

Bounding obesity rates in the presence of self-reporting errors

Donal O'Neill¹ · Olive Sweetman¹

Received: 26 June 2014 / Accepted: 20 March 2015 / Published online: 14 May 2015
© Springer-Verlag Berlin Heidelberg 2015

Abstract We examine what, if anything, we can learn about obesity rates using self-reported BMI once we allow for possible measurement error. We use self-reported obesity rates, along with estimates of misclassification rates, to derive upper and lower bounds for the true population obesity rate. These bounds are then used as the basis for obesity rankings. Our results show, that once measurement error is taken into account, it is difficult to obtain meaningful rankings across European countries. However, our analysis shows that it is still possible to rank US states by obesity status using only minimal assumptions on the nature of the error process. As a result, cross-state variation in self-reported BMI, when used in conjunction with our bounds, may still provide a useful source of information for understanding the causes and consequences of obesity in the USA.

Keywords Obesity · Self-reporting errors · Misclassification · Bounds

JEL Classification C13 · C14 · I1

1 Introduction

Obesity is an important cause of morbidity, disability and premature death and increases the risk of a wide range of chronic diseases (WHO 2009; Antonanzas and

✉ Donal O'Neill
donal.oneill@nuim.ie

Olive Sweetman
olive.sweetman@nuim.ie

¹ Department of Economics, National University of Ireland Maynooth, Maynooth, Co. Kildare, Ireland

Rodriguez 2010). Devising effective policies to tackle the growing obesity problem is a health priority for many countries. One popular strategy used in helping to devise such strategies is to use cross-country or cross-state variation in obesity rates to identify effective policies (Loureiro and Nayga 2005; O'Neill and O'Neill 2007; Costa-Font et al. 2009, 2010; Nanney et al. 2010; Taber et al. 2011; Cawley et al. 2013). However, due to data limitations, comparisons of this nature often rely on self-reported obesity levels. There is a large body of evidence that suggests that individuals underreport their weight and overstate their height, which leads to biased estimates of obesity rates (Gorber et al. 2007).¹ This in turn may limit the usefulness of these data when informing policy makers.

A number of correction strategies have been proposed to deal with the problem of measurement error in self-reported height and weight (Plankey et al. 1997; Giacchi et al. 1998; Ezzati et al. 2006; Nyholm et al. 2007; Dauphinot et al. 2009; Stommel and Schoenborn 2009; Shields et al. 2011). While such corrections tend to reduce misclassification rates, they do not eliminate it (Visscher et al. 2006; Nyholm et al. 2007; Shields et al. 2011). Rather than trying to obtain corrected measures of obesity, we examine what, if anything, one can learn about obesity using self-reported data and imposing only minimal assumptions on the error process. In particular, we use self-reported obesity rates, along with estimates of the misclassification rates, to derive upper and lower bounds for the population obesity rate in ten European countries and each of the US states. 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 makes it difficult to obtain meaningful rankings across European countries, the estimated bounds are narrow enough to rank US states by the level of obesity.

2 Methods

Obesity is typically measured using body mass index (BMI), where $BMI = \text{weight} / \text{height}^2$. Individuals are classified as obese if their BMI exceeds 30. There have been a number of recent papers examining the use of BMI as a measure of obesity. For instance, Burkhauser and Cawley (2008) and O'Neill (2015) examine the sensitivity and specificity of BMI-based tests for obesity, relative to alternative measures such as bioelectrical impedance analysis or waist circumference-based measures. The results in both these papers suggest that BMI-based measures of obesity may suffer from high rates of false negatives (or low sensitivity); there is a danger that BMI (even when based on clinically measured height and weight) may underreport true obesity. In addition, some papers have discussed the appropriateness of a uniform threshold

¹ As well as making it difficult to determine true underlying prevalence rates, measurement error also causes problems when trying to estimate relationships between the mismeasured variable and other correctly measured outcomes. Carroll et al. (2006) provide an excellent general summary of work in this area, while O'Neill and Sweetman (2013) provide a recent discussion and analysis in the specific context of mismeasured obesity. However, this is not the focus of the current paper. We also follow the vast majority of research and public discussion in this area and treat body mass index as a satisfactory measure of obesity (for some discussion of this issue, see Burkhauser et al. 2009).

for BMI-based obesity measures when applied to different groups (Hu 2008). We do not focus on either of these issues in this paper. Instead, we take the standard BMI measure and threshold, which despite its potential weaknesses is still by far the most popular measure of obesity, and examine the informational content of self-reported obesity rates in this context. To do this, we use bounds developed in Molinari (2008) and Nicoletti et al. (2011).

Let X^* denote the true measure of BMI and D_X^* a true obesity indicator equal to one if $X^* > 30$ and zero otherwise. The true obesity rate is given by $\Pr(D_X^* = 1) = \Pr(X^* > 30)$. Typically, we do not have access to X^* and must rely on a self-reported measure, X . The observed obesity indicator D_X is equal to one if $X > 30$ and 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. Molinari (2008) provides direct bounds in the presence of misclassification. The simplest bounds are obtained under the assumption that

Assumption 1

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

Under this assumption, tight bounds on $\Pr(D_{X^*} = 1)$ are given by

$$\begin{aligned} \text{UB}_1 &= \min \{ \Pr(D_X = 1) + \lambda_1, 1 \} \\ \text{LB}_1 &= \max \{ \Pr(D_X = 1) - \lambda_1, 0 \} \end{aligned} \tag{1}$$

Alternative bounds follow from the imposition of alternative restrictions. 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 by manipulating the bounds reported in (3.7) and (3.8) of Molinari (2008), we can establish the following bounds for the obesity rate²

$$\begin{aligned} \text{UB}_2 &= \min \left\{ \frac{\Pr(D_X=1)}{1-\lambda_2}, 1 \right\} \\ \text{LB}_2 &= \max \left\{ \frac{\Pr(D_X=1)-\lambda_2}{1-2\lambda_2}, 0 \right\} \end{aligned} \quad \text{if } \Pr(D_X = 1) < .5 \quad \text{and} \quad \lambda_2 < .5. \tag{2a}$$

$$\begin{aligned} \text{UB}_2 &= \left\{ \frac{\Pr(D_X=1)}{1-\lambda_2} \right\} \\ \text{LB}_2 &= 0 \end{aligned} \quad \text{if } 1 - \Pr(D_X = 1) > \lambda_2 \geq 5. \tag{2b}$$

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.

Nicoletti et al. (2011) consider restrictions on the indirect misclassification probabilities, $\Pr(D_{X^*} = x^* | D_X = x)$. Using

² Additional bounds are possible to cover other configurations of $\Pr(D_X = 1)$ and λ_2 , but the bounds in (2a) and (2b) cover all the cases observed in our later empirical analyses.

Assumption 3

$$\Pr(D_{X^*} = 0 | D_X = 1) \leq \Pr(D_{X^*} = 1 | D_X = 0) \leq \lambda_3 < 1,$$

they derive the following bounds:

$$\begin{aligned} \text{UB}_3 &= \Pr(D_X = 1)(1 - \lambda_3) + \lambda_3 \\ \text{LB}_3 &= \Pr(D_X = 1) \end{aligned} \quad (3)$$

If we assume that Assumptions 1–3 all hold, we can combine the information from the three individual bounds. The resulting identification interval is given by $\{\text{LB}^*, \text{UB}^*\}$, where LB^* is the maximum between $\{\text{LB}_1, \text{LB}_2, \text{LB}_3\}$ and UB^* is the minimum between $\{\text{UB}_1, \text{UB}_2, \text{UB}_3\}$. We check the validity of all these assumptions later.

3 Data

To estimate the misclassification rate, we use two data sets: the National Health and Nutrition Examination Survey (NHANES) for the USA and the Surveys of Lifestyle Attitudes and Nutrition (SLAN) for Europe. The NHANES is a programme of studies designed to assess the health and nutritional status of adults and children in the USA. Earlier waves of the NHANES data have been used previously to examine the extent of misclassification error in self-reported BMI (Villanueva 2001; Kuczmarski et al. 2001). We use the NHANES 2009–2010 to derive our misclassification rates, as the timing is consistent with the 2010 Behavioural Risk Factor Surveillance System (BRFSS) used to derive obesity rates across US states later in the paper. The NHANES 2009–2010 data set contains data for 10,537 individuals of all ages. Interviews were carried out between 1 January 2009 and 31 December 2010. When we restrict attention to white individuals aged between 18 and 65, with both self-reported and recorded BMI, we have 1055 women and 1030 men.

When estimating misclassification rates for the European analysis, we use the Irish SLAN data. The SLAN data are cross-sectional surveys of Irish men and women conducted in 1998, 2002 and 2007. These data have been used to examine trends in obesity in Ireland (Shiely et al. 2010) and provide key inputs into health policy making in Ireland (National Taskforce on Obesity 2005). We use the 2002 SLAN data to estimate European misclassification rates so as to be consistent with the timing of the European Community Household Panel data used for the European analysis.³

In addition to self-reported measures of height and weight, both the NHANES and SLAN data contain independent measures of the respondent's height and weight. Comparing obesity status on the basis of self-reported and recorded measures of BMI allows us to estimate the underlying misclassification rates in both data sets. In addition, the availability of two independent auxiliary data sources also allows us check the robustness of the assumptions imposed on the misclassification rates in deriving our bounds.

³ None of the substantive results in our paper change if we use the 1998 or 2007 SLAN data instead of the SLAN 2002 data or the NHANES III data instead of NHANES 2009–2010.

Table 1 Summary statistics for ECHP data

| Country | Total sample size | Overall obesity rate | Male obesity rate | Female obesity rate |
|----------|-------------------|----------------------|-------------------|---------------------|
| Italy | 10,866 | .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 |

In order to compare obesity rates across Europe, we use the European Community Household Panel (ECHP). The ECHP provides self-reported BMI for ten European countries from 1998 to 2001.⁴ We use data for 2001. In keeping with the earlier restrictions, we limit our analysis to those aged between 18 and 65. Summary statistics are provided in Table 1. In general, self-reported obesity rates are higher for men than for women. In keeping with Brunello and d'Hombres (2007), we find that self-reported obesity rates vary across countries with Italy having the lowest rate and Finland the highest.⁵

When comparing obesity rates across US states, we use the 2010 Behavioural Risk Factor Surveillance System (BRFSS). The BRFSS is a telephone survey conducted by the Centers for Disease Control and Prevention in the USA. There are 451,075 total records for 2010. Such a large sample size makes it possible to compute reliable state-level descriptive statistics. We limit our analysis to whites aged between 18 and 65.⁶ Summary statistics for men and women, which are given in Table 2, show that Mississippi has the highest rate of obesity for white men and that West Virginia had the highest rate for white women. Although the BRFSS has the advantage of large sample sizes, it is similar to the ECHP in that only self-reported weight and height are available. In this paper, we examine the extent to which meaningful rankings of obesity across countries and states persist after accounting for misclassification error.

⁴ France, Germany, the Netherlands, the UK and Luxembourg also participated in the ECHP, but the height and weight data needed to construct BMI were not available for these countries.

⁵ Our obesity rates differ to those reported in Brunello and d'Hombres (2007) because we look at all respondents, whereas they focus on employees working at least 15 h. 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.

⁶ We have also carried out the analysis for all races together. This had no effect on the key findings of the paper.

Table 2 Summary statistics for BRFSS data

| State | Total sample size | Overall obesity rate | Male obesity rate | Female obesity rate |
|----------------------|-------------------|----------------------|-------------------|---------------------|
| Alabama | 3183 | .32 | .33 | .31 |
| Alaska | 1088 | .30 | .29 | .31 |
| Arizona | 2250 | .28 | .28 | .28 |
| Arkansas | 1754 | .33 | .37 | .30 |
| California | 6866 | .24 | .25 | .23 |
| Colorado | 6507 | .22 | .25 | .20 |
| Connecticut | 3284 | .24 | .28 | .21 |
| Delaware | 2048 | .29 | .30 | .28 |
| District of Columbia | 1295 | .11 | .12 | .11 |
| Florida | 15,880 | .31 | .33 | .29 |
| Georgia | 2576 | .29 | .30 | .28 |
| Hawaii | 1448 | .20 | .21 | .18 |
| Idaho | 3960 | .28 | .30 | .27 |
| Illinois | 2502 | .28 | .30 | .26 |
| Indiana | 5354 | .33 | .35 | .31 |
| Iowa | 3505 | .31 | .34 | .29 |
| Kansas | 4723 | .33 | .35 | .32 |
| Kentucky | 4777 | .35 | .36 | .34 |
| Louisiana | 3128 | .30 | .34 | .28 |
| Maine | 5079 | .30 | .30 | .30 |
| Maryland | 4446 | .28 | .33 | .25 |
| Massachusetts | 8544 | .26 | .30 | .23 |
| Michigan | 4216 | .33 | .34 | .32 |
| Minnesota | 5566 | .26 | .29 | .24 |
| Mississippi | 2978 | .33 | .38 | .30 |
| Missouri | 2722 | .34 | .34 | .34 |
| Montana | 4069 | .26 | .28 | .24 |
| Nebraska | 9270 | .32 | .34 | .31 |
| Nevada | 2007 | .26 | .30 | .23 |
| New Hampshire | 3749 | .29 | .30 | .27 |
| New Jersey | 5882 | .25 | .30 | .21 |
| New Mexico | 2363 | .24 | .25 | .24 |
| New York | 4032 | .26 | .30 | .23 |
| North Carolina | 5754 | .28 | .30 | .27 |
| North Dakota | 2875 | .31 | .34 | .29 |
| Ohio | 5423 | .32 | .34 | .31 |
| Oklahoma | 3430 | .33 | .32 | .33 |
| Oregon | 2564 | .30 | .30 | .30 |

Table 2 continued

| State | Total sample size | Overall obesity rate | Male obesity rate | Female obesity rate |
|----------------|-------------------|----------------------|-------------------|---------------------|
| Pennsylvania | 5992 | .31 | .32 | .30 |
| Rhode Island | 3619 | .28 | .30 | .27 |
| South Carolina | 3515 | .31 | .33 | .30 |
| South Dakota | 3450 | .31 | .34 | .28 |
| Tennessee | 2936 | .34 | .33 | .34 |
| Texas | 6097 | .30 | .32 | .28 |
| Utah | 6517 | .25 | .28 | .22 |
| Vermont | 4312 | .25 | .26 | .25 |
| Virginia | 2701 | .30 | .31 | .28 |
| Washington | 10,520 | .30 | .30 | .29 |
| West Virginia | 2651 | .36 | .37 | .35 |
| Wisconsin | 2587 | .30 | .31 | .28 |
| Wyoming | 3412 | .27 | .29 | .25 |

Table 3 Misclassification rates from NHANES 2010 and SLAN 2002 data

| | Men | | Women | |
|------------------------------|----------------------------------|--------------------------------|----------------------------------|--------------------------------|
| | NHANES 2010 estimated value (SE) | SLAN 2002 estimated value (SE) | NHANES 2010 estimated value (SE) | SLAN 2002 estimated value (SE) |
| $\Pr(D_{X^*} \neq D_X)$ | .07 (.008) | .116 (.026) | .047 (.006) | .103 (.022) |
| $\Pr(D_X = 1 D_{X^*} = 0)$ | .02 (.005) | 0 | .01 (.004) | .0357 (.0157) |
| $\Pr(D_X = 0 D_{X^*} = 1)$ | .17 (.019) | .40 (.075) | .121 (.017) | .318 (.07) |
| $\Pr(D_{X^*} = 0 D_X = 1)$ | .035 (.009) | 0 | .021 (.008) | .143 (.058) |
| $\Pr(D_{X^*} = 1 D_X = 0)$ | .086 (.01) | .139 (.03) | .058 (.0097) | .094 (.0238) |

The misclassification rates for the SLAN data are based on the numbers reported in Table 2 of [Shiely et al. \(2010\)](#). Their report is based on individuals aged 18 and older. Unfortunately, the public release version of the 2002 data does not contain the clinical results, so we cannot directly calculate the misclassification rate for the 18–65 group. However, we have used later waves of the SLAN data to examine the sensitivity of our findings to the inclusion of the older workers when calculating the rate and found that it made little difference to the results

4 Results

Table 3 reports the estimated misclassification rates from NHANES and SLAN. The estimated misclassification rate is approximately 7 and 5 % for men and women in the NHANES data and 10–11 % in the SLAN data. However, the Irish and US misclassification rates estimates are not statistically significantly different from each other.

Both auxiliary data sets provide clear support for the direct monotonicity assumption (Assumption 2). Very few people report BMI's above the obesity threshold when

Table 4 Estimated bounds by country for men and women

| Country | Men | | Women | |
|----------|---|--------|--|---------|
| | $(\lambda_1 = .168, \lambda_2 = .55$ and $\lambda_3 = .199)$ | | $(\lambda_1 = .147, \lambda_2 = .458$ and $\lambda_3 = .142)$ | |
| | LB* | UB* | LB* | UB* |
| Denmark | .09143 | .20319 | .09125 | .168371 |
| Belgium | .11701 | .26003 | .08480 | .15646 |
| Ireland | .08485 | .18857 | .08509 | .15699 |
| Italy | .08460 | .18801 | .06559 | .12103 |
| Greece | .09845 | .21878 | .08816 | .16267 |
| Spain | .13552 | .30117 | .10975 | .202492 |
| Portugal | .08778 | .19509 | .10261 | .189321 |
| Austria | .10903 | .24229 | .09936 | .183324 |
| Finland | .12305 | .27346 | .13039 | .240583 |
| Sweden | .09986 | .22192 | .08232 | .151884 |

For each country, the estimates of the lower and upper bounds are reported

their true BMI is below 30. In contrast, the proportion of the NHANES sample who report a BMI below 30 when their recorded measure exceeds the obesity threshold is 12 % for women and 17 % for men. The corresponding estimates based on the Irish data are 32 and 40 %, respectively. The data also provide support for the indirect monotonicity assumption (Assumption 3). The condition is violated in only one of the four samples we consider (women in the SLAN data).^{7,8}

In the analysis that follows, we use the misclassification rates for the Irish SLAN data when estimating the bounds for the ten European countries and the misclassification rates from the NHANES data when estimating the bounds for the US states. We follow Nicoletti et al. (2011) and set the bounds on the misclassification probabilities equal to the estimated values plus twice their standard errors. We then use these misclassification rates to estimate upper and lower obesity bounds as described in Sect. 2.

Table 4 reports the upper and lower obesity bounds {LB*, UB*} for our ten European countries.⁹ We first use these bounds to compare the male and female obesity rates

⁷ 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 1998 and 2007 SLAN data. Assumptions 2 and 3 hold for both men and women in both these data sets.

⁸ We have also used the NHANES data to check whether the misclassification rates reported for the USA vary by region. Unfortunately, the public release file of the NHANES data does not include state identifiers. However, we can compare misclassification rates by census region and fail to reject the null hypothesis that the misclassification rates are equal across regions.

⁹ The bounds for both men and women make use of all three assumptions. Although the point estimates for the 2002 SLAN data suggest that Assumption 3 may be violated for women, we cannot reject the hypotheses that the two rates are equal. The effect of relaxing this assumption when calculating the bounds for women is to widen the interval even further [the lower bound becomes zero for all countries]. This does nothing to change the substantive findings of the paper.

within countries. Despite the general tendency for male obesity rates to be higher than female obesity rates, we see that the bounds for men and women overlap in every country. Consequently, it is not possible to make any comparisons across gender once the potential loss in information arising from measurement error is accounted for.

We next determine the extent to which one can make rankings across these countries. To distinguish between countries, we require the upper bound for one country to be less than the lower bound for another country. Once again meaningful rankings are very difficult given the width of the bounds. Of all the comparisons available across European countries, the only comparison that is robust to the adjustment for measurement error is that the obesity rate for women in Finland is higher than that for women in Italy. It is important to emphasise that this is a problem of identification resulting from measurement error rather than one of sample size. Once we allow for the loss of information inherent in self-reported obesity measured using the SLAN misclassification rates, there is simply not enough information in the data to make meaningful distinctions across countries. Simply obtaining larger samples (or more of the same type of data) will not solve this problem.

In the remainder of this section, we report the results when our bounds are applied to differences in obesity rates across US states. We begin by looking at the US gender gap in obesity. The obesity rates reported in Table 2 show that the raw self-reported obesity rate is higher for men than for women in 49 of the 51 cases analysed. However, once again, the results in Table 5 reveal that it is more difficult to make comparisons across gender once we allow for measurement error. The gender gap is robust to our adjustment for measurement error in only eleven of the states. These states are Arkansas, Colorado, Connecticut, Louisiana, Maryland, Massachusetts, Mississippi, Nevada, New Jersey, New York and Utah. In each of these states, the bounds indicate that the obesity rate for white males is higher than that for white females, even when the error in self-reported BMI is taken into account. In all the other states, the obesity bounds for men and women overlap.

However, in contrast to the earlier European analysis, the results in Table 5 show that meaningful rankings across US states based on self-reported obesity are still possible with only minimal assumptions on the error process. For men, one can distinguish between a set of high-obesity states consisting of {Arkansas, Indiana, Iowa, Kentucky, Louisiana, Michigan, Mississippi, Missouri, Nebraska, North Dakota, Ohio, South Dakota, Tennessee and West Virginia} relative to a grouping of low-obesity states consisting of {California, Colorado, District of Columbia, Hawaii, New Mexico and Vermont}. For women, the number of meaningful rankings is even larger though the groupings are quite similar to the male classes. The high-obese states for females include Alabama, Indiana, Kentucky, Michigan, Nebraska, Ohio, Oklahoma, Pennsylvania, Tennessee and West Virginia, and the low-obese states include {California, Colorado, Connecticut, District of Columbia, Hawaii, Maryland, Massachusetts, Minnesota, Montana, Nevada, New Jersey, New Mexico, New York, Utah, Vermont and Wyoming}.

These groupings are very similar to the unadjusted groupings that appear in the popular press. Since 2008, Gallup has published a well-being index for each state in the USA which includes the state obesity rates as a component. These indices are constructed using data from telephone interviews in which data on health perceptions,

Table 5 Estimated bounds by state for men and women

| State | Men ($\lambda_1 = .086, \lambda_2 = .2096$ and $\lambda_3 = .106$) | | Women ($\lambda_1 = .069, \lambda_2 = .16$ and $\lambda_3 = .077$) | |
|----------------------|--|-------|--|-------|
| | LB* | UB* | LB* | UB* |
| Alabama | .3321 | .4029 | .3088 | .3620 |
| Alaska | .2888 | .3642 | .3089 | .3621 |
| Arizona | .2830 | .3580 | .2752 | .3276 |
| Arkansas | .3661 | .4333 | .3027 | .3564 |
| California | .2534 | .3207 | .2274 | .2707 |
| Colorado | .2429 | .3073 | .1962 | .2336 |
| Connecticut | .2747 | .3475 | .2049 | .2440 |
| Delaware | .2993 | .3735 | .2792 | .3325 |
| District of Columbia | .1167 | .1476 | .1094 | .1302 |
| Florida | .3336 | .4042 | .2916 | .3461 |
| Georgia | .3034 | .3772 | .2806 | .3340 |
| Hawaii | .2123 | .2685 | .1810 | .2155 |
| Idaho | .2989 | .3732 | .2740 | .3262 |
| Illinois | .2967 | .3712 | .2642 | .3146 |
| Indiana | .3495 | .4184 | .3139 | .3667 |
| Iowa | .3406 | .4105 | .2933 | .3478 |
| Kansas | .3493 | .4183 | .3153 | .3680 |
| Kentucky | .3621 | .4298 | .3433 | .3939 |
| Louisiana | .3395 | .4095 | .2806 | .3340 |
| Maine | .3017 | .3757 | .3009 | .3547 |
| Maryland | .3247 | .3964 | .2520 | .3000 |
| Massachusetts | .2938 | .3687 | .2278 | .2712 |
| Michigan | .3431 | .4127 | .3139 | .3738 |
| Minnesota | .2852 | .3609 | .2425 | .2887 |
| Mississippi | .3771 | .4432 | .2989 | .3529 |
| Missouri | .3419 | .4116 | .3377 | .3887 |
| Montana | .2760 | .3492 | .2422 | .2884 |
| Nebraska | .3375 | .4078 | .3074 | .3608 |
| Nevada | .3010 | .3752 | .2330 | .2774 |
| New Hampshire | .3021 | .3761 | .2744 | .3267 |
| New Jersey | .2982 | .3726 | .2048 | .2438 |
| New Mexico | .2449 | .3098 | .2421 | .2883 |
| New York | .2942 | .3690 | .2323 | .2766 |
| North Carolina | .3023 | .3762 | .2692 | .3205 |
| North Dakota | .3392 | .4092 | .2941 | .3484 |
| Ohio | .3374 | .4076 | .3029 | .3566 |
| Oklahoma | .3189 | .3910 | .3287 | .3804 |

Table 5 continued

| State | Men ($\lambda_1 = .086, \lambda_2 = .2096$ and $\lambda_3 = .106$) | | Women ($\lambda_1 = .069, \lambda_2 = .16$ and $\lambda_3 = .077$) | |
|----------------|--|-------|--|-------|
| | LB* | UB* | LB* | UB* |
| Oregon | .2985 | .3728 | .2950 | .3493 |
| Pennsylvania | .3210 | .3930 | .3033 | .3569 |
| Rhode Island | .3038 | .3776 | .2684 | .3195 |
| South Carolina | .3308 | .4017 | .2902 | .3448 |
| South Dakota | .3353 | .4058 | .2839 | .3379 |
| Tennessee | .3323 | .4031 | .3422 | .3929 |
| Texas | .3219 | .3939 | .2789 | .3319 |
| Utah | .2826 | .3576 | .2201 | .2621 |
| Vermont | .2596 | .3284 | .2502 | .2979 |
| Virginia | .3122 | .3851 | .2829 | .3368 |
| Washington | .3024 | .3764 | .2886 | .3434 |
| West Virginia | .3656 | .4329 | .3513 | .4013 |
| Wisconsin | .3138 | .3865 | .2841 | .3383 |
| Wyoming | .2949 | .3696 | .2530 | .3012 |

For each state, the estimates of the estimates of the lower and upper bounds are reported

including self-reported height and weight, are gathered. Using these raw data, states are grouped according to obesity rates. Typically in the reported analysis, Gallup distinguishes between the ten states with highest obesity and the ten states with the lowest obesity rates. When we compare the Gallup results, based on unadjusted self-reported data¹⁰, with our results that take measurement error into account, we find a substantial overlap. Nine of the ten states identified by Gallup as low-obesity states appear in one or the other of our low-obesity classes (either the male or the female class), the exception being Rhode Island. In addition, 10 out of the top 11 most obese states in the Gallup report are listed in our high-obesity states (the exception being South Carolina).

We can illustrate the geographical distribution of adjusted obesity by plotting the state rankings after accounting for measurement error. The results for women are shown in Fig. 1. Here, we distinguish between low-obesity states, high-obesity states and states for which no clear ranking is possible. Even allowing for the potential loss of information due to measurement error, there is still an obvious grouping of high-obesity states in the mid-west and south of the US, with lower-obesity states located in the east and west. These findings indicate that cross-state variation in self-reported BMI in the USA, when used in conjunction with our bounds, may still provide a useful source of information for understanding the causes and consequence of obesity despite potential misreporting of obesity.

¹⁰ <http://www.gallup.com/poll/146534/obesity-lowest-colorado-highest-west-virginia.aspx>.

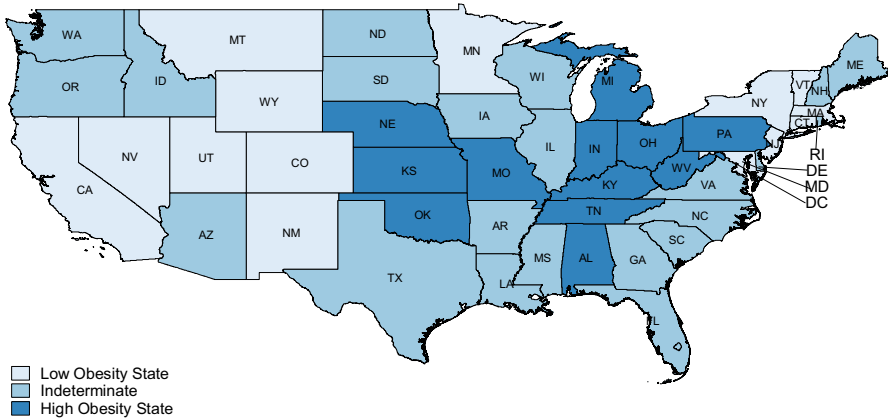


Fig. 1 Rankings of US states by female obesity rates allowing for potential loss of information from self-reported BMI

5 Discussion

The results from Sect. 4 suggest that meaningful obesity rankings across European countries are likely to be difficult using only self-reported data. This is because the estimated misclassification rate is higher in Europe than in the USA and also because the variation in obesity rates across countries is relatively small to begin with. Consequently, the rankings are not robust to adjustments which account for the loss of information arising from errors in self-reported obesity status. When confronted with cases like this, we would urge readers to be very cautious about making rankings. This is not a problem of imprecise estimates but rather one of identification, reflecting the cost of measurement error. In these cases, identification of meaningful rankings requires extra information over and above the minimal assumptions adopted in our analysis. Researchers are free to try alternative adjustments in an attempt to obtain narrower bounds or more detailed rankings. However, since our bounds are tight, given our assumptions, our findings make explicit the fact that any more detailed rankings can only be secured with additional information or assumptions. In this way, our analysis forces researchers to be explicit about these restrictions and to justify why they may be appropriate.

Our analysis of US obesity rates on the other hand shows that self-reported obesity rates may still form the basis of meaningful rankings, when used in conjunction with our bounds, provided that the misclassification rate is sufficiently low and the variation across observations is sufficiently large. Ideally, we suggest that readers use estimates of the misclassification rate in self-reported obesity (either those we provide or perhaps estimates of their own), along with our bounds, to derive interval estimates of the true obesity rate—this is not difficult in practice. However, if readers are looking for a quick rule of thumb, then we draw attention to the fact that, in every case we considered, the estimated lowest bound given by the actual reported obesity rate. In light of this, we would encourage readers to think of self-reported obesity rates as lower bounds on the true rate. In addition, the upper bounds for

the US states are all about 15–20% higher than the reported obesity rate. A quick rule of thumb for researchers interested in comparing obesity rates across states would therefore be to think in terms of intervals of the form $\{OB_{SR} : (1.20) * OB_{SR}\}$ for each state, rather than point estimates, where OB_{SR} is the self-reported obesity rate. If these intervals do not overlap, then you can rank the two states using only minimal assumptions on the error process. However, if the intervals overlap, then further restrictions would be required. For example, suppose that we wished to compare Arkansas (reported obesity rate of 36.61) with New Mexico (reported obesity rate of 29.95). Using our suggested rule of thumb rather than comparing the point estimates, we compare the interval (36.61:43.93) for Arkansas, with the interval (29.95:35.94) for New Mexico. Since these do not overlap, we remain confident that Arkansas has a higher obesity rate than New Mexico, even allowing for the self-reported nature of the data. However, a comparison of Arkansas and Louisiana would involve a comparison of Arkansas interval of (36.61:43.93) with the interval for Louisiana of (33.95:40.74). Since these overlap, we would caution against drawing inferences from a comparison of these two states even though the reported obesity rate in Arkansas is almost three percentage points higher than in Louisiana.

6 Conclusion

We examine the robustness of obesity rankings across ten European countries and fifty US states, taking into account any potential loss of information resulting from measurement error in self-reported BMI. In doing so, we place only minimal restrictions on the error process. Our analysis shows that a meaningful ranking of countries in Europe is difficult once measurement error in the European data is taken into account. The inability to rank European countries is a reflection of the cost of measurement error. Researchers who feel that the loss of rankings is too heavy a price to pay for adopting only minimal assumptions are free to adopt additional restrictions (or equivalently incorporate additional information about the error process) into their analysis in a quest for meaningful rankings; however, our results force such researchers to be explicit about what these restrictions are and why they are appropriate.

However, our analysis also shows that the use of self-reported obesity rates, in and of itself, need not eliminate all information. We illustrate this using data on obesity rates across the US states. In contrast to the cross-country European results where obtaining meaningful rankings was difficult, the measurement error-adjusted bounds we derive for the USA are still informative about the rankings of US states by obesity. This reflects the greater variation in obesity rates across US states, as a result of which the US rankings are more robust to the adjustments for measurement error. The US rankings that emerge, after allowing for measurement error, are consistent with rankings typically reported in the popular press based on raw uncorrected BMI. These results show that cross-state variation in self-reported obesity rates in the US, used in conjunction with the bounds we report, may still provide a useful input into an analysis of the causes and consequences of obesity in the USA.

References

- Antonanzas F, Rodriguez R (2010) Feeding the economics of obesity in the EU in a healthy way. *Eur J Health Econ* 11:351–353
- Brunello G, d'Hombres B (2007) Does body weight affect wages? evidence from Europe. *Econ Hum Biol* 5:1–19
- Burkhauser R, Cawley J (2008) Beyond BMI: the value of more accurate measures of fatness and obesity in social science research. *J Health Econ* 27:519–529
- Burkhauser R, Cawley J, Schmeiser M (2009) The timing of the rise in US obesity varies with measure of fatness. *Econ Hum Biol* 7:307–318
- Cawley J, Frisvold D, Meyerhoefer C (2013) The impact of physical education on obesity among elementary school children. *J Health Econ* 32:743–755
- Carroll RJ, Ruppert D, Stefanski LA, Crainiceanu C (2006) Measurement error in nonlinear models: a modern perspective. Chapman & Hall/CRC, Boca Raton
- Costa-Font J, Fabbri D, Gil J (2009) Decomposing body mass index gaps between Mediterranean countries: a counterfactual quantile regression analysis. *Econ Hum Biol* 7:351–365
- Costa-Font J, Fabbri D, Gil J (2010) Decomposing cross-country differences in levels of obesity and overweight: does the social environment matter? *Soc Sci Med* 70:1185–1193
- Dauphinot V, Wolff H, Naudin F, Gueguen R, Sermet C, Gaspoz JM, Kossovsky M (2009) New obesity body mass index threshold for self-reported data. *J Epidemiol Community Health* 63:128–132
- Ezzati M, Martin H, Sjkold S, Vander Hoorn S, Murray C (2006) Trends in national and state level obesity in the USA after correcting for self-report bias: analysis of health survey. *J R Soc Med* 99:25–257
- Giacchi M, Mattei R, Rossi S (1998) Correction of the self-reported BMI in a teenage population. *Int J Obes* 30:644–651
- Gorber SC, Temblay M, Moher D, Gorver B (2007) A comparison of direct vs self-reported measures for assessing height, weight and body mass index: a systematic review. *Obes Rev* 8:307–326
- Hu F (2008) Measurements of adiposity and body composition. In: Hu F (ed) *Obesity epidemiology*. Oxford University Press, New York, pp 53–83
- Kuczmarski M, Kuczmarski R, Najja M (2001) 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. *J Am Diet Assoc* 101:28–34
- Loureiro M, Nayga R (2005) International dimensions of obesity and overweight related problems: an economic perspective. *Am J Agric Econ* 87:1147–1153
- Molinari F (2008) Partial identification of probability distributions with misclassified data. *J Econ* 144:81–117
- Nanney M, Nelso T, Wall M, Haddad T, Kubik M, Nelson Laska M, Story M (2010) State school nutrition and physical activity policy environments and youth obesity. *Am J Prev Med* 38:9–16
- National Taskforce on Obesity (2005) *Obesity: the policy challenges. The report of the national task force on obesity*. Department of Health and Children, Dublin. <https://www.safefood.eu/SafeFood/media/SafeFoodLibrary/Documents/Professional/All-island%20Obesity%20Action%20Forum/obesity-taskforce.pdf>. Accessed 25 Mar 2015
- Nicoletti C, Peracchi F, Foliano F (2011) Estimating income poverty in the presence of missing data and measurement error. *J Bus Econ Stat* 29:61–72
- Nyholm M, Gullberg B, Merlo J, Lundqvist-Persson C, Rastam L, Lindblad U (2007) The validity of obesity based on self-reported weight and height: implications for population studies. *Obesity* 15:197–208
- O'Neill J, O'Neill D (2007) Health status, health care and inequality: Canada vs the US. In: Cutler D, Garber A, Goldman D (eds) *Frontiers in health policy research*. MIT Press, Cambridge
- O'Neill D, Sweetman O (2013) The consequences of measurement error when estimating the impact of obesity on income. *IZA J Labor Econ* 2:3. doi:10.1186/2193-8997-2-3
- O'Neill D (2015) Measuring obesity in the absence of a gold standard. *Econ Hum Biol* 17:116–128. doi:10.1016/j.ehb.2015.02.002
- Plankey MW, Stevens J, Flegal KM, Rust PF (1997) Prediction equations do not eliminate systematic error in self-reported body mass index. *Obes Res* 5:308–314
- Shields M, Gorber SC, Janssen I, Temblay M (2011) Bias in self-reported estimates of obesity in Canadian health surveys: an update on correction equations for adults. *Health Rep* 22:35–45

- Shiely F, Perry I, Lutomski J, Harrington J, Kelleher C, McGee H, Hayes K (2010) Temporal trends in misclassification patterns of measured and self-report based body mass index categories-findings from three population surveys in Ireland. *BMC Public Health*. doi:[10.1186/1471-2458-10-560](https://doi.org/10.1186/1471-2458-10-560)
- Stommel M, Schoenborn C (2009) Accuracy and usefulness of BMI measures based on self-reported weight and height: findings from Nhanes and NHIS 2001–2006. *BMC Public Health*. doi:[10.1186/1471-2458-9-421](https://doi.org/10.1186/1471-2458-9-421)
- Taber D, Chirqui J, Chaloupka F (2011) Geographic disparities in state and district policies targeting youth obesity. *Am J Prev Med* 41:407–414
- Villanueva E (2001) The Validity of self-reported weight in US adults: a population based cross-sectional study. *BMC Public Health*. doi:[10.1186/1471-2458-1-11](https://doi.org/10.1186/1471-2458-1-11)
- Visscher T, Viet A, Kroesbergen I, Seidel J (2006) Underreporting of BMI in adults and its effect on obesity prevalence estimations in the period 1998 to 2001. *Obesity* 14:2054–2063
- WHO (2009) Global health risks: mortality and burden of disease attributable to selected major risks. World Health Organisation, Geneva