On constructing a composite indicator with multiplicative aggregation and the avoidance of zero weights in DEA

Chris Tofallis

Reader/Associate Professor of Decision Science

Hertfordshire Business School, UK

c.tofallis@herts.ac.uk

Published in *Journal of the Operational Research Society* (2014) 65, 791–792.
doi:10.1057/jors.2013.137

Recently there has been interest in combining the use of multiplicative aggregation together with data envelopment analysis (DEA). For example Blancas et al (2013), Cook and Zhu (2013), Giambona and Vassallo (2013); the first two of these are JORS papers. The purpose of this note is to highlight differences in the way multiplicative DEA is being applied and to draw attention to the fact that a units-invariant (i.e. scale-invariant) form is available. Moreover, this model avoids the 'zero weight problem' in DEA (where criteria are effectively ignored).

Composite indicators, indices or scores based on multiple criteria are widely used for ranking purposes. Historically, the weighted sum approach was the first to be used, and remains the most widely known. This involves normalising the data and then attaching weights to produce an index score. The purpose of normalisation is to ensure that variables with numerically larger magnitude do not dominate, and also to render the variables dimensionless so that they can be added together, thus avoiding the addition of quantities measured in different units. Criticism of such approaches often focused on the subjective weights that were chosen. What was not generally appreciated was that the particular normalisation selected by the analyst also affected the rankings. The fact is that a variety of normalisations is available, and they do *not* lead to the same results (Pomerol and Barba-Romero 2000, p.78*ff*). The great advantage of the multiplicative approach is that it enables variables to be aggregated without need for normalisation. This makes it more objective because there is less potential for the analyst to influence the results, be it intended or unintentional. It also simplifies the analysis. If the data are measured on a ratio scale then the rankings will be be scale-invariant. A ratio scale has a true or absolute zero value (e.g. mass, length, duration, cost) so that the ratio of two values remains the same if the data are rescaled. Physical laws typically relate quantities measured in different units, and this is done by multiplication of variables.

The papers of Blancas et al (2013) and Giambona and Vassallo (2013) have the admirable aim of constructing composite indicators which possess the advantages of multiplicative aggregation, which they describe, whilst also avoiding the subjective selection of weights. However, in progressing from additive to multiplicative aggregation they have continued to apply a normalisation step. Both of these papers have employed forms of normalisation which removes proportionality and so destroys the ratio scale property. In Blancas et al, for each variable the lowest observed value was allocated a value of 10 and the largest observed value was allocated a value of 100 (personal communication). Whereas, in Giambona and Vassallo, this range was restricted to be from 2 to 10. Such transformations convert a ratio scale to an interval scale, in a similar way that temperatures are converted from the absolute scale to the Fahrenheit scale. It is incumbent on those who employ normalisation to demonstrate that their rankings and conclusions are not a result of the particular data transformation they happen to have selected.

The key to ensuring scale invariance in multiplicative DEA is to have a score function which includes an overall scale factor. For example in the case of two component criteria this would appear as w_0 in the following score formula:

$S = w_0 (y_1)^{W1} (y_2)^{W2}$

The presence of this factor allows the DEA score to adjust correctly if any variables are re-scaled.

As one might expect, multiplicative DEA leads to an additive DEA model after taking logs. The log of w_0 can be negative (when $w_0 < 1$) hence it is a variable that is unrestricted in sign when it appears as a term in the objective function of the additive model. This implies an equality constraint in the dual: $\Sigma\lambda_i = 1$. (Cooper et al, 2006, p.114).

A rescaling of the data (multiplication by a positive constant, c) corresponds to a shift or translation of the data by log(c) in log-space. On a graph plot, the *relative* positions of the data points in log space (which includes the frontier) remain the same after such a translation; this can also be described as a shift of the origin. If our multiplicative DEA model is to be scale-invariant then our additive DEA model must be translation invariant. The objective function in the additive envelopment model is simply the sum of slacks, which is precisely the L_1 distance (also known as rectilinear or Manhattan distance) to the frontier from the point being assessed. Since the frontier does not move relative to the other data points when there is a translation of the origin, it follows that this distance remains unchanged and hence the additive model is translation invariant.

Additive DEA model to assess an entity with criteria values y_{01} and y_{02} [note Y = Log(y)]:

$\text{Min} - s_1 - s_2$ [or $\text{Max } s_1 + s_2$]

Subject to:

$\Sigma\lambda_i Y_{i1} - s_1 = Y_{01}$ (i.e. efficient value – slack = current value)

$\Sigma\lambda_i Y_{i2} - s_2 = Y_{02}$

$\Sigma\lambda_i = 1$

$\lambda_i, s_1, s_2 \geq 0$ i = 1...n (n is the number of entities being compared)

The dual of the above linear programme is:

$\text{Max } W_0 + W_1 Y_{01} + W_2 Y_{02}$ i.e. max Log(S)

Subject to

$W_0 + W_1 Y_{i1} + W_2 Y_{i2} \leq 0$ [since Log(1) = 0] i = 1...n

$-W_1 \le -1$

$-W_2 \le -1$

$W_1, W_2 \ge 0$, $W_0 = \text{Log}(w_0)$ and is unrestricted.

Looking at this dual, one notices a wonderful and unexpected benefit of this DEA model, and one which does not seem to have been remarked upon previously. The lower limit on weights is unity, which means that none of the criteria can be ignored in the analysis. There is no longer any fear of zero (or infinitesimal epsilon) weights - something which has often been an annoyance with other DEA models. [In fact a separate 'weight restriction' literature has arisen to deal with that problem.]

Unfortunately, Blancas *et al* and Giambona and Vassallo (2013) do not quite use the above DEA model and so lose its valuable advantages. Instead they employ a form used by Zhou et al which differs in two important ways: it does not avoid infinitesimal weights, and it does not contain the scale factor w_0, and so is not units invariant. Indeed, in a later paper Zhou et al (2010, p.177) actually stated "the composite indicator values based on the proposed multiplicative optimization approach are not invariant to the measurement units of the sub-indicators". Giambona and Vassallo end up applying arbitrary restrictions to avoid having zero weights.

Presumably, the reason that Blancas et al (2013) and Giambona and Vassallo (2013) chose their respective normalisations was because their data contained some zero values, which would cause the aggregated multiplicative score to become zero for those cases. There may be situations where a zero overall score would be appropriate because a zero on any dimension is a clear indication that this performance is not acceptable. But in situations where one does not wish that to happen then it may be more appropriate to employ a different DEA model, such as that of Kerstens and Van de Woestyne (2010) which allows zero and negative data.

We hope the points highlighted in this note will encourage users to choose the above form of multiplicative DEA so as to benefit from its advantages.

References

Blancas, FJ, Contreras, I, and Ramirez-Hurtado, JM. (2013). Constructing a composite indicator with multiplicative aggregation under the objective of ranking alternatives. *Journal of the Operational Research Society*, 64, 668-678.

Cook, WD and Zhu, J. (2013). DEA Cobb–Douglas frontier and cross-efficiency. *Journal of the Operational Research Society* advance online publication, doi:10.1057/jors.2013.13. http://www.palgrave-journals.com/jors/journal/vaop/ncurrent/full/jors201313a.html

Cooper, WW, Seiford, L, and Tone, K. (2006). *Introduction to data envelopment analysis and its uses*. Springer, NY.

Giambona, F and Vassallo, E. (2013). Composite Indicator of Social Inclusion for European Countries. *Social Indicators Research* advance online publication. DOI 10.1007/s11205-013-0274-2. http://link.springer.com/article/10.1007%2Fs11205-013-0274-2

Kerstens, K., and Van de Woestyne, I. (2010). Negative data in DEA: A simple proportional distance function approach. *Journal of the Operational Research Society*, 62(7), 1413-1419.

Pomerol, J-C and Barba-Romero, S. (2000). *Multicriterion Decision in Management - Principles and Practice.* Kluwer Academic.

Zhou, P., Ang, B. W., & Zhou, D. Q. (2010). Weighting and aggregation in composite indicator construction: a multiplicative optimization approach. *Social Indicators Research*, *96*(1), 169-181.