Elsevier Editorial System(tm) for European Journal of Operational Research
Manuscript Draft

Manuscript Number: EJOR-D-17-01749

Title: A Benefit-of-the-Doubt Model with Reverse Indicators

Article Type: SI: Advances in DEA

Section/Category: (D) Data envelopment analysis

Keywords: Data envelopment analysis; benefit of the doubt; diet problem; health index

Corresponding Author: Professor Dimitris Margaritis, PhD

Corresponding Author's Institution: University of Auckland

First Author: Rolf Fare

Order of Authors: Rolf Fare; Giannis Karagiannis; Maryam Hassanasab; Dimitris Margaritis, PhD

Abstract: Recognizing the popularity of healthier food options, we reformulate the diet problem as a linear optimization program with desirable and undesirable food nutrients. We then show how the dual formulation of this diet problem is equivalent to a new Benefit-of-the-Doubt (BoD) model with forward and reverse indicators and with a wide range of applications in the construction of composite indicators. As an illustration, we use the BoD model to construct a composite index of public health for 180 countries.

Dear Editors

We would like you to consider our paper for inclusion in the special EJOR issue.

Thank you.

D Margaritis on behalf of co-authors

Highlights (for review)

We extend the diet problem to include desirable and undesirable food nutrients

The dual of diet problem is equivalent to BoD model with forward and reverse indicators

We show how the BoD model relates to the Lewis and Sexton (2004) model

We use the new BoD model formulation to construct public health index for 180 countries

A Benefit-of-the-Doubt Model with Reverse Indicators

Rolf Färe[a], Giannis Karagiannis[b], Maryan Hassanasab[c], Dimitris Margaritis[c]
[a]Oregon State University, [b]University of Macedonia, [c]University of Auckland

Abstract: Recognizing the popularity of healthier food options, we reformulate the diet problem as a linear optimization program with desirable and undesirable food nutrients. We then show how the dual formulation of this diet problem is equivalent to a new Benefit-of-the-Doubt (BoD) model with forward and reverse indicators and with a wide range of applications in the construction of composite indicators. As an illustration, we use the BoD model to construct a composite index of public health for 180 countries.

Keywords: Data envelopment analysis; benefit of the doubt; diet problem; health index

1. Introduction

Recently, Färe and Karagiannis (2014) examined the relation between the diet problem, one of the first linear optimization problems (see Stigler, 1945), and the benefit-of-the-doubt (BoD) model (see Cherchye *et al.*, 2007), one of the currently widely employed data envelopment analysis (DEA) models for constructing composite indicators.[1] It was shown by Färe and Karagiannis that the diet problem and the BoD model are linear programming (LP) duals (i.e., the primal formulation of the diet problem is equivalent to the dual formulation of the BoD model and *vice versa*) as long as food prices are equal to one. Given this specification of the BoD model, it implies that the diet problem and the radial input-oriented DEA model with a single constant input are linear programming duals. In addition, the diet problem and the inverted BoD model (see Caporaletti *et al.*, 1999; Zhou *et al.*, 2007) are linear programming equivalent. The primal formulation of the diet problem is equivalent to the primal formulation of the inverted BoD model and the same holds true for their dual formulations as long as nutritional requirements are set equal to one. Given this specification of the inverted BoD model, it implies that the diet problem and the radial output-oriented DEA model with a single constant output are linear programming equivalent.

In this paper, we explore further the relationship between the BoD model and an extended formulation of the diet problem considering not only desirable nutrients, such as calories, proteins, vitamins and minerals, but also undesirable ones, such as saturated fats (see e.g. the discussion in Lancaster, 1992, Garille and Gass, 2001). From this, we derive a novel BoD model that can incorporate reverse indicators, namely indicators that are not isotonic and their increasing values are considered as unfortunate events. A large number of indicators fall in this category including:

- income inequality and unemployment rates in assessing economic performance and sustainability;
- child mortality and teen fertility rates in evaluating child well-being;
- homicide rates and road fatalities in gauging efficiency of safety and security;
- the infant mortality rate, share of population with non-communicable diseases, and years lost to diseases in constructing a Health Status Index (see e.g., Larson, 1994; Klomp and deHaan, 2010; Tikunov and Chereshnya, 2016);
- air pollution, crime rate, traffic accidents and solid waste for assessing livability indices for cities or countries (see e.g. Hashimito and Kodama, 1997; Zanella *et al*., 2015b);
- air and acoustic pollution, commuting time, and unemployment in constructing a Quality of Life Index (see e.g. Gonzalez *et al*., 2011); and
- failure rates and average time to repair in assessing hydropower plants' overall quality of services (see Zanella *et al*., 2015a).

Previous attempts to deal with reverse indicators include the use of the inverse value (e.g., Hashimoto and Kodama, 1997), normalization with the sample minimum value (e.g., Zhou *et al*., 2007; Wu *et al*., 2011), rescaling or ranging normalizations inspired from the MCDM literature where each reverse indicator is subtracted from its sample maximum and then is divided by the difference between its sample maximum and minimum values (e.g., Reig-Martinez *et al*., 2011), and more recently by using the directional BoD model (Fusco, 2015; Vidoli *et al*., 2015; Zanella *et al*., 2015a,b; Charles *et al*., 2016). Such transformation attempts, even though simple, are not attractive since the BoD model is not translation invariant in the output variables (i.e., indicators) as for example a DEA model with constant returns to scale.

Consequently, transformation of reverse indicators will affect the estimated value of composite indicators. On the other hand, the directional BoD model treats reverse indicators as undesirable outputs by means of weak disposability. Implicit in this modeling choice is the assumption of null-jointness, namely that desirable outputs cannot be produced without the production of undesirable outputs, which is a rather reasonable assumption for conventional production processes albeit less justifiable in the context of the BoD model. In other words, as a Koopman's "helmsman" who has at his disposal a unitary quantity of an aggregate input attempts to steer all of forward indicators toward their maximum levels.

Our proposed BoD model treats non-isotonic indicators as reverse rather than undesirable outputs. The main difference is that reverse outputs might not be accompanied by desirable outputs. That is, the presence of forward indicators does not imply nor is implied by the presence of reverse indicators. Thus instead of the common case having only forward indicators one can also treat with the proposed model cases where there are only reverse indicators. More interestingly, the proposed BoD model is the single-constant-input version of Lewis and Sexton (2004) input-oriented, constant-returns-to-scale DEA model with forward inputs and forward and reverse outputs as the conventional BoD model is the single-constant-input version of Charnes *et al*. (1978) input-oriented DEA model.

2. The Proposed Model

The extended diet problem, with both desirable and undesirable nutrients, may be formulated in term of the following Tableau:

1	2	……………	…………	K	
y_{11}	y_{21}	…………………	…………………	y_{K1}	$\geq y_{k'1}$
:	:			:	:
:	:			:	:
y_{1m}	y_{2m}	……………	…………………	y_{Km}	$\geq y_{k'm}$
y_{1m+1}	y_{2m+1}	………………	…………………	y_{Km+1}	$\leq y_{k'm+1}$
:	:			:	:
:	:			:	:
y_{1J}	y_{2J}	…………………	………………	y_{KJ}	$\leq y_{k'J}$
p_1	p_2	…………………	………………	p_K	

where *p* refers to food prices, *y* to the amount of (desirable and undesirable) nutrients, $y_{k'}$ to (desirable and undesirable) nutritional standards, and there are *K* foods, *(1,..,m)* desirable nutrients, and *(m+1,…,J)* undesirable nutrients. The main difference with the conventional formulation of the diet problem is that here we assume that foods may contain both desirable and undesirable nutrients, where for the former there are low limits (desirable nutritional standard) and for the latter upper limits (undesirable nutritional standard). Similar to the conventional formulation of the diet problem it is implicitly assumed that there are no interactions between foods and between nutrients (Garille and Gass, 2001), regardless of whether nutrients are considered as desirable or undesirable. Consequently, the quantity of a nutrient consumed by eating a specified amount of a certain food is exactly the quantity of that nutrient (desirable or undesirable) that will be used by the human body. This assumption allows us to write the above diet problem in a linear programming format.

In a linear programming format, the revised diet problem, with z denoting food quantities, is given as:

$$
\begin{aligned}
&\min_{z_k} \sum_{k=1}^{K} p_k z_k \\
&st \sum_{k=1}^{K} z_k y_{kj} \geq y_{k'j} \qquad \forall\, j = 1, \dots, m \\
&\sum_{k=1}^{K} z_k y_{kj} \leq y_{k'j} \qquad \forall\, j = m+1, \dots, J \\
&z_k \geq 0 \qquad\qquad\qquad \forall\, k = 1, \dots, K
\end{aligned}
\tag{1}
$$

which equivalently may be written as:

$$
\begin{aligned}
&\min_{z_k} \sum_{k=1}^{K} p_k z_k \\
&st \sum_{k=1}^{K} z_k y_{kj} \geq y_{k'j} \qquad \forall\, m = 1, \dots, m \\
&-\sum_{k=1}^{K} z_k y_{kj} \geq -y_{k'j} \qquad \forall\, j = m+1, \dots, J \\
&z_k \geq 0 \qquad\qquad\qquad \forall\, k = 1, \dots, K
\end{aligned}
\tag{1$'$}
$$

Then its dual is given as:

$$\max_{\lambda_j} \sum_{j=1}^{m} \lambda_j y_{k\prime j} - \sum_{j=m+1}^{J} \lambda_j y_{k\prime j}$$

$$st \sum_{m=1}^{m} \lambda_j y_{k\prime j} - \sum_{j=m+1}^{J} \lambda_j y_{k' j} \leq p_k \quad \forall\, k = 1, \dots, K \tag{2}$$

$$\lambda_j \geq 0 \qquad \forall\, j = 1, \dots, J$$

where λ refers to the (shadow) prices of nutrients. An intuitive interpretation of the dual problem, given by Berg and Ehtano (2010) for the conventional diet problem, is that it refers to a firm that instead of producing foods is manufacturing nutrient pills. The difference though with the conventional model where these pills-producing firms maximize revenue is that here they maximize profit as they produce pills with both desirable and undesirable nutrients and the latter are not freely disposable. Then the problem of the firm is to choose the unit prices of the pills that maximize its profit, given by the difference between revenue from pills with desirable nutrients and cost from pills with undesirable nutrients, where the nutrient standards can be interpreted as the demand for (desirable and undesirable) nutrients. The constraints of the dual problem indicate that pills should be competitive against the real foods in the sense that the price of artificial foods made out of the pills should after accounting for the costly disposability of undesirable nutrients be less than or equal to the price of the relevant food. In this context, the shadow price λ_j measures by how much the optimal diet cost increases when a component in the vector of nutrient standard increases.

On the other hand, as in Färe and Karagiannis (2014), one can verify that as long as $p_k = 1$ for all k, the above dual formulation of the revised diet problem is equivalent to a modified BoD model containing both forward (i.e., capturing positive aspect) and reverse (i.e., capturing negative aspect) indicators, where the λ's are now interpreted as aggregation weights, the $y_{k\prime j}$'s (*j=1,…,m*) as forward indicators, the $y_{k\prime j}$'s (*j=m+1,…,J*) as reverse indicators, and there are *K* evaluated units. That is,

$$\max_{\lambda_j} \sum_{j=1}^{m} \lambda_j y_{k\prime j} - \sum_{j=m+1}^{J} \lambda_j y_{k\prime j}$$

$$st \sum_{j=1}^{m} \lambda_j y_{k'j} - \sum_{j=m+1}^{J} \lambda_j y_{k'j} \leq 1^k \quad \forall\, k = 1, \dots, K \tag{3}$$

$$\lambda_j \geq 0 \qquad \forall\, j = 1, \dots, J$$

The dual of problem (3), which can also be obtained by setting $p_k = 1$ (k= 1, 2, …, K) in (1), is given as:

$$\min_{z_k} \sum_{k=1}^{K} z_k$$

$$st \sum_{k=1}^{K} z_k y_{kj} \geq y_{k'j} \quad \forall\, j = 1, \dots, m \tag{4}$$

$$\sum_{k=1}^{K} z_k y_{kj} \leq y_{k'j} \quad \forall\, j = m+1, \dots, J$$

$$z_k \geq 0 \qquad \forall\, k = 1, \dots, K$$

The main difference between the proposed formulation and the conventional BoD model is that in (3) we seek to maximize the weighted average of both forward and reverse indicators, with the latter being subtracted from the former. This difference is then reflected in the second inequality constraints in (4), which are absent from the dual formulation of the conventional BoD model. These constraints have the reverse inequality sign to reflect that increasing values are considered as unfortunate events. Lastly, one can verify that (3) or (4) reduce to the conventional formulation of the BoD model when there are no reverse indicators.

The above formulation could have resulted by assuming a single input with a unitary value for all the evaluated units as in Lewis and Sexton (2004) input-oriented model with forward inputs, forward and reverse outputs, and constant returns to scale, which is the same as model ‘A’ in Korhonen and Luptacik (2004). To verify this consider its dual formulation:

$$\min_{z_k} E_{k'}$$

$$st \sum_{k=1}^{K} z_k x_{ki} \leq E_{k'} x_{k'i} \quad \forall\, i = 1, \dots, I$$

$$\sum_{k=1}^{K} z_k y_{kj} \geq y_{k'j} \qquad \forall\, j = 1, \dots, m \tag{5}$$

$$\sum_{k=1}^{K} z_k y_{kj} \leq y_{k'j} \qquad \forall\, j = m+1, \dots, J$$

$$z_k \geq 0 \qquad \forall\, k = 1, \dots, K$$

where x refers to input quantities and E to the Farrell input-oriented technical efficiency score. Then by considering a single input (i.e., $i=1$) and setting its value equal to one (i.e., $x_{k1} = 1$) for all evaluated units the first set of inequality constraints in (5) is reduced to $\sum_{k=1}^{K} z_k \leq E_{k'}$. Since $\sum_{k=1}^{K} z_k$ is less than $E_{k'}$ and we seek to minimize $E_{k'}$ we may substitute $\sum_{k=1}^{K} z_k$ instead of $E_{k'}$ in the objective function of (5). Then (5) is reduced to (4). This is similar to deriving the conventional BoD model from the input-oriented, constant returns to scale DEA model of Charnes *et al.* (1978) (see Karagiannis, 2017).

3. Application

As an illustration, we consider the problem of constructing a public health index. Similar to Tikunov and Chereshnya (2016) we opt for an index which integrates some of the most objective indicators of public health such as infant mortality rate (the reverse indicator in our model), life expectancy for both men and women at birth and immunization coverage among 1-year-olds (the forward indicators). As noted by Tikunov and Chereshnya (2016) such indicators have several important advantages, namely they are readily available for almost all countries, they do not require expert assessment, and they are quite reliable. Our data source is the World Health Organization (WHO). We use the indicator values in (4) above to construct country-specific indexes of public health (PHI) for 180 countries as well as averages for four groups of countries using UN/IMF economic development indicators. They include 34 developed countries (DC2), 58 emerging economies (DC1), 45 less developed countries (LDC2), and 43 least developed countries (LDC1). We also look at changes over time calculating the index for 4 periods, 2001, 2005, 2009 and 2014. We assess health status with reference to both, a global frontier and two sub-frontiers one for developed and emerging economies combined and one for the 88 less developed economies.[2]

Descriptive statistics of the two public health forward indicators and one reverse health indicator as well as statistics of our public health index are given in Table 1. They show a clear pattern of strong association between health status and the level of economic development. The differences across country groups are statistically significant. More specifically, the differences between LDC1 and LDC2, LDC2 and DC1, DC1 and DC2 are statistically significant at the 5% level in all but three cases (LDC2 and DC1 in 2001 and 2005, LDC1 and LDC2 in 2009) where they are significant at the 10% level.

Table 2 presents the list of the 10 countries with the lowest values of the public health index across the four periods. As expected, the group of the least developed countries dominates the list. The notable exception is South Africa in 2009 and 2014.

We turn next to address the question of convergence, in particular whether public health programs, typically those of the LDCs with international aid to Government, are helping to close the gaps. The results of Table 3 show that there is statistical evidence in support of the convergence hypothesis, demonstrated by the negative and significant values of the lagged PHI coefficients. There is no evidence that group effects are significant and hence they have been omitted from the regressions.

The last question we assess empirically involves the relation between the health index and the percentage of public money that goes into the health sector. The results of Table 4 indicate that there is indeed a positive and significant relationship. More specifically, we find that on average one percentage point increase in the public health budget is associated with 0.25 points increase in the public health index, after we control for group effects and period fixed effects.[3]

4. Conclusion

Recognizing the increasing focus on healthier lifestyles, we have proceeded to update the diet problem, one of the earliest LP problems, with the inclusion of both desirable and undesirable food nutrients in the specification of the model. Following, Färe and Karagiannis (2014) we have then shown how to relate the diet problem to its linear programming dual, the BoD model. This results into a novel BoD model that can incorporate reverse indicators, namely indicators that are not isotonic and their increasing values are considered as unfortunate events. As an application, we have shown how to use the new BoD model formulation to construct a public health index. We are not suggesting that we have come up with a proper public health index as this is beyond the scope of this study. What we offer instead is a new way to think about a very important in terms of its policy implications albeit highly controversial area of research. Our approach has far wider applications in the area of composite indicator construction, with obvious relevance to among others, the field of nutrition.

Table 1: Descriptive Statistics

Year			2001				2005				2009				2014		
								Life expectancy at birth (years)									
GROUP	Obs.	Mean	Max	Min.	Std. Dev.	Mean	Max	Min.	Std. Dev.	Mean	Max	Min.	Std. Dev.	Mean	Max	Min.	Std. Dev.
LDC1	43	54.8	69.1	40.1	6.7	56.9	69.9	43.3	6.5	59.5	70.8	47.1	5.8	61.7	71.7	48.1	5.5
LDC2	45	65.4	76.7	46.7	8.6	66.3	77.2	46.0	8.4	68.1	78.1	51.0	7.2	69.7	79.0	52.8	6.5
DC1	58	70.7	77.5	45.3	5.3	71.3	78.6	44.6	5.7	72.4	79.3	50.0	5.0	73.6	80.7	59.2	4.3
DC2	34	77.4	81.5	69.9	2.9	78.4	82.0	70.6	2.9	79.5	83.0	72.2	2.7	80.6	83.5	73.4	2.5
All	180	66.8	81.5	40.1	10.0	67.9	82.0	43.3	9.7	69.6	83.0	47.1	8.7	71.1	83.5	48.1	8.1
								Immunization coverage among 1-year-olds (%)									
GROUP	Obs.	Mean	Max	Min.	Std. Dev.	Mean	Max	Min.	Std. Dev.	Mean	Max	Min.	Std. Dev.	Mean	Max	Min.	Std. Dev.
LDC1	43	62.6	92.0	26.0	19.5	72.9	97.0	25.0	17.4	78.7	98.0	24.0	15.7	79.5	99.0	46.0	14.6
LDC2	45	81.1	99.0	27.0	18.5	84.9	99.0	36.0	16.3	88.5	99.0	43.0	12.8	87.4	99.0	20.0	15.5
DC1	58	91.4	99.0	59.0	9.3	92.0	99.0	65.0	9.4	91.9	99.0	70.0	8.4	90.5	99.0	23.0	13.3
DC2	34	94.0	99.0	84.0	3.9	94.7	99.0	86.0	3.2	94.8	99.0	73.0	5.1	95.9	99.0	90.0	2.6
All	180	82.4	99.0	26.0	18.7	86.1	99.0	25.0	15.3	88.5	99.0	24.0	12.7	88.1	99.0	20.0	14.1
								Neonatal mortality rate (per 1000 live births)									
GROUP	Obs.	Mean	Max	Min.	Std. Dev.	Mean	Max	Min.	Std. Dev.	Mean	Max	Min.	Std. Dev.	Mean	Max	Min.	Std. Dev.
LDC1	43	38.1	58.2	12.0	9.3	34.5	56.1	12.0	9.0	31.2	53.4	12.1	8.7	27.6	49.6	11.8	8.3
LDC2	45	20.6	49.3	4.1	11.0	18.8	46.1	3.4	10.3	17.2	42.6	2.8	9.5	15.2	38.6	2.4	8.6
DC1	58	14.9	59.0	3.1	10.1	12.8	53.4	2.4	9.1	11.2	50.6	1.5	8.6	9.4	46.6	1.1	7.8
DC2	34	4.1	19.0	1.5	3.1	3.5	14.5	1.3	2.3	3.0	10.8	1.2	1.8	2.5	7.5	1.0	1.3
All	180	19.8	59.0	1.5	14.8	17.7	56.1	1.3	13.6	15.9	53.4	1.2	12.6	13.9	49.6	1.0	11.4
								Government expenditure on health as % of total govt exp.									
GROUP	Obs.	Mean	Max	Min.	Std. Dev.	Mean	Max	Min.	Std. Dev.	Mean	Max	Min.	Std. Dev.	Mean	Max	Min.	Std. Dev.
LDC1	43	9.9	34.4	1.4	5.2	11.0	23.9	1.9	4.8	10.6	21.1	1.5	4.5	9.5	17.9	2.4	3.9
LDC2	45	10.4	18.0	4.2	3.4	11.0	28.6	5.1	4.7	10.9	20.0	4.6	3.7	11.4	24.0	4.3	4.5
DC1	58	9.8	21.8	1.7	3.9	10.4	23.6	3.2	4.2	10.6	30.6	4.1	4.9	10.9	26.6	3.9	4.6
DC2	34	13.1	17.7	4.7	3.2	14.0	18.5	6.6	3.2	14.8	22.5	7.2	3.4	15.5	23.4	7.6	3.8
All	180	10.6	34.4	1.4	4.2	11.4	28.6	1.9	4.5	11.5	30.6	1.5	4.5	11.6	26.6	2.4	4.7
								Public Health Index									
GROUP	Obs.	Mean	Max	Min.	Std. Dev.	Mean	Max	Min.	Std. Dev.	Mean	Max	Min.	Std. Dev.	Mean	Max	Min.	Std. Dev.
LDC1	43	0.73	0.93	0.49	0.11	0.79	0.98	0.56	0.12	0.83	0.99	0.59	0.11	0.84	1.00	0.61	0.10
LDC2	45	0.87	1.00	0.58	0.11	0.89	1.00	0.60	0.10	0.91	1.00	0.64	0.09	0.92	1.00	0.64	0.09
DC1	58	0.94	1.00	0.72	0.07	0.94	1.00	0.69	0.07	0.94	1.00	0.71	0.07	0.94	1.00	0.74	0.06
DC2	34	0.97	1.00	0.89	0.02	0.98	1.00	0.91	0.02	0.98	1.00	0.95	0.02	0.98	1.00	0.93	0.02
All	180	0.88	1.00	0.49	0.12	0.90	1.00	0.56	0.11	0.92	1.00	0.59	0.10	0.92	1.00	0.61	0.09

Notes: LCD1 denotes the group of least developed countries, LCD2 is the group of less developed countries, DC1 denotes emerging economies and DC2 the group of developed economies.

Table 2: Public Health Index Bottom 10 Countries

Country	Year		Year	Country	Year	Country	Year
	2001		2005		2009		2014
Sierra Leone	0.49	Central African Republic	0.56	Central African Republic	0.59	Central African Republic	0.61
Central African Republic	0.56	Angola	0.58	Angola	0.61	Chad	0.63
Angola	0.56	Chad	0.59	Chad	0.61	Nigeria	0.64
Nigeria	0.58	Nigeria	0.60	Nigeria	0.64	Angola	0.65
Uganda	0.59	Uganda	0.65	Equatorial Guinea	0.67	Equatorial Guinea	0.69
Chad	0.59	Niger	0.65	Guinea	0.69	Guinea	0.70
Mali	0.62	Sierra Leone	0.66	Afghanistan	0.71	Liberia	0.70
Niger	0.62	Democratic Republic of Congo	0.66	South Africa	0.71	Niger	0.74
Burkina Faso	0.63	Equatorial Guinea	0.66	Niger	0.72	South Africa	0.74
Liberia	0.63	Guinea	0.67	Democratic Republic of Congo	0.73	Papua New Guinea	0.75

Table 3: Convergence in Public Health Performance

Dependent Variable PHI-PHI(-3)					Dependent Variable PHI-PHI(-1)				
Variable	Coefficien	Std. Error	t-Statistic	Prob.	Variable	Coefficien	Std. Error	t-Statistic	Prob.
PHI(-3)	-0.424	0.032	-13.186	0	PHI(-1)	-0.749	0.044	-17.107	0
C	0.413	0.029	14.467	0	C	0.686	0.039	17.435	0
R-squared	0.494				R-squared	0.61			
Adj R-sq	0.491				Adj R-sq	0.418			
					Cross-section fixed (dummy variables)				
Total panel (balanced) observations: 180					Total panel (balanced) observations: 540				

Notes: PHI is the Public Health Index. The dependent variable on the left side is $PHI_{t=2014}$-$PHI_{t=2001}$; the dependent variable on the right side is PHI_t-PHI_{t-1}, with t = 2005, 2009, 2014 and t-1 = 2001, 2005, 2009, respectively.

Table 4: Public Health and Health Expenditure

			Dependent Variable Public Health Index						
Variable	Coefficien	Std. Error	t-Statistic	Prob.	Coefficien	Std. Error	t-Statistic	Prob.	
HEXP	0.285	0.073	3.878	0.0001	0.254	0.073	3.489	0.0005	
LDC2	0.099	0.009	11.196	0	0.100	0.009	11.381	0	
DC1	0.145	0.008	17.301	0	0.145	0.008	17.555	0	
DC2	0.172	0.010	17.186	0	0.173	0.010	17.554	0	
C	0.767	0.010	77.952	0	0.771	0.010	79.196	0	
R-squared	0.394	Mean dep var		0.903	R-squared	0.413	Mean dep var		0.903
Adj R-sq	0.391	S.D. dep var		0.106	Adj R-sq	0.408	S.D. dep var		0.106
					Period Fixed Effects				
Total panel (balanced) observations: 720					Total panel (balanced) observations: 720				

Notes: PHI is the Public Health Index; HEXP is Government health expenditure as a percentage of total Government expenditure. LDC2, DC1 and DC2 are dummy variables capturing group fixed effects for less developed, emerging and developed countries, respectively. The constant captures the least developed countries effect. The panel regression on the right also included period fixed effects. The panel includes data for 4 periods, 2001, 2005, 2009, 2014.

References

Berg, K. and H. Ehtamo. Interpretation of Lagrange Multipliers in Nonlinear Pricing Problem, *Optimization Letters*, 2010, 4, 275-285.

Caporaletti, L.E., Dula, J.H. and N.K. Womer. Performance Evaluation based on Multiple Attributes with Nonparametric Frontiers, *Omega*, 1999, 27, 637-645.

Charles, V., Färe, R. and S. Grosskopf. A Translation Invariant Pure DEA Model, *European journal of Operational Research*, 2016, 249, 390-392.

Charnes, A., Cooper, W.W. and E. Rhodes. Measuring the Efficiency of Decision Making Units, *European Journal of Operational Research*, 1978, 2, 429-444.

Cherchye, L., Moesen, W., Rogge, N. and T. van Puyenbroeck. An Introduction to "Benefit of the Doubt" Composite Indicators, *Social Indicators Research*, 2007, 82, 111-145.

Färe, R., Grossskopf, S. and D. Margaritis. The Diet Problem and DEA, *Journal of the Operational Research Society*, 2011, 62, 1420-1422.

Färe, R., Grosskopf, S., Karagiannis, G. and D. Margaritis. Data Envelopment Analysis and its Related Linear Programming Models, *Annals of Operational Research*, 2017, 250, 37-43.

Färe, R. and G. Karagiannis. Benefit-of-the-doubt Aggregation and the Diet Problem, *Omega*, 2014, 47, 33-35.

Fusco, E. Enhancing Non-compensatory Composite Indicators: A Directional Proposal, *European Journal of Operational Research*, 2015, 242, 620-630.

Garille, S.G. and S.I Gass. Stigler's Diet Problem Revisited, *Operations Research*, 2001, 49, 1-13.

Gonzalez, E., Carcaba, A. and J. Ventura. The Importance of the Geographical Level of Analysis in the Assessment of the Quality of Life: The Case of Spain, *Social Indicators Research*, 2011, 102, 209-228.

Hashimoto, A. and M. Kodama. Has Livability of Japan Gotten Betteer for 1956-1990? A DEA Approach, *Social Indicators Research, 1997*, 40, 359-373.

Karagiannis, G. On Aggregate Composite Indicators, *Journal of Operational Research Society*, 2017, 68, 741-746.

Klomp, J. and J. deHaan. Measuring Health: A Multivariate Approach, *Social Indicators Research*, 2010, 96, 433-457.

Korhonen, P.J. and M. Luptacik. Eco-efficiency Analysis of Power Plants: An Extension of Data Envelopment Analysis, *European Journal of Operational Research*, 2004, 154, 437-446.

Lancaster, L.M. The History of the Application of Mathematical Programming to Menu Planning, *European Journal of Operational Research*, 1992, 57, 339-347.

Larson, J.S. The Weighting of an International Health Status Index, *Social Indicators Research*, 1994, 31, 265-275.

Lewis, H.F. and T.R. Sexton. Data Envelopment Analysis with Reverse Inputs and Outputs, *Journal of Productivity Analysis*, 2004, 21, 113-132.

Masoumzadeh, A., Toloo, M. and A. Amirteimoori. Performance Assessment in Production Systems without Explicit Inputs: An Application to Basketball Players, *IMA Journal of Management Mathematics Advance*, 2014.

Reig-Martinez, E., Gomez-Limon, J.A. and A.J. Picazo-Tadeo. Ranking Farms with a Composite Indicator of Sustainability, *Agricultural Economics*, 2011, 42, 561-575.

Stigler, G.J. The Cost of Subsistence, *Journal of Farm Economics*, 1945, 27, 303-314.

Tikunov, V.S. and O.Y. Chereshnya. Public Health Index in Russian Federation from 1990 to 2012, *Social Indicators Research*, 2016, 129, 775-786.

Vidoli, F., Fusco, E. and C. Mazziota. Non-compensability in Composite Indicators: A Robust Directional Frontier Model, *Social Indicators Research*, 2015, 122, 635-652.

Wu, J., Sun, J., Liang, L. and Y.C. Zha. Determination of Weights for Ultimate Cross Efficiency using Shannon Entropy. *Expert Systems with Applications*, 2011, 38, 5162-5165.

Yang, C.C., Hsiao, C.K. and M.M. Yu. Technical Efficiency and Impact of Environmental Regulations in Farrow-to-finish Swine Production in Taiwan, *Agricultural Economics*, 2008, 39, 51-61.

Zanella, A., Camanho, A.S. and T.G. Dias. Undesirable Outputs and Weighting Schemes in Composite Indicators based on Data Envelopment Analysis, *European Journal of Operational Research*, 2015a, 245, 517-530.

Zanella, A., Camanho, A.S. and T.G. Dias. The Assessment of Cities' Livability Integrating Human Wellbeing and Environmental Impact, *Annals of Operation Research*, 2015b, 226, 695-726.

Zhou, P., Ang, B.W. and K.L. Poh. A Mathematical Programming Approach to Constructing Composite Indicators, *Ecological Economics*, 2007, 62, 291-297.

Footnotes

[1] Other BoD applications include multiple-criteria decision-making (MCDM) problems such as inventory classification, supplier selection, etc. Färe, Grosskopf and Margaritis (2011) have explored the relationship between the diet problem and DEA. For other interesting relationships between well-known linear programming problems, see Färe *et al*. (2017).

[2] We only report results for the single global frontier. Results using two frontiers were only marginally higher for the KDC1/LDC2 frontier and roughly the same for the DC1/DC2 frontier.

[3] We also included interaction variables between country groups and health expenditure but none was statistically significant.