Obesity and its relation to employment income: Does the bias in self-reported BMI matter?

Thomas Alexander Perks

Abstract

This study explores what difference, if any, the bias in self-reported body mass index (BMI) has on our understanding of the relationship between body size and income attainment. To accomplish this, aggregated data from Cycle 1 and Cycle 2 of the Canada Health Measures Survey, in which information on both self-reported and measured BMI was collected, are used. Based on subsamples of female and male employees, OLS regression analyses contrasting the effect of self-reported and measured BMI on income show that for women, self-reported BMI leads to underestimates of a negative body size effect, whereas for men, self-reported BMI leads to overestimates of a positive body size effect. Additional analyses examining the appropriateness of correction factors to improve the accuracy of self-reported BMI effect estimates suggest correction factors do little to reduce these systematic errors.

Keywords: body mass index, bias, income attainment, measured and self-reported, correction factors.

Résumé

Cette étude porte sur l’importance éventuelle du signalement de son propre indice de masse corporelle (IMC) pour comprendre la relation entre la taille du corps humain et le potentiel de revenu. Pour cela, on a utilisé les données regroupées des Cycles 1 et 2 du sondage sur les mesures de Santé Canada, où figure l’information sur l’IMC mesuré et signalé par le client. Selon les sous-échantillons des employés hommes et femmes, les analyses de régression MCO qui contrastent l’effet de l’IMC mesuré et signalé par le client sur le revenu, indiquent que pour les femmes, le résultat est un effet négatif sur la taille du corps, alors que pour les hommes, il en résulte un effet positif sur la taille du corps. Des analyses additionnelles examinant la pertinence des facteurs de correction dans le but de réduire les erreurs systémiques indiquent que les facteurs de correction n’ont que peu d’effet sur la réduction des erreurs systémiques.

Mots-clés: indice de masse corporelle, biais, potentiel de revenu, mesuré et signalé, facteurs de correction.

Canadian population rates of obesity have seen a steady increase over the past few decades, with one study suggesting that obesity rates in Canada have tripled since the early 1980s (Twells et al. 2014). This steady growth in the average “waistlines” of Canadians has paralleled the global growth (WHO 2015), and has given rise to numerous scholarly works examining obesity and its increasing salience in our lives. While much of this work has focused on the health risks associated with being overweight, other studies, although comparatively less common, have focused on the social consequences of obesity. For example, a number of studies published in Canada and elsewhere suggest...
that body size has a significant bearing on an individual’s employment income (Register and Williams 1990; Gortmaker et al. 1993; Loh 1993; Sargent and Blanchflower 1994; Averett and Korenman 1996; Puhl and Brownell 2001; Baum and Ford 2004; Cawley 2004; Cawley et al. 2005; Conley and Glauber 2007; Perks 2012; Mason 2012). With few exceptions, these studies show either significant negative effects of body size on income for both women and men, with the negative effect being stronger for women than for men, or significant negative effects for women only. These patterns have typically been understood as being a consequence of gendered forms of weight discrimination in the workplace. In other words, it appears that overweight people, and especially overweight women, are more likely to be overlooked for workplace opportunities and pay raises than their slimmer counterparts (Perks 2012). Indeed, there have been a number of human rights cases in Canada in recent years claiming discrimination based on body size (Luther 2010).

In these studies examining the body size–income relationship, it is a common practice for body size to be quantified using the body mass index (BMI), calculated as weight in kilograms divided by height in metres squared. Because of the practical difficulties as well as relatively high costs associated with collecting actual measurements of weight and height, it is typical for researchers to use BMI estimates that are derived from self-reports. However, numerous studies have shown that due to cultural norms about what constitutes ideal weight and height, respondents tend to underestimate their weight and overestimate their height when self-reporting, which, in turn, leads to underestimates of BMI and subsequent underestimates of obesity prevalence (Plankey et al. 1997; Connor Gorber et al. 2007; Shields et al. 2008; Stommel and Schoenborn 2009; Connor Gorber and Tremblay 2010; Krul et al. 2010). Although the extent of the self-report bias varies across studies (Connor Gorber et al. 2007), a study from Canada suggests that the discrepancy between self-reported and measured BMI underestimated the measured prevalence of obesity by as much as 8 per cent, with the same study reporting a much lower discrepancy (3 per cent) in the US (Connor Gorber and Tremblay 2010).

The purpose of this study is to examine the bias in self-reported BMI and assess its impact on our understanding of the association between body size and income, using aggregated data from representative samples of Canadians in which information on both self-reported and measured weight and height was collected. While other studies have explored the consequences of the bias in self-reported BMI on estimates of the health risks that are associated with obesity (Shields et al. 2008; Stommel and Schoenborn 2009), few if any studies have looked at what effect this bias has on our understanding of the body size–income relationship. Specifically, the present study seeks to: (1) determine the magnitude of the bias in self-reported BMI relative to BMI estimates based on measured weight and height; (2) assess what effect this bias has on our understanding of the relationship between body size and income attainment; and (3) explore the appropriateness of using correction factors to adjust self-reported BMI in order to account for these reporting biases and improve the accuracy of the estimated relationship between self-reported BMI and income.

Method

Data Source and Procedure

The data for these analyses come from the aggregated microdata files of the 2007–08 (Cycle 1) and 2009–10 (Cycle 2) Canada Health Measures Survey (CHMS), conducted by Statistics Canada and accessed through the Canadian Research Data Centre Network. The CHMS is especially suited for the present research purpose because, in addition to yielding a range of demographic questions, the
survey is, to my knowledge, the only recent Canadian survey data currently available that provides information on both self-reported and measured BMI. Consisting of an in-home general health interview, followed by a visit to a mobile examination centre, the survey collected health information from a representative sample of Canadians aged 6 to 79 years and aged 3 to 79 years in Cycle 1 and Cycle 2, respectively. In total, 5,604 people were sampled in Cycle 1 and 6,395 people were sampled in Cycle 2, resulting in a total aggregated sample of 11,387 individuals (for more information on the CHMS, see Statistics Canada 2011 and 2012). I chose as my working sample only those respondents aged 20–64 who reported being an employee and earning an income at the time of the survey. These restrictions were used to help ensure that the majority of respondents had completed secondary schooling, were of an age group typically associated with employment, and were presently working in the employ of others (i.e., not self-employed). The final sample restriction was that I only selected those respondents with valid values on all variables used in the analysis. These selection criteria yielded a working sample of 5,299 employees, with all data weighted and then rounded down by a constant in order to ensure an \(N\) equal to that of the original unweighted sample size. For my primary statistical procedure, OLS regression was used, with separate analyses carried out for female \((N = 2,657)\) and male \((N = 2,642)\) subsamples.

**Variables**

*Self-reported BMI* was calculated based on two questions that asked respondents: “How tall are you without shoes on?” and “How much do you weigh?” (in both cases, responses in imperial units were converted to metric). *Measured BMI* was based on actual measurements of respondents taken by trained Statistics Canada employees.

The dependent variable was personal income, which was logged due to its non-normal distribution, and was based on a question in the survey asking respondents about their annual personal income from all sources. Specifically, the question was phrased as follows: “What is your best estimate of your total personal income, before taxes and other deductions, from all sources in the past 12 months?” Notably, this question includes both workplace income and income from other sources, such as investment income or government transfers. As such, compared to measures of income specific to the workplace, the use of this question may lead to the understatement of the body size effect, since it is in the workplace where body size discrimination, and its impact on income, is often theoretically argued to operate. Unfortunately, this possibility cannot be modeled in the present analysis, since personal income questions specific to the workplace were not asked in the CHMS. As well, in anticipation of my findings, given the possibility of bias in the self-reporting of personal income (i.e., the tendency of some respondents to underreport/overreport their income), it is assumed that any bias in the reporting of personal income is unrelated to body size.

The following control variables were included in the analyses: height (mean centred and coded in centimetres), age (mean centred and coded in years), age-squared, marital status (dummy coded as ‘divorced/separated/widowed’ and ‘single, never married,’ with ‘married/common-law’ as the reference category), education level (dummy coded as ‘secondary school graduate,’ ‘some post-secondary,’ ‘post-secondary certificate or diploma,’ ‘bachelor’s degree,’ and ‘university degree or certificate

---

2. Residents living in the three territories, persons living on reserves or other Aboriginal settlements, full-time members of the Canadian Armed Forces, those residing in institutions, and residents of certain remote regions of Canada were excluded from the sample.
3. Respondents aged 3 to 5 years in Cycle 2 were excluded from the aggregated data file.
4. The mean age of respondents in the working sample was 40.4 years for females and 39.5 years for males.
above bachelor’s degree,’ with ‘less than secondary school graduation’ as the reference category), immigration status (coded as ‘non-immigrant’ and ‘immigrant’), working status (coded as ‘less than 30 hours per week’ and ‘30 or more hours per week’), level of physical activity (dummy coded as ‘moderately active’ and ‘active,’ with ‘non-active’ as the reference category), health status (dummy coded as ‘fair,’ ‘good,’ ‘very good,’ and ‘excellent,’ with ‘poor’ as the reference category), and occupation (dummy coded based on the 10 broad occupational categories of the 2011 National Occupational Classification, with ‘management occupations’ as the reference category). Also included in the analysis, for both self-reported and measured BMI, was BMI². This was added to account for any non-linearity in the relationship between BMI and income, as suggested by the literature (Godley and McLaren 2010; Perks 2012).

To address the question of whether the bias in self-reported BMI can be adjusted to improve the accuracy of self-reported BMI as a predictor of income, two new variables, called adjusted BMI1 and adjusted BMI2, were created based on estimates from two different regression equations. Specifically, for adjusted BMI1, measured weight was regressed on self-reported weight and measured height was regressed on self-reported height. Using the estimated coefficients from these regression equations, the self-reported values of weight and height were adjusted and used to calculate adjusted BMI1. For adjusted BMI2, a similar procedure was used, except that in addition to self-reported weight and height, education level and age were also included in the prediction equations, as preliminary investigation showed that of the control variables included in the analysis, age and education level were among the strongest predictors of the bias in self-reported weight and height for both women and men. The regression equations used for the adjusted BMI variables can be found in Table 1.

Table 1. Regression equations used to calculate adjusted measures of BMI, by gender.

<table>
<thead>
<tr>
<th>Gender</th>
<th>Adjusted BMI1</th>
<th>Adjusted BMI2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female</td>
<td>Adjusted height = 18.132 + (self-reported height × 0.886)</td>
<td>Adjusted height = 19.247 + (self-reported height × 0.876) + (age × −0.029) + (educdummy1 × 0.262) + (educdummy2 × 0.391) + (educdummy3 × 0.062) + (educdummy4 × 0.222) + (educdummy5 × 0.198)</td>
</tr>
<tr>
<td></td>
<td>Adjusted weight = −2.597 + (self-reported weight × 1.061)</td>
<td>Adjusted weight = −0.654 + (self-reported weight × 1.063) + (age × −0.003) + (educdummy1 × 1.043) + (educdummy2 × 0.541) + (educdummy3 × 1.100) + (educdummy4 × 0.613) + (educdummy5 × 1.053)</td>
</tr>
<tr>
<td>Male</td>
<td>Adjusted height = 12.523 + (self-reported height × 0.923)</td>
<td>Adjusted height = 13.772 + (self-reported height × 0.916) + (age × −0.024) + (educdummy1 × −0.129) + (educdummy2 × 0.027) + (educdummy3 × −0.119) + (educdummy4 × 0.144) + (educdummy5 × −0.122)</td>
</tr>
<tr>
<td></td>
<td>Adjusted weight = −2.426 + (self-reported weight × 1.033)</td>
<td>Adjusted weight = −3.368 + (self-reported weight × 1.035) + (age × −0.002) + (educdummy1 × 0.704) + (educdummy2 × 1.123) + (educdummy3 × 0.830) + (educdummy4 × 1.024) + (educdummy5 × 1.260)</td>
</tr>
</tbody>
</table>

Results

For descriptive purposes, measured BMI scores for females and males by age group (20–34, 35–49, and 50–64) are presented in Table 2. The table shows that females have lower BMI scores, on average, relative to males across all age cohorts, and that average BMI scores increase progressively among older age cohorts. This latter finding, based on cross-sectional data, is consistent with
longitudinal studies showing that average BMI increases as individuals age (Sheehan et al. 2003; Baum and Ruhm 2009).

Table 2. Measured BMI for females and males, by age group.

<table>
<thead>
<tr>
<th>Age group</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>20–34 (N = 926/1,010)*</td>
<td>20–64 (N = 2,657/2,642)</td>
</tr>
<tr>
<td>35–49 (N = 1,022/1,019)</td>
<td>50–64 (N = 709/612)</td>
</tr>
<tr>
<td><strong>Female</strong></td>
<td></td>
</tr>
<tr>
<td>25.47</td>
<td>27.32</td>
</tr>
<tr>
<td>25.88</td>
<td>28.16</td>
</tr>
<tr>
<td><strong>Male</strong></td>
<td></td>
</tr>
<tr>
<td>26.11</td>
<td>28.08</td>
</tr>
</tbody>
</table>

*Reported Ns refer to females and males, respectively.

In Table 3, differences in the average height, weight, and BMI derived from the measured, self-reported, and adjusted data are reported for women and men separately. As the table shows, for the self-reported data, there are systematic biases for both women and men across all three variables. Specifically, for women, height was, on average, overreported by 0.49 cm, weight was underreported by 1.61 kg, and the corresponding BMI based on self-reported weight and height was 0.76 kg/m² lower than the BMI based on measured values. For men, these patterns of under- and overreporting on weight and height, along with the corresponding difference in BMI, were similar to those of women, with men overreporting their height by 1.09 cm, underreporting their weight by 0.39 kg, and BMI being 0.46 kg/m² higher based on measured versus self-reported data. Notably, the average difference between measured and self-reported height was slightly higher among men compared to women, and the average difference between measured and self-reported weight was slightly higher among women compared to men—likely due to the gender stereotypes that are associated with weight and height—with these disparities resulting in a slightly higher BMI difference among women relative to men.

Also included in Table 3 are the average height, weight, and BMI derived from the adjusted data. As the table shows, the adjusted data offer much more accurate estimates of average height, weight, and BMI relative to the self-reported data, with the average discrepancies between the adjusted and measured data for either women or men being no more than 0.06, 0.03, and 0.02 units for height, weight, and BMI, respectively, with the discrepancies for self-reported data for either women or men being no lower than 0.49, 0.39, and 0.46 units, again for height, weight, and BMI, respectively.

Table 3. Differences in mean height, weight, and BMI for measured, self-reported, and adjusted data, by gender.

<table>
<thead>
<tr>
<th>Female</th>
<th>Measured</th>
<th>Self-reported</th>
<th>Diff.*</th>
<th>Adjusted BMI1</th>
<th>Diff.</th>
<th>Adjusted BMI2</th>
<th>Diff.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Height (cm)</td>
<td>162.47</td>
<td>162.96</td>
<td>−0.49</td>
<td>162.51</td>
<td>−0.04</td>
<td>162.44</td>
<td>0.03</td>
</tr>
<tr>
<td>Weight (kg)</td>
<td>70.87</td>
<td>69.26</td>
<td>1.61</td>
<td>70.89</td>
<td>−0.02</td>
<td>70.84</td>
<td>0.03</td>
</tr>
<tr>
<td>BMI (kg/m²)</td>
<td>26.81</td>
<td>26.05</td>
<td>0.76</td>
<td>26.79</td>
<td>0.02</td>
<td>26.80</td>
<td>0.01</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Male</th>
<th>Measured</th>
<th>Self-reported</th>
<th>Diff.*</th>
<th>Adjusted BMI1</th>
<th>Diff.</th>
<th>Adjusted BMI2</th>
<th>Diff.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Height (cm)</td>
<td>175.91</td>
<td>177.00</td>
<td>−1.09</td>
<td>175.89</td>
<td>0.02</td>
<td>175.85</td>
<td>0.06</td>
</tr>
<tr>
<td>Weight (kg)</td>
<td>84.77</td>
<td>84.38</td>
<td>0.39</td>
<td>84.74</td>
<td>0.03</td>
<td>84.80</td>
<td>−0.03</td>
</tr>
<tr>
<td>BMI (kg/m²)</td>
<td>27.35</td>
<td>26.90</td>
<td>0.46</td>
<td>27.34</td>
<td>0.01</td>
<td>27.38</td>
<td>−0.02</td>
</tr>
</tbody>
</table>

*Calculated as measured minus self-reported and measured minus adjusted.

†All differences in the table are statistically significant at p≤0.05.
Table 4 assesses the question of what effect the systematic biases in self-reported weight and height have on the estimated relationship between BMI and income, as well as what difference the adjustment of BMI makes to these estimates. Presented in the table are the regression estimates, after controls, for BMI and BMI\(^2\) from the four regression models (i.e., the fully controlled models with measured BMI, self-reported BMI, adjusted BMI\(_1\), and adjusted BMI\(_2\), along with their corresponding quadratic, as the independent variables). Focusing first on self-reported BMI, we see that for women, there is a negative BMI effect, although this negative effect lessens as the BMI increases. The opposite is true for men, where there is evidence of a positive BMI effect—although, like women, this effect diminishes as the BMI increases. Importantly, although the general patterns of a negative BMI effect for women and positive BMI effect for men are also evident for measured BMI, the estimated coefficient associated with self-reported BMI is smaller compared to the coefficient for measured BMI among women, whereas for men, the opposite pattern is evident, with the coefficient associated with self-reported BMI being larger compared to the coefficient associated with measured BMI. What these results by gender suggest is that the negative body size–income relationship for women is stronger than we would expect, or underestimated, based on self-reported BMI, and for men, the positive body size–income relationship is weaker than we would expect, or overestimated, based on self-reported BMI. Regarding the regression estimates associated with the adjusted BMI measures, the results from the table suggest that adjusting BMI does little to improve the accuracy of the self-reported BMI estimate, since the coefficients based on the adjusted BMI variables closely approximate the coefficients associated with self-reported BMI.

Table 4. OLS regression coefficients (in log dollars) estimating the relationship between the different measures of BMI (measured, self-reported, adjusted BMI\(_1\), and adjusted BMI\(_2\)) and income, with controls,\(^a\) by gender.

<table>
<thead>
<tr>
<th></th>
<th>Measured BMI(^b)</th>
<th>Self-reported BMI</th>
<th>Adjusted BMI(_1)</th>
<th>Adjusted BMI(_2)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Females</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>11.052</td>
<td>10.694</td>
<td>10.698</td>
<td>10.700</td>
</tr>
<tr>
<td>BMI</td>
<td>−0.077</td>
<td>−0.055</td>
<td>−0.054</td>
<td>−0.054</td>
</tr>
<tr>
<td>BMI(^2)</td>
<td>0.001</td>
<td>0.001</td>
<td>0.001</td>
<td>0.001</td>
</tr>
<tr>
<td><strong>Males</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BMI</td>
<td>0.039</td>
<td>0.061</td>
<td>0.055</td>
<td>0.053</td>
</tr>
<tr>
<td>BMI(^2)</td>
<td>−0.001</td>
<td>−0.001</td>
<td>−0.001</td>
<td>−0.001</td>
</tr>
</tbody>
</table>

\(^a\)Includes controls for height, age, age-squared, marital status, education level, immigrant status, working status, physical activity, health status, and occupation.

\(^b\)All coefficients in the table are statistically significant at \(p \leq 0.05\).

As a supplement to the regression results presented in Table 4, Figures 1 and 2 provide the reader with a visual sense of what these different regression estimates—based on self-reported and measured BMI, along with the adjusted BMI estimates—mean to our understanding of the body size–income relationships for women and men. To improve the intuitive meaning of the estimates, the figures are based on the antilog of predicted log income estimates, in order to provide actual estimated dollar values that are associated with the coefficients. They are especially useful for showing the different conclusions one would reach based on which BMI measure is used, particularly with respect to the degree of advantage/disadvantage in income attainment. As Figure 1 shows, for women the measured BMI regression is visually more pronounced (i.e., steeper) relative to the regression line based on the self-reported measure, whereas in Figure 2 it is the opposite case for men, where the
regression line for measured BMI is noticeably more gradual relative to the regression line for self-reported BMI. It is also evident in the figures that any improvement to the accuracy of self-reported BMI by adjusting the self-reported values is, at best, minimal for both women and men.

Figure 1. Relationship between different measures of BMI (measured, self-reported, adjusted BMI1, and adjusted BMI2) and income based on the regression estimates from Table 4, females only.

Figure 2. Relationship between different measures of BMI (measured, self-reported, adjusted BMI1, and adjusted BMI2) and income based on the regression estimates from Table 4, males only.
Conclusion

To summarize, the findings of this study are threefold. First, there is a systematic bias in the self-reporting of weight and height among employed women and men that results in biased BMI estimates, and the extent of the bias differs by gender. Second, for women the findings suggest that using self-reported BMI to assess the body size–income relationship leads to underestimates of the negative effect of body size on income. In contrast, for men the findings suggest that self-reported BMI leads to overestimates of the positive effect of body size on income. And third, the findings suggest that using correction factors to “improve” the accuracy of self-reported BMI does little, if anything, to improve body size effect estimates.

How should these findings be interpreted? One interpretation, at least for women, is that if we consider that almost all prior research examining the body size–income relationship has been based on self-reported data, and that most prior research has found a significant negative effect among women, the negative body size effect on income is larger than what prior work suggests. This is an important finding for those interested in combating social inequality, as it speaks to the possibility that the degree of inequality, at least with respect to the possibility of discrimination due to body size, is more consequential than past work shows. For men, the positive body size effect is smaller than what one might expect based on self-reported data. It should be noted, however, that the finding of a positive BMI effect on income among men is somewhat unique relative to most other studies of the body size–income relationship, which have tended to find either marginally negative body size effects or no significant effects of body size on income at all among male employees. At the same time, as has been noted elsewhere (Perks 2012), much of this prior work on the body size–income relationship has examined BMI as a dichotomous measure, in which the effect of body size is assessed by comparing the earnings of non-obese and obese men. Although this design is a reasonable one for assessing income disparities between groups, it cannot discern the potentially more nuanced relationship between body size and income, at least among males, that is found in the present study. In other words, the use of a dichotomous measure of BMI may be masking what appear to be important non-linear and significant positive effects among men.

Still, the question remains as to what can be done about this bias in self-reported BMI. As has been noted, the use of correction equations, based on a regression procedure, was unable to compensate for these reporting errors. Equally discouraging, perhaps, is that the BMI effect estimates from the more complicated regression model (i.e., adjusted BMI2) did no better than the estimates from the simpler model (i.e., adjusted BMI1), suggesting that even more elaborate models than the ones used here, too, may do little to address the biased BMI effect estimates. In fact, additional regression analyses (not shown) with all of the control variables included in the prediction equation did no better in creating a more accurate predictive tool. What this means for future studies examining the body size–income relationship in which self-reported BMI is used, as well as the conclusions that are derived from them, is unclear.

One alternative, perhaps, is that other measures may be needed in order to ensure that measures of body size are able to more accurately represent the true relationship between body size and income. For example, the use of waist circumference as a measurement of obesity has garnered some attention in the medical literature as an alternative to BMI, and has in some cases been found to offer a more accurate assessment of the health risks associated with obesity (Ashwell et al. 2011). This may be true of the assessment of the body size–income relationship as well. Of course, the issue of the relative difficulty of collecting information on waist circumference is similar to that of measured
BMI, where the costs associated with collecting actual measurements of people are considerable, and in many cases prohibitive. As such, it remains an open research objective to come up with solutions that would reduce the reporting bias in self-reported BMI, in order to ensure that these estimates accurately represent the true relationship between body size and income.

Acknowledgments

The analysis was conducted at the University of Lethbridge Branch RDC, which is part of the Canadian Research Data Centre Network (CRDCN). The services and activities provided by the CRDCN are made possible by the financial or in-kind support of SSHRC, CIHR, CFI, Statistics Canada, and participating universities, whose support is gratefully acknowledged. The views expressed in this paper do not necessarily represent the CRDCN’s or that of its partners.

References


