that our findings are useful. The overwhelming majority of studies continue to use BMI to measure of obesity. For this reason a detailed empirical analysis of the informational content of self-reported BMI is of considerable value.

Conclusion

Obesity is an important cause of morbidity, disability and premature death and increases the risk for a wide range of chronic diseases [1-3]. As result there are substantial direct and indirect costs associated with obesity that put a strain on national welfare systems. There have also been a number of studies that examine the impact of obesity on individual outcomes, so that the costs of obesity are borne at the individual as well as the national level [4, 36, 38, 41-42]. Reliable measures of obesity are essential in order to develop effective policies aimed at reducing the substantial costs associated with obesity. However, the vast majority of obesity statistics are based on self-reported data which is known to be subject to error. In this paper we examine what can be learned from self-reported BMI when one makes allowance for the possibility of measurement error. Despite the presence of measurement error our analysis reveals that, for males, self-reported measures of BMI may still be used to rank countries by obesity levels. However, it is more difficult to obtain meaningful rankings for females. The informational content of self-reported BMI for women is more limited and tighter bounds require additional restrictions to be placed on the data generating process. These restrictions may be difficult to validate given existing data. Thus despite the costs involved in obtaining clinical measures of height and weight our analysis suggests that such measures may be required in order to make meaningful comparisons of obesity rates both within and between countries. The ease of obtaining self-reported measures of BMI must be weighed against the biases and subsequent loss of information associated with such measures.

References

1. WHO. Global Health Risks: Mortality and Burden of Disease Attributable to Selected Major Risks. World Health Organisation 2009; WHO Press.
2. Antonanzas, F., Rodriguez R. feeding the Economics of Obesity in the EU in a Healthy Way. *European Journal of Health Economics* 2010; 11:351-353.
3. Konnopka A, Bodemann M, Konig H. Health Burden and Costs of Obesity and Overweight in Germany. *European Journal of Health Economics* 2011; 12:345-352.
4. Brunello G, d'Hombres B. Does Body Weight Affect Wages? Evidence from Europe. *Economics and Human Biology* 2007; 5:1-19.
5. Gorber SC, Temblay M, Moher D, Gorver B. A Comparison of Direct vs Self-Reported Measures for Assessing Height, Weight and Body Mass Index: A Systematic Review. *Obesity Reviews* 2007; 8:307-326.
6. Krul A, Daanen H, Choi J. Self-Reported Weight, Height and Body Mass index (BMI) in Italy, the Netherlands and North America. *European Journal of Public Health* 2010; 21:414-419.
7. Shiely F, Perry I, Lutomski J, Harrington J, Kelleher C, McGee H, Hayes K. Misclassification Patterns of Measured and Self-Report Based Body Mass Index Categories - Findings from Three Population Surveys in Ireland. *BMC Public Health* 2010; 10, 56.
8. Carroll RJ, Ruppert D, Stefanski LA, Crainiceanu C. *Measurement Error in Nonlinear Models: A Modern Perspective*. Chapman & Hall, CRC Press, Boca Raton, 2006.
9. O'Neill D, Sweetman O. The Consequences of Measurement Error when Estimating the Impact of Obesity on Income. *IZA Journal of Labor Economics*, 2013; **2**:3 doi:10.1186/2193-8997-2-3.
10. Dauphinot V, Wolff H, Naudin F, Gueguen R, Sermet C, Gaspoz JM, Kossovsky M. New Obesity Body Mass Index Threshold for Self-Reported Data. *Journal of Epidemiology and Community Health* 2009; 63:128-132.
11. Madden D. Adjusting the obesity thresholds for self-reported BMI in Ireland: A Cross-Sectional Analysis *BMJ Open* 2013; 3:e002865 doi:10.1136/bmjopen-2013-002865.
12. Plankey MW, Stevens J, Flegal KM, Rust PF. Prediction equations do not eliminate systematic error in self-reported body mass index. *Obesity Research* 1997; 5:308-14.
13. Giacchi M, Mattei R, Rossi S. Correction of the Self-Reported BMI in a teenage population, *International Journal of Obesity* 1998; 30:644-651.
14. Ezzati M, Martin H, Sjkold S, Vander Hoorn S, Murray C. Trends in National and state Level Obesity in the USA after correcting for self-report bias: analysis of health Survey. *Journal of the Royal Society of Medicine* 2006; 99:25-257.

15. Nyholm M, Gullberg B, Merlo J, Lundqvist-Persson C, Rastam L, Lindblad U. The validity of obesity based on self-reported weight and height: Implications for population studies. *Obesity* 2007; 15(1):197-208.
16. Faeh D, Bopp M. Underestimation of obesity prevalence in Switzerland: comparison of two methods for correction of self-report. *Swiss Medical Weekly* 2009; 139(51-52): 752-756.
17. Stommel M, Schoenborn C. Accuracy and Usefulness of BMI Measures based on Self-Reported weight and height: findings from Nhanes and NHIS 2001-2006. *BMC Public Health* 2009; 9:421, doi:10.1186/1471-2458-9-421.
18. Gorber SC, Shields M, Temblay M, McDowell I. The Feasibility of Establishing Correction Factors to Adjust Self-Reported Estimates of Obesity. *Health Reports* 2008; 19(3):71-82.
19. Shields M, Gorber SC, Janssen I, Temblay M. Bias in Self-Reported Estimates of Obesity in Canadian Health Surveys: An update on Correction Equations for Adults. *Health Reports* 2011; 22(3):35-45.
20. Hayes A, Clarke P, Lung T. Change in Bias in Self-reported Mass Index in Australia between 1995 and 2008 and the evaluation of correction equations. *Population Health Metrics* 2011; 25;9:53. doi: 10.1186/1478-7954-9-53.
21. Visscher T, Viet A, Kroesbergen I, Seidel J. Underreporting of BMI in Adults and Its Effect on Obesity Prevalence Estimations in the Period 1998 to 2001. *Obesity* 2006; 14(11):2054–2063.
22. Biemer P. *Latent Class Analysis of Survey Error*, Wiley and Sons, New Jersey, 2011.
23. Biemer P, Wiesen C. Measurement Error evaluation of self-reported drug use: a latent class analysis of the US National Household Survey on Drug Abuse. *Journal of the Royal Statistical Association A* 2002; 165, Part 1:97-119.
24. Gastwirth J, Sinclair M. Estimates of the Errors in Classification in the Labour Force Survey and Their Effects on the Reported Unemployment Rate. *Survey Methodology* 1998; 24(2):157-169.
25. Breen R, Moisio P. Poverty Dynamics corrected for Measurement Error. *Journal of Economic Inequality* 2004; 2(3):171-191.
26. Goodman L. Exploratory Latent Structure Analysis Using Both Identifiable and Unidentifiable Models. *Biometrika* 1974; 61(2):215-231.
27. Albert P, Dodd L. A Cautionary Note on the Robustness of Latent Class Models for Estimating Diagnostic Error without a Gold Standard, *Biometrics* 2004; 60:427-435.
28. Hiu S, Walter S. Estimating the Error Rates of Diagnostic Tests. *Biometrics* 1980; 36: 167-171.
29. Hausman J, Abrevaya J, Scott-Morton FM. Misclassification of the Dependent Variable in a Discrete Response Setting. *Journal of Econometrics* 1998; 87:239-269.
30. Bergin A. *Job Mobility and Measurement Error* PhD Thesis, NUI Maynooth, Ireland, 2011.
31. Molinari F. Partial Identification of Probability Distributions with Misclassified Data. *Journal of Econometrics* 2008; 144:81-117.

32. Nicoletti C, Peracchi F, Foliano F. Estimating Income Poverty in the Presence of Missing Data and Measurement Error. *Journal of Business & Economic Statistics* 2011; 29(1):61-72.
33. Gustafson P, Le N. Comparing the Effects of Continuous and Discrete Covariate Mismeasurement, with Emphasis on the Dichotmization of Mismeasured Predictors. *Biometrics* 2002; 58:878-887.
34. Villanueva E. The Validity of Self-reported Weight in US adults: a Population Based Cross-Sectional Study. *BMC Public Health* 2001; 1(11).
35. Kuczmarski M, Kuczmarski R, Najja M. Effects of Age on Validity of self-reported height, weight and body mass index: findings from the third National Health and Nutrition Examination Survey, 1988-1994. *Journal of the American Dietetic Association* 2001; 101:28-34.
36. Cawley J. The Impact of Obesity on Wages. *Journal of Human Resources* 2004; 39: 451-474.
37. Burkhauser R, Cawley J. Beyond BMI: The Value of More Accurate Measures of Fatness and Obesity in Social Science Research. *Journal of Health Economics* 2008; 27:519-529.
38. Lindeboom M, Lundborg P, van der Klaauw B. Assessing the Impact of Obesity on Labor Market Outcomes. *Economics and Human Biology* 2010; 8:309–319
39. National Task Force on Obesity. *Obesity: The Policy Challenges* http://www.dohc.ie/publications/pdf/report_taskforce_on_obesity.pdf, 2005.
40. Gorber SC, Temblay M, Moher D, Gorver B. The Bias in Self-Reported Obesity from 1976 to 2005: a Canada-US Comparison. *Obesity* 2010; 18(2):354-361.
41. Mora, T. BMI and Spanish Labour Status: Evidence by Gender from the City of Barcelona. *European Journal of Health Economics* 2010; 11:239-253.
42. Olivia-Moreno J, Giil-Lacruz. Body Weight and Health Related Quality of Life in Catalonia, Spain. *European Journal of Health Economics* 2013; 14:95-105.

Table 1.

Summary statistics for ECHP data

Country	Total Sample Size	Overall Obesity Rate	Male Obesity Rate	Female Obesity Rate
Italy	10866	.075	.085	.066
Ireland	3142	.085	.085	.085
Sweden	4406	.091	.099	.082
Denmark	3109	.091	.091	.091
Greece	6817	.093	.099	.088
Portugal	8270	.095	.088	.103
Belgium	3338	.100	.117	.085
Austria	4331	.104	.109	.099
Spain	8897	.123	.136	.110
Finland	4433	.127	.123	.130
Average		.098	.103	.093

Table 2
Misclassification Rates from NHANES III and SLAN data[11]

	Women		Men	
	NHANES Estimated Value (SE)	Slan 2002 Estimated Value (SE)	NHANES Estimated Value (SE)	Slan 2002 Estimated Value (SE)
$\Pr(D_{X^*} \neq D_X)$	.067 (.005)	.103 (.022)	.06 (.005)	.116 (.026)
$\Pr(D_X = 1 \mid D_{X^*} = 0)$	.007 (.002)	.0357 (.0157)	.012 (.002)	0
$\Pr(D_X = 0 \mid D_{X^*} = 1)$	.268 (.009)	.318 (.07)	.248 (.0095)	.40 (.075)
$\Pr(D_{X^*} = 0 \mid D_X = 1)$	.0285 (.003)	.143 (.058)	.0595 (.005)	0
$\Pr(D_{X^*} = 1 \mid D_X = 0)$	.075 (.005)	.094 (.0238)	.061 (.005)	.139 (.03)

[11] The misclassification rates for the SLAN data are based on the numbers reported in table 2 of [7].

Table 3

Estimated Bounds by Country for Females and Males. For each country the estimates of the lower and upper bounds are reported in the first row, while corresponding lower and upper limits of the bootstrap 95% confidence interval are reported in the second row.

Country	Women (λ_1=.077, λ_2=.288 and λ_3=.085) LB*	UB*	Men (λ_1=.07, λ_2=.267 and λ_3=.071	
Denmark	.09125	.12817	.09143	**.12474**
	.07664	.14869	.07696	.14448
Belgium	.08480	.11910	.11701	.15963
	.07195	.13715	.10101	.18146
Ireland	.08509	.11951	.08485	**.11576**
	.07168	.13835	.07071	.13506
Italy	.06559	**.09213**	.08460	**.11542**
	.05887	.10157	.07720	.12551
Greece	.08816	.12383	.09845	**.13431**
	.07885	.13690	.08806	.14849
Spain	**.10975**	.15414	**.13552**	.18489
	.10027	.16744	.12571	.19828
Portugal	.10261	.14411	.08778	**.11976**
	.09325	.15726	.07912	.13159
Austria	.09936	.13955	.10903	.14874
	.08705	.15683	.09609	.16639
Finland	**.13039**	.18314	.12305	.16787
	.11675	.20229	.10978	.18598
Sweden	.08232	.11561	.09986	.13623
	.07082	.13176	.08716	.15356

Table 4

Estimated Bounds by Country for Females with alternative misclassification bounds. For each country the estimates of the lower and upper bounds are reported in the first row, while corresponding lower and upper limits of the bootstrap 95% confidence interval are reported in the second row.

Country	Women (λ_1=.07, λ_2=.267 and λ_3=.071 LB*	UB*	Women (λ_1=.06, λ_2=.2247 and λ_3=.066		Women (λ_1=.05, λ_2=.1872 and λ_3=.055	
Denmark	.08232	.11230	.08232	.11770	.08232	**.10617**
Belgium	.08480	.11569	.08480	**.10938**	.08480	**.10433**
Ireland	.08509	.11608	.08509	.10975	.08509	**.10469**
Italy	.06559	**.08949**	.06559	**.08460**	.06559	**.08070**
Greece	.08816	.12028	.08816	.11372	.08816	**.10847**
Spain	**.10975**	.14972	**.10975**	.14155	**.10975**	.13502
Portugal	**.10261**	.13998	.10261	.13235	.10261	.12624
Austria	**.09936**	.13555	.09936	.12815	.09936	.12224
Finland	**.13039**	.17789	**.13039**	.16818	**.13039**	.16042
Sweden	.08232	.11230	.08232	**.10617**	.08232	**.10128**