UNIVERSITY OF SOUTHERN DENMARK

COHERE - Centre of Health Economics Research, Department of Business and Economics
Discussion Papers, No. 2014:5
ISSN: 2246-3097

Refining Population Health Comparisons: A Multidimensional First Order Dominance Approach

by
M. Azhar Hussain[1], Mette Møller Jørgensen[2] and Lars Peter Østerdal[3]
[1] Roskilde University
[2] Harvard University
[3] University of Southern Denmark

COHERE, Department of Business and Economics
Faculty of Business and Social Sciences
University of Southern Denmark
Campusvej 55
DK-5230 Odense M
Denmark
www.cohere.dk

Refining Population Health Comparisons: A Multidimensional First Order Dominance Approach

M. Azhar Hussain[1], Mette Møller Jørgensen[2], Lars Peter Østerdal[3]

February 20, 2014

Abstract: How to determine if a population group has better overall (multidimensional) health status than another is a central question in the health and social sciences. We apply a multidimensional first order dominance concept that does not rely on assumptions about the relative importance of each dimension or the complementarity/substitutability across dimensions. In particular, we suggest that one can explore the "depth" of dominances by sequentially refining the health dimensions to see which dominances persist. Using The Danish National Health Interview Survey, we conduct dominance comparisons between population groups based on education, gender, marital status, and ethnicity for given age intervals. Our empirical illustration shows that it is possible to operationalize and meaningfully apply the multidimensional first order dominance concept with sequential refinements of health status to as much as ten health dimensions.

Keywords: Multidimensional first order dominance; population health comparisons; refinement; inequalities in health; The Danish National Health Survey.
JEL codes: D63, I14.

1. Introduction

How to determine if a defined population has overall better health status than another is a fundamental question in health economics, epidemiology, and public health sciences. The health sciences have a long tradition of recording multiple indicators covering the different

[1] Department of Society and Globalisation, Roskilde University.
E-mail: azharh@ruc.dk

[2] John F. Kennedy School of Government, Harvard University.
E-mail: mette_moeller_joergensen@hks.harvard.edu

[3] Department of Business and Economics, and Centre of Health Economics Research (COHERE), University of Southern Denmark.
E-mail: lpro@sam.sdu.dk

aspects of individual health status, and it is well-recognized that health status is a complex multi-facetted phenomenon. For relevant population health comparisons we thus need methods that deal appropriately with the multidimensional nature of health status. In this way, the problem of comparing multidimensional population health echoes the problem of making multidimensional comparisons of poverty or social welfare (e.g., Townsend 1979, Sen 1999, Sen 2009).

One approach to dealing with the multidimensionality of outcome is to introduce a weight to each indicator and then calculate an overall health index for each individual, which is again aggregated into an overall population health indicator. Indeed, many clinical studies and health economics evaluations have followed this route (e.g., Drummond et al. 2005). While this approach is operationally very tractable once indicators and weights have been established, it is based on strong underlying assumptions about the way each indicator contributes to overall health, and the procedure for establishing such weights raises methodological, practical, and ethical concerns.

A particular sort of population health comparisons are those that intend to give complete rankings of countries/health care systems (e.g., Feeny et al. 2002, UHF 2012, Kohn 2012, HCP 2012, WHO 2013). Such population health studies typically overcome the problems of multidimensionality by analyzing one dimension at a time (thus disregarding interaction between dimensions and giving as many rankings as there are dimensions), or a "counting approach" is applied where a person is classified as having a bad overall health status when he or she experiences a bad outcome in at least x out of a possible n dimensions (e.g., Alkire and Foster 2011). This counting approach assumes given (often equal) weights to different health indicators and also implicitly assumes substitutability in health indicators. The analysis is therefore sensitive to indicator weights, which means that ranking can change depending on the weighting scheme.

In this paper, we propose another approach to multidimensional population health comparisons. The main feature of our proposal is that it is truly multidimensional and neither relies on weighting schemes nor on assumptions about the possible substitutability/complementarity relationships between defined health dimensions. A potential drawback of our proposal is that we are not always able to determine a ranking when two given populations are compared – but the cases where a ranking cannot be determined are precisely those in which a ranking would have relied on certain weight ranges or underlying substitutability/complementarity assumptions.

Our methodology builds upon the growing literature developing techniques for comparing different groups' multidimensional social welfare or poverty, which are methodologically robust in the sense that they do not rely on particular weighting or counting schemes (e.g., Atkinson and Bourguignon 1982, Bourguignon and Chakravarty 2003, Duclos and Makdissi 2005, Duclos et al. 2007, Cowell and Victoria-Feser 2007, Gravel et al. 2009, Gravel and Mukhopadhyay 2010, Batana and Duclos 2010, Duclos and Echevin 2011, Duclos et al. 2011). However, as in Arndt et al. (2012, 2013), we apply a multidimensional first order dominance (FOD) comparison technique that differs from the other robust welfare/poverty comparisons methods in that it avoids implicit assumptions about the substitutability/complementarity relationships between the dimensions chosen.[4] According to the FOD methodology, population group A dominates another population group B if the distribution of B's individuals on all possible outcome combinations can be obtained from A's by moving population shares from better to worse outcomes (e.g., Østerdal 2010). Thus, the FOD concept compares two given multidimensional distributions without specifying weights to the different dimensions and is the natural way to determine whether one group is unambiguously better off than another group.

Considering a large number of health dimensions comes at a cost: The more refined dimensions we apply, the more demanding the FOD criterion is, and therefore there will be fewer dominances (as explained in detail below). We tackle this issue by applying a sequential refinement strategy. We initially look at only a single health dimension encompassing all underlying indicators and then sequentially refine dimensions ending up with a 10-dimensional picture of individual health. In the initial 1-dimensional case (where all indicators are collapsed into a single health dimension) for each individual we simply see if he or she has reported a health problem for *some* indicator. When we take a multidimensional view, we see if he or she has reported a health problem within each specified dimension (described by one or more health indicators). By sequentially exploring dominances for increasingly refined health dimensions it is possible to see whether dominances remain when more nuanced descriptions of individual health are considered.

[4] That is, we do not make other assumptions on an underlying population health evaluation function than the (trivial structural) assumption that better individual health yields better population health.

We analyze a comprehensive Danish health survey covering 22 aspects of health.[5] These health indicators are first collapsed into a single health dimension and then refined to four, seven, and ten health dimensions, for which the joint health status distribution of different groups is compared. To our knowledge, this is the first study that empirically has analyzed FOD with as many as ten dimensions (of any type).

Overall we find that younger age groups dominate older age groups in up to four dimensions (1-dimensional and 4-dimensional cases), but no dominations in higher dimensions (7-dimensional and 10-dimensional cases) are present when only age groups are compared. All other comparisons are done within age groups. We find that groups with long education dominate groups with short education across age groups. In one case we see dominance in ten dimensions, and more often we see domination in four and seven dimensions, but generally the depth of the dominance varies and, hence, when we take a more refined view of individual health, groups with long education are not unambiguously better off. In addition to education we also consider gender, marital status, region, and ethnicity. Here we find less dominances in four dimensions and only one in seven dimensions.

The remainder of the article is organized as follows. The next section reviews the literature. Section 3 presents the multidimensional FOD methodology in greater detail. Sections 4 and 5 present results from applying the methods to Danish national health survey data. Finally, section 6 summarizes and concludes.

2. Earlier studies

Ranking studies of population health and health systems are regularly conducted. The World Health Organization's (WHO) annual ranking of its 194 member countries is the most comprehensive study in terms of global coverage. The latest ranking (WHO 2013) includes multiple indicators per health dimension (count in parentheses): life expectancy and mortality (2 and 5); cause-specific mortality and morbidity (6 and 2); selected infectious diseases (18); health service coverage (18); risk factors (17); and health systems with three sub-dimensions representing health workforce, infrastructure and technologies, and essential medicines (7, 5, and 2). In total, 82 indicators are applied in the WHO ranking. All indicators are presented

[5] More indicators have been created from these data (see Koch et al. 2012). For example a number of indicators have been constructed referring to occupational health, which is obviously not relevant for all. We focus in our empirical section on a selection of indicators that is suitable for general population health comparisons.

nationally and separately for females and males. Although multiple dimensions are the basis for the analysis, the ranking of countries is based on a given single indicator at a time. Thus, there could be as many rankings as there are indicators (82) since no index is created to make an overall rank. Furthermore, the joint distribution of health characteristics at the individual level is not considered.

An overall ranking of U.S. states applying the United Health Foundation's (UHF) 24 health indicators is carried out by constructing one composite health index (UHF 2012). In order to make the composite health index, a weighting scheme is applied with a 75% weight on health determinants (behaviors, community and environment, policy, and clinical care) and a 25% weight on health outcomes. 16 indicators cover the determinants and eight cover outcomes. For each of the 24 indicators a Z score is calculated representing how many standard deviations a state is above the national average. The composite Z score is a weighted average of indicator-specific Z scores using the weighting scheme. In contrast to WHO (2013), UHF (2012) explicitly presents weights specifying the relative importance of given health dimensions. Although the weighting scheme is carefully composed, the ranking of U.S. states can nevertheless change if the weight composition is changed. Also, the joint distribution of indicators at the individual level is not taken into account since the composite index is based on state level aggregate incidence rates for the 24 indicators. An advantage of the UHF measure is however that we can compare the state population's health status according to multiple indicators of health.

In the EU, systematic health rankings are conducted by the Health Consumer Powerhouse (HCP). Their study includes five dimensions (sub-disciplines) aggregated from 42 indicators (HCP 2012). The five dimensions are (number of indicators and relative weight in parentheses): patient rights/information/e-Health (12, 0.175), accessibility/waiting times for treatment (5, 0.25), outcomes (8, 0.3), prevention/range and reach of services provided (10, 0.175), and pharmaceuticals (7, 0.1). A given indicator for a country gets a score of 1 (worst), 2, or 3 (best). Apart from dimensions, indicator content, and the specific weighting scheme, the UHF and HCP basically use the same methodology to construct rankings based on a composite health indicator. The HCP ranking thus also potentially suffers from the problems encountered by the UHF ranking. The problem of lacking robustness of rankings is not addressed methodologically, although HCP (2012) states that "final ranking of countries presented by the EHCI [Euro Health Consumer Index] 2012 is remarkably stable if the weight coefficients are varied within rather wide limits". But this seeming stability in the EHCI does not preclude future instabilities or instabilities caused by including new countries (enlargement of EU) or limiting analysis to smaller socio-economic population groups.

Another strand of the literature has focussed on constructing individual level health utility indices, which are convenient for comparing the overall health of individuals and groups in clinical trials and assessments of health care programs. For example, Feeny et al. (2002) construct a single health utility index from eight indicators of health (vision, hearing, speech, ambulation, dexterity, emotion, cognition, and pain) assessed at a five or six level Likert scale representing highly impaired to normal. The methodology transforms multiple individual health questionnaire answers into a single index using a utility function whose parameters are estimated from the general population. For a survey on health utility functions we refer to Drummond et al. (2005). Other researchers have applied statistical techniques to get from multiple underlying health indicators to a single unidimensional measure. Kohn (2012), for example, applies multiple correspondence analysis (MCA) on British data to reduce multiple discrete indicators to a continuous variable, which is shown to perform better (higher explanatory power) than self-assessed health in econometric wage equations. Seventeen indicators are used in the areas of mental and emotional health, smoking, disability, accidents, and reported health problems. In this case, the author does not specify indicator weights, but instead the MCA methodology itself suggests weights depending on the data.[6]

All of the above-mentioned studies explore important avenues in the search for aggregation of multiple health indicators (at the individual or aggregate level) into a single index which can be used for the ranking of groups of people. The approaches involve an implicit or explicit value judgement when aggregating underlying indices. In one way or the other, all the studies apply a weighting scheme specifying the importance of health indicators in order to arrive at a single unidimensional measure. None of the studies fully take the joint health distribution of individuals into account; for example, the accumulation of problematic health at the individual level is not considered. The approach taken in the following explicitly takes the individual joint health distribution into account, and at the same time it does not,

[6] Although the literature on concentration curves and concentration indices (O'Donnell et al. 2008, Wagstaff and van Doorslaer 2000) also involves dominance comparisons between population subgroups, it is not discussed further here since it is based on a unidimensional health measure, while the focus of this paper is on multidimensional health comparisons.

implicitly or explicitly, specify any weight to any one indicator. Instead, a robust method is developed where dominance conclusions are independent of any weighting scheme.

3. The multidimensional first order dominance methodology

As mentioned in the Introduction, according to the FOD concept for two (multidimensional discrete) population distributions A and B we have that A dominates B if and only if it is possible to obtain A from B by moving probability mass from worse to better outcomes.[7] Mosler and Scarsini (1991) and Dyckerhoff and Mosler (1997) describe a method based on linear programming for checking first order dominance in the general multivariate finite case (see Range and Østerdal 2013 for further details). An empirical implementation of a linear programming-based method for checking multidimensional FOD was provided in Arndt et al. (2012) and applied to binary poverty indicators. The five-dimensional coding is displayed in Arndt et al. (2013). We use up to ten binary dimensions in this study, which means that a total of 1,024 (=2^{10}) health status combinations are possible for each individual.

Suppose there are n dimensions and that a number of 0/1 (unhealthy/healthy) valued binary indices are defined – one for each of the dimensions. Let a_i and b_i be the shares of population A and B, respectively, with health status i. Let the variable x_{ij} represent probability mass transfer from health status i to health status j. Define Z as the set of source-destination pairs ij that move probability from a health status i to a less preferred health status j (i.e., i is at least as good as j in all dimensions). Under these conditions, population A dominates population B if and only if there exists x_{ij} with $x_{ij} \geq 0, x_{ii} = 0$, such that $a_i + \sum_{ji \in Z} x_{ji} - \sum_{ij \in Z} x_{ij} = b_i$ for all i.

For a specified set of dimensions, a good outcome in all indicators of a given dimension is required in order for an individual to be classified as being healthy in that dimension. Since we have binary indicators, in the 1-dimensional case the FOD condition degenerates to comparing the prevalence of good health (i.e., comparing two fractions) guaranteeing a

---

[7] The concept of multidimensional first order dominance (FOD) applied in this paper is also known simply as *dominance*, or the *usual (stochastic) order* (e.g., Lehmann 1955, Levhari et al. 1975). Note that FOD is more demanding, and less easy to check, than the multidimensional dominance concepts applied by Atkinson and Bourguignon (1982) and subsequent work referenced in the Introduction. For a general treatment of stochastic dominance theory, we refer to Shaked and Shanthikumar (2007).

dominance for any two distributions that are compared. When we increase the number of dimensions, the FOD criterion becomes more demanding. More precisely, whenever a specified dimension is refined into two or more sub-dimensions, dominances will be fewer or the same:

Lemma: *Consider two population distributions A and B with n health dimensions (described by n binary 0-1 indicators), and suppose that A dominates B. Let A' and B' be two distributions obtained from A and B, respectively, by collapsing a given subset of the dimensions consisting of m ≤ n indicators into a single dimension (such that this new indicator takes the value 1 for an individual if and only if the individual has value 1 for all the m collapsed indicators). Then A' dominates B'.*

Proof: Recall that by the definition of FOD, we have that any non-decreasing social welfare function defined on $(0,1)^n$ provides at least as high welfare for distribution A as for distribution B. Specifically, we observe that for any non-decreasing social welfare function defined on $(0,1)^n$ that takes the same function value for any two outcomes which are identical on the *n*-*m* indicators (those that are not collapsed) and 0 for at least one of the remaining *m* indicators (those that are collapsed), we also have that the function value is at least as high for A as for B. For distributions A' and B' (with *n*-*m*+1 indicators), we have, by definition, that A' dominates B' if and only if any non-decreasing social welfare function on $(0,1)^{n-m+1}$ provides at least as high social welfare for A' as for B'. It follows from the observation above that indeed A' dominates B', since the class of *n*-*m*+1 dimensional non-decreasing social welfare functions corresponds to the class of *n* dimensional social welfare functions such that the social welfare level is the same for any two outcomes which are equal on the *n*-*m* indicators (those that are not collapsed) and 0 for at least one of the remaining *m* indicators (those that are collapsed). To see this more precisely, note that an *n*-*m*+1 dimensional non-decreasing social welfare function gives at least as much welfare to A' than to B' if any *n* dimensional non-decreasing social welfare function such that the function value is the same for any two outcomes which are equal on the *n*-*m* indicators (those that are not collapsed) and 0 for at least one of the remaining *m* indicators (those that are collapsed) gives at least as much welfare at A that at B which is indeed the case.

4. Data

Data is provided by the Danish National Institute for Public Health through "The National Health Interview Survey 2010". The sample is representative of the Danish adult (16+) population as a whole as well as for the five regions of Denmark. The sample size is 15,165 individuals. Some questions were not answered by all individuals, and thus for our purpose the effective sample size is 11,433 individuals. We apply weights for non-response provided by Statistics Denmark. See Christensen et al. (2012) and Koch et al. (2012) for more information on "The National Health Interview Survey 2010". The basic indicators applied in this study are selected from an even broader set of indicators from which we have chosen 22 that are well-suited to describe the individuals' general health status. We use a combination of indicators covering self-reported health, pain and discomfort, chronic diseases, own health efforts, and risk factors.[8] The basic 22 indicators are shown in the far right side of Figure 1. The different sub-indicators and the way the indicators are grouped with different numbers of dimensions are also presented in Figure 1. To be classified as being healthy in a given dimension the individual needs to be free from problems with respect to all indicators in that dimension. For instance having good health in the 1-dimensional case means the individual reported no problems in all 22 underlying indicators. Similarly, good health, for instance in dimension IV in the 7-dimensional case, means the individual neither reported problems with asthma, allergy, or migraines/frequent headaches.

When increasing the number of dimensions, we simultaneously refine several dimensions. In the four dimensions case we look at health consisting of self-reported health, pain and discomfort, chronic conditions, and behavioral factors. In seven and ten dimensions we refine these categories further and thus consider physical, sleeping, and psychological discomforts separately. The prevalence of being free from different health problems is also presented in Figure 1. For example, using the basic indicators, 94.7% do not report asthma as a problem, 82.2% do not report allergy as a problem, and 89.2% are free from migraine or frequent headaches. Using the 10-dimensional case, 79.8% report neither asthma nor allergy as a problem. Using the 7-dimensional case, 72.4% do not suffer from either asthma, allergy or headaches, etc. Because of the aggregation procedure the prevalence is non-decreasing when indicators are aggregated (moving from right to left in Figure 1).

---

[8] The left-out items include individual behavior and use of health services, social relations, and working environment – the last item is only relevant for employed people.

Figure 1. Aggregation rules and prevalence of good health in each dimension. 2010. %

1 dimension		4 dimensions		7 dimensions		10 dimensions		Free from problem with respect to …	
All	28.1	i	87.7	I	87.7	1	87.7	Poor or fair self-reported health	89.1
								Never or hardly never feels well enough to do what one wants to do	94.5
		ii	72.3	II	77.3	2	77.3	Pain or discomfort in the shoulder or neck	89.4
								Pain or discomfort in the arms, hands, legs, knees, hips or joints	89.3
								Pain or discomfort in the back or lower back	89.1
								Headaches	94.8
				III	89.0	3	91.9	Sleeping problems or insomnia	91.9
						4	94.8	Melancholy, depression, unhappiness	95.7
								Anxiety, nervousness, restlessness or apprehension	96.6
		iii	60.2	IV	72.4	5	79.8	Asthma	94.7
								Allergy	82.2
						6	89.2	Migraine or frequent headaches	89.2
				V	82.5	7	82.5	Diabetes	96.3
								Hypertension (high blood pressure)	86.4
								Chronic bronchitis	96.7
		iv	54.8	VI	87.5	8	87.5	Do not believe that individual efforts to stay healthy are important	99.4
								Do not do anything to stay healthy or improve own health	87.7
				VII	57.5	9	72.2	Daily tobacco use	79.1
								Heavy tobacco use	89.1
								Transcending Board of Health high-risk limit for alcohol consumption	89.1
						10	78.3	Sedentary leisure activity	88.0
								Obesity	87.5

Note: From left to right the aggregation rules from 1 to 10 dimensions are seen. To the right each basic health indicator is shown. The numbers indicate the percentage of respondents not suffering from any of the health problems in the given dimension.

Source: Own calculations based on "The National Health Interview Survey 2010", The National Institute for Public Health, Denmark. Indicators inspired by Koch, Davidsen, and Juel (2012).

Analysis variables and descriptive statistics

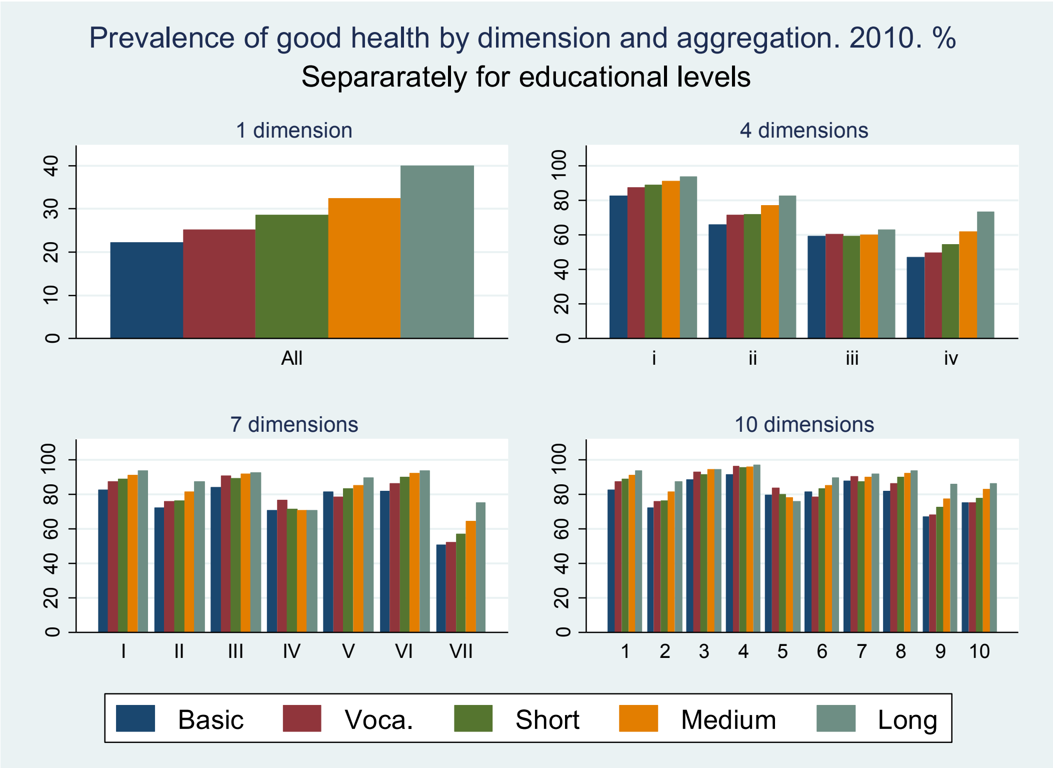
FOD is investigated for different population groups which in all cases are conditioned upon age (group). As a measure for social-economic status we use the educational level. Educational attainment is irreversible and thus not conversely correlated with disease, which is in contrast to, for instance, labor market attachment or income (e.g., Juel et al. 2006). The association between education and health is also analyzed in Kunst et al. (2005) for 10 European countries, in Jørgensen et al. (2013) for Denmark regarding self-reported health, and in Dalstra et al. (2006) regarding chronic diseases. Other considered variables are gender, ethnicity, marital status, and area of residence. The age conditional distribution of the survey is displayed in Table 1.

Table 1. Characteristics distribution and sample size. By age group. Denmark. 2010. %

	16-24	25-44	45-64	65+	Total	Sample
Education:						
Basic	78.8	14.9	18.6	30.1	27.7	2,866
Vocational	10.2	28.0	32.9	38.4	28.5	3,354
Short higher	4.0	12.9	15.1	7.8	11.6	1,384
Medium higher	4.8	24.6	22.4	14.9	19.6	2,400
Long higher	2.3	19.5	11.1	8.9	12.7	1,429
Gender:						
Male	53.1	51.3	51.8	51.9	51.8	5,500
Female	46.9	48.7	48.2	48.1	48.2	5,933
Etnicity:						
Danish	86.0	88.7	93.3	94.6	90.7	10,786
Western immi.	4.7	4.8	3.2	4.1	4.2	297
Non-West. immi.	9.4	6.5	3.4	1.3	5.2	350
Marital status:						
Widow(er)	3.8	52.5	68.4	58.0	51.5	6,679
Divorced	21.7	23.2	7.9	3.3	15.0	1,479
Couple	1.8	5.1	14.0	32.7	11.4	1,198
Cohabiting	70.5	18.0	8.7	4.4	20.7	1,932
Missing	2.3	1.3	1.0	1.7	1.4	145
Region:						
Northern Jutland	11.8	9.6	10.7	8.8	10.2	1,203
Middle Jutland	24.2	24.2	22.0	20.4	22.9	2,726
Southern Denmark	19.1	20.2	22.8	21.5	21.1	2,521
Capital Area	32.0	34.1	28.6	32.1	31.7	3,428
Zealand (rest)	12.9	11.9	16.0	17.1	14.1	1,555
Total	100	100	100	100	100	
Sample	1,393	3,636	4,493	1,911		11,433

Source: Own calculations based on "The National Health Interview Survey 2010" by the National Institute for Public Health, Denmark.

Figure 2.



Below we describe the data in more detail and provide a qualification for the use of FOD based on this.

The 1-dimensional case

Being healthy with regard to all indicators is clearly linked to educational level (Figure 2). People with university degrees have a 40% probability of not reporting any health issues, while this probability is 22% among people with basic education. But already at a refinement with four indicators the pattern becomes blurred; particularly regarding indicator iii, we see the same prevalence around 60% for all educational groups. For the other indicators, i, ii, and iv, we still see a pattern of higher education being linked to better health, albeit in a weaker relation compared to the 1-dimensional case. Refinement to 7 and 10 dimensions further weakens the relationship between education and health. In one instance we nearly see a negative correlation: The risk of asthma and allergy increases with educational attainment (except for basic education).

Table 2. Prevalence of good health wrt. all indicators (1-dimension). By characteristic given age. Denmark. 2010. %

	16-24	25-44	45-64	65+	Total
Education:					
Basic	31	19	15	14	22
Vocational	24	30	23	19	25
Short higher	25	32	27	24	29
Medium higher	29	35	31	26	32
Long higher	34	45	38	19	40
Gender:					
Male	30	32	27	19	28
Female	29	33	25	19	28
Etnicity:					
Danish	31	33	26	19	28
Western immi.	28	31	30	14	28
Non-West. immi.	20	27	17	9	22
Marital status:					
Widow(er)	26	35	29	22	30
Divorced	29	32	21	15	29
Couple	17	23	21	13	18
Unmarried	30	30	17	23	28
Missing	28	38	6	23	26
Region:					
Northern Jutland	35	30	25	20	28
Middle Jutland	26	36	27	22	30
Southern Denmark	29	30	25	20	27
Capital Area	32	34	26	18	29
Zealand (rest)	27	32	25	16	26
Total	30	33	26	19	28

Source: Own calculations based on "The National Health Interview Survey 2010" by the National Institute for Public Health, Denmark.

When conditioning on age, we still often see that higher education is associated with better health regarding all health indicators, except for basic education in the youngest age group, and for the best educated in the oldest age group. For the young group this is explained by the fact that many of these individuals have not yet reached their final education level. The pattern for the oldest group is harder to explain, but might partly be due to less reliability of the educational variable for older people.

The lowest health disparity is found between females and males (Table 2). Marital status and region show more health disparities. A factor with even more health disparities than education is ethnicity. This is mainly due to the higher level of good health among residents with a Danish or Western background compared to residents with a non-Western background. The largest difference is between Danes in the age group 25-44 years and non-Western immigrants 65+ years, where the former has a 3.6 times higher chance of good health.

Joint distribution

Although the 1-dimensional view provides some information, our focus is on the joint distribution which is required to test FOD. The joint distribution becomes quite intractable graphically when there are a large number of dimensions. With ten dimensions there are 1024 possible outcomes, while with four dimensions there are sixteen outcomes (2^4). The joint distribution is therefore exemplified with the 4-dimensional case (Table 3).

Table 3. Joint distribution of indicators by education. 25-44 years. 4-dimensional case. 2010

Combination type	Indicator combination i	ii	iii	iv	Educational attainment, % Basic	Vocat.	Short	Medium	Long
Worst→	0	0	0	0	9.12	2.54	1.89	2.20	0.96
Intermediate health combinations => not always internally rankable	0	0	0	1	1.44	1.64	1.00	0.60	1.03
	0	0	1	0	3.84	2.48	2.24	1.59	0.32
	0	0	1	1	1.83	0.94	0.00	0.41	0.95
	0	1	0	0	0.92	0.16	0.38	0.62	0.11
	0	1	0	1	0.22	0.14	0.39	0.28	0.23
	0	1	1	0	2.17	0.52	0.49	0.45	0.00
	0	1	1	1	0.38	0.45	0.40	0.32	0.74
	1	0	0	0	4.15	3.69	5.12	2.38	0.65
	1	0	0	1	4.05	3.37	6.03	5.17	4.76
	1	0	1	0	7.90	5.53	4.52	4.07	0.91
	1	0	1	1	5.55	5.54	4.08	4.85	6.38
	1	1	0	0	9.30	8.10	8.14	6.69	5.28
	1	1	0	1	6.79	11.62	13.58	18.26	20.67
	1	1	1	0	23.37	22.94	19.95	17.09	12.32
Best→	1	1	1	1	18.96	30.34	31.80	35.01	44.68
Total					100.00	100.00	100.00	100.00	100.00
Avg. no. good outcomes, equal					2.45	2.88	2.89	3.01	3.25
Avg. no. good outcomes, unequal					2.51	2.82	2.71	2.78	2.95

Note: In "Avg. no. good outcomes" in bottom of table, "equal" refers to equal weights across the four indicators (weight=1 each), while "unequal" refers to high weight to indicator iii (weight=2.5), and low weight to the remaining indicators i, ii, and iv (weight=0.5 each).

Source: Own calculations based on "The National Health Interview Survey 2010" by the National Institute for Public Health, Denmark.

From Table 3 we see that among people aged 25-44 years with medium higher education, 35% have good health with respect to all four dimensions, and 2.2% have bad outcome in all four indicators. The similar percentage among people with vocational training is 30% and 2.5%. Thus, the better educated group have (relatively) more people with the best outcome and fewer people with the worst outcome. Thus, regarding the two most extreme cases, we see medium educated perform better than people with vocational training. We even see that using a summary measure (bottom of table) taking all sixteen outcomes into account, medium education outperforms vocational training, when equal weights are assumed across the four dimensions. But if we change the weights and give much higher weight to dimension iii then people with vocational training have better multidimensional health than people with medium training. In other words, in this case the best performing of the two educational groups depends on the applied weighting scheme. This is because in the intermediate indicator combinations (Table 3) we cannot in many instances unambiguously classify one outcome as being better than the other. For instance, vocational has 22.9% in combination 1110 (good health regarding indicators i, ii and iii, and bad regarding iv), but that is not unambiguously better than 1101 or 1011 or 0111, each of which also has three good health outcomes and one bad outcome. Note that combination 1110 is not unambiguously better than 0001 (with only one good outcome) or 0011, 0101, or 1001 (with two good outcomes). For these intermediate cases the attractiveness of the combinations depends on the weighting applied to the dimensions. Thus, applying measures where the ranking of population subgroups depends on the weight distribution across dimensions involves the risk of rank reversal when changing the weights. This can be avoided by using the robust FOD method that does not depend on the weighting scheme. These FOD results are presented in the next section.

5. Results

FOD comparisons

Since we consider binary variables the FOD criteria in the 1-dimensional case simplifies to comparing the share of respondents having good health in all indicators. Hence, in every comparison each group either dominates another group, or is dominated by another group in the 1-dimensional case. Note that, as shown in the Lemma in Section 3, FOD in a higher dimension implies FOD in (all) lower dimension(s).

Figure 3. FOD test for four age groups and FOD table explanantion. 2010

	16-24	25-44	45-64	65+	Total
16-24			1 4	1 4	2 2 0 0
25-44	1		1 4	1 4	3 2 0 0
45-64				1 4	1 1 0 0
65+					0 0 0 0
Total	1 0 0 0	0 0 0 0	2 2 0 0	3 3 0 0	6 5 0 0

In the diagram, 1 (upper left in a cell) indicates that the row age group dominates the column age group in the 1-dimensional case. Likewise, 4 (upper right in a cell) indicates that the row age group FOD the column age group in the 4-dimensional case. Interpretations regarding 7 and 10 dimensions are similar. The total in the row indicates the number of times the age group in the column is dominated by others. The total in the column indicates the number of times the age group in a row dominates others.

Figure 3 illustrates the unconditional age FOD comparisons. Generally the younger age groups dominate the older age groups in the 1- and 4-dimensional case with the exception of the group 16-24, which does not dominate 25-44 but is instead dominated by this group in the 1-dimensional case. There are no dominances in the 7- and 10-dimensional cases. As we refine the dimensions it is shown that even though age is important when comparing health, it is not possible to show that the younger are unambiguously better off when we refine the dimensions and consider health as a more multidimensional phenomenon. Since age influences the distribution of health indicators, and educational levels are unevenly distributed across age groups (see Table 1), the analysis is made for separate age groups to minimize this effect.

FOD comparisons between educational levels for a given age group are shown in Figure 4. Generally, the higher educational levels dominate lower educational levels; however, the depth of dominance varies across educational groups and age.

For the 25-44 year age group, the groups with vocational and long education dominate the groups with basic education up to the 7-dimensional case. The group with long higher education additionally dominates the groups with short and medium higher education up to four dimensions. In the 1-dimensional case the picture is clear showing that the higher the educational level is the larger the share not having a single bad outcome in any health indicator will be; for example, any given educational level dominates each of the lower educational levels.

Figure 4. Educational FOD. Displayed separately for age groups.

<u>Age 16-24</u>

	Basic	Vocat	Short	Medium	Long	Total
Basic		1	1	1		3 0 0 0
Vocat						0 0 0 0
Short		1				1 0 0 0
Medium		1	1			2 0 0 0
Long	1	1	1	1		4 0 0 0
Total	1 0 0 0	4 0 0 0	3 0 0 0	2 0 0 0	0 0 0 0	10 0 0 0

<u>Age 25-44</u>

	Basic	Vocat	Short	Medium	Long	Total
Basic						0 0 0 0
Vocat	1 4 7					1 1 1 0
Short	1	1				2 0 0 0
Medium	1	1	1			3 0 0 0
Long	1 4 7	1	1 4	1 4		4 3 1 0
Total	4 2 2 0	3 0 0 0	2 1 0 0	1 1 0 0	0 0 0 0	10 4 2 0

<u>Age 45-64</u>

	Basic	Vocat	Short	Medium	Long	Total
Basic						0 0 0 0
Vocat	1 4 7 10					1 1 1 1
Short	1 4 7	1				2 1 1 0
Medium	1 4 7	1	1 4			3 2 1 0
Long	1 4 7	1 4	1 4	1 4		4 4 1 0
Total	4 4 4 1	3 1 0 0	2 2 0 0	1 1 0 0	0 0 0 0	10 8 4 1

<u>Age 65+</u>

	Basic	Vocat	Short	Medium	Long	Total
Basic						0 0 0 0
Vocat	1 4				1	2 1 0 0
Short	1 4	1			1	3 1 0 0
Medium	1 4	1 4	1		1	4 2 0 0
Long	1 4 7					1 1 1 0
Total	4 4 1 0	2 1 0 0	1 0 0 0	0 0 0 0	3 0 0 0	10 5 1 0

Note and source: See Figure 1.

For the 45-64 year age group, only the group with basic education is unambiguously worse off than any other group. The group with basic education is dominated by all other educational groups in the 7-dimensional case (and thus for the 1- and 4-dimensional case as well; cf. the Lemma in Section 3), and for the vocational level also in the 10-dimensional case. This indicates a very broad-based inequality between the group with only basic education and any of the higher educational groups. The group with long higher education is better off than vocational, short, and medium higher educational groups up to the 4-

dimensional case. Furthermore, the group with medium higher education dominates the group with short higher education in the 4-dimensional case as well.

In the 65+ age group, people with basic education are again worse off than any other (higher) educational group. The group with long higher education dominates the group with basic education in the 7-dimensional case, while for the groups with vocational, short, and medium higher education the domination of basic education is only present in the 4-dimensional case. The picture is somewhat surprising when we consider the group with long higher education. Although only in the 1-dimensional case, the highest educated are dominated by people with either vocational training, or short to medium higher education.

FOD tests for individuals aged 16-24 years are included for completeness, but it should be kept in mind that this age group has rarely passed higher education. The number of people having completed higher education is thus obviously very low. Since only 1-dimensional dominances occur, none of the groups are unambiguously better off in the multidimensional case.

Analyses similar to education have been made for gender, marital status, region, and ethnicity. We generally find few dominances in the 4-dimensional case, a single dominance in the 7-dimensional case, and none in the 10-dimensional case. See Appendix A for all results.

Bootstrapping

Since sample data are used, it is relevant to check if the observed FODs in Figure 4 are knife edge cases; i.e., if small changes in the data could have resulted in different results than those presented. Bootstrapping (see Efron 1979), with a hundred replications of data with re-sampling over groups is applied to investigate this issue further, and the results from the bootstrap analysis can be interpreted as a kind of sensitivity analysis. We only present the bootstrap results for education of the age group 25-44 years, see Table A2 in Appendix B where the static (actual data) is compared with the bootstrap results. In this case we see that observed FOD is associated with a bootstrap empirical probability of FOD between 67% and 97%. No observed FOD is usually associated with a zero bootstrap probability but is otherwise between 2% and 43%. The generally low empirical probability in case of no observed FOD and the generally high empirical probability in the case of observed FOD indicate that the observed FOD results are not based on a shaky underlying distribution of health outcomes.[9] For the other age groups we have somewhat similar results with correlations between 0.87 and 0.96 (results available from authors upon request).

Empirical summary

To summarize, we are often able to detect FOD in the 4-dimensional case, sometimes in the 7-dimensional case, but only once in the 10-dimensional case, which shows that it is difficult to conclude robustly about which group dominates another when more and more indicators are included, since with an increasing number of dimensions the prevalence of dominances decreases drastically (Figure A1 in the Appendix). This is in sharp contrast to traditional multidimensional methods that have no "problems" ranking groups regardless of the number of indicators. But, although these methodologies can always rank groups, as we have discussed it will not be a ranking which is robust to the weighting scheme. Thus, the pattern emerging here with only a few observed dominances in the 7+ dimensional cases reflects the fact that in an analysis of multidimensional health we can actually often not rank groups unambiguously when there are a lot of indicators – a lack of dominances should thus not be interpreted as a methodological flaw in the FOD technique. Also, the results here are of course dependent on the data we have. The data are from Denmark, where the socio-economic differences are relatively small. Application of this methodology to other countries with (much) higher disparities may produce other and more frequent group dominances in health.

Regarding the other background variables (gender, ethnicity, marital status, and region) we see the same pattern as for education with respect to number of dominances when the number of dimensions increases, but the decrease in the number of dominances is much higher than for education. Gender has the lowest number of dominances, but the number is also small for region. It is also not high for ethnicity although there was an indication in the descriptive part that many dominances could be expected given that this variable had the highest disparity in the 1-dimensional case.

6. Discussion and conclusion

To overcome the lack of rank robustness when applying traditional methods of analyzing multidimensional health indicators, we have applied a new multidimensional FOD methodology to health indicators. In the FOD approach, when a group dominates another

[9] The correlation between the actual and bootstrap data FOD is 0.89.

group then that domination is robust to any weighting of included indicators. Although FOD is based on a demanding set of conditions, we are nevertheless able to detect FOD in many instances when applying the method to multidimensional health disparity across educational and demographic population subgroups. We have presented a refinement strategy starting out with one single health aggregate summarizing the health status of individuals based on a set of 22 basic indicators describing important facets of health. The single aggregate health dimension is successively refined to 4, 7, and 10 health dimensions. To our knowledge, this is the first empirical illustration of truly multidimensional population health comparisons involving as many as 10 dimensions that do not rely on ad-hoc weighting or counting procedures. Of course, the number of dominances drops when the health dimensions become more refined. But the speed with which the number of dominances changes may be different for other indicators, other countries, or other analysis variables. There are thus many possible new paths to pursue in empirical FOD group (health) comparisons.

Another methodological path to pursue is the question about a complete ranking of all included population subgroups. In principle, the present FOD methodology could be used to produce a complete ranking of groups, but in practice that is unlikely since the methodology is designed for the comparison of two groups. Complete ranking procedures based on FOD information (with bootstrapping) are possible and have been applied (Arndt et al. 2013); however, as any other complete ranking procedure with multidimensional outcomes, one should bear in mind that it is based on stronger and more controversial assumptions and procedures.

Acknowledgements

We thank Mette Bjerrum Koch for providing the data and helpful comments.

Literature

Alkire S and Foster J (2011). Counting and multidimensional poverty measurement. Journal of Public Economics, 95, 476-487.

Arndt C, Distante R, Hussain MA, Østerdal LP, Huong PL, and Ibraimo M (2012). Ordinal welfare comparisons with multiple discrete indicators: a first order dominance approach and application to child Poverty. World Development, 40 (11), 2290-2301.

Arndt C, Hussain MA, Salvucci V, Tarp F, and Østerdal LP (2013). Advancing small area estimation. UNU-WIDER Working Paper No. 2013/053. United Nations University World Institute for Development Economics Research.

Atkinson A and Bourguignon F (1982). The comparison of multi-dimensioned distributions of economic status. The Review of Economic Studies, 49(2), 183-201.

Batana Y and Duclos J-Y (2010). Comparing multidimensional poverty with qualitative indicators of well-being. CIRPEE Working Paper 10-04.

Bourguignon F and Chakravarty S (2003). The measurement of multidimensional poverty. Journal of Economic Inequality, 1(1), 25-49.

Christensen AI, Ekholm O, Davidsen M, and Juel K (2012). Sundhed og sygelighed i Danmark 2010 – og udviklingen siden 1987 (Health and Morbidity in Denmark 2010 – and the Development since 1987). National Institute of Public Health, University of Southern Denmark.

Cowell FA and Victoria-Feser M-P (2007). Robust stochastic dominance: a semi-parametric approach. Journal of Economic Inequality, 5(1), 21-37.

Dalstra JAA, Kunst AE, and Mackenbach JP (2006). A comparative appraisal of the relationship of education, income and housing tenure with less than good health among the elderly in Europe. Social Science and Medicine, 62 (8), 2046-2060.

Drummond MF, Sculpher MJ, Torrance GW, O'Brien BJ, and Stoddart GL (2005). Methods for the Economic Evaluation of Health Care Programmes. OUP Oxford, 3rd edition.

Duclos J-Y and Echevin D (2011). Health and income: a robust comparison of Canada and the US. Journal of Health Economics, 30(2), 293-302.

Duclos J-Y and Makdissi P (2005). Sequential stochastic dominance and the robustness of poverty orderings. Review of Income and Wealth, 51(1), 63-87.

Duclos J-Y, Sahn D, and Younger S (2007). Robust multidimensional poverty comparisons with discrete indicators of well-being. In: Jenkins SP and Micklewright J (ed.), Inequality and poverty reexamined, Oxford, UK: Oxford University Press.

Duclos J-Y, Sahn D, and Younger S (2011). Partial multidimensional inequality orderings. Journal of Public Economics, 95, 225-238.

Dyckerhoff R and Mosler K (1997). Orthant orderings of discrete random vectors. Journal of Statistical Planning and Inference, 62, 193-205.

Efron B (1979). Bootstrap methods: Another look at the jackknife. Annals of Statistics, 7(1), 1-26.

Feeny D, Furlong W, Torrance GW, Goldsmith CH, Zhu Z, DePauw S, Denton M, and Boyle M (2002). Multiattribute and single-attribute utility functions for the health utilities index mark 3 system. Medical Care, 40(2), 113-28.

Gravel N, Moyes P, and Tarroux B (2009). Robust international comparisons of distributions of disposable income and regional public goods. Economica, 76, 432-461.

Gravel N and Mukhopadhyay A (2010). Is India better off today than 15 years ago? A robust multidimensional answer. Journal of Economic Inequality, 8, 173-195.

HCP (2012). Euro Health Consumer Index 2012. Report. Health Consumer Powerhouse AB. Sweden.

Juel K, Sørensen J, and Brønnum-Hansen H (2006). Risikofaktorer og folkesundhed i Danmark (Risk Factors and Public Health in Denmark). National Institute of Public Health, University of Southern Denmark.

Jørgensen MM, Koch MB, and Østerdal LP (2013). Ulighed i selvvurderet helbred i Danmark (Inequality in self-assessed health in Denmark). Nationaløkonomisk Tidsskrift, 151, 55-88.

Koch MB, Davidsen M, and Juel K (2012). Social ulighed i sundhed, sygelighed og trivsel 2010 og udviklingen siden 1987 (Social Health Inequality, Morbidity and Well-Being 2010 and the Development since 1987). National Institute of Public Health, University of Southern Denmark.

Kohn JL (2012). What is Health? A Multiple Correspondence Health Index. Eastern Economic Journal, 38, 223-250.

Kunst AE, Bos V, Lahelma E, Bartley M, Lissau I, Regidor E, Mielck A, Cardano M, Dalstra JAA, Geurts JJM, Helmert U, Lennartsson C, Ramm J, Spadea T, Stronegger WJ, and Mackenbach JP (2005). Trends in socioeconomic inequalities in self-assessed health in 10 European countries. International Journal of Epidemiology, 34(2), 295-305.

Lehmann E (1955). Ordered families of distributions. The Annals of Mathematical Statistics, 26, 399–419.

Levhari D, Paroush J, and Peleg B (1975). Efficiency analysis for multivariate distributions. Review of Economic Studies, 42, 87–91.

Mosler K and Scarsini M (1991). Some theory of stochastic dominance. In: Mosler K and Scarsini M (eds.), Stochastic Orders and Decision under Risk. Institute of Mathematical Statistics, Heyward, California.

O'Donnell O, van Doorslaer E, Wagstaff A, and Lindelow M (2008). Analyzing Health Equity Using Household Survey Data: A Guide to Techniques and their Implementation. The World Bank, Washington DC.

Range TM and Østerdal LP (2013). Checking bivariate first order dominance. Discussion Papers on Business and Economics, SDU, No. 9/2013.

Sen AK (1999). Development as Freedom. Oxford University Press, Oxford.

Sen AK (2009). The Idea of Justice. Allen Lane, London.

Shaked M and Shanthikumar J (2007). Stochastic Orders. Springer.

Townsend P (1979). Poverty in the United Kingdom. A Survey of Household Resources and Standards of Living. Penguin Books, Middlesex.

UHF (2012). America's Health Rankings® - 2012 Edition. United Health Foundation, USA.

Wagstaff A and van Doorslaer E (2000). Measuring and testing for inequity in the delivery of health care. Journal of Human Resources, 35(4), 716-733.

WHO (2013). World Health Statistics 2013.World Health Organization. Geneva.

Østerdal LP (2010). The mass transfer approach to multivariate discrete first order stochastic dominance: direct proof and implications. Journal of Mathematical Economics, 46, 1222-1228.

Appendix

A. FOD for ethnicity, marital status, and region

Table A1. FOD for ethnicity, marital status, and region in the 4-dimensional case. Denmark.

Characteristic		Dominating	Dominated	Age group
Ethnicity		Western	Non-Western	16-24
		Danish	Non-Western	45-64
Marital status		Unmarried	Widow(er)	16-24
		Widow(er)	Couple	45-64
		Widow(er)	Unmarried	"
		Divorced	Couple	"
		Divorced	Unmarried	"
		Couple	Unmarried	"
		Widow(er)	Couple	65+
	7-dim →	Widow(er)	Unmarried	45-64
Region		Capital Area	Middle Jutland	16-24
		Northern Jutland	Capital Area	65+
		Northern Jutland	Zealand (rest)	"

Note: Gender is excluded since it has no FOD.

Source: Own calculations based on "The National Health Interview Survey 2010" by the National Institute for Public Health, Denmark.

We present the 4-dimensional FOD (and the single 7-dimensional FOD also occurring) for ethnicity, marital status, and region in Table A1. Gender is not included in the table since that does not have any FOD beyond the 1-dimensional case.[10] From Table A1 we see that non-Western immigrants are dominated by both Danes (16-24 years) and other Westerners (45-64 years), but not in all age groups. Thus, we cannot generally conclude that Danes have better health than immigrants when we take a refined multidimensional view, but we can say that when there is domination, it only goes in one direction. Note that although ethnicity had the largest prevalence disparity (in Table 2), it is one of the characteristics with fewest dominances. Most dominances, next to education, are found for marital status. Particularly for the 45-64 year age group a lot of dominances are found. Generally, we see that widow(er)s tend to have better multi-dimensional health than many of the other marital status groups. We also generally observe that unmarried people are more often dominated by other groups. The

[10] The 1-dimension case FOD are easily extractable from Table 2 since it is merely a matter of ranking a uni-variate series of prevalence.

only 7-dimensional FOD is found for the marital status characteristic; widow(er)s dominate unmarried people for the 45-64 year age group. The results for regions are mixed, but we can say that Northern Jutland is the frequent one to dominate others (the Capital Area and the rest of Zealand for the 65+ age group). Middle Jutland together with the rest of Zealand are the only regions who are only dominated and do not dominate others. Southern Denmark is the only region which is neither dominated or dominates other regions.

B. Bootstrap

Table A2. Static and bootstrap FOD in the 4-dimensional case. Displayed separately for educational levels. 25-44 years. Denmark. 2010

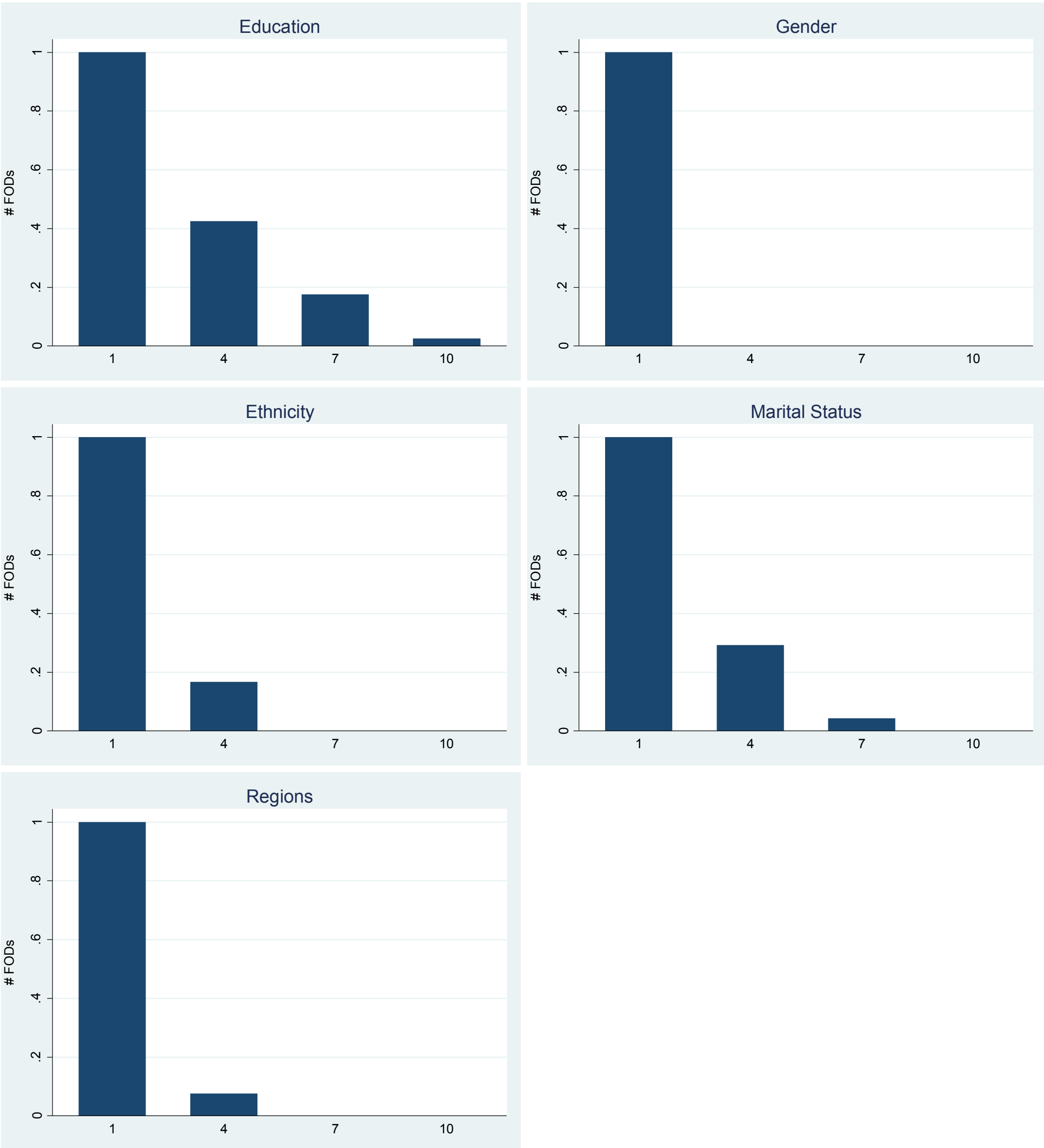
	Basic		Vocational		Short higher		Medium higher	
	Static	Bootst.	Static	Bootst.	Static	Bootst.	Static	Bootst.
Basic								
Vocational	1	0.97						
Short higher		0.43						
Medium higher		0.43		0.02		0.08		
Long higher	1	0.80		0.20	1	0.67	1	0.68

Note: Under 'static' any FOD in actual data is presented, e.g. '1' appears in the same places as 4 appears in the upper right diagram of Figure 4. The empirical probability of FOD using k =100 bootstrap replications is presented under the heading 'Bootst.'.

Source: Own calculations based on "The National Health Interview Survey 2010" by the National Institute for Public Health, Denmark.

C. Total number of dominances

Figure A1. Number of dominances for different characteristics given age group. Fraction of the maximum number of dominances that always exists in the 1-dimensional case.



Note: The maximum possible number of dominances is $4k(k\text{-}1)/2$, where k is the number of categories in the characteristics variable (displayed in Table 3), and the 4-factor originates from the number of age groups.

Source: Own calculations based on "The National Health Interview Survey 2010", the National Institute for Public Health, Denmark.