



## Reply

## A reply to Gandjour and Gafni

Han Bleichrodt<sup>a,\*</sup>, Martin Filko<sup>b</sup><sup>a</sup> Erasmus School of Economics & iMTA/iBMG, Netherlands<sup>b</sup> Institute of Public Policy and Economics, Faculty of Social and Economic Sciences, Comenius University, Bratislava, Slovakia

## ARTICLE INFO

## Article history:

Received 7 January 2010

Accepted 8 January 2010

Available online 18 January 2010

## Keywords:

QALYs

Utility theory

## ABSTRACT

Gandjour and Gafni (in press) criticize our paper on two counts. Their first point of criticism is ill-founded and results from many mathematical mistakes that they make. The second is due to a lack of understanding of the general principles of empirical research.

© 2010 Elsevier B.V. All rights reserved.

In Bleichrodt and Filko (2008) we performed new tests of the QALY model when health varies over time. The novelty of our tests is that they control for violations of expected utility. It is well known that people do not behave according to expected utility (Starmer, 2000) and these violations may have confounded previous tests of the QALY model. Our experimental data supported QALYs at the aggregate level, but not at the individual level.

In a comment, Gandjour and Gafni (in press) criticize our paper on two grounds. First, they argue that it is possible that the condition we tested, *generalized marginality*, is not sufficient to imply the QALY model. In other words, subjects may simultaneously satisfy generalized marginality and violate the QALY model. Second, Gandjour and Gafni argue that we cannot make generalized statements about preferences because our sample is not representative. Related to this, they argue that we cannot conclude in support of a particular model based on a limited number of tests because the variety of health profiles is essentially endless. In this reply we will show that Gandjour and Gafni's first point of criticism is wrong. Their arguments contain many mathematical mistakes implying that their counterexamples are wrong and, therefore, that all their corresponding speculations are irrelevant. Their other points of criticism are completely standard (Popper, 1934, 1963) and reflect a lack of understanding of the general principles underlying all empirical studies in all fields of science. They also reflect poor reading as these points have actually been acknowledged and discussed in our paper.

### 1. First criticism: support for generalized marginality and violations of the QALY model can coexist

There are three principal problems with Gandjour and Gafni's first point of criticism, which we will outline below.

#### 1.1. First problem: Eq. (1) is ambiguous and ill-defined<sup>1</sup>

The first fundamental problem is that the model of Gandjour (2008), which underlies Gandjour and Gafni's (in press) analysis and is stated in their Eq. (1), is not well-defined. A problem that recurs throughout their comment is that even though Gandjour and Gafni (in press) use mathematical derivations, they do not follow the rules and logic of mathematics (Suppes, 1957).

According to the left hand side of Eq. (1)  $u$  depends only on health states  $a$ ,  $b$ , and  $c$ . However, on the right hand side of Eq. (1) the distributions  $L(b)$  and  $L(c)$  also appear. If these distributions play a role then they should also appear in the argument of the function. Then the utility of an outcome depends not only on the outcome itself, but on the whole distribution that it is part of. The model then loses all its tractability and becomes completely general without any predictive power. In particular, it is unclear how the formula should be applied when computing probability weighted averages such as in expected utility or its generalizations. Gandjour (2008) claims that expected utility should not be used to compute probability weighted averages but in Eq. (4) of their comment Gandjour and Gafni (in press) do use expected utility to compute probability weighted averages.

\* Corresponding author at: Erasmus University, Dept. of Applied Economics, H13-27, P.O. Box 1738, 3000 DR Rotterdam, Netherlands.

E-mail address: [bleichrodt@ese.eur.nl](mailto:bleichrodt@ese.eur.nl) (H. Bleichrodt).

<sup>1</sup> Throughout this reply, the equation numbers refer to the equations in Gandjour and Gafni (in press).

It is further a complete mystery where the functions  $L$  come from. Are these population statistics, marginal distributions or are they specific to the prospects that are being considered? Moreover, given the many parameters in Eq. (1) and their unclear nature, identifiability of the model is also a problem.

There are two additional inaccuracies related to Eq. (1). Gandjour and Gafni (in press) call Eq. (1) additive, which it is not. A function is normally called additive if it is additively decomposable, which implies strong separability. It is obvious that Eq. (1) does not satisfy strong separability. Apparently the authors use the term additive each time they discern an additive operation amidst other mathematical operations. We will ignore their claims about additivity in what follows.

A second inaccuracy is that Gandjour and Gafni use the same symbol  $u$  for several different things. In Eq. (1)  $u$  is used both as a function of a sequence of health states, as a function of the single-period health states, and as a function of the function  $L$ , representing the distribution of the health states within a period. This ambiguity about what the functions represent makes it hard to discuss the theory.

Because Gandjour's (2008) model is ill-defined and ambiguous, it is impossible to understand exactly what Gandjour and Gafni mean. Nevertheless we will try our best to interpret their writings as good as we can.

### 1.2. Second problem: Eq. (2) is wrong

The second problem is that their claims made in Eq. (2) are unsubstantiated and wrong. Gandjour and Gafni claim that if strong separability, or additive utility independence as they call it, is imposed on top of Eq. (1) then Eq. (2) results. No proof is given for this claim and we will show that it is wrong. Gandjour and Gafni claim that additive separability implies that all  $\lambda$ 's must be equal to zero. Suppose, in contrast with Gandjour and Gafni's claim, that at least one of the  $\lambda$ 's is unequal to zero. Say  $\lambda(a) = 1$ . Suppose also that  $u(L(b)) = u(L(c)) = 0$  for all  $b$  and  $c$ . Then Eq. (1) in Gandjour and Gafni becomes

$$u(a, b, c) = u(a) + u(b) + u(c) + u(b) + u(c) = u(a) + 2u(b) + 2u(c),$$

which is an additively decomposable form and which satisfies strong separability and additive utility independence. Hence, it is not true that strong separability or additive utility independence implies Eq. (2). This simple counterexample shows that Eq. (2) in Gandjour and Gafni (in press) is wrong, that their claims about the  $\lambda$ 's being equal to zero are wrong, and that all the claims made later in the paper about generalized marginality and Eq. (2) are wrong.

### 1.3. Third problem: Eq. (4) and, hence, Gandjour and Gafni's counterexample against generalized marginality, is wrong

We finally show that Eq. (4) in which Gandjour and Gafni (in press) derive what they believe generalized marginality tests is wrong. Before we do so, we must correct two mistakes in their Eq. (3), which describes our test of generalized marginality. A first problem with Eq. (3) is that Gandjour and Gafni, once again, violate the rules of logic and use different symbols to denote identical things. According to the rules of logic it is possible, for example, that  $a_I$  and  $a_{II}$  in Gandjour and Gafni's Eq. (3) are different. In the definition of generalized marginality they have to be identical. We will therefore ignore subscripts in what follows and simply write  $a_I = a_{II} = a$ ,  $c_I = c_{II} = c$ , etc.

A second problem with Eq. (3) is that in the prospects on the right hand sides of the two indifference signs  $a''$  appears twice. This is wrong. In each of these two prospects, the second term  $a''$

has to be different from the first. We assume that this is a typo and that the authors had in mind to write  $a'''$  for the second terms.

Let us now explain the problems with Eq. (4). A first problem is that Gandjour and Gafni use expected utility. In our paper we use a much more general model than expected utility and Gandjour and Gafni should have shown that their conclusion holds under this more general model. A second problem is that Eq. (4) contains a term  $L(b_I - b_{II})$ . Why does  $b_{II}$  suddenly appear within brackets? This can only be if  $L$ , whatever it is, is linear (Aczel, 1966, Theorem 1, p. 34). But such linearity has never been assumed. Moreover,  $b_{II}$  does not appear in Eq. (3) so where does it come from?

However, the fundamental problem with Eq. (4) is that it is wrong. Assuming Eq. (1) and expected utility as Gandjour and Gafni do, and following the same line of analysis as they do we obtain that the difference between the prospects on the left hand sides of the indifference signs is equal to:

$$\begin{aligned} p[u(b) - u(b'') + \lambda(a)(u(b) - u(b'')) + u(c)(\lambda(b) - \lambda(b'')) \\ + \lambda(a)(u(L(b)) - u(L(b'')))] + u(L(c))(\lambda(b) - \lambda(b'')) \\ + (1 - p)[u(b') - u(b''') + \lambda(a')(u(b') - u(b''')) \\ + u(c)(\lambda(b') - \lambda(b''')) + \lambda(a')(u(L(b')) - u(L(b''')))] \\ + u(L(c))(\lambda(b') - \lambda(b'''))]. \end{aligned}$$

And the difference between the prospects on the right hand sides of the indifference signs is equal to:

$$\begin{aligned} u'' + \lambda(a'')(u(b) - u(b'')) + u(c)(\lambda(b) - \lambda(b'')) \\ + \lambda(a'')(u(L(b)) - u(L(b'')))] + u(L(c))(\lambda(b) - \lambda(b'')) \\ + (1 - p)[u(b') - u(b''') + \lambda(a''')(u(b') - u(b''')) \\ + u(c)(\lambda(b') - \lambda(b''')) + \lambda(a''')(u(L(b')) - u(L(b''')))] \\ + u(L(c))(\lambda(b') - \lambda(b'''))]. \end{aligned}$$

Deleting common terms this implies that generalized marginality tests whether

$$\begin{aligned} p[\lambda(a)(u(b) - u(b'') + u(L(b)) - u(L(b'')))] + (1 - p)[\lambda(a')(u(b') \\ - u(b''') + u(L(b')) - u(L(b''')))] = p[\lambda(a'')(u(b) - u(b'')) \\ + (u(L(b)) - u(L(b'')))] + (1 - p)[\lambda(a''')(u(b') - u(b''')) \\ + u(L(b')) - u(L(b'''))]. \end{aligned}$$

This is clearly different from what Gandjour and Gafni obtain. Contrary to what Gandjour and Gafni claim the terms involving  $L(\cdot)$  do not cancel. Having shown that Gandjour and Gafni's derivations and, hence, their counterexample, are wrong, all their speculations that follow Eq. (4) become irrelevant and we can safely ignore them.

## 2. Second criticism: no general statements are possible

Regarding their second point of criticism we can be short: their expressed concerns are completely standard and have actually been acknowledged in our paper.

The issue of representativeness is discussed in the third paragraph on page 1247. It is common to use convenience samples such as students to first test new decision concepts. Consider, for example, Kahneman and Tversky (1979), the second most cited paper in economics since 1970 (Kim et al., 2006), which introduced prospect theory, the theory for which Kahneman was awarded the Nobel prize in economics in 2002. Kahneman and Tversky (1979)

is entirely based on the responses of students and university faculty. Later studies then tested these new concepts in more general samples. In our case the new decision principle was generalized marginality. Because the test was new it made sense to first employ a convenience sample. Future studies should try to replicate our findings in more general samples.

Regarding the limited number of tests, it is well-known (Popper, 1934, 1963) that a hypothesis can never be proved right and can only be shown to be false. The classical example is the hypothesis “all swans are white.” This hypothesis can never be proved right but can be falsified by observing one single black swan. Gandjour and Gafni have nothing new to add here. According to Popper data that are in line with the theory “corroborate” the theory. Our general conclusion, repeated below, is entirely consistent with Popper (1934, 1963):

“Our results provide support for the QALY model at the aggregate level. It should be pointed out though that this conclusion is based on three tests only. It should also be kept in mind that we only used mild to moderate health states to avoid considerations like maximal endurable time. Our conclusions may no longer hold when more severe health states are involved. More evidence is needed and we invite other researchers to try and replicate our findings using other experimental designs.” (p. 1247)

Let us end by correcting one final mistake in Gandjour and Gafni’s (in press) comment. They imply that we cite Spencer and Robinson (2007) as providing support for generalized marginality at the aggregate level. Once again, they did not read carefully. On page 1246 we wrote: “our aggregate findings on *utility independence* [emphasis added] are consistent with the findings of Spencer and Robinson (2007).” As we point out in our paper, if utility independence holds but generalized marginality is violated then period-specific utilities can still be defined and utility remains tractable.

### 3. Conclusion

Gandjour and Gafni (in press) criticize our paper on two counts. Their first point of criticism is ill-founded and results from many mathematical mistakes that they make. The second is due to a lack of understanding of the general principles of empirical research.

### Acknowledgements

We are grateful to Jason N. Doctor and Peter P. Wakker for helpful comments. Han Bleichrodt’s research was made possible through a grant from the Netherlands Organization for Scientific Research (NWO).

### References

- Aczel, J., 1966. Lectures on Functional Equations and Their Applications. Academic Press, New York.
- Bleichrodt, H., Filko, M., 2008. New tests of QALYs when health varies over time. *Journal of Health Economics* 27, 1237–1249.
- Gandjour, A., 2008. Incorporating feelings related to the uncertainty about future health in utility measurement. *Health Economics* 17, 1207–1213.
- Gandjour, A., Gafni, A., 2010. The additive utility assumption of the QALY model revisited. *J. Health Econ.* 29, 325–328.
- Kahneman, D., Tversky, A., 1979. Prospect theory: an analysis of decision under risk. *Econometrica* 47, 263–291.
- Kim, E.H., Morse, A., Zingales, L., 2006. What has mattered to economics since 1970. *Journal of Economic Perspectives* 20, 189–202.
- Popper, K.R., 1934. *Logik der Forschung*. Julius Springer, Verlag, Vienna.
- Popper, K.R., 1963. *Conjectures and Refutations: The Growth of Scientific Knowledge*. Routledge & Kegan Paul, London.
- Spencer, A., Robinson, A., 2007. Test of utility independence when health varies over time. *Journal of Health Economics* 26, 1003–1013.
- Starmer, C., 2000. Developments in non-expected utility theory: the hunt for a descriptive theory of choice under risk. *Journal of Economic Literature* 28, 332–382.
- Suppes, P., 1957. *Introduction to Logic*. Van Nostrand, New York.