Is body weight better distributed among men than among women? A robust normative analysis for France, the UK and the US

Fatiha Bennia, Nicolas Gravel, Brice Magdalou, Patrick Moyes

Is body weight better distributed among men than among women? A robust normative analysis for France, the UK and the US*

Fatiha Bennia

Aix-Marseille Univ., Laboratoire de Santé Publique, F-13385 Marseille, France
fatiha.bennia@univ-amu.fr

Nicolas Gravel

Centre de Sciences Humaines & Aix-Marseille Univ., CNRS, AMSE, 11 0011 Delhi, India
nicolas.gravel@csh-delhi.com

Brice Magdalou

CEE-M, Univ. Montpellier, CNRS, INRAE, Institut Agro, F-34960 Montpellier, France
brice.magdalou@umontpellier.fr

Patrick Moyes

Univ. Bordeaux, GREThA, UMR 5113, F-33600 Pessac, France
patrick.moyes@u-bordeaux.fr

Abstract

We compare distributions of Body Mass Index (BMI) categories among genders in France, the US and the UK on the basis of efficiency and inequality considerations. The new normative criteria that we propose are well-suited to the ordinal nature of this variable. Our empirical results, which are supported by robust statistical inference, are twofolds. First, BMI categories are better distributed in France than in the UK, and in the UK than in the US for the two genders. Second, BMI categories happen to be more equally distributed among men than among women in all three countries.

Keywords: body mass index; equality; efficiency; gender; ordinal
JEL classification: D63, I14, I31

"To lose confidence in one's body is to loose confidence in oneself"

Simone de Beauvoir

*This paper has benefited from comments and suggestions made by the editor and three referees. We also acknowledge financial support from the French Agence Nationale de la Recherche (ANR) through three contracts: Measurement of Ordinal and Multidimensional Inequalities (ANR-16-CE41-0005), Preference for Redistribution (ANR-15-CE26-0004) and Challenging Inequalities: A Indo-European perspective (ANR-18-EQUI-0003).

I. Introduction

It is widely acknowledged that body weight is an important contributor to well-being. The adverse effect of overweight and obesity on health is largely documented. It is also well-known that the prevalence of obesity and overweightedness has been growing in OECD countries in the last 40 years (see e.g. Sassi (2014)). Less spectacular and discussed, but nonetheless present, is the small, but significant, fraction of the population found in developed countries that is considered pathologically underweight. A significant body of epidemiological evidence (see e.g. Cao *et al.* (2014) or Flegal *et al.* (2005)) suggests that being underweight can also be associated with a significant increase in the probability of death in the next 5 years as compared to having a "normal" weight.

It is also largely documented - and recalled to our attention by the quotation from Simone de Beauvoir above - that body weight may impact individual well-being in a way that is not reducible to its health consequences, however severe these may be. Abnormal body weight may indeed affect self esteem and happiness (Oswald and Powdthavee (2007) and Stutzer and Meier (2016)), lead to social stigma (Carr and Friedman (2005), Roberts *et al.* (2000), Roberts *et al.* (2002) and Mooney and El-Sayed (2016)) or to an unfavorable image of one self when comparing with others (see e.g. Blanchflower *et al.* (2010)). As it happens (see e.g. Furnham *et al.* (2002) or Feingold and Mazzella (1998)), the impact of BMI on individual happiness through self-esteem and stereotypes is likely to differ between men and women. Another important difference between men and women is that - as suggested by Khlat *et al.* (2009) or Wells *et al.* (2012) - obesity and overweightedness tend to be more influenced by socioeconomic origins for women than for men. Furthermore, it is widely recognized (see for example Kanter and Caballero (2012)) that women are in average more subject to obesity then men.

This article proposes a novel methodology for comparing *distributions* of individual body weights from a normative standpoint. It focuses specifically on the measurement of body weight provided by the Body Mass Index (BMI), defined to be the ratio of the individual weight (in kilograms) over the "surface body" (in squared meters). Despite its limitations as a predictor of health hazards and indicator of excess adiposity (see e.g. Keys *et al.* (1972) or Gray and Fujioka (1991)), the BMI remains one of the most common index for measuring body weight. As such, it has been used in thousands of scientific studies, and serves as the reference for defining the various weight *categories* that are considered medically meaningful. According to the World Health Organization (WHO), these are, for the adult population above the age of 20:

Underweight (BMI below 18.5)
Normal (BMI between 18.5 and 25)
Overweight (BMI between 25 and 30)
Grade 1 obesity (BMI between 30 and 35)
Grade 2 obesity (BMI between 35 and 40)
Grade 3 obesity (BMI above 40)

Suppose that we accept this WHO categorization. How could we then compare alternative distributions of these BMI categories? When can we say that one distribution of BMI categories is "better" than another? *Two* principles are usually invoked for evaluating distributions of a contributor to individual well-being, like body weight, from a normative standpoint.

The *first* is the "efficiency" principle according to which it is *ceteris paribus* desirable to move individuals from bad to good categories. Such efficiency considerations clearly underlie the largely discussed concern for the increase in the prevalence of obesity in the adult population observed in OECD countries in the last 30 years. There are many ways of appraising efficiency improvements in the distribution of an attribute. All of them produce rankings of distributions that are compatible with *first order dominance* (see e.g. Lehmann (1955), Quirk and Saposnik (1962) or Hadar and Russell (1974)). According to this criterion, any change that decreases the fraction of the population falling in a worse category than any given category is worth doing. The use of first order dominance clearly requires that the categories be unambiguously ordered from worst to best. As discussed in the epidemiological literature (see e.g. Cao *et al.* (2014) or Flegal *et al.* (2005)), there is no such clear ranking of the six aforementioned BMI categories in terms of their impact on individual well-being. While it is hardly disputable that individual well-being *decreases* with the value of the BMI in ranges where this value is above normal, it is not clear how the underweight status compares with, say, the overweight one. As we shall see, the empirical comparisons of BMI distributions performed in this paper are not overly sensitive to the assumptions made on the rankings of the BMI categories.

The *second* principle, that may be open to discussion in the case of an attribute such as body weight, is an "equity" one that captures a concern for "equalizing" BMI categories across individuals. Implementing such a principle requires of course an operational definition of what it means for a distribution of BMI categories to be "more equal" than another. The meaning of "more equal" is somewhat consensual when applied to distributions of a *cardinally measurable* attribute like income. The view here is, indeed, to consider a *transfer* of a *given quantity* of income from a richer individual to a poorer one as a clear *inequality-reducing operation*. Transfers of this kind are called *Pigou-Dalton transfers*. An equalization of an income distribution is therefore defined as the fact of performing such Pigou-Dalton transfers a finite number of times. The theory of income inequality measurement (see e.g. Kolm (1969), Atkinson (1970), Dasgupta *et al.* (1973), Sen (1973) and Fields and Fei (1978)) has established tight connections between this theoretical notion of equalization and implementable criteria such as Lorenz or second order stochastic dominance.

The difficulty in applying this notion of equalization to BMI categories lies, of course, in the *ordinal* nature of the information on individual well-being that these categories provide. The fact for an individual to belong to the "grade 3 obese" category is certainly indicative of a lower well-being than the fact of belonging to the "grade 2 obese" one. However, one would hesitate in quantifying further this assessment. In particular, one would hesitate in comparing the difference in well-being between "grade 3" and "grade 2" obese categories with the difference in well-being between the "overweight" and the "normal weight". This difficulty of defining equalization for an ordinally measurable attribute is often ignored by researchers. For instance Etile (2014) examines the impact of education on the Gini coefficient of the distribution of BMI in France over the period 1981-2003. Yet, it is fair to say that a growing number of researchers, especially in the fields of health and development economics where categorical or ordinal data are common, have recognized it. Examples of contributions that have proposed specific methods for that purpose are Allison and Foster (2004), Abul-Naga and Yalcin (2008), Apouey (2007), Kobus and Milós (2012), Zheng (2008), Zheng (2011), Cowell and Flachaire (2017) and Sonne-Schmidt *et al.* (2016).

Some forty years ago, Peter J. Hammond (1976) has suggested a "minimal equity principle" that is explicitly concerned with distributions of an ordinally measurable attribute. According to Hammond's principle, a reduction in the gap between two individuals unequally endowed with the ordinal attribute is a good thing, irrespective of whether or not what is given to the poor is equal to what is taken from the rich. A Pigou-Dalton transfer is just a particular case of a Hammond transfer that imposes on the later the additional requirement - meaningless for an ordinal variable - that the "amount" given by the rich should equal the "amount" received by the poor. A Hammond transfer seems to be a plausible definition of an ordinal inequality reduction.

In a recent paper, Gravel *et al.* (2020) have provided an operational test that enables one to identify when a distribution of BMI categories can be considered unequivocally better than another by either - or both - these equity and efficiency considerations. In this paper, we put this criterion to work by evaluating the distributions of BMI categories in three developed countries: France, the UK, and the US. The criterion used to make these comparisons is based on a curve that we call the H^+- curve, by reference to the Hammond principle of transfer to which it is closely related. This curve is defined by first assigning to the worst category the fraction of the population falling in that category and by then assigning to every category above the worst one the fraction of the population falling in this category plus *twice* the value assigned by the curve to the preceding category. Just like for the first order stochastic dominance, the definition of this curve requires that BMI categories be unambiguously ordered from the worst to the best. The criterion requires the dominating distribution to have a H^+-curve nowhere above and somewhere below that of the dominated one. Gravel *et al.* (2020) have shown that observing H^+-dominance between two distributions is *equivalent* to the possibility of going from the dominated to the dominating distribution by a finite sequence of Hammond transfers and/or increments of the attribute. Gravel *et al.* (2020) have also identified another test - which combines the H^+- curve and a dual H^-- curve - that captures equalization considerations *only.*

We find worth emphasizing that concerns for equalization in BMI categories may not seem as appealing as similar concerns applied to income. There is a widespread opinion in favour of reducing the income of "rich" people in exchange of an increase in the income of poorer ones. However, the support for deterorating the BMI status of people who are in the normal category in exchange of improving the BMI category of, say, obese people may seem less prevalent. We therefore complement our analysis with a questionnaire survey proposed to a representative sample of 1005 French adults that investigates the extent to which "ordinary people" favour equality considerations - in the form of Hammond transfers - when applied to BMI categories. The main result of our survey is that a significant faction of them happen to do so.

Our empirical use of the dominance criteria, that stands on a statistical inference methodology inspired by Davidson and Duclos (2000), reveals two noteworthy features, among many others. First, and non surprisingly, the distribution of BMI categories appears to be better in France than in the UK and the US by first order dominance for both males and females. This is at least true for the welfare rankings of the BMI categories that consider the "underweight" status to be better than the grade 3 obese one. Such a dominance does not hold between countries however when one assumes that being underweight is worse than being severely obese. This latter absence of dominance comes from the fact that the fraction of the population falling in the underweight category is higher in France than in the US and the UK. Interesting also are the cross gender comparisons of distributions of BMI in each of the three countries. As it turns out, BMI categories appear to be more equally distributed among men than among women in all three countries as per Hammond transfers. However, for France, this dominance is only observed for the rankings that consider underweight to be a worse status than overweight.

The plan of the rest of the paper is as follows. The next section introduces the normative criteria and their logical connections, and briefly discusses the results of our questionnaire survey on attitude to inequalities in BMI categories. The third section presents the data and the results of the comparisons of distributions of those categories in France, the UK and the US. The fourth section concludes.

II. Comparing distributions of BMI categories

We consider a setting in which every individual can fall into one out of k different categories. While $k = 6$ is going to be assumed here so as to stick to the well-established WHO categorizations of BMI mentioned above, we find useful to discuss the criteria in the more general case of an arbitrary number of categories. Important for the analysis is the assumption that these categories be unambiguously ordered from the worst (e.g. being obese of grade 3) to the best (e.g. having a normal weight). As discussed above, while this assumption is plausible when applied to the various overweight categories, it is not so clear how the underweight category compares with the overweight categories. In the empirical analysis, we consider several alternative assumptions concerning the rankings of these categories that are consistent with the assumption of a reverse U-shaped relationship between BMI and welfare that peaks somewhere in the normal BMI range (see e.g. Cao *et al.* (2014)).

Efficiency and inequality considerations

A distribution of BMI categories is an ordered list $d = (d_1, ..., d_k)$ of k numbers that specify, for each category j - ordered from the worst to the best - the fraction d_j of the population falling in category j. It is therefore understood that $d_j \in [0, 1]$ for every category j and

$d_1 + ... + d_k = 1$. The basic question that we address is: when can a distribution of BMI categories be said to be unquestionably better than another?

The first principle that can be invoked for answering this question is *efficiency.* Given the ordering of the categories from the worst to the best, there seems to be a clear formulation of this efficiency principle here: that of improving the category of some people, without reducing that of anyone else. We define such an *elementary increment* as follows.

Definition 1 *(Elementary increment) We say that distribution d is obtained from distribution d' by means of an elementary increment if and only if there is a category $j \in \{1, \ldots, k-1\}$ and a (positive) fraction ε such that $d_h = d'_h$ for all $h \neq j, j+1$, and $d_j = d'_j - \varepsilon$, and $d_{j+1} = d'_{j+1} + \varepsilon$.*

In words, a distribution d of BMI categories is obtained from a distribution d' by an elementary increment if d results from d' by a transfer of a fraction of the population from a category j to the immediately superior category $j+1$.

As is well-known (see e.g. Lehmann (1955), Quirk and Saposnik (1962)) this notion of elementary increment happens to be tightly related to the criterion of *first order (stochastic) dominance*. The formal definition of first order stochastic dominance in the current context makes use of the *Cumulative Distribution Function* (CDF) associated to d, which is denoted, for every category j, by $F(j; d)$ and which is defined by $F(j; d) = \sum_{h=1}^{j} d_h$. Hence $F(j; d)$ is the fraction of the population who are in a weakly worse BMI category than j. With this notation, one can define first order dominance as follows.

Definition 2 *(1st order dominance) We say that distribution d first order dominates distribution d' if and only if $F(j; d) \leq F(j; d')$ holds for all $j = 1, ...k$.*

In words, a distribution d first order dominates a distribution d' if, for every reference category, the fraction of the population falling in a weakly worse category than the reference is no greater in d than in d'. The following well-known theorem (proved in Lehmann (1955) or Quirk and Saposnik (1962)) connects first order stochastic dominance to elementary increments.

Theorem 1 *For any two distributions d and d' of BMI categories, the following two statements are equivalent:*

(a) *d is obtained from d' by means of a finite sequence of elementary increments,*

(b) *d first order dominates d'.*

As we shall see below, this criterion alone enables one to compare a sizeable number of distributions of BMI.

Yet, efficiency may not be the only principle for comparing distributions of contributors to well-being. Concerns for *reducing inequalities* in those contributors may also be considered important. To motivate these concerns in our context, consider our sixth BMI categories and assume - in line with Cao *et al.* (2014) - that they are ordered as in the table below (from the top (best) to the bottom (worst)). Consider also the two theoretical distributions d and d' as follows:

BMI Category	d	d'
Normal	0.4	0.5
Overweight	0.3	0.2
Underweight	0.05	0.05
Grade 1 obesity	0.2	0.1
Grade 2 obesity	0.05	0.05
Grade 3 obesity	0	0.1

As can be seen, none of the two distributions first order dominates the other. Indeed, nobody in d falls in the (worst) grade 3 obese category while 10% of the population in d' have this misfortune. On the other hand there are 60% of the population in d with non-normal BMI category while only 50% of the population in d' are in this situation. However, in some sense, the distribution of BMI categories in d can be considered "less spread out" than that of d'. Indeed, one can describe the move from d' to d as the result of a few "equalizing transfers" of BMI categories between individuals. Specifically this move can be viewed as resulting from the gift of "one unit" of BMI category by 10% of the population belonging to the normal category. Who are the beneficiaries of these "gifts"? They are all extremely obese individuals who were representing 10% of the population in

d' and who have disappeared in d. Observe that all these extremely obese people have received a gift of "2 units" of BMI category since the fraction of grade 1 obese people in the population has grown from 10% in d' to 20% in d. These transfers are illustrated on Figure 1.

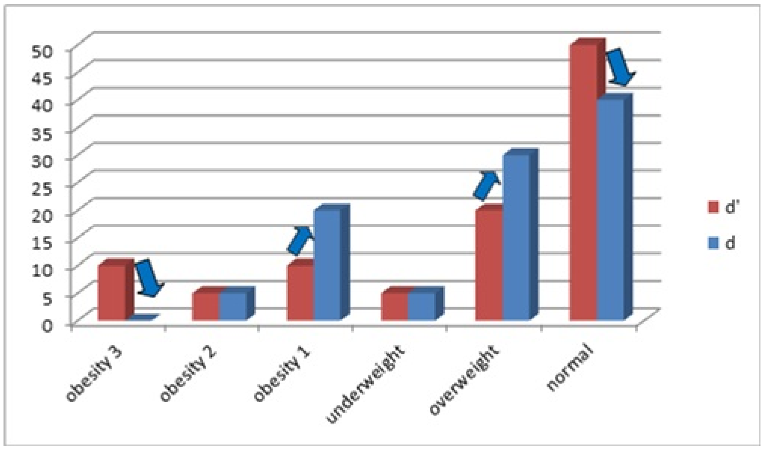


Figure 1: A Hammond transfer of 10% of the population.

Transfers of this kind have been proposed in the mid seventies by Peter J. Hammond (1976) as the analogues, for an ordinally measurable attribute, to the usual Pigou-Dalton transfers. A Hammond transfer is indeed like a Pigou-Dalton transfer, but without the (meaningless for an ordinal attribute) requirement that the quantity transferred from the rich to the poor be "given". Indeed, as in the example above, a Hammond transfer allows a healthy individual to fall "one step" to the immediately inferior overweight category in exchange for a Grade 3 obese person to rise "two steps" to the Grade 1 obese one. Conversely, a Hammond transfer may be "leaky", by taking "a lot" of attribute away from a "rich" person in exchange of giving "just a little" to a poorer one. The precise definition of a Hammond transfer is as follows.

Definition 3 (Hammond's transfer) *We say that distribution d is obtained from distribution d' by means of a Hammond' equalizing transfer if and only if there are four categories $1 \leq g < h \leq i < j \leq k$ and a (positive) fraction ε such that $d_l = d'_l$ for all $l \neq g, h, i, j$, and $d_g = d'_g - \varepsilon$, and $d_h = d'_h + \varepsilon$, and $d_i = d'_i + \varepsilon$, and $d_j = d'_j - \varepsilon$.*

Suppose that we accept Hammond transfer as the appropriate notion of equalization in an ordinal context. If one is interested in both equality and efficiency, it is tempting to say that a distribution of BMI categories is better than another if it has been obtained from the later by a finite sequence of Hammond transfers and/or elementary increments. It would be nice to dispose of an easy diagnostic tool for identifying, given any two distributions, wether or not one has been obtained from the other by a finite sequence of equalizing Hammond transfers and/or elementary increments. As shown in Gravel *et al.* (2020), it turns out that a particular curve - that we call here the H^+- curve[1] - provides such a diagnostic. For any distribution d, its H^+- curve is defined by $H^+(1; d) = F(1; d) = d_1$ and, for all $j = 2, 3, \ldots, k$, recursively by $H^+(j; d) = 2\, H^+(j-1; d) + d_j$.

This curve is easy to draw and implement. It is initiated by plotting, for the worst category, the fraction of the population in this category. It is then iterated by plotting, for every other category above the worst one, the fraction of the population lying in this category plus twice the value assigned by the curve to the immediately preceding category.

Definition 4 *(H^+- dominance) We say that distribution d H^+ dominates distribution d', which we write $d \succeq_{H^+} d'$, if and only if $H^+(j; d) \leq H^+(j; d')$ for all $j = 1, 2, \ldots, k$.*

It is easy to see that first order dominance entails H^+- dominance. The converse however does not hold. The following theorem, proved in Gravel *et al.* (2020), shows that H^+- dominance provides indeed an exact diagnostic tool for identifying, between any two distributions of BMI categories, whether or not one is *more equal and/or more efficient* than the other.

Theorem 2 *For any two distributions d and d' of BMI categories, the following two statements are equivalent:*

(a) *d is obtained from d' by means of a finite sequence of Hammond's transfers and/or elementary increments,*

(b) *$d \succsim_{H^+} d'$.*

[1]The curve is referred to as the H-curve in Gravel *et al.* (2020).

In essence, H^+- dominance is similar to the generalized Lorenz dominance *à la* Shorrocks (1983) used in standard income inequality measurement. Just like this latter criterion, H^+-dominance combines efficiency (increments) and equality (Hammond) considerations, but by giving priority the former notion. Indeed, the best conceivable distribution of BMI categories for either first order or H^+- dominance is obtained when everyone belongs to the normal BMI category. Similarly, the ideal distribution from the view point of generalized Lorenz dominance is that where everyone is getting the largest conceivable income. Of course, one may consider that the situation where everyone has a normal body weight is less out-of-reach than that where everybody earns the income of the world's best paid CEOs. This, again, suggests that concerns for reducing inequality - lowering down the category of "top" individuals in exchange of increasing that of "bottom" ones - may be less appealing for BMI than for income. But leaving this question aside for a moment, it could be also useful to have a diagnostic tool for identifying whether or not a distribution is *more equal only* than another (without being more efficient).

Results in Gravel *et al.* (2020) show that such a diagnostic tool uses, along with the H^+- curve, the somewhat dual H^-- curve.[2] The formal definition of this dual curve makes use of the *Survival* function associated to a distribution d denoted, for every category $j = 1, ..., k$, by $S(j; d)$ and defined by $S(j; d) = \sum_{h=j}^{k} d_j$. In plain English, $S(j; d)$ is the fraction of the population in d who are in a weakly better category than j. The H^- curve is defined under the same recursive principle than the H^+ one, but starting with the best category k (rather than the worst category 1) and iterating with the survival function (rather than the cumulative one). The H^- curve therefore starts at category k by $H^-(k; d) = S(k; d) = d_k$. Then, it is defined recursively, for categories $j = 1, 2, \ldots, k-1$, by $H^-(j; d) = 2\, H^-(j+1; d) + d_j$. This curve gives rise to the following definition of H^--dominance.

Definition 5 *(H^-- dominance) We say that distribution d H^- dominates distribution d', which we write $d \succeq_{H^-} d'$, if and only if $H^-(j; d) \leq H^-(j; d')$ for all $j = 1, 2, \ldots, k$.*

[2]This curve is called the $\overline{H}$-curve in Gravel *et al.* (2020).

It is shown in Gravel *et al.* (2020) that the fact, for a distribution d, to H^-- dominate a distribution d' is equivalent to the possibility of moving from d' to d by a finite sequence of *elementary decrements* (the converse of increments) and/or Hammond transfers. Hence, if a distribution d has been obtained from distribution d' by a finite sequence of Hammond transfers *only*, it follows from theorem 2 that distribution d both H^-- dominates and H^+- dominates distribution d'. To the contrary, observing both H^+- and H^-- dominance between two distributions, implies the possibility of going from the dominated to the dominating distribution by a finite sequence of Hammond transfers only.

Are these criteria consistent with people attitude to BMI inequality?

Is equalization of BMI categories - in the sense of Hammond transfers - really desirable from a normative standpoint? To address this question, we have borrowed from the methodology of Amiel and Cowell (1992) and conducted a questionnaire survey through internet in July 2020 on a sample of 1005 individuals that is representative of the adult French population (18 years and more).[3] A description of the sample and the main socioeconomic information gathered from the survey is provided in Appendix A, while the written instructions given to the surveyed individuals are reproduced in Appendix B. In order to avoid the aforementioned difficulty of ordering the “underweight” category *vis-à-vis* the overweight ones, we have excluded the former from the list presented to the subjects. [4]

Just like Amiel and Cowell (1992) in the case of income inequality, each surveyed individual was provided with a set of questions involving the comparisons of *two* stylized distributions of BMI categories made of three pictured individuals. In each question, one of the two distribution is obtained from the other by one or many Hammond transfers.

[3]The survey was conducted in the experimental laboratory of Montpellier (LEEM). The sampling and the data collection have been realised by the Viavoice Institute (`http://www.institut-viavoice.com`). Every surveyed individual has received a lump sum monetary compensation for his/her effort.

[4]We acknowledge that this simplification makes the survey imperfect in capturing the preferences for equalizing BMI categories considered in this paper.

Thirty two such questions were presented to the subjects, the list of which is provided in Appendix C. These 32 pairs of distributions were themselves obtained from eight initial distributions, to which one or several Hammond transfers were applied. The set of questions obtained from the same initial distribution were proposed to the surveyed individuals in a random order and the positioning - at the left or right of the screen - of the more equal and the less equal distribution was also made randomly (see Appendix B). In order to make the distributions more palatable to the surveyed persons, the three pictured individuals involved in the Hammond transfers were depicted either as females or as males. The same gender were used in the whole questionnaire given to any given surveyed person. However, the use of a male or a female figure in any questionnaire was made at random. In Table 1, we provide the percentage of the subjects preferring the more equal to the less equal distribution (rate of acceptance) for both the whole sample, and for each of the two gender subsamples.

	Global	Women pictured	Men pictured
Number of subjects	1005	508	497
Rate of acceptance	46.74%	47.38%	46.08%

Table 1: Global results of the survey

The fraction of interviewed persons who happen to prefer the more equal to the least equal distribution of BMI is a little less than 50%, which is quite significant. Of course this preference is expressed by a little less than half the subjects, which may be interpreted - from a majoritarian perspective - as a rejection of egalitarianism. Yet, one should bear in mind that in the context of income distributions - in which egalitarianism is a widely held value - Amiel and Cowell (1992) find that only 35% of their surveyed individuals express a preference for a distribution obtained from another by a Pigou-Dalton transfer (in their numerical example). Hence, egalitarian preferences, as captured by these surveys, appear more prevalent for BMI than for income distribution.

We find also worth noticing that the “rate of acceptance” of Hammond transfers among the surveyed subjects depends significantly upon *both* the initial distribution in which the transfers are performed *and* the type of Hammond transfers proposed. For the first type of dependency, Table 2 shows that when Hammond transfers are performed on initially very unequal distributions (B7 and B8 questions, see Appendix C), more than 60% of the interviewed persons express a preference in favour of making those transfers.

	Global	B1	B2	B3	B4	B5	B6	B7	B8
Rate of acceptance	46.74%	38.07%	36.87%	39.80%	32.44%	37.41%	27.66%	60.62%	60.42%

Table 2: Results of the survey by sets of questions

Concerning the second type of dependency, the rate of acceptance of Hammond transfers happens to be increasing with respect to the impact of the proposed transfer on the mean of the distributions. As emphasized in this paper, the "mean" has little meaning for an ordinal variable such as BMI. However, it is possible that subjects express a stronger preference for Hammond transfers that increase what they perceive as the mean - for example a person with normal BMI category becomes overweight (drop of one category) in exchange for a morbidly obese person to become mildly obese (gain of two categories) - than for Hammond transfers that are "leaky". As shown in Table 3, that distinguishes between mean reducing, mean preserving, and mean increasing transfers - we get a significant fraction of the subjects who express preference for performing efficient Hammond transfers (this rate is even increased to more than 70% if the transfers are made from initially unequal distributions, as in B7 and B8 questions). Hence, these survey findings suggest that equalization in BMI categories is a value judgement that do receive some support.

	Global	mean-reducing	mean-preserving	mean-improving
Rate of acceptance	46.74%	38.43%	41.96%	62.49%

Table 3: Results of the survey by type of transfers

In the next section, we therefore use the criteria presented in this section to compare the distributions of BMI categories between genders in France, UK and the US on the basis of both equality and efficiency.

III. Data and Empirical Results

The data for France are taken from the *Enquête sur la Santé et la Protection Sociale (ESPS)*. The ESPS is a panel survey created in 1998. It follows a sample of about 8000 households (roughly 22 000 individuals) that is representative of 95% of the inland European France (excluding outer sea regions such New Caledonia, Reunion Island, etc.). In 2006, the sample has been slightly enlarged to provide a better representation of the segment of the French population covered by the "*Couverture Maladie Universelle*" programme (which provide basic free medical care to poor French households who are not medically covered by the national social security regime). Each individual in the sample is interrogated every two years in a comprehensive manner about a large spectrum of health characteristics of every member of his or her household. It is important to notice that the health information of this data set is *reported* by the person surveyed rather than being measured or appraised by an external health expert. We restricted our sample to the 11 255 individuals aged above 20 who report information on their height and weight for the year 2008 (the most recent one for which we had data on all three countries).

Data on United States are taken from the 2007-2008 issue of the *National Health And Nutrition Examination Survey* (NHANES). This survey, initiated in 1960, became regular in 1999. The survey collects information by both interviews (conducted at the place of residence of the individual) and physical examinations (including weighting and height measurement) performed at mobile centers by professional health technicians. Every year, 5 000 persons (different from year to year) are interviewed in 15 states. In order to be representative of the whole US population, there is an over sampling of individuals above sixty, and of members of the Afro-American and the Latin American community. From the initial 10 000 individuals surveyed in the years 2007 and 2008, we ended up with a sample of 5 607 individuals who were above 20 and for which information on height, weight and gender was available. It is also important to notice that each person interviewed through this survey receives a monetary compensation as well as a complete report of his/her health status.

Data on United Kingdom come from the *National Diet and Nutrition Survey* (NDNS) and the *Rolling Programme* (RP) for the 2008-2012 period. The NDNS survey was initiated in 1992 and became, in 2008, the RP survey that covered both adults and children. This survey is ran by three governmental bodies: the Ministry of Agriculture, the Public Health of England and the Food Standards Agency. The survey covers all four members of the United Kingdom (England, Wales, Scotland and Northern Ireland). As in the US, information on nutrition and health is collected from two complementary modes: interview, and physical examination (including height and weight measurement) by health technicians. Individuals who participate to the survey are selected by stratified sampling, based on a random selection from postal codes taken from all parts of the UK, with the codes grouped by broad geographical sectors. The sample size is somewhat small however since, during the 4 years, 2 083 adults and 2 073 children and teenagers have been sampled. our aim of working with adults with at least 20 years old on which information on gender, weight and height is available led us to a sample of 1 912 individuals for the United Kingdom.

A summary description of the samples is given in Table 4, while Table 5 reports the distributions of the BMI categories for the female and male (respectively) samples in France, the UK and the US for the considered periods.

A few noticeable features emerge from these tables. First, the fraction of the sample that suffers from grade 3 obesity is somewhat small. Indeed less than 1% of the adult male population in France are concerned. The corresponding fraction for the female population is slightly higher (1.14 %). Non-surprisingly, the figures are somewhat larger for the UK and the US where grade 3 obesity concerns, respectively, 4.16% and 7.25 % of the female

	France (2008)	UK (2008-2012)	US (2007-2008)
Nb. of individuals	11 255	1 912	5 607
Nb. of men	5 369	906	2 747
Nb. of women	5 886	1 006	2 860

Table 4: Sample description

	Women			Men		
BMI Category	France	UK	US	France	UK	US
Grade 3 obesity	0.0114	0.0416	0.0786	0.0047	0.0121	0.0440
Grade 2 obesity	0.0311	0.0951	0.1125	0.0211	0.0623	0.0736
Grade 1 obesity	0.0910	0.1681	0.2045	0.1030	0.1990	0.2163
Overweight	0.2490	0.3019	0.3016	0.3841	0.4536	0.3896
Underweight	0.0467	0.0166	0.0213	0.0132	0.0084	0.0116
Normal	0.5707	0.3767	0.2814	0.4739	0.2642	0.2647

Table 5: distributions of BMI categories among women and men

population and 1.21% and 4.4% of the male population. It is noteworthy that, in every country, a larger fraction of women than men suffer from extreme obesity. Another BMI category that concerns only a rather small fraction of the population is the underweight one. About 4.6% of the French female were having a BMI that enters in this category in 2008. The corresponding fractions for the French male population (1.32%) is significantly smaller. Contrary to what was the case for grade 3 obesity, the UK and the US are less affected than France by the prevalence of underweightedness.

The other BMI categories concern a much larger fraction of the relevant population. Grade 2 obesity for instance concerns 11.25% of the US women. Yet similar qualitative patterns can be found with respect to the comparative prevalence of grade 2 obesity between the different population. In all three countries, women are relatively more affected than men by grade 2 obesity. Moreover, for both men and women, the fraction of the population that is obese of grade 2 is larger in the US than in the UK and the UK than in France.

As for grade 1 obesity and overweightedness, the comparative exposures of men and women are markedly different than what they are for the three previous ones. Indeed, in all three countries, one finds a *lower* fraction of women than men who are *either* overweight *or* grade 1 obese. Just like for obesity of grades 2 and 3, and for both men and women, the fraction of the sample that is either overweight or obese of grade 1 is smaller in France than in the UK and is smaller in the UK than in the US. The only exception to this is the overweight category and the male population in the US and UK. As it happens indeed, the fraction of men who are overweight is much larger in the UK (45.36%) than in the US (38.96%). While females are more affected than males by severe obesity and underweight, they are more likely than men to have a normal weight. One observes also that the fraction of the male population with a normal weight is slightly larger in the US than in the UK.

Let us now try to get somewhat more palatable normative conclusions from these observations. Important for this endeavour is a ranking of the BMI categories in terms of their impact on individual well-being. As discussed earlier, the relative ranking of the underweight category *vis-à-vis* any of the overweight and obese ones is not clear. Because of this ambivalence, we find safe to consider the five rankings of BMI categories - from the worst (bottom) to the best (top) - described in Table 6.

Ranking 1	Ranking 2	Ranking 3	Ranking 4	Ranking 5
Normal	Normal	Normal	Normal	Normal
Underweight	Overweight	Overweight	Overweight	Overweight
Overweight	Underweight	Grade 1 obesity	Grade 1 obesity	Grade 1 obesity
Grade 1 obesity	Grade 1 obesity	Underweight	Grade 2 obesity	Grade 2 obesity
Grade 2 obesity	Grade 2 obesity	Grade 2 obesity	Underweight	Grade 3 obesity
Grade 3 obesity	Grade 3 obesity	Grade 3 obesity	Grade 3 obesity	Underweight

Table 6: possible welfare rankings of BMI categories

To the very best of our understanding of the impact of body weight on individual welfare, rankings 1, 2 and 3 are the most plausible of the five, at least for developed countries such as France, UK and the US. But we perform the analysis with all of them so as to illustrate the dependency of the conclusion on the assumptions made about the rankings.

Cross-country comparisons

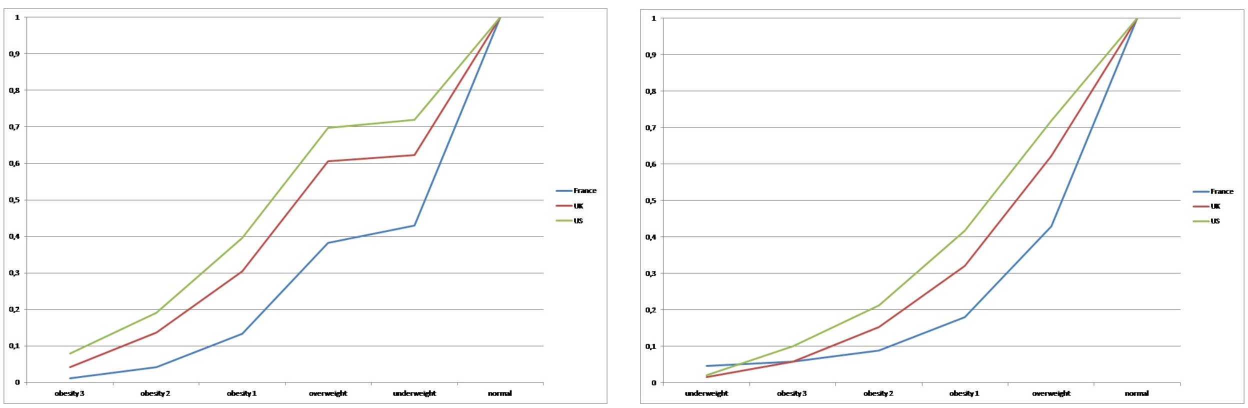
We illustrate the comparisons that we perform by showing the graph of some sample curves. However the fact that the curves are constructed from (small) samples should make one careful in deriving from their visual examination definite conclusions concerning the distributions of BMI categories in the populations of interest. This is especially the case given the fact that some of the compared curves are very close to each other, making the interpretation of their crossing/non-crossing rather difficult. We have therefore found useful to secure the statistical significance of our conclusions.[5]

For this sake, we have adapted to H^+- dominance and to the intersection of H^+- and H^-- dominance the general methodology proposed by Davidson and Duclos (2000) (Theorem 1) for performing statistical inference on standard stochastic dominance criteria. This methodology is based on the law of large numbers and the central limit theorem, and is akin to testing differences in sample means tests, where the "means" in our case are the values of the expressions $F(j;d)$, $H^+(j;d)$ and $H^-(j;d)$ (for $j = 1, ..., 6$) involved in the various criteria. However, instead of testing one difference between the means of two distributions, the dominance criteria require that we test jointly a collection of differences in those means. This gives rise to at least two possible ways of specifying the null hypothesis of the non-dominance of a distribution d over a distribution d'. The first one, that is somewhat "liberal", is referred to by Bishop *et al.* (1992) as the *Union-Intersection* (UI) test. It rejects the null hypothesis of non-dominance of d over d' (hence concluding in the dominance of d over d') if at least one of the inequalities that would give rise to the dominance of d over d' is significantly negative and none of them are significantly positive. This test is "liberal" because it could conclude in dominance even if two sample curves were "crossing a bit", provided that they don't cross "too much". By contrast the conservative *Intersection-Union* (IU) test proposed by Howes (1994) would conclude in the dominance of d' by d only if all inequalities that are required for the dominance criterion

[5]We are grateful to a referee for pushing us in that direction.

are "sufficiently" negative. The exact formula and definitions of the two UI and IU tests are provided in Appendix D.

We first examine cross-country comparisons of distributions of BMI categories for each gender. We of course keep in mind that calculated BMI are based on declarative information in France, while they are the results of measurement of height and weight by health professionals in the US or the UK. Figure 2 shows the CDF of BMI categories for women in France, UK and the US for the first ranking of Table 6 (the pattern is similar for the rankings 2-4).



1st ranking *5th ranking*

Figure 2: CDF of BMI categories among women, France, UK and US, 2008

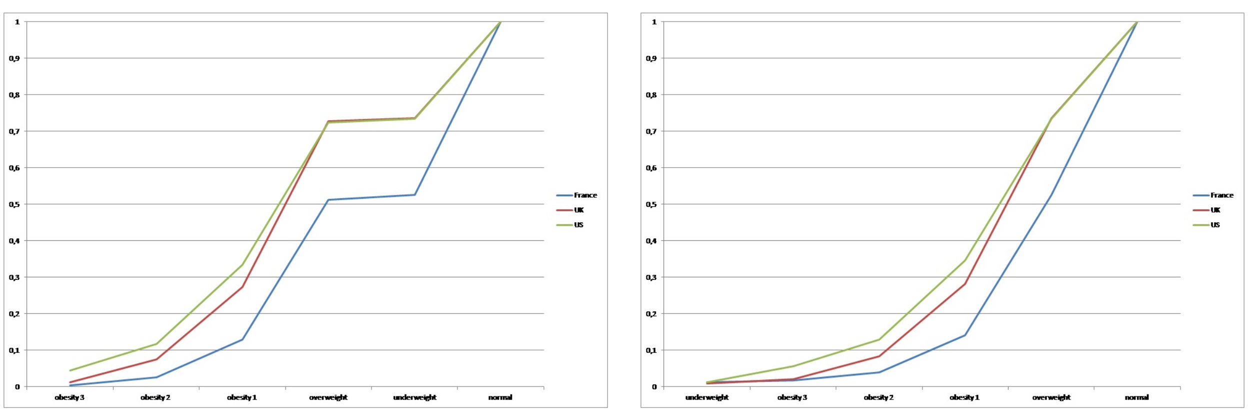
The conclusion obtained from these comparisons, confirmed by Tables 9-13 (in Appendix E) showing the relevant first-order dominance tests based on the statistical inference protocol discussed above,[6] is crystal clear. The distribution of BMI categories is unambiguously better - as per the first order dominance criterion - in France than in the UK and is better in the UK than in the US for the two genders. The only exception to this concerns the French women, who do not dominate their British and American counterpart if the 5th ranking of Table 6 is used, as shown on Figure 2. This absence of dominance

[6]In every matrix of Appendix E, a dominance relation (if any) is indicated with its most stringent degree of significance. The significance of these mentions is as follows:
IU* : all differences that define dominance are negative at a significance level of 99%.
IU : all differences are negative at a significance level of 95%.
UI* No difference is positive at a 95% and at least one difference is negative at 99%
UI No difference is positive at 99% and at least one difference is negative at 99%
The symbol "?" indicates that the two distributions are non-comparable by the considered criterion.

comes from the fact the fraction of underweight women is larger in France than in the UK or the US.

It is also worth observing, on Table 18, that the lack of dominance of French women over their British and American counterparts for the fifth ranking transforms itself into a reverse dominance of the British women over the French ones if one uses the H^+-dominance criterion, that assigns more importance to the underweight category when the latter is considered the worst of all. However, the resort to H^+- dominance does not lead to a reversal of dominance in the case of French and American women, who remain non-comparable if the 5th ranking is used. Observe moreover that the use of the 5th ranking does not affect the dominance of the British women over the American ones. It only weakens its statistical significance, that only satisfies the liberal requirement of the UI methodology. Hence the distribution of BMI categories among women is unambiguously better in the UK than in the US. With the exception of the fifth ranking, the same conclusion holds for the comparison of French women with their British and American counterparts.

The situation is somewhat similar for men. Figure 3 depicts the CDF of the distributions of BMI categories for the first and fifth rankings of Table 6 in each of the three countries.



1st ranking *5th ranking*

Figure 3: CDF of BMI categories among men, France, UK and US, 2008

The figures show a very similar standing of France *vis-à-vis* the US or the UK for the male population than what is observed for the female one. Indeed, the French distribution of BMI categories among men dominates at the first order the British and the American ones for all but the 5th ranking of Table 6. However, and contrary to what was the case for women, the distribution of BMI categories among French males first order dominates those of British and American male even for the fifth ranking where the underweight BMI category is considered to be the worst of all. Indeed, even thought French males tend to be slightly more prone to be underweight than their British or American counterparts, the difference does not have a statistical significance that makes it a breaker of dominance. As shown on Table 18, this is true even if one makes the comparisons by using the H+ criterion. It is also worth noticing that the ranking of the US and the UK distributions of BMI categories among men shows a dominance of the UK over the US that is less solid, from a statistical point of view, than what is observed for women. The only welfare ranking where this difference in significance is not observed is the 5th one. The reason for this "weaker" dominance of British men over American ones come from the observation made above that the fraction of overweight men is slightly larger in the UK than in the US. However, this difference in overweightness in favour of the US does not appear to be statistically significant to the extent that it could break the overall first-order dominance verdict by the liberal UI criterion.

Cross-gender comparisons

As we have seen casually, men and women are affected differently by anomalies in body weight. Women are more prone to severe obesity or excessive thinness than men. However, a larger fraction of women than men have a normal BMI category. It is, for this reason, interesting to investigate whether or not BMI categories are "better" distributed - as per H^+- dominance - among men than among women. As it happens, they tend to be even thought the gender differences are somewhat different in the US than what they are in France or in the UK.

In the US, the male distribution of BMI happens to first-order dominate that of the females for all five rankings of Table 6. This is at least so if we use the liberal UI methodology for performing statistical inference. While a visual examination of the two sample CDF curves (depicted on Figure 4) would suggest that they cross, this crossing is considered insignificant by the UI criterion. What happens here is that the lower fraction of

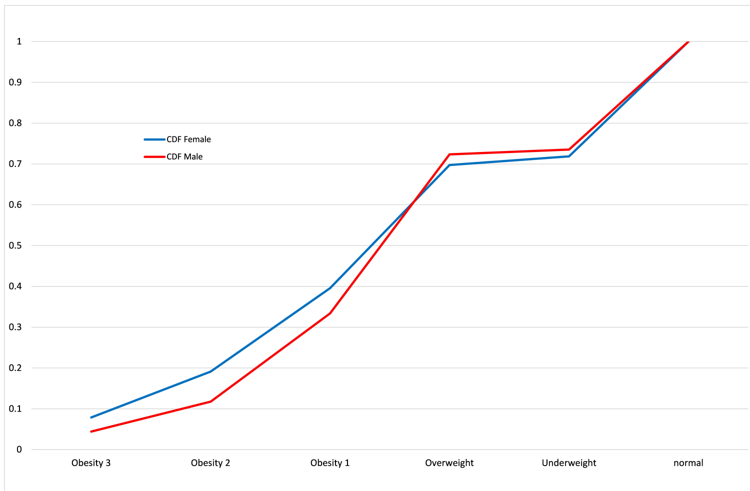


Figure 4: CDF of US males and females, 1st ranking

men - as compared to women - who fall in very unfavorable BMI categories - no matter how these are defined - is statistically more important than the reversal in the ordering of these fractions that takes place in the more favorable BMI categories. We view this first order dominance of men over women in the US, observed for all welfare rankings, as providing a rather strong evidence that women are more adversely affected than men in terms of the distribution of BMI that they face. However, one must notice, from looking at Tables 19-22, that the US distribution of BMI among male dominates the female distribution *also* for the intersection of the H^+ and H^- dominance criteria, which corresponds to the possibility of moving from the female to the male distribution by a finite sequence of Hammond transfers. This may seem contradictory with first order dominance, which is associated with increments only. The reason for this comes from the fact that the BMI distributions among men and among women in the US are very close to each other, and that their only statistically significant differences happen at the bottom of distribution, irrespective of how this bottom is defined. This appears particularly clearly on Figure 5, where the female and male H^- curves are literally confounded, while the H^+ curves, which give priority to the bottom categories, are significantly distinguished. But a similar

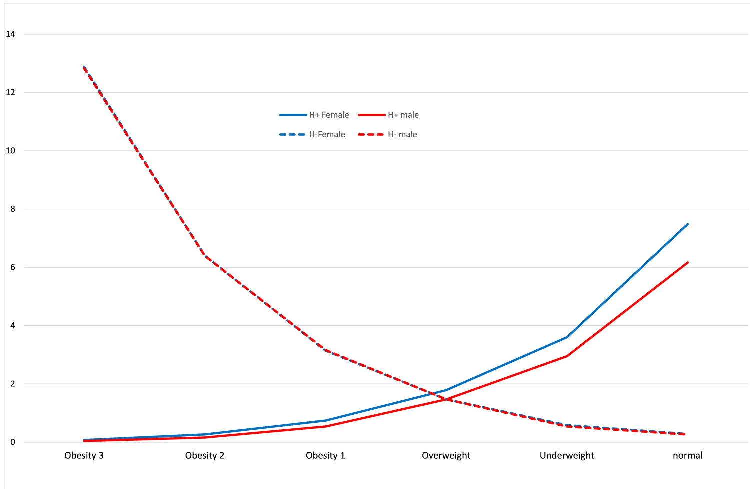


Figure 5: H+ and H- curves of US males and females, 1st ranking

pattern is observed in all the other four rankings. This feature of the US distributions across genders of BMI categories makes the distinction between efficiency and equality in BMI category somewhat difficult to identify from a statistical point of view. Hence, we find safe to consider that efficiency - especially concentrated on the bottom categories - is at stake here.

The situation is quite different in France and the UK where, with one exception, there are no first order dominance relations between the two genders but where the distribution of BMI categories happens to be more equally distributed among men than among women. The exception to this concerns France for the first ranking of Table 6. Indeed, one sees a statistically significant - under the liberal UI rule - first order dominance of female over male for that ranking. The reason for this, illustrated in Figure 6, is somewhat the opposite to what is observed in the US. For the first ranking, the male and the female CDF curves in France are almost confounded at the bottom categories, where there is a slightly lower faction of men than women, but are significantly different for the upper categories, especially the overweight one, where is concentrated a large fraction of men. This therefore

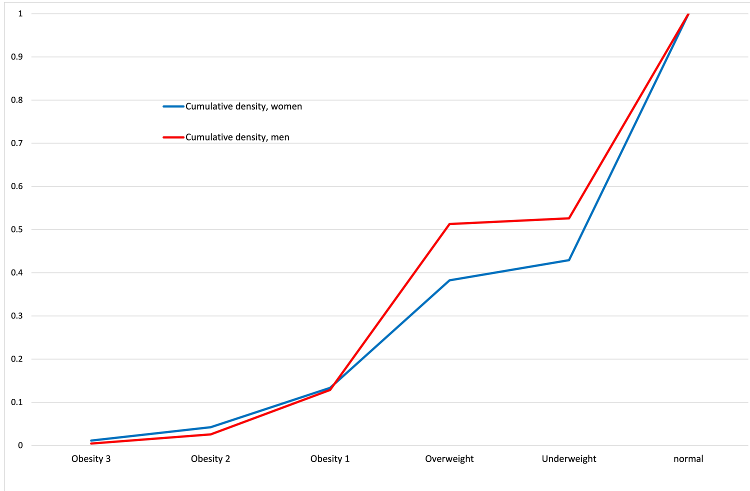


Figure 6: CDF of France male and female BMI categories, 1st ranking

leads to the conclusion that only the differences between the two curves in the upper categories are statistically significant, thus yielding the verdict of first order dominance of the women over the men distribution. However, the lack of significance of the CDF differences in the bottom categories is not robust to the use of the H^+ curves, because of the larger weight assigned by these curves to the bottom categories. This explains why one also observes, in an apparently contradictory fashion, that the men distribution dominates the women distribution for the H^+ criterion (as well as for the H^- one). But if we except this exceptional first order dominance of female over male in France for the first ranking, we otherwise observe for all rankings of Table 6 a more equal - as per Hammond transfer - distribution of BMI categories among men than among women in the UK and in France.

The main forces underlying state of affairs is clear enough. A larger fraction of women than men suffer from severe grade 2 or 3 obesity. Similarly, the fraction of underweight individuals is larger among women than among men. On the other hand men are more represented than women in the mildly obese or overweight categories. Hence, and depending upon the ranking of the BMI categories, these states of affair suggest that the distribution of BMI categories "is more in the middle" among men than among women. This statement obviously depend upon where this "middle" is, and of the exact distribution of the population in the different categories.

IV. Conclusion

Two clear normative conclusions about the cross-gender distribution of individual body weights can be taken out of this paper.

First, for almost all rankings of Table 6, the distribution of BMI is better in France than the UK and in the UK than the US for both males and females for the very robust first order dominance criterion. The only exception to this concerns women, under the (unlikely) assumption that excessive thinness is the worst possible BMI category. In that case, it happens that the fraction of excessively thin women is larger in France than in the US and, somewhat surprisingly, is larger in the US than in the UK. Because this relative standing of the three countries in terms of the fraction of excessively thin women is the opposite to what is observed in terms of excessive weight, one is led to the conclusion that French women have a BMI distribution that is dominated by that of the UK non-comparable to that in the US. Of course this conclusion may be discounted by the fact that measured BMI results from self-declaration in France, while it has been obtained from objective measurement of height and weight by health professionals in the UK and the US.

The second - but in our view most important - conclusion concerns the clear verdict of a more equal distribution of body weights among men than among women in France and the UK. For the case of France, this gender differential is only observed under the assumption that excessive thinness is worse than non-obese overweightedness. In these two countries - the only exception being France for the first ranking of Table 6 - one finds therefore that the male distribution of BMI categories can be obtained from the female one by a finite sequence of Hammond transfers. On the other hand, in the US, there is even first-order dominance of men over women. Hence, in some sense, it can be said that the gender differential in the distribution of BMI categories is worse in the US than in the UK or in France. We believe that the fact that body weight is less favorably distributed among women than among men should be a serious source of concern for health and nutrition policies, especially in view of the important deterioration of body weight observed in the last 15 years in many parts in the world. We also believe that such a conclusion illustrates the usefulness of evaluating distributions of BMI categories by sound, explicit, and robust ethical principles.

A. Survey details

Variables		Individuals	
Name	Value	Nb.	(%)
Gender	Woman	537	(53.43%)
	Man	468	(46.57%)
	All	1005	(100%)
Age	18 − 29 years.	175	(17.41%)
	30 − 49 years.	342	(34.03%)
	50 − 64 years	251	(24.98%)
	65 − 84 years	234	(23.28%)
	≥ 85 years	3	(3.00%)
Marital status	Maried/Registred partnership	495	(49.25%)
	In consensual union	133	(13.23%)
	Single parent	37	(3.68%)
	Widower	39	(3.88%)
	Single	301	(29.95%)
Employment status	Employed	502	(49.95%)
	Unemployed	87	(8.66%)
	Student	63.	(6.27%)
	Other inactivity situation	353	(35.12%)
Education	Primary education	23	(2.29%)
	Lower secondary education	101	(10.05%)
	Upper secondary education	309	(30.75%)
	Short cycle tertiary education	288	(28.66%)
	Bachelors	110	(10.95%)
	Masters	140	(13.93%)
	Doctoral	34	(3.38%)
Net monthly income	< €1400	292	(29.05%)
	€1400 − €1600	122	(12.14%)
	€1601 − €2000	155	(15.42%)
	€2001 − €2700	170	(16.92%)
	≥ €2700	116	(11.54%)
	Does not wish to answer	150	(14.93%)

Table 7: Sociodemographic variables and descriptive statistics

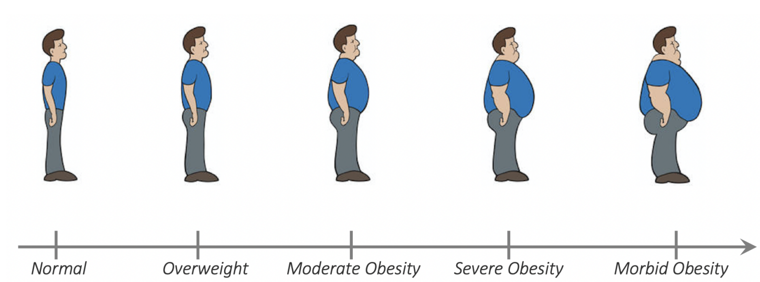
B. Presentation of the survey to the subjects

According to the World Health Organization (WHO), overweight and obesity are defined as abnormal or excessive fat accumulation that may impair health, and can have a negative impact on self-esteem. Overweight and obesity are risk factors from some chronic diseases, such as diabetes, cardiovascular diseases and cancer. Once considered a high-income country problem, overweight and obesity are now on the rise in low- and middle-income countries, particularly in urban settings. The Body mass index (BMI) is a simple index that is commonly used to classify overweight and obesity in adults: It is defined as a person's weight in kilograms divided by the square of his height in meters (kg/m2). For instance a woman 1.65m tall weighting 63kg has a BMI equals to 23.14 (63/1.652).

The WHO proposes a classification for an adult's nutritional status, with categories from undernutrition to obesity. The ObEpi French national epidemiological investigation on overweightand obesity provides, for adults in France in 2012, the following distribution:

Body mass index (BMI)	Nutritional status of the individual	Percentage of adults in France in the category
18.5 – 25	Normal	49%
25 – 30	Overweight	32%
30 – 35	Moderate obesity (Grade 1)	11%
35 – 40	Severe obesity (Grade 2)	3%
More than 40	Morbid obesity (Grade 3)	1%

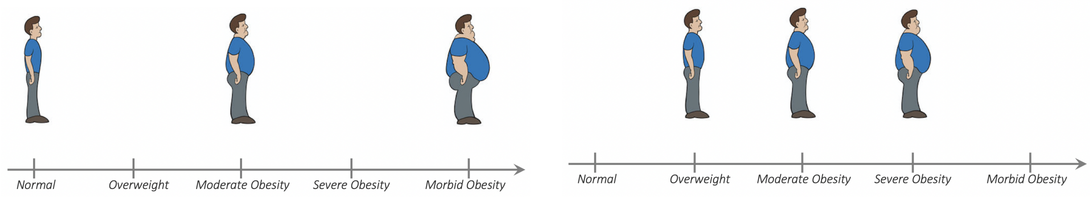
Considering the normal weight category as reference, any BMI increase implies increased health risks, as well as a possible deterioration of self-esteem. The purpose of this survey is to analyze the distribution of the adults within the differentcategories described above, in order to define proper health policies. More precisely, we want to compare such distributions for several countries (France, UK, USA), taking into account possible differences between women and men. One can represent, for instance, the different BMI categories along a straight line, from the less risky category for health (normal weight) to the most risky (morbid obesity). Then, consider a fictive society consisting of 5 adult men, uniformly distributed on these categories. We obtain the following distribution:



In all the following questions, you are asked to compare 2 possible distributions for a society consisting of 3 individuals of the same gender. We note that we don't have any information for those individuals, as regard to their eating patterns, their social origins or their incomes. Below an example of a pair of distributions you will have to compare:

'***Which distribution do you consider as better for the society?***'

Distribution A Distribution B



C. List of questions

Table 8: List of questions

Blocks	Questions	Distribution A					Distribution B				
		Normal	Overweight	Grade 1 obesity	Grade 2 obesity	Grade 3 obesity	Normal	Overweight	Grade 1 obesity	Grade 2 obesity	Grade 3 obesity
B1	B1–Q1	1	-	1	-	1	-	1	1	1	-
	B1–Q2	1	-	1	-	1	-	1	2	-	-
	B1–Q3	1	-	1	-	1	-	-	2	1	-
	B1–Q4	1	-	1	-	1	-	-	3	-	-
	B1–Q5	1	-	1	-	1	-	2	-	1	-
	B1–Q6	1	-	1	-	1	-	1	-	2	-
	B1–Q7	1	-	1	-	1	-	2	-	-	1
	B1–Q8	1	-	1	-	1	1	-	-	2	-
B2	B2–Q1	-	1	1	-	1	-	1	-	2	-
	B2–Q2	-	1	1	-	1	-	-	3	-	-
	B2–Q3	-	1	1	-	1	-	-	1	2	-
	B2–Q4	-	1	1	-	1	-	-	2	1	-
B3	B3–Q1	1	-	1	1	-	-	2	-	1	-
	B3–Q2	1	-	1	1	-	-	-	3	-	-
	B3–Q3	1	-	1	1	-	-	2	1	-	-
	B3–Q4	1	-	1	1	-	-	1	2	-	-
B4	B4–Q1	1	1	1	-	-	-	3	-	-	-
B5	B5–Q1	-	1	1	1	-	-	-	3	-	-
B6	B6–Q1	-	-	1	1	1	-	-	-	3	-
B7	B7–Q1	2	-	-	-	1	1	-	-	2	-
	B7–Q2	2	-	-	-	1	-	2	-	1	-
	B7–Q3	2	-	-	-	1	-	1	2	-	-
	B7–Q4	2	-	-	-	1	-	-	2	1	-
	B7–Q5	2	-	-	-	1	-	1	1	1	-
	B7–Q6	2	-	-	-	1	-	-	3	-	-
	B7–Q7	2	-	-	-	1	1	-	2	-	-
	B7–Q8	2	-	-	-	1	-	2	1	-	-
	B7–Q9	2	-	-	-	1	1	2	-	-	-
	B7–Q10	2	-	-	-	1	1	1	-	1	-
	B7–Q11	2	-	-	-	1	-	3	-	-	-
	B7–Q12	2	-	-	-	1	1	1	1	-	-
	B7–Q13	2	-	-	-	1	1	-	1	1	-
B8	B8–Q1	1	-	-	-	2	-	2	-	-	1
	B8–Q2	1	-	-	-	2	-	1	-	2	-
	B8–Q3	1	-	-	-	2	-	-	2	1	-
	B8–Q4	1	-	-	-	2	-	1	2	-	-
	B8–Q5	1	-	-	-	2	-	1	1	1	-
	B8–Q6	1	-	-	-	2	-	-	3	-	-
	B8–Q6	1	-	-	-	2	-	-	2	-	1
	B8–Q8	1	-	-	-	2	-	-	1	2	-
	B8–Q9	1	-	-	-	2	-	-	-	2	1
	B8–Q10	1	-	-	-	2	-	1	-	1	1
	B8–Q11	1	-	-	-	2	-	-	-	3	-
	B8–Q12	1	-	-	-	2	-	-	1	1	1
	B8–Q13	1	-	-	-	2	-	1	1	-	1

D. Statistical Inference

The dominance criteria described in section 2 involves the verification of k inequalities ($k = 6$ for first order and H^+ dominance and $k = 12$ for the intersection of H^+ and H^- dominance). Each of these inequalities can be seen as a statistical hypothesis and the collection of those inequalities (one such collection for every dominance criterion) can also be seen as a statistical hypothesis. To be specific, consider the following notation, for all $j = 1, \ldots, k$:

$$\begin{aligned} H_0^j &: \gamma_j^d \geqslant \gamma_j^{d'} \\ H_A^j &: \gamma_j^d < \gamma_j^{d'} \\ \overline{H}_0^j &: \gamma_j^{d'} \geqslant \gamma_j^d \\ \overline{H}_A^j &: \gamma_j^{d'} < \gamma_j^d \end{aligned}$$

where γ_j^δ (for $\delta = d, d'$) can be either $F(j;\delta)$, $H^+(j;\delta)$ or $H^-(j;\delta)$ (depending upon the dominance test that one is interested in), H_0^j (resp. $\overline{H}_0^j$) is the null subhypothesis that γ_j^d (resp. $\gamma_j^{d'}$) is no smaller than $\gamma_j^{d'}$ (resp. γ_j^d) and H_A^j (resp. $\overline{H}_A^j$) is the alternative to the null subhypothesis. According to the conservative *Intersection Union* (IU) rule proposed by Howes (1994) (but similar in spirit to the rule proposed by Kaur *et al.* (1994)), the *rejection region* of the global null hypothesis is the union of the k subhypothesis and the *non-rejection region* of the global null hypothesis is the intersection of the non-rejection regions of the k subhypothesis. With this inference rule, we accept dominance of d over d' (resp. of d' over d) if we fail to reject all k null subhypothesis $\overline{H}_0^i$ (resp. H_0^j) and we reject dominance if we reject any one of the k null subhypothesis.

Bishop *et al.* (1992) (see also Bishop and Formby (1999)) have suggested an alternative *Union Intersection* (UI) rule for which the rejection region of the null hypothesis is the intersection of the rejection of the k relevant subhypothesis and the non-rejection region of the null hypothesis is the union of the non-rejection regions of the k subhypothesis. With such a UI rule, we accept dominance of d over d' (resp. of d' over d) if if we fail to reject one of the k null subhypothesis $\overline{H}_0^i$ (resp H_0^j) and we reject dominance if we reject all the k null subhypothesis.

For either inference rule, one needs to construct a test statistics for the sample values γ_j^δ involved in the dominance criteria. To this aim, define the statistic T_j by:

$$T_j = \frac{\widehat{\gamma}_j^d - \widehat{\gamma}_j^{d'}}{\left(\frac{\widehat{Var}(\widehat{\gamma}_j^d)}{N^d} + \frac{\widehat{Var}(\widehat{\gamma}_j^{d'})}{N'} \right)^{\frac{1}{2}}}$$

where $\widehat{\gamma}_j^\delta$ is the observed sample value of γ_j^δ , $\widehat{Var}(\widehat{\gamma}_j^\delta)$ is the sample variance of $\widehat{\gamma}_j^\delta$ and N^δ the sample size of distribution $\delta = d, d'$. The sample variances are calculated as follows for $\widehat{F(j;\delta)}$, $\widehat{H^+(j;\delta)}$ or $\widehat{H^-(j;\delta)}$ (exploiting the detailed definitions of the H^+ and H^-curves provided in Gravel *et al.* (2020) - notably Equations (9) and (14) - and using the Davidson and Duclos (2000) formula for $\widehat{F(j;\delta)}$):

$$\begin{aligned}
\widehat{Var}(\widehat{F(j;\delta)}) &= \widehat{F(j;\delta)} - (\widehat{F(j;\delta)})^2 \ , \\
\widehat{Var}(\widehat{H^+(j;\delta)}) &= \sum_{h \le j} 2^{2(j-h)} \widehat{\delta}_h - (\widehat{H^+(j;\delta)})^2 \text{ and,} \\
\widehat{Var}(\widehat{H^-(j;\delta)}) &= \sum_{h \ge j} 2^{2(h-j)} \widehat{\delta}_h - (\widehat{H^-(j;\delta)})^2
\end{aligned}$$

where $\widehat{\delta}_h$ denote the observed fraction of the sample in category h in distribution δ.

The IU rule as applied to any two sampled distributions d and d' is then defined by :

$$\begin{aligned}
d \text{ dominates } d' &\Leftrightarrow \max(T_1, ..., T_k) < -Z_\alpha \\
d' \text{ dominates } d &\Leftrightarrow \min(T_1, ..., T_k) > Z_\alpha
\end{aligned}$$

while the UI rule is:

$$\begin{aligned}
d \text{ dominates } d' &\Leftrightarrow \min(T_1, ..., T_k) < -C_\alpha \text{ and } \max(T_1, ..., T_k) < C_\alpha \\
d' \text{ dominates } d &\Leftrightarrow \max(T_1, ..., T_k) > C_\alpha \text{ and } \min(T_1, ..., T_k) > -C_\alpha
\end{aligned}$$

where Z_α is the critical value for a significance level of α (α is the probability of rejecting the null when the null is true) derived from the Student distribution and C_α is the critical value for a significance level α determined from the Student Maximum Modulus (SMM) distribution with a degree of freedom of infinity provided by Stoline and Ury (1979).

E. Comparison of distributions by pairs

	FR women	FR men	UK women	UK men	US women
FR men	FR women (UI)				
UK women	FR women (IU*)				
UK men		FR men (IU)	?		
US women	FR women (IU*)		UK women (IU*)		
US men		FR men (IU*)		UK men (UI*)	US men (UI*)

Table 9: 1st order dominance matrix, 1st ranking

	FR women	FR men	UK women	UK men	US women
FR men	?				
UK women	FR women (IU*)				
UK men		FR men (IU)	?		
US women	FR women (IU*)		UK women (IU*)		
US men		FR men (IU*)		UK men (UI*)	US men (UI*)

Table 10: 1st order dominance matrix, 2nd ranking

	FR women	FR men	UK women	UK men	US women
FR men	?				
UK women	FR women (IU*)				
UK men		FR men (IU)	?		
US women	FR women (IU*)		UK women (IU*)		
US men		FR men (IU*)		UK men (UI*)	US men (UI*)

Table 11: 1st order dominance matrix, 3rd ranking

	FR women	FR men	UK women	UK men	US women
FR men	?				
UK women	FR women (UI*)				
UK men		FR men (UI*)	?		
US women	FR women (IU*)		UK women (IU*)		
US men		FR men (IU*)		UK men (UI*)	US men (UI*)

Table 12: 1st order dominance matrix, 4th ranking

	FR women	FR men	UK women	UK men	US women
FR men	?				
UK women	?				
UK men		FR men (UI*)	?		
US women	?		UK women (UI*)		
US men		FR men (UI*)		UK men (UI*)	US men (UI*)

Table 13: 1st order dominance matrix, 5th ranking

	FR women	FR men	UK women	UK men	US women
FR men	FR men (UI)				
UK women	FR women (IU*)				
UK men		FR men (IU)	UK men (IU)		
US women	FR women (IU*)		UK women (IU*)		
US men		FR men (IU*)		UK men (IU*)	US men (IU*)

Table 14: H+ dominance matrix, 1st ranking

	FR women	FR men	UK women	UK men	US women
FR men	FR men (IU)				
UK women	FR women (IU*)				
UK men		FR men (IU)	UK men (IU*)		
US women	FR women (IU*)		UK women (IU*)		
US men		FR men (IU*)		UK men (IU*)	US men (IU*)

Table 15: H+ dominance matrix, 2nd ranking

	FR women	FR men	UK women	UK men	US women
FR men	FR men (IU*)				
UK women	FR women (IU*)				
UK men		FR men (IU)	UK men (IU*)		
US women	FR women (IU*)		UK women (IU*)		
US men		FR men (IU*)		UK men (IU*)	US men (IU*)

Table 16: H+ dominance matrix, 3rd ranking

	FR women	FR men	UK women	UK men	US women
FR men	FR men (IU*)				
UK women	FR women (IU*)				
UK men		FR men (UI*)	UK men (IU*)		
US women	FR women (IU*)		UK women (IU*)		
US men		FR men (IU*)		UK men (IU*)	US men (IU*)

Table 17: H+ dominance matrix, 4th ranking

	FR women	FR men	UK women	UK men	US women
FR men	FR men (IU*)				
UK women	UK women (UI*)				
UK men		FR men (UI*)	UK men (IU)		
US women	?		UK women (IU*)		
US men		FR men (UI*)		UK men (IU*)	US men (IU*)

Table 18: H+ dominance matrix, 5th ranking

	FR women	FR men	UK women	UK men	US women
FR men	FR men (UI*)				
UK women	?				
UK men		?	UK men (IU)		
US women	?		?		
US men		?		UK men (UI*)	US men (UI*)

Table 19: Hammond Equality dominance matrix, 1st ranking

	FR women	FR men	UK women	UK men	US women
FR men	FR men (IU)				
UK women	?				
UK men		?	UK men (IU)		
US women	?		?		
US men		?		UK men (UI*)	US men (UI*)

Table 20: Hammond Equality dominance matrix, 2nd ranking

	FR women	FR men	UK women	UK men	US women
FR men	FR men (IU*)				
UK women	?				
UK men		?	UK men (IU)		
US women	?		?		
US men		?		UK men (UI*)	US men (UI*)

Table 21: Hammond Equality dominance matrix, 3rd and 4th rankings

	FR women	FR men	UK women	UK men	US women
FR men	FR men (IU*)				
UK women	UK women (UI*)				
UK men		?	UK men (IU)		
US women	?		?		
US men		?		UK men (UI*)	US men (UI*)

Table 22: Hammond Equality dominance matrix, 5th ranking

References

Abul-Naga, R. and Yalcin, T. (2008) Inequality measurement for ordered response health data, *Journal of Health Economics*, **27**, 1614–1625.

Allison, R. A. and Foster, J. E. (2004) Measuring health inequality using qualitative data, *Journal of Health Economics*, **23**, 505–524.

Amiel, Y. and Cowell, F. (1992) Inequality measurement: An experimental approach, *Journal of Public Economics*, **69**, 3–26.

Apouey, B. (2007) Measuring health polarization with self-assessed health data, *Health Economics*, **16**, 875–894.

Atkinson, A. B. (1970) On the measurement of inequality, *Journal of Economic Theory*, **2**, 244–263.

Bishop, C. M. and Formby, J. P. (1999) Test of significance for lorenz partial orders, in *Handbook of Inequality Measurement* (Ed.) J. Silber, Kluwer Academic Press.

Bishop, C. M., Formby, J. P. and Thistle, P. (1992) Convergence of the south and the non-south income distributions, 1969-1979, *American Economic Review*, **82**, 262–272.

Blanchflower, D. G., Landeghem, B. V. and Oswald, A. J. (2010) Imitative obesity and relative utility, *Journal of the European Economic Association*, **7**, 528–538.

Cao, S., Moineddin, R., Urquia, M. L., Razak, F. and Ray, J. G. (2014) J-shapedness: an often missed, often miscalculated relation: the example of weight and mortality, *Journal of Epidemiology and Public Health*, **68**, 683–690.

Carr, D. and Friedman, M. A. (2005) Is obesity stigmatizing? body weight, perceived discrimination, and psychological well-being in the united states, *Journal of Health and Social Behavior*, **246**, 244–259.

Chen, Y. M., Ho, S. C., Lam, S. S. and Chan, S. S. (2006) Validity of body mass index and waist circumference in the classification of obesity as compared to percent body fat in chinese middle-aged women, *International Journal of Obesity*, **22**, 423–443.

Cowell, F. A. and Flachaire, E. (2017) Inequality with ordinal data, *Economica*, **84**, 290–321.

Dasgupta, P., Sen, A. K. and Starrett, D. (1973) Notes on the measurement of inequality, *Journal of Economic Theory*, **6**, 180–187.

Davidson, R. and Duclos, J. Y. (2000) Statistical inference for stochastic dominance and for the measurement of poverty and inequality, *Econometrica*, **58**, p.1435–1465.

Etile, F. (2014) Education policies and health inequalities: Evidence from changes in the distribution of body mass index in france, *Economics and Human Biology*, **13**, 46–65.

Feingold, A. and Mazzella, R. (1998) Gender differences in body image are increasing, *Psychological Science*, **9**, 190–195.

Fields, G. and Fei, J. (1978) On inequality comparisons, *Econometrica*, **46**, 305–316.

Flegal, K. M., Graubard, B. I., Williamson, D. F. and Gail, M. H. (2005) Excess deaths associated with underweight, overweight, and obesity, *Journal of the American Medical Association*, **293**, 1861–1867.

Furnham, A., Badmin, N. and Sneade, I. (2002) Body image dissatisfaction: Gender differences in eating attitudes,self-esteem, and reasons for exercise, *The Journal of Psychology*, **136**, 581–596.

Gravel, N., Magdalou, B. and Moyes, P. (2020) Ranking distributions of an ordinal variable, *Economic Theory*, forthcoming.

Gray, D. S. and Fujioka, K. (1991) Use of relative weight and body mass index for the determination of adiposity, *Journal of Clinical Epidemiology*, **44**, 545–590.

Hadar, J. and Russell, W. (1974) Stochastic dominance in choice under uncertainty, in *Essays on Economics Behavior under Uncertainty* (Eds.) M. S. Balch, D. L. McFadden and S. Wu, North Holland, Amsterdam, UK.

Hammond, P. J. (1976) Equity, arrow's conditions and rawls's difference principle, *Econometrica*, **44**, 793–803.

Howes, S. (1994) Testing for dominance: Inferring population rankings from sample data, unpublished paper, policy research department, World Bank.

Kanter, R. and Caballero, B. (2012) Global gender disparities in obesity: A review, *Advances in Nutrition*, **3**, 491–498.

Kaur, A., Prakasa, R. and Singh, H. (1994) Testing for second-order stochastic dominance for two distributions, *Econometric Theory*, **10**, 849–866.

Keys, A., Fidanza, F., Karvonen, T. M. J., Kimura, N. and Taylor, H. L. (1972) Indices of relative weight and obsesity, *Journal of Chronic Diseases*, **25**, 329–343.

Khlat, M., Jusot, F. and Ville, I. (2009) Social origins, early hardship and obesity: A strong association in women, but not in men?, *Social Science and Medicine*, **68**, 1692–1699.

Kobus, M. and Milós, P. (2012) Inequality decomposition by population subgroups for ordinal data, *Journal of health Economics*, **31**, 15–21.

Kolm, S. C. (1969) The optimal production of social justice, in *Public Economics* (Eds.) H. Guitton and J. Margolis, Macmillan, London.

Lehmann, E. L. (1955) Ordered family of distributions, *The Annals of Mathematics and Statistics*, **26**, 399–419.

Mooney, S. J. and El-Sayed, A. M. (2016) Stigma and the etiology of depression among the obese: An agent-based exploration, *Social Science and Medicine*, **148**, 1–7.

Oswald, A. J. and Powdthavee, N. E. (2007) Obesity, unhappiness, and the challenge of affluence: Theory and evidence, *Economic Journal*, **117**, 441–454.

Quirk, J. D. and Saposnik, R. (1962) Admissibility and measurable utility functions, *Review of Economic Studies*, **29**, 140–146.

Roberts, R. E., Kaplan, G. A., Shema, S. J. and Strawbridge, W. J. (2000) Are the obese at greater risk of depression?, *American Journal of Epidemiology*, **152**, 163–170.

Roberts, R. E., Strawbridge, W. J., Deleger, S. and Kaplan, G. A. (2002) Are the fat more jolly?, *Annals of Behavioral Medicine*, **24**, 169–180.

Sassi, F. (2014) Obesity update, Tech. rep., OECD.

Sen, A. K. (1973) *On Economic Inequality*, Oxford, Clarendon.

Shorrocks, A. F. (1983) Ranking income distributions, *Economica*, **50**, 3–17.

Sonne-Schmidt, C., Tarp, F. and Osterdal, L. P. (2016) Ordinal bivariate inequality: Concepts and application to child deprivation in mozambique, *Review of Income and Wealth*, **62**, 559–573.

Stoline, M. R. and Ury, H. K. (1979) Tables of the studentized maximum modulus distributions and an application to multiple comparisons among means, *Technometrics*, **21**, 87–93.

Stutzer, A. and Meier, A. N. (2016) Limited self control, obesity, and the loss of happiness, *Health Economics*, **25**, 1409–1424.

Wells, J. C., Marphatia, A. A., Cole, T. and McCoy, D. (2012) Association of economic and gender inequality with global obesity prevalence: understanding the female excess, *Social Science and Medicine*, **75**, 482–490.

Zheng, B. (2008) A note on measuring inequality with ordinal data, in *Research in Economic Inequalities, vol. 16* (Eds.) J. Bishop and B. Zheng, Emerald, pp. 177–188.

Zheng, B. (2011) A new approach to measure socioeconomic inequality in health, *Journal of Economic Inequality*, **9**, 555–577.