Article:
Davillas, Apostolos and Jones, Andrew Michael (2021) The implications of self-reported body weight and height for measurement error in BMI. Economics Letters. 110101. ISSN: 0165-1765

https://doi.org/10.1016/j.econlet.2021.110101

The implications of self-reported body weight and height for measurement error in BMI*

Apostolos Davillas†
Andrew M Jones‡

Abstract

We designed an experiment to explore the extent of measurement error in body mass index (BMI), when based on self-reported body weight and height. We find that there is a systematic age gradient in the reporting error in BMI, while there is limited evidence of systematic associations with gender, education and income. This is reassuring evidence for the use of self-reported BMI in studies that use it as an outcome, for example, to analyse socioeconomic gradients in obesity. However, our results suggest a complex structure of non-classical measurement error in BMI, depending on both individuals' and within-household peers' true BMI. This may bias studies that use BMI based on self-reported data as a regressor. We also observe non-classical reporting error in height and weight — taller people seem to report their height more accurately, while a nonlinear relationship is evident for weight, with a sharper increase in reporting errors for those of greater weight. Common methods to mitigate reporting error in BMI using predictions from corrective equations do not fully eliminate reporting heterogeneity associated with individual and within-household true BMI. Overall, the presence of non-classical error in BMI highlights the importance of collecting measured body weight and height data in large social science datasets.

Keywords: BMI; Experiment; Measurement error; Reporting bias.

JEL codes: I10, C18, C50

---

* We are grateful to David Madden, Donal O'Neill, Nicolas Ziebarth, Robert Sugden and Stefan Penczynski for their comments on an earlier version of this paper. Understanding Society is an initiative funded by the Economic and Social Research Council (ESRC) and various Government Departments, with scientific leadership by the Institute for Social and Economic Research, University of Essex, and survey delivery by NatCen Social Research and Kantar Public. This paper makes use of data from the Understanding Society Innovation Panel administered by ISER, University of Essex, and funded by the Economic and Social Research Council. The research data are distributed by the UK Data Service. Apostolos Davillas and Andrew M. Jones acknowledge financial support from the ESRC (ES/V003356/1). The funders, data creators and UK Data Service have no responsibility for the contents of this paper.

† Health Economics Group, Norwich Medical School, University of East Anglia.

‡ Department of Economics and Related Studies, University of York; Centre for Health Economics, Monash University.

1 Introduction

Obesity is associated with increased risks of morbidity and mortality. This has led to a plethora of studies on the socio-economic consequences of obesity, such as labour market outcomes (Cawley, 2015). Because of the absence of measured anthropometric data in large-scale datasets[1], many existing studies are based on self-reports (Cawley, 2015; Cawley et al., 2015; Gil and Mora, 2011). The reliability of these measures in social science datasets is therefore of critical importance for obesity research.

We designed an experiment to explore the extent of measurement error in body mass index (BMI), when based on self-reported body weight and height, in the context of a multi-purpose survey. We collected information on self-reported body weight and height data immediately before the relevant physical measurements were taken.[2] The limited number of existing econometric analyses that examine measurement errors in anthropometrics mostly compare self-reports and measured anthropometric data that were collected with a considerable time difference and/or respondents were informed about the subsequent physical measurements (Cawley et al., 2015; Gil and Mora, 2011); these are also studies that are based on selected population samples (O'Neill and Sweetman, 2013). Moreover, the majority of the medical literature confirms the presence of measurement error in self-reported anthropometrics, although it is often based on selected age groups, non-representative samples and does not aim to explore the potential implications of the measurement error for econometric modelling (e.g., Engstrom et al., 2003, Gorber et al., 2007, Keith et al., 2011).

The implications of measurement error are different depending on whether BMI is to be used as an outcome or as an explanatory variable. We explore whether the implied measurement error in BMI is systematically associated with socio-economic variables used in inequalities research. This is relevant for studies that use BMI as an outcome, modelled as a function of socioeconomic status (SES), and where measurement error contributes to the error term of the BMI regression equation. In addition, we explore whether the measurement error in BMI is non-classical, i.e., systematically associated with the measured values, and whether this association varies depending on the BMI of other household members. Non-classical measurement error may cause bias in regression models for other outcomes (e.g., earnings, health care costs) that use BMI as a regressor, even when instrumental variable methods are used to deal with endogeneity or errors-in-variables (O'Neill and Sweetman, 2013; Cawley et al., 2015). To

[1] For example, the Survey of Health, Ageing and Retirement in Europe (SHARE), the European Community Household Panel (ECHP), the British Household Panel Survey (BHPS), the German Socio-Economic Panel (GSOEP) as well as the National Longitudinal Survey of Youth (NLSY), the Medical Expenditure Panel Survey (MEPS) and the Behavioral Risk Factor Surveillance System (BRFSS) in the US are datasets that are frequently used for economics of obesity research but do not collect physical measures of body weight and height.

[2] The questionnaire design is structured so that respondents' consent on the measurement followed their self-reports of weight/height and, thus, the latter is not contaminated by their informed consent to have their anthropometric measured.

explore the underlying sources of the observed reporting error in BMI, reporting error in body weight and height are also analysed separately in our study.

As an extension, we revisit existing practices on using corrective equations[3] to partially address reporting error in weight and height in the absence of measured data (Cawley et al., 2015). We show that the predicted BMI values from these corrective equations still suffer from measurement error that depends on an individual's own and within household peer's measured BMI.

2 Data

Understanding Society is a UK nationally representative household panel survey. One of its features is a sub-panel, the Innovation Panel (UKHLS-IP), reserved for experimental work.[4]

As part of the UKHLS-IP wave 12, we designed an experiment on the survey measurement of anthropometrics. Respondents were first asked for their self-reported body weight and height, followed by physical measurements[5].The respondents gave their informed consent for these measurements (that follow conventional best practices on measurement of anthropometrics) at the point they were collected, which follows their self-reports of body weight/height. We focus on adults (aged 20+) here to eliminate any puberty-related body-size changes.

Two BMI measures are calculated, as the body weight (Kg) divided by the square of height (metres), separately for the measured and the self-reported data. To facilitate interpretation of results, the absolute differences between the BMI based on self-reports and measured body weight and height data is used in our analysis (Cawley et al., 2015; Gil and Mora, 2011)[6]. We also implemented separate analyses using absolute differences between the self-reported and measured body weight and height data.

[3] Corrective equations for self-reported body weight and height data rely on the external validity of the association between measured and reported weight/height values from one dataset to another. Corrections are based on the association in an external dataset that is translated to the main analysis dataset ("transferability" assumption).
[4] The UKHLS-IP sample covers England, Wales and Scotland south of the Caledonian Canal.
[5] Households were randomly allocated to two different survey modes to collect self-reports of body weight and height: a self-completion and an open interview mode. As we found no differences in reporting error by interview mode, these samples are pooled for our analysis.
[6] A limitation of raw reporting error is that under- and over-reports may cancel each out and, thus, creating a misleading impression on the error's magnitude. One may argue that under-reporting in BMI may be more important for a public health perspective as it may have more serious health consequences and result in an underestimation of the true overweight and obesity prevalence. However, the main scope of our analysis is to explore whether measurement error is non-classical. Given that the presence of non-classical error matters for models that use BMI as an explanatory variable (for example, wage equations, health care demand and costs), both under-reports and over-reports are of equal importance to get an unbiased estimate of the effect of adiposity on the outcome of interest.

Our regression models for the absolute reporting error account for gender and age (in years divided by 10) ─ polynomials in age are used where they improve the statistical fit of the models. In the BMI absolute measurement error models, a cubic polynomial in age was included as this is supported by Wald tests appropriate to compare nested models (linear, quadratic, cubic and quartic associations with BMI absolute measurement error)[7]. Analogously, quadratic and cubic age associations with absolute reporting error in body weight and height are employed, respectively.

Our SES measures are collected at UKHLS-IP wave 11: the highest educational qualification that a respondent had ever achieved[8] (degree/post-secondary; A-level/equivalent; GCSE/equivalent; basic/no qualification) and household income (equivalised and log transformed). To explore the role of within-household peer effects, we use a dummy variable for being part of a household with low/moderate BMI levels ("low_hh_BMI"), defined as having an average BMI for all other adult household members, apart from the respondent, that is below the obesity threshold (<30kg/m^2)[9]. Descriptive statistics for all explanatory variables used in our analysis are presented in Table A.1 (Appendix).

After excluding information on all variables used in our analysis and focusing on adults (aged 20+), our working sample restricted to 873 individuals, from a potential sample of 1,058 respondents. To allow for our results to be generalised to the population of Great Britain, we use sample weights that account for differential nonresponse, unequal selection during the sampling and non-response to our experiment. These weights are calculated by adjusting the UKHLS-IP wave 11 weights using a backward stepwise probit model on predictors from UKHLS-IP wave 11.

[7] Compared to the linear and quadratic associations with reporting error in BMI, a cubic polynomial in age is selected as the estimated coefficients of the cubic polynomial are statistically significant (jointly) as well as the age-cubed coefficient itself, following standard procedures on comparing nested models. Moreover, higher order age polynomials (such as, quartic associations) result in non-significant age coefficients (for the higher order age polynomials) and no improvements in statistical fit of the BMI absolute measurement error models.

[8] It should be explicitly noted that UKHLS does not provide/collect years of schooling as a (derived) variable that covers all qualifications that individuals may have obtained – instead, respondents are asked to report the highest educational or school qualification they have obtained. Although crude and incomplete (as it does not account multiple degrees of the same level, studies that did not reach the level of a formal qualification and differences in the length of study for degrees obtained overseas) we have created a continuous years of schooling variable based on estimates of years of schooling for each highest educational level reported from each respondent; these results (available upon request) shows that our conclusions about the absence of any systematic associations between reporting error in BMI and education remained robust.

[9] Children are excluded here as (age and gender-specific) BMI in childhood should be interpreted differently than adult obesity. BMI values above the obesity threshold may be of particular interest here as they are more visible in people's silhouettes and, thus, more likely to exert peer-effects on reporting behaviour (Lønnebotn et al., 2018).

3 Methods

Absolute reporting error in BMI is modelled by linear regression. Regression models are first estimated using the set of demographics and SES. To explore whether measurement error is non-classical, we add an individual's own BMI based on their measured data. This specification is augmented by adding BMI information for the other household members and its interaction with an individual's own (measured) BMI[10]. Although BMI is of the main measure of interest in the majority of the relevant existing research in the economics of obesity (and, thus, its reporting error of particular importance), BMI is itself derived from body weight and height; as an additional analysis, we also explore the potential underlying factors for the absolute error in body weight and height as well as to what extent measured weight and height affect reporting behaviour in the corresponding self-reported measures (non-classical measurement error).

As an extension, we test whether the conventional method of using corrective equations for self-reports of body weight/height is sufficient to mitigate reporting error in BMI and, more, importantly its systematic association with covariates (Cawley, 2015).[11] Availability of self-reported and measured data allow us to estimate analogous corrective equations by regressing measured weight and height data on self-reports and a vector of demographics. To mimic correction procedures for self-reported anthropometrics in the existing studies, the predictions from these equations are used to calculate self-reports of body weight and height that are corrected for reporting error. To explore the remaining reporting error following this correction procedure we compute the absolute difference between the corrected and measured BMI. This measure of the remaining reporting error is regressed on our set of demographics, SES and individual's own and within-household peer's (measured) BMI to explore whether there are still systematic associations with these factors.

4 Results

Figure 1 shows that there is a high correlation between reported and measured BMI data. However, there is not a perfect match — reporting error is more likely to result in under-reporting of BMI than over-reporting; more of the data points are concentrated above, rather than below, the 45-degree equality line. Despite the small differences that initial visualisation in Figure 1 shows, obesity prevalence is systematically higher when

---

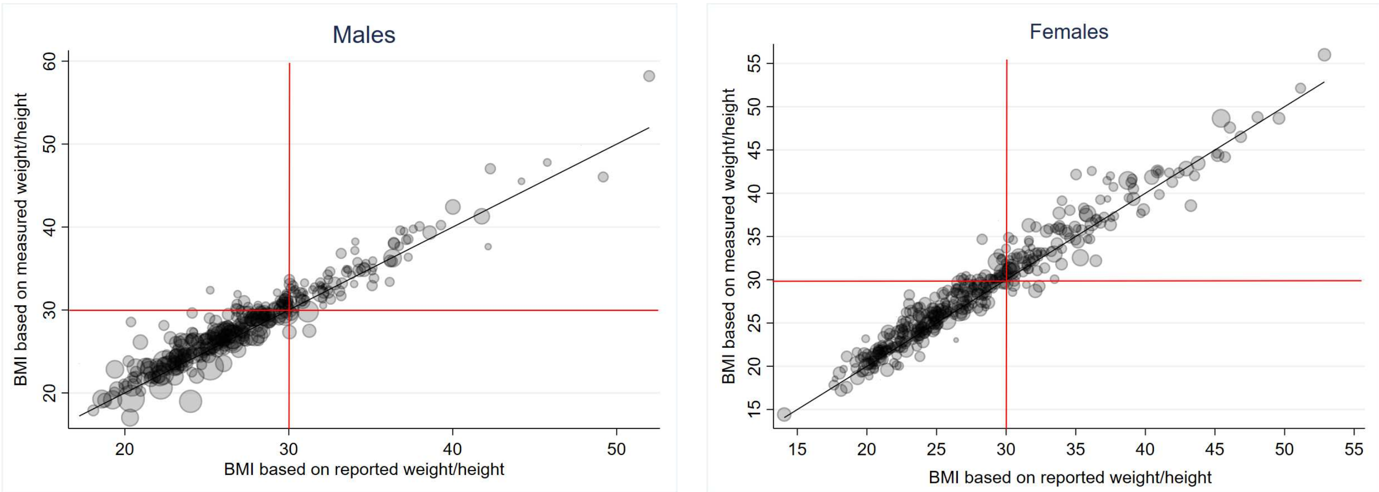
[10] We use interaction terms as a way to explore the presence of potentially heterogenous effects in the role of measured BMI on reporting error subject to within-household peers' adiposity levels ("low_hh_BMI").

[11] In the existing economic studies of obesity that rely on self-reports only, corrective equations — based on the relationship between measured and self-reported body weight and height data — are estimated using alternative, rather than the main analysis, data source. Then, the coefficients from these equations are transferred to the analysis sample and, after multiplying the coefficients by the self-report values, they obtain measures of weight and height corrected for the reporting error (Cawley, 2015).

based on measured (36% and 32% for females and males) as opposed to self-reported data (32% and 26% for females and males); this is evident in Table A.2 (Appendix).

To further quantify the magnitude of reporting error, the mean of raw reporting errors (defined as reported minus measured BMI) show that, on average, respondents over-report their height (by 1.196 cm), under-report their weight (by 0.941 kg) and, consequently, BMI is underestimated by 0.738kg/m^2 (Tables 1 and Table A.3, Appendix). Our absolute measure of the reporting error shows that both under- and over-reports result in an average total error in BMI of 1.3kg/m^2, i.e., 4.4% of measured BMI (Table 1). Graphs of the distributions of reporting error in weight, height and BMI are presented in Figure A.1 (Appendix).

Figure 1. Scatter plots of measured versus reported BMI.



Notes: Markers are scaled to reflect sample weights. Darker regions representing more concentrated data points. The black line is a 45-degree line.

Table 1. Summary statistics for measurement error in BMI(kg/m^2).

	Mean	Std. Dev.
Raw error	-0.738	1.541
Absolute error	1.266	1.147
Absolute error (%measured)	4.418	3.927
Sample size	873	

Note: Sample weights are accounted for.

4.1. Regression analysis

Table 2 presents regression analyses of the absolute BMI reporting error. We find a non-linear and systematic association between age and the absolute reporting error in BMI across all model specifications. Figure 2 presents the adjusted predictions at representative values (APRs), i.e., the predicted absolute reporting error in BMI across selected age values with all the other variables kept at their initial values; this graph shows a steep increase in predicted errors in BMI for those aged 70 and above (Figure 2). No systematic associations are evident for gender and SES. If BMI is the outcome of interest, for example, in an analysis of socioeconomic inequality, then we do not find systematic reporting error by SES.

Specification 2 shows that measurement error in BMI is non-classical, with the respondent's own measured BMI being positively associated with the absolute error in BMI. Conditional on the individuals' own BMI, their within-household peers BMI also plays an important role (specification 3). Specifically, as illustrated in Figure 3, although the predicted absolute error in self-reported BMI increases in magnitude for every unit increase in an individual's measured BMI *ceteris paribus*, there is heterogeneity related to household-peer effects as suggested by the interaction term (Table 2). Respondents with measured BMI of around 31 and above, which coincides with the obesity threshold, reported anthropometrics more accurately (lower reporting error in BMI) when living in households with other members having low or moderate BMI values as opposed to those living in households with excess BMI levels.[12]

Sensitivity analysis shows that the time of anthropometric measurement during the day does not affect our results in Table 2; main effect of the time of the day and its interaction terms with measured BMI are not statistically significant (p-values>0.10).Our results on the role of within household peer-effects in absolute reporting error in BMI (as graphically illustrated in Figure 3) remain robust to a sensitivity analysis that restricts our sample to those households with two and more household members (about 60% of our working sample) who responded and provided body weight and height data (Figure A.2., Appendix).

We have also implemented separate analyses on the absolute measurement error in body weight and height (Tables A.4 and A.5, Appendix). Our results show that there is no systematic SES gradient in absolute reporting error in weight (Table A.4, specification 2), while there is some weak evidence of lower absolute reporting error in height for those with higher income (Table A.5, specification 2). We also observe a U-shaped association between age and reporting error in weight and a steep increase in reporting error in height for those of 80 years old and above (Figure A.3, Appendix)[13].

[12] Overall, this suggests the presence of another source of non-classical measurement error, known as differential measurement error (O'Neill and Sweetman, 2013). Typically, differential measurement error is defined as another type of non-classical measurement error that arises when the reported measurement error in BMI is correlated with the error term in econometric models for which BMI is the independent variable of interest (O'Neill and Sweetman, 2013). For example, assume that the within-household peers' BMI matters for the reporting error in individual's BMI, as we will explore here. If the within-household peers' BMI is directly or indirectly correlated with the outcome of interest (often, in our case, labour marker outcomes, healthcare demand etc.) and, thus, captured by the regressions' residuals, then, differential reporting error exists.

[13] One may argue that the observed steep increase in reporting error in height for the elderly may be due to height loss related to physical ageing due to changes in the bones, muscles, and joints rather than because of any genuine reporting pattern by age (or due to cognitive

The combination of these patterns results in the nonlinear association with age observed for the absolute reporting error in BMI (Figure 3). Moreover, our results on the presence of non-classical measurement error in body height, i.e., those of higher height report their height more accurately (Table A.5), and weight, where a nonlinear relationship is evident, with a steeper increase in reporting error for those of higher measured weight (Table A.4 and Figure A.4), confirms and disentangles the observed non-classical error for the derived BMI measure (Figure 2).

4.2. Correction equations

Table 3 presents regression analysis on the absolute difference between BMI based on predictions from the corrective equations (Table A.7, Appendix) and measured BMI. Measured BMI still plays a systematic role in the remaining error in BMI (after the correction). Moreover, we observe similar patterns (but with lower magnitude of predicted errors) for the heterogenous role of individuals' measured BMI (on the remaining reporting error in BMI) related to household-peer effects (Figure 4) to those observed in Figure 3, without adjustments using the corrective equations.

---

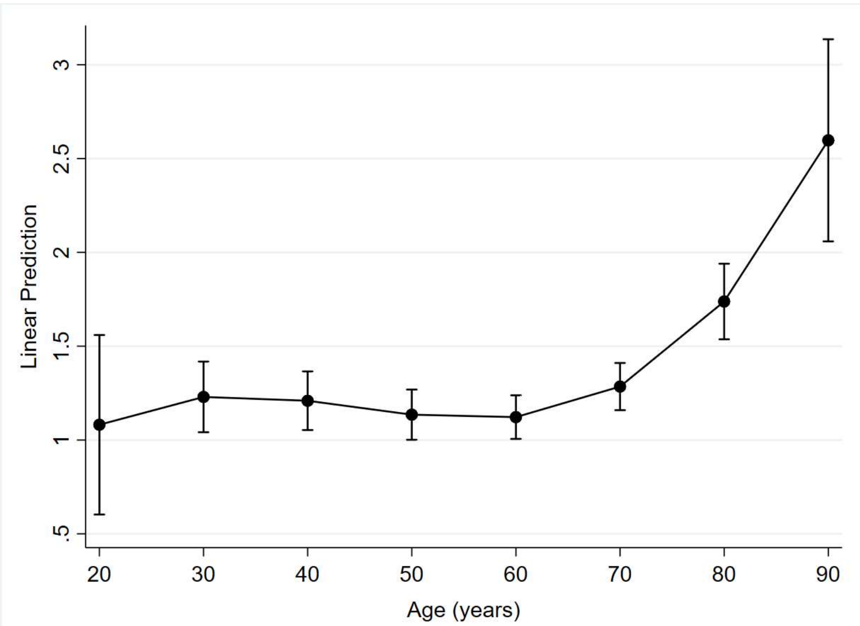
impairments among the elderly). To explore the extent to which age-related body shrinkage may contaminate our results on reporting error in BMI we repeated analysis of Table 2 restricting our sample to adults below the age of 70. Compared to our base case results (Table 2), and apart from the observed differences in the age profile of reporting error in BMI (as expected given the restricted sample with respect to age), our conclusions remain unchanged when restricted the sample to adults below 70. Table A.6 shows no systematic SES gradients in reporting error in BMI, and the coefficients for measured BMI of the respondent and BMI levels of other household members (as well as the relevant graphical predictions, Figure A.5, Appendix) are practically identical to our base-case results (Table 2 and Figure 3). Overall, this sensitivity analysis shows that restricting the sample to adults below 70 affects the age profile of reporting behaviour relevant to BMI (as expected), but not all other conclusions of the study.

Table 2. Regression analysis of absolute BMI reporting error.

	Specification 1	Specification 2	Specification 3
Age	1.040	0.860	1.067
	(0.651)	(0.703)	(0.686)
Age squared	-0.247**	-0.216	-0.257**
	(0.126)	(0.133)	(0.130)
Age cubed	0.018**	0.017**	0.019**
	(0.008)	(0.008)	(0.008)
Male	-0.019	0.007	0.006
	(0.090)	(0.086)	(0.085)
Degree/post-secondary	-0.277*	-0.208	-0.232
	(0.156)	(0.156)	(0.156)
A-level/equivalent	-0.023	0.018	0.011
	(0.185)	(0.182)	(0.180)
GCSE/equivalent	-0.109	-0.161	-0.170
	(0.165)	(0.160)	(0.157)
Income	-0.003	-0.024	-0.021
	(0.076)	(0.074)	(0.074)
BMI measured		0.049***	0.066***
		(0.010)	(0.011)
Low_hh_BMI			1.586**
			(0.621)
BMI measured*Low_hh_BMI			-0.051**
			(0.022)
R-squared	0.048	0.117	0.137
Sample size		873	

*p<0.10;**p<0.05;***p<0.001. Sample weights are accounted for.

Figure 2. Prediction (based on specification 3, Table 2) of the absolute BMI error by age.



Note: Adjusted predictions at representative values (APRs) are presented here — i.e., the predicted absolute reporting error in BMI across selected age values, with all the other variables kept fixed.

Figure 3. Prediction (based on specification 3, Table 2) of the absolute error in self-reported BMI by measured BMI values and household BMI levels.

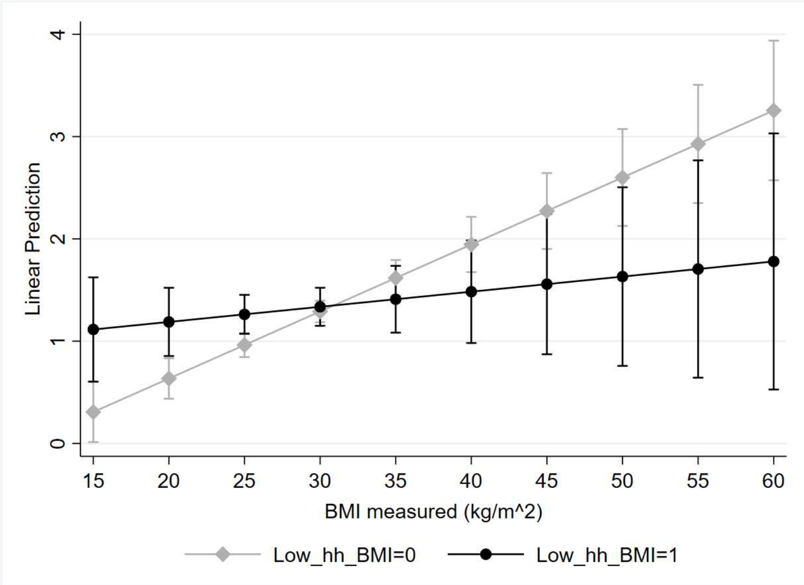
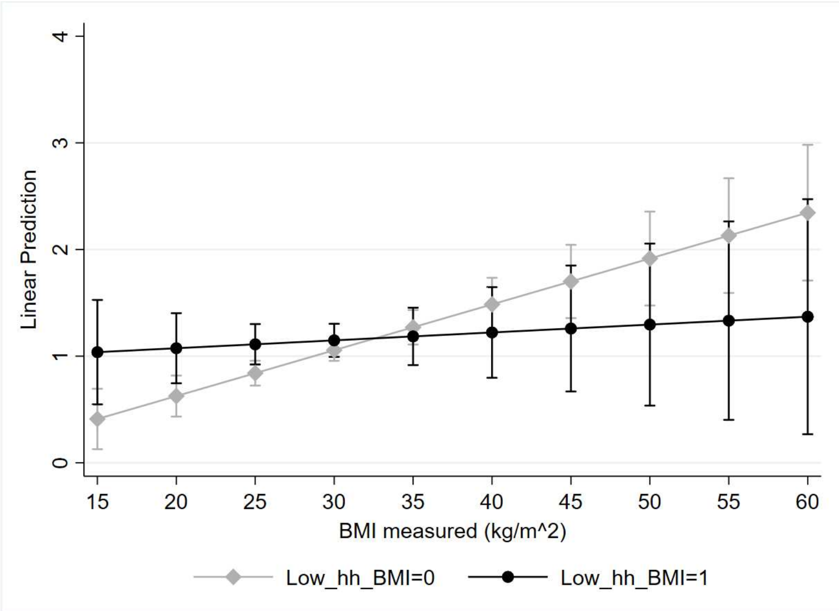


Table 3. Regression analysis of the absolute remaining BMI measurement error following adjustments from corrective equations.

	Coeff. (std. error)
Age	0.285
	(0.615)
Age squared	-0.061
	(0.119)
Age cubed	0.004
	(0.007)
Male	0.026
	(0.078)
Degree/post-secondary	-0.098
	(0.140)
A-level/equivalent	0.105
	(0.156)
GCSE/equivalent	-0.054
	(0.138)
Income	-0.106
	(0.070)
BMI measured	0.043***
	(0.010)
Low_hh_BMI	1.157**
	(0.584)
BMI measured*Low_hh_BMI	-0.036*
	(0.020)
R-squared	0.061
Sample size	873

*p<0.10;**p<0.05;***p<0.001. Sample weights are accounted for.

Figure 4. Prediction (based on Table 3) of the remaining BMI error by measured BMI values and household BMI levels.



5 Conclusion

We designed an experiment to measure reporting error in BMI. In accordance with existing results our analysis shows the presence of reporting error in self-reported body weight, height and consequently in BMI (e.g., Cawley et al., 2015, Gorber et al., 2007; Keith et al., 2011). We find a systematic age gradient in the reporting error in BMI, while there is limited evidence of systematic associations with gender and SES. This is reassuring evidence for the use of self-reported BMI in studies that use it as an outcome, for example, to analyse socioeconomic gradients in obesity.

Reporting error in BMI is associated with individual's measured BMI. This is driven by the non-classical reporting error in body weight and height — taller people report their height more accurately, while a nonlinear relationship is observed for body weight with a sharper increase in reporting errors for those of higher body weight. Focusing on BMI, the measure of adiposity that is most commonly used in the literature, the role of an individual's measured BMI on reporting error varies as a result of within-household peer-effects: for individual's with measured BMI values above the obesity threshold, measurement error is higher for those living in households with other members having high BMI levels. The latter is broadly consistent with the role of social norms on health reporting (Gil and Mora, 2011) and challenges the between-households (or and above between-individuals) reliability of the self-reported data. This complex structure of non-classical measurement error may be an issue in studies that use self-reported BMI as an explanatory variable.

Common methods to mitigate reporting error in BMI using corrective equations fail to fully eliminate systematic associations with individual and within-household BMI. Overall, as the error in anthropometrics is non-classical our results highlight the importance of collecting measured body weight and height data in large social science datasets where possible.

References

Cawley, J., Maclean, J.C., Hammer, M., Wintfeld, N.(2015). Reporting error in weight and its implications for bias in economic models. *Economics & Human Biology*, 19, 27-44.

Cawley, J.(2015). An economy of scales: a selective review of obesity's economic causes, consequences, and solutions. *Journal of Health Economics*, 43, 244-268.

Engstrom, J.L., Paterson, S.A., Doherty, A., Trabulsi, M., Speer, K.L. (2003). Accuracy of self-reported height and weight in women: an integrative review of the literature. *Journal of Midwifery & Women's Health*, 48(5), 338-345.

Gil, J., Mora, T.(2011). The determinants of misreporting weight and height: the role of social norms. *Economics & Human Biology*, 9, 78-91.

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

Keith, S.W., Fontaine, K.R., Pajewski, N.M., Mehta, T., Allison, D.B. (2011). Use of self-reported height and weight biases the body mass index–mortality association. *International Journal of Obesity*, 35(3), 401-408.

Lønnebotn, M., Svanes, C., Igland, J., Franklin, K. A.,… & Demoly, P. (2018). Body silhouettes as a tool to reflect obesity in the past. *PloS ONE*, e0195697.

O'Neill, D., Sweetman, O.(2013). The consequences of measurement error when estimating the impact of obesity on income. *IZA Journal of Labor Economics*, 2, 1-20.

Appendix

Table A.1 Summary statistics for the explanatory variables used in our analysis.

	Mean	Std. Dev.
Age (years)	52.74	18.20
Male†	0.50	0.50
Degree/post-secondary†	0.41	0.49
A-level/equivalent†	0.21	0.41
GCSE/equivalent†	0.22	0.42
No qualification†	0.15	0.36
Log household income (equivalized/deflated)	7.51	0.55
BMI measured (kg/m^2)	28.6	6.21
Weight measured (kg)	81.3	19.10
Height measured (cm)	168.4	9.89
Low_hh_BMI†	0.37	0.48
Sample size	873	

Notes: Sample weights are accounted for.
† Dummy variable

Table A.2 Classification of obesity using self-reported and measured BMI.

	Females	**Males**
Obesity prevalence (measured BMI)	36.43	32.16
Obesity prevalence (reported BMI)	32.42	26.10
P-value (difference)	0.003	0.000
Percentage classified as:		
True positive	30.51	24.37
False positive	1.92	1.73
True negative	61.65	66.11
False negative	5.93	7.79
Total	100.0	100.0
Sensitivity	83.7	75.8
Specificity	97.0	97.4
Sample size	469	404

Note: Sample weights are accounted for.

Table A.3 Reporting error in body weight and height.

	Mean	Std. Dev.
Body height (in cm)		
Raw error	1.196	2.767
Absolute error	2.243	2.013
Absolute error (% of measured)	1.344	1.228
Body weight (in kg)		
Raw error	-0.941	3.440
Absolute error	2.310	2.716
Absolute error (% of measured)	2.889	3.380
Sample size	873	

Note: Sample weights are accounted for.

Table A.4 Regression analysis of absolute reporting error in weight.

	Specification 1	Specification 2
Age	-0.888*	-0.998**
	(0.488)	(0.465)
Age squared	0.077*	0.088**
	(0.044)	(0.042)
Male	0.385*	0.172
	(0.230)	(0.287)
Degree/post-secondary	-0.510	-0.530
	(0.363)	(0.369)
A-level/equivalent	-0.073	-0.249
	(0.475)	(0.461)
GCSE/equivalent	-0.167	-0.243
	(0.395)	(0.390)
Income	0.107	0.109
	(0.189)	(0.189)
Body weight measured (in 10's of kg)		3.679***
		(0.778)
Body weight measured squared		-0.402***
		(0.074)
Body weight measured cubed		0.014***
		(0.002)
R-squared	0.024	0.093
Sample size	873	

*p<0.10;**p<0.05;***p<0.001. Sample weights are accounted for.

Table A.5 Regression analysis of absolute reporting error in height.

	Specification 1	Specification 2
Age	2.504**	2.460**
	(1.115)	(1.097)
Age squared	-0.537**	-0.533**
	(0.221)	(0.217)
Age cubed	0.038***	0.037***
	(0.014)	(0.013)
Male	0.238	0.631***
	(0.155)	(0.219)
Degree/post-secondary	-0.507*	-0.439
	(0.303)	(0.301)
A-level/equivalent	-0.492	-0.434
	(0.325)	(0.322)
GCSE/equivalent	-0.474	-0.452
	(0.323)	(0.320)
Income	-0.311**	-0.275*
	(0.157)	(0.160)
Body height measured (in cm)	-	-0.030**
		(0.012)
R-squared	0.094	0.103
Sample size	873	

*p<0.10;**p<0.05;***p<0.001. Sample weights are accounted for.

Table A.6 Regression analysis of absolute BMI reporting error: Adults below the age of 70.

	Specification 1	Specification 2	Specification 3
Age	-0.008	-0.009	-0.005
	(0.040)	(0.040)	(0.039)
Male	-0.032	0.013	0.016
	(0.106)	(0.101)	(0.099)
Degree/post-secondary	-0.198	-0.161	-0.186
	(0.225)	(0.224)	(0.225)
A-level/equivalent	0.018	0.035	0.029
	(0.250)	(0.248)	(0.246)
GCSE/equivalent	-0.003	-0.098	-0.114
	(0.233)	(0.229)	(0.227)
Income	-0.056	-0.068	-0.054
	(0.083)	(0.080)	(0.081)
BMI measured		0.045***	0.062***
		(0.012)	(0.012)
Low_hh_BMI			1.594**
			(0.689)
BMI measured*Low_hh_BMI			-0.051**
			(0.024)
R-squared	0.011	0.077	0.100
Sample size		664	

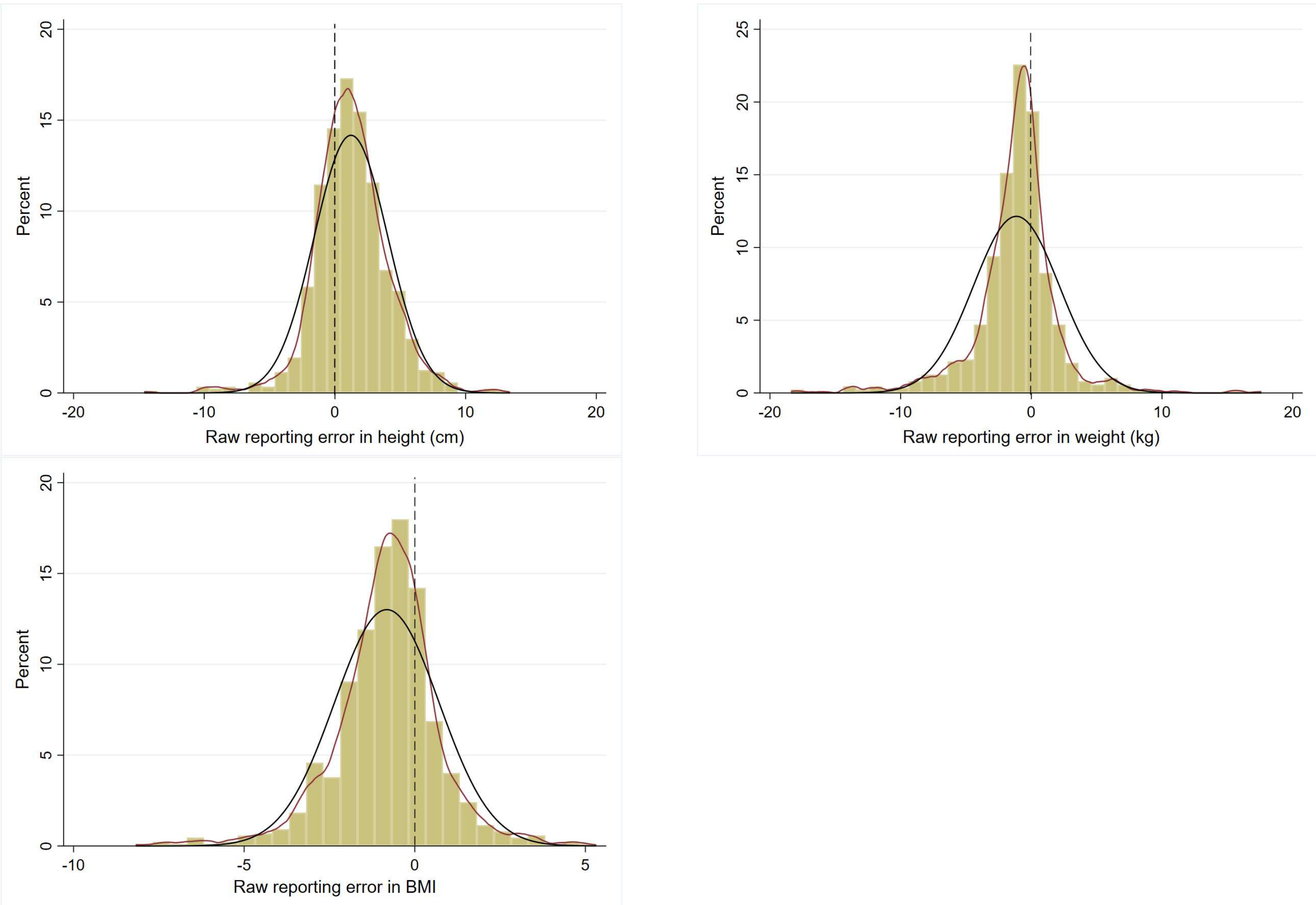
*p<0.10;**p<0.05;***p<0.001. Sample weights are accounted for.

Table A.7 Prediction equation for measured body weight and height.

	Body weight (measured)	Body height (measured)
Body weight reported	1.008***	–
	(0.012)	
Body height reported	–	0.889***
		(0.017)
Age†	-2.590	0.983***
	(2.691)	(0.333)
Age squared†	0.371	-0.143***
	(0.509)	(0.031)
Age cubed†	-0.015	–
	(0.030)	
Male	-18.293**	0.898***
	(7.176)	(0.323)
Male*Age	10.102**	–
	(4.449)	
Male* Age squared	-1.772**	–
	(0.872)	
Male*Age cubed	0.098*	–
	(0.054)	
Constant	5.814	16.453***
	(4.463)	(3.018)
R-squared	0.969	0.937
Sample size	873	

Notes: Robust standard errors in parentheses. Sample weights are accounted for.
†Age is divided by 10.
*p<0.10; **p<0.05; ***p<0.001.

Figure A.1 Raw reporting error in height, weight, and BMI.



Notes: The red lines show the distribution of the raw reporting error in height, weight and BMI. Superimposed in each histogram is the corresponding normal distribution (black lines). In each graph, the vertical axis is the percentage of the sample and the horizontal axis represents units of raw measurement error.

Figure A.2 Prediction of the absolute error in self-reported BMI by measured BMI values and household BMI levels: BMI information (low_hh_BMI) for more than 2 household members

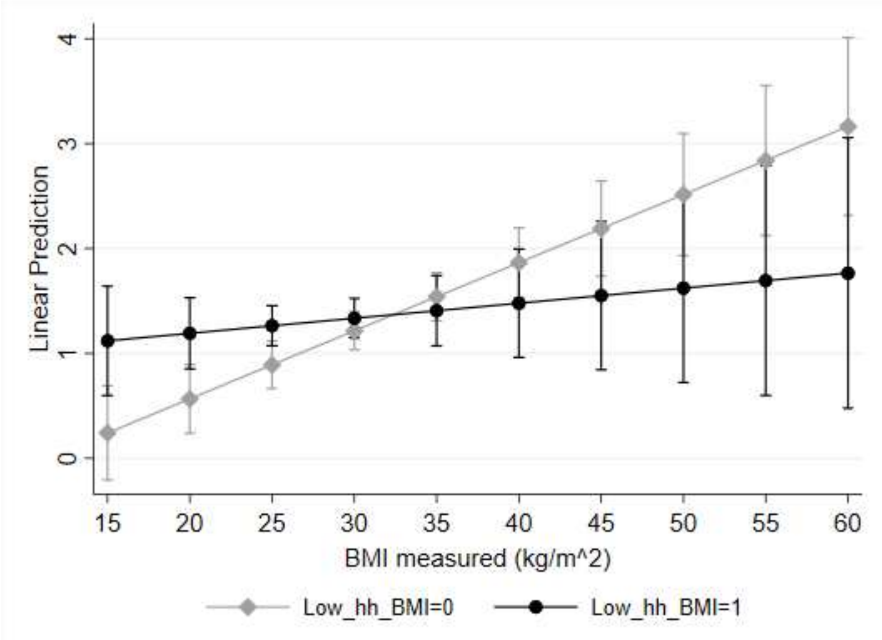
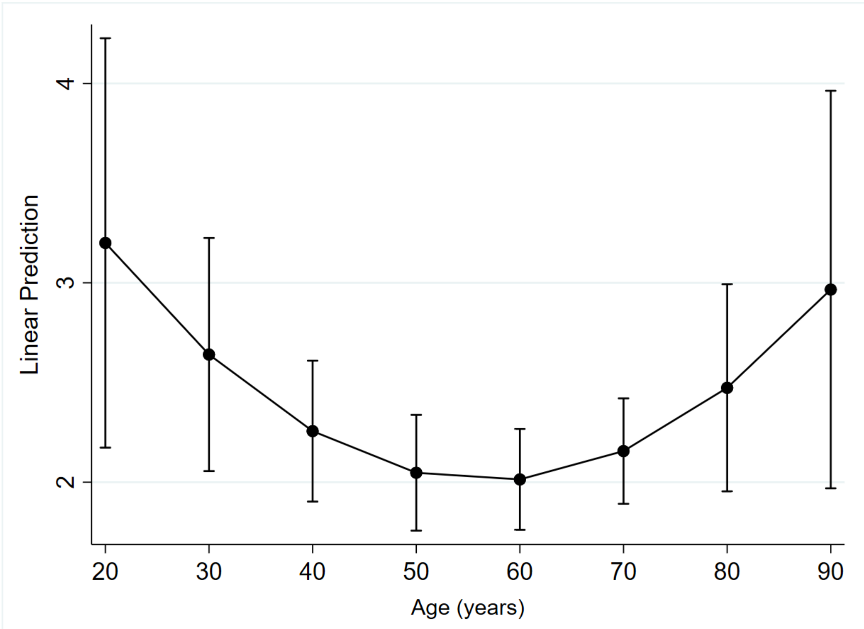
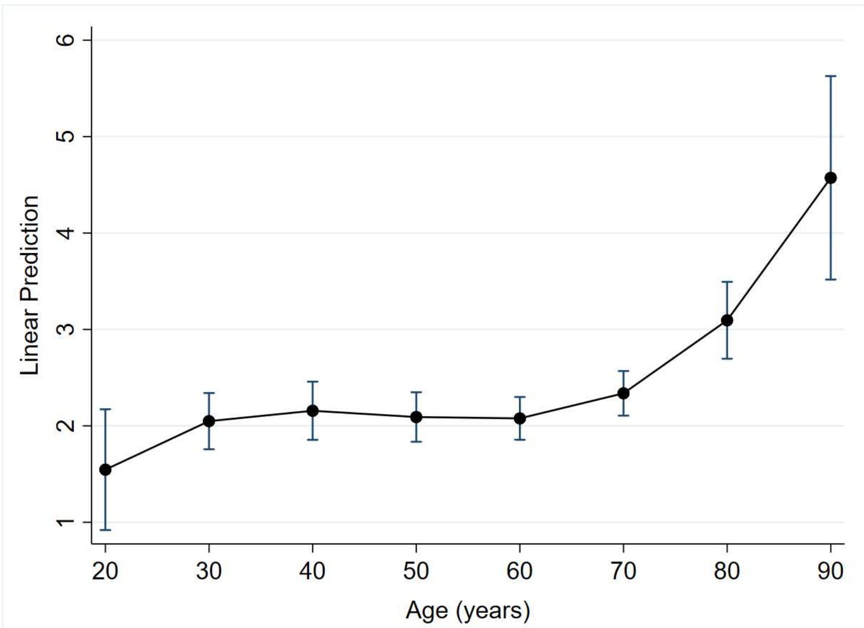


Figure A.3. Prediction (based on specification 2, Tables A.4-A.5) of the absolute error in weight and height by age.

Panel A: Weight



Panel B: Height



Note: Adjusted predictions at representative values (APRs) are presented here — i.e., the predicted absolute reporting error in weight and height across selected age values, with all the other variables kept at their initial values.

Figure A.4 Prediction (based on specification 2, Table A.4) of the absolute error in weight by measured weight.

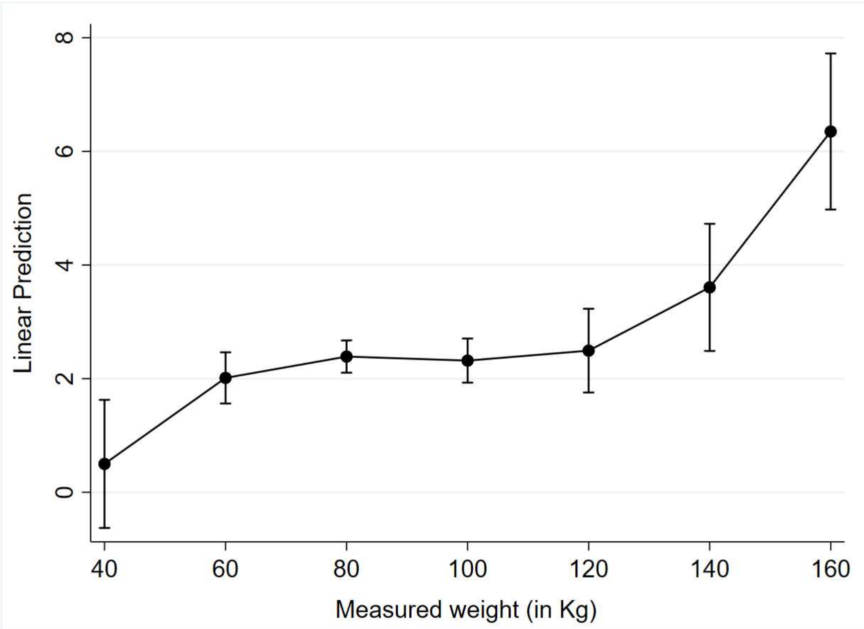


Figure A.5 Prediction (based on specification 3, Table A.6) of the absolute error in self-reported BMI by measured BMI values and household BMI levels:
Adults below the age of 70