White Rose Research Online

eprints@whiterose.ac.uk
https://eprints.whiterose.ac.uk
Universities of Leeds, Sheffield and York

Deposited via The University of Sheffield.

White Rose Research Online URL for this paper:
https://eprints.whiterose.ac.uk/id/eprint/175351/

Version: Accepted Version

Article:
Dolan, P. and Tsuchiya, A. (2003) The person trade-off method and the transitivity principle : an example from preferences over age weighting. Health Economics, 12 (6). pp. 505-510. ISSN: 1057-9230

https://doi.org/10.1002/hec.731

This is the peer reviewed version of the following article: Dolan, P. and Tsuchiya, A. (2003), The person trade-off method and the transitivity principle: an example from preferences over age weighting. Health Econ., 12: 505-510., which has been published in final form at https://doi.org/10.1002/hec.731. This article may be used for non-commercial purposes in accordance with Wiley Terms and Conditions for Use of Self-Archived Versions.

Reuse
Items deposited in White Rose Research Online are protected by copyright, with all rights reserved unless indicated otherwise. They may be downloaded and/or printed for private study, or other acts as permitted by national copyright laws. The publisher or other rights holders may allow further reproduction and re-use of the full text version. This is indicated by the licence information on the White Rose Research Online record for the item.

Takedown
If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

UNIVERSITY OF LEEDS
University of Sheffield
UNIVERSITY of York

The person trade-off method and the transitivity principle: an example from preferences over age weighting

Paul Dolan[1,2] and Aki Tsuchiya[1,3]

[1] Sheffield Health Economics Group, University of Sheffield,
[2] Department of Economics, University of Oslo, Norway

[3] Corresponding author: 30 Regent Street, Sheffield, S1 4DA;
tel 0114.222.0710 fax 0114.272.4095

Email: P.Dolan@shef.ac.uk **or** A.Tsuchiya@sheffield.ac.uk

ACKNOWLEDGEMENTS

This study was funded by the Economic and Social Research Council (Award No: L128251050). We are grateful to Rebecca Shaw, Peter Smith, Alan Williams, and all the respondents who agreed to take part. The usual disclaimers apply.

The person trade-off method and the transitivity principle: an example from preferences over age weighting

ABSTRACT

The person trade-off (PTO) is increasingly being used to elicit preferences in health. This paper explores the measurement properties of the PTO method in the context of a study about how members of the public prioritise between patients of different ages. In particular, it considers whether PTO responses satisfy the *transitivity principle*; that is, whether one PTO response can be inferred from two other PTO responses. The results suggest that very few responses to PTO questions satisfy cardinal transitivity condition. However, this study has produced results that suggest that *cardinal* transitivity will hold, on average, when respondents who fail to satisfy the *ordinal* transitivity condition have been excluded from the analysis. This suggests that future PTO studies should build in checks for ordinal transitivity.

Key words: person trade-off, transitivity

INTRODUCTION

There is a rapidly growing literature within health economics that addresses the issue of what patient characteristics should be taken into account when setting health care priorities.[1-7] A number of studies have elicited the preferences of the general public about whether characteristics such as the severity of the condition, age, lifestyle, or family responsibilities should be taken into account, and if so, how they should be traded-off against the maximisation of health gain. A method that is often used in such empirical studies is the *person trade-off* (PTO) method, in which respondents are asked to specify the number of people with one set of characteristics that would make them indifferent between treating that group and a different number of people with another set of characteristics.[6-12] The method has also been used to quantify preferences regarding the relative value of different health states[9, 13] and is used by the World Bank and others to value disease states in terms of DALYs (Disability Adjusted Life Years).[14, 15]

This paper explores the measurement properties of the PTO method in the context of an in interview-based study about how members of the public prioritise between patients of different ages. In particular, it considers whether PTO responses satisfy the *transitivity principle*; that is, whether one PTO response can be inferred from two other PTO responses. To our knowledge, there have only been two other published studies – both using convenience samples – that have addressed this issue, and which found that transitivity did not hold.[16, 17]

METHODS

Five groups of different ages – 5 year-olds, 20 year-olds, 35 year-olds, 55 year-olds and 70 year olds – were chosen to represent different stages in the life cycle. Respondents were told that each group would die in a few days, and were asked to rank the groups in the order in which they would choose to give an extra five years of life. After this ranking exercise, respondents were asked three PTO questions that asked them to consider pairs of programmes that targeted different groups. The age groups used were those that the respondent had ranked first, third and fifth in the previous question. They were asked to indicate how many people would have to be treated by the programme aimed at the more preferred age group (relative to a fixed number of people treated by the programme aimed at the less preferred age group) for the two programmes to be of equal social value. See Figure 1 for an example of the layout of the questions.

If responses to PTO questions are to be used to reflect the relative importance of one patient group as compared to another, then the results should have ratio scale properties, and will thus satisfy *cardinal transitivity*. In other words, the marginal rate of substitution (MRS) between, for example, the fifth and third ranked ages multiplied by the MRS between the third and first ranked ages should be equal to the MRS between the fifth and first ranked ages i.e. $MRS_{5v3} \times MRS_{3v1} = MRS_{5v1}$. Thus, by examining the relationship between the actual MRS_{5v1} and the predicted MRS_{5v1}, obtained via the other two MRSs, the ratio scale (or cardinal transitivity) property of PTO responses can be tested. The calculation of the predicted MRS_{5v1} is then analogous to the calculation of "chained" values for health states, where intermediate

health states are evaluated against some severe state, which in turn is evaluated against death, so that the valuation of the initial state against full health and death becomes indirect, or "chained".

If respondents are also concerned about the *difference* between the number of people in the two groups, as well as about the ratio between them, then $MRS_{5v3} \times MRS_{3v1} > MRS_{5v1}$ i.e. the respondent is willing to trade-off fewer people in the fifth versus first ranked ages question than would be implied by their responses to the two other questions. If respondents do focus on the absolute difference, then the gap between the predicted and actual MRS_{5v1} will increase as the baseline number of people in the PTO question increases. For this reason, the interviews were carried out in two rounds. The starting point was 100 people in each age group in the first round of interviews (R1), and 1000 people in the second round (R2). Further, to determine whether there were any ordering effects, half the respondents in each round of interviews received the order MRS_{5v1}, MRS_{5v3}, then MRS_{3v1} (O1), and half received the order MRS_{3v1}, MRS_{5v1}, then MRS_{5v3} (O2). Therefore, there were four experimental cells.

Scatterplots of the relationship between the actual MRS_{5v1} and the MRS_{5v1}, predicted from responses to the two other PTO questions, are produced for all four variants. The predicted MRS_{5v1} is regressed on the actual MRS_{5v1} with no intercept to test whether or not these are equal to each other. The Chi^2 test is used to test for whether certain respondents are located above or below the 45º ray of the scatterplot (i.e. whether their predicted MRS_{5v1} is larger or smaller than the actual MRS_{5v1}). Regression analysis using the ratio of the predicted and actual MRS_{5v1} is employed to

see whether certain respondents are nearer to or further away from the 45º ray. In each case, the explanatory variables are: the variant of the questionnaire respondents were given; their background characteristics (age, sex, employment, dependants, smoking status, illness); and the "pattern" of age preference they have (a pattern is called "linear" when the first, third, and fifth preferred ages are in ascending age order, and "non-linear" otherwise). A significance level of 0.1 is used throughout.

The three PTO questions can also be used to test the *ordinal transitivity* of responses. Since the distance, in terms of strength of preference, between the fifth and first ranked ages should be greater than each of the distances between the fifth and third and third and first ranked ages, there are two consistency conditions that each respondent's PTO responses should satisfy: namely, $MRS_{5v1} > MRS_{5v3}$ and $MRS_{5v1} > MRS_{3v1}$.

Letters of invitation were sent out to 1,500 people on the electoral register in three wards in York (this amounted to every eighth person). They were invited to attend an interview for which they would be paid £15. Of these, 467 people (31%) agreed to participate. These potential respondents, who provided information about their sex and age on their reply slips, were more likely to be female and in the older age groups. Therefore, men and the younger age groups were over-sampled so that the final sample would be representative of the wider population. Of the 140 people selected for interview, 130 turned up and these were indeed broadly representative of the general population in terms of their sex and age distribution. The interviews were conducted by three trained interviewers and took place at the University of York. The

full interview lasted for about one hour, of which about fifteen minutes was spent on the questions analysed in this paper.

RESULTS

Figures 2(a)-(d) show scatterplots of the relationship between the actual MRS_{5v1} and predicted MRS_{5v1}. For all four variants, there are very few respondents on, or close to, the 45º ray, suggesting that very few respondents satisfy cardinal transitivity. The side of the 45º ray a respondent lies was found to be related to the variant of the questionnaire they were given, with predicted values higher than actual ones more likely in R1 than R2, and in O2 than O1. Respondent age was also found to have a significant effect, with only 3% of respondents in the 65+ age group having predicted values higher than actual ones compared to 30% in the under 65s. While 55% of respondents had a linear pattern of preferences and 45% had a non-linear pattern, this had no effect on which side of the 45º ray a respondent is located. In terms of the distance from the 45º ray, there were no statistically significant differences by questionnaire variant, respondent characteristics, or pattern of preference.

Table 1 presents aggregate level results. For three of the four questionnaire variants, the mean actual MRS_{5v1} is greater than the mean predicted MRS_{5v1}, as shown by the proportion of those with ratios > 1 and β coefficients that are significantly different from 1.00. The exception is variant R1-O2, where the actual and predicted values are very close to one another. On the whole, then, respondents are trading off less on average in the fifth-ranked versus first-ranked age PTO question than they would if they were satisfying cardinal transitivity. For example, take R1-O1 where, on

average, $MRS_{5v3} = 0.59$ and $MRS_{3v1} = 0.52$. To satisfy cardinal consistency, the average respondent should set MRS_{5v1} equal to $0.59 \times 0.52 = 0.31$. However, average $MRS_{5v1} = 0.43$, suggesting that they are taking account of the absolute difference in the number of people in the two groups. Changing the starting point from 100 people in round one to 1000 people in round two has the predicted result: that is, the actual value exceeds the predicted value by an even greater amount.

Overall, the results do not look very good for the cardinal transitivity condition, but what about the weaker condition of ordinal transitivity? The rate of ordinal intransitivity was related to questionnaire variant, being higher for R2 than R1, and for O1 than O2. Ordinal intransitivity was also related to the age of the respondent, with the highest rate (40%) found in the 65+ age group and the lowest rate (17%) in the 16-34 age group. The remaining background variables and the pattern of preferences had no effect.

Table 2 replicates Table 1, but after excluding those respondents who violate at least one of the two ordinal transitivity conditions. Notice the large variation in the exclusion rate, reflecting the difference in ordinal intransitivity rates across the questionnaire variants. Compared to Table 1, there is a marked improvement in the match between the actual and predicted values for MRS_{5v1}. None of the regression coefficients are statistically significantly different from 1.00, implying that, on average, the actual MRS_{5v1} is approximately equal to that predicted from responses to the other two PTO questions.

DISCUSSION

It is clear from the scatterplots in Figure 2 that very few responses to PTO questions in the study satisfy the cardinal transitivity condition. Many respondents have a MRS between the most and least preferred ages that is lower than would be predicted from the other two PTO responses, particularly when the starting number of people in the PTO is 1000 as opposed to 100. Similar violations of cardinal transitivity were reported in [16] where about 75% of respondents had a lower MRS (across different conditions) than would have been predicted from other responses. The authors conclude that, whilst "this inconsistency should be of great concern to proponents of the person trade-off method", it "may be correctable by modifying the way person trade-offs are elicited". In the other study that tested for – and found similar evidence of – violations of cardinal consistency, the authors pointed out the inconsistencies to respondents and this reduced the number of violations in PTO responses.[17]

This study has produced results that suggest that *cardinal* transitivity will hold, on average, when respondents who fail to satisfy the *ordinal* transitivity condition have been excluded from the analysis. Of course, this means using the data from only a subset of respondents and, depending on the variant used, could mean excluding as many 40% of respondents. O2 performs much better in this respect (only 3 respondents are excluded from R1-02). O2 also performs better in terms of cardinal transitivity, at least at the aggregate level, and so these results may have implications for chained PTO studies. O1 is the ordering where the PTO question involving the most preferred and least preferred ages comes before PTO questions involving the middle-ranked age, while O2 is the ordering where the intermediate combinations

precedes the extreme combination. This implies that, when PTO studies chain the responses of less severe states to the worst state, transitivity may be more closely approximated when the extreme combination (i.e. the valuation of the worst state against full health and death) comes after the evaluations of the intermediate states.

There is at least one caveat that is worth adding; namely, that respondents may change their minds regarding the rank ordering of the five ages as they go through the PTO exercise. In this case, the issue is not one of intransitivity, but one of instability. However, this is a complicated issue to address since, even if the ranking exercise is repeated after the PTO, it will be difficult to determine at which point the respondent changed his mind, or indeed, how many times he changed his mind. In any event, it is important that the robustness of findings reported here is tested in future studies. If they are found to be robust, then future PTO studies, especially those that involve a chaining procedure, should build in checks for ordinal transitivity.

REFERECES

1. Abellan Perpinan J, Pinto Prades J.L. Health state after treatment: a reason for discrimination? Health Economics. 1999;**8**:701-707.

2. Cropper ML, Aydede SK, Portney PR. Preferences for life saving programs: How the public discounts time and age. Journal of Risk and Uncertainty. 1994;**8**:243-265.

3. Neuberger J, Adams D, Maidment A, et al. Assessing priorities for allocation of donor liver grafts: Survey of the public and clinicians. British Medical Journal. 1998;**317**:172-175.

4. Nord E. The relevance of health state after treatment in prioritising between different patients. Journal of Medical Ethics. 1993;**19**:37-42.

5. Nord E, Richardson J, Street A, et al. The significance of age and duration of effect in social evaluation of health care. Health Care Analysis. 1996;**4**:103-111.

6. Ratcliffe J. Public preferences for the allocation of donor liver grafts for transplantation. Health Economics. 2000;**9**:137-148.

7. Rodriguez E, Pinto JL. The social value of health programmes: Is age a relevant factor? Health Economics. 2000;**9**:611-621.

8. Dolan P, Green C. Using the person trade off approach to examine differences between individual and social values. Health Economics. 1998;**7**:307-312.

9. Pinto Prades JL. Is the person trade-off a valid method for allocating health care resources? Health Economics. 1997;**7**:71-81.

10. Richardson J, Nord E. The importance of perspective in the measurement of quality-adjusted life years. Medical Decision Making. 1997;**17**:33-41.

11. Ubel PA, Loewenstein G. Distributing scarce livers: The moral reasoning of the general public. Social Science and Medicine. 1996;**42**:1049-1055.

12. Ubel PA, Richardson J, Pinto Prades J-L. Life saving treatments and disabilities: Are all QALYs created equal? International Journal of Technology Assessment in Heath Care. 1999;**15**:738-748.

13. Nord E. The Person Trade-Off approach to valuing health care programes. Medical Decision Making. 1995;**15**:201-208.

14. Murray CJL. Rethinking DALYs. In: Murray CJL, Lopez AD, eds). The Global Burden of Disease: Harvard University Press; 1996.

15. Stouthard MEA, Essink-Bot M-L, Bonsel GJ, et al. Disability Weights for Diseases in the Netherlands. Erasmus University Rotterdam, Department of Public Health, 1997

16. Ubel PA, Kamlet M, Scanlon D, Loewenstein G. Individual utilities are inconsistent with rationing choices: a partial explanation of why Oregon's cost-effectiveness list failed, Medical Decision Making. 1996;**16**:108-119.

17. Baron J, Wu Z, Brennan DJ, Weeks C, Ubel P. Analog scale, magnitude estimation and person trade-off as measures of health utility: biases and their correction, Journal of Behavioral Decision Making. 1999;**14**:17-34.

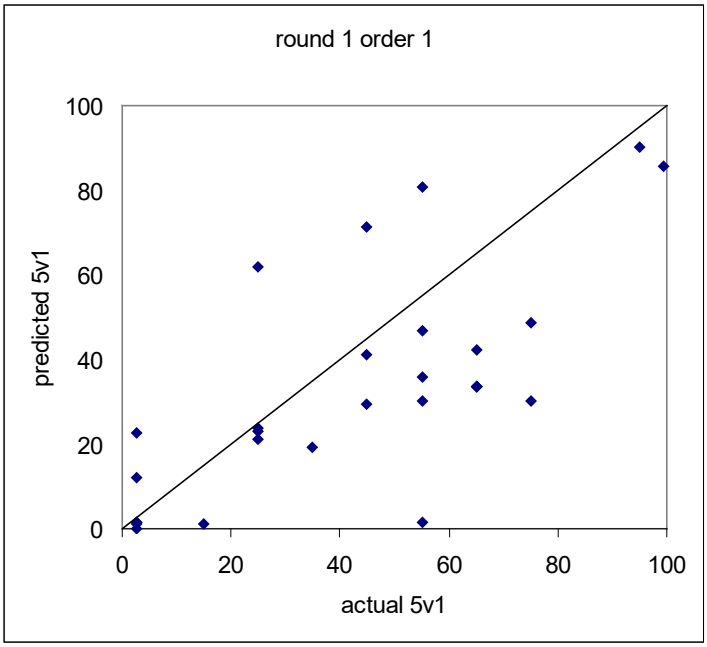
Figure 1 – An example of the PTO question

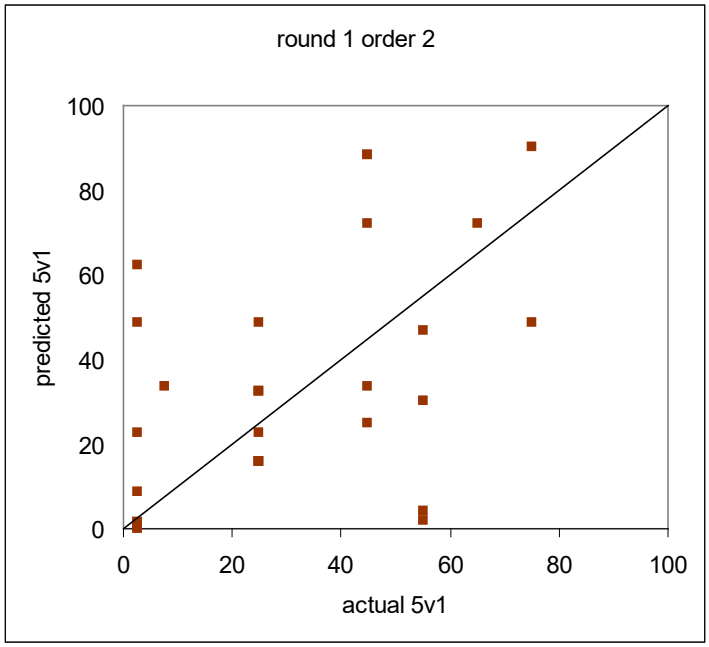
The interviewer wrote in the respondent's first, third, or fifth ranked age in the space provided.

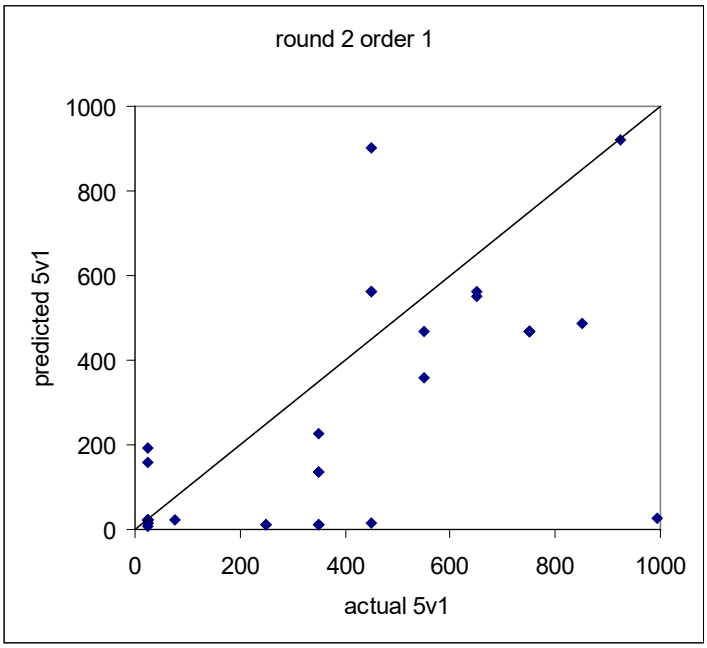
Programme A		**Programme B**	
Number of ___ year olds who will live for five years		Number of ___ year olds who will live for five years	
	✓ for **A**	✓ for **B**	
1000			1000
1000			990
1000			980
1000			970
1000			960
1000			950
1000			900
1000			800
1000			700
1000			600
1000			500
1000			400
1000			300
1000			200
1000			100
1000			50

Figure 2 – Scatterplots of individual responses

the numbers on the x-axis refer to the actual number of first-ranked age that is equivalent to 100 (or 1000) fifth-ranked age, and the numbers on the y-axis refer to the response that is predicted from the two other PTO questions.







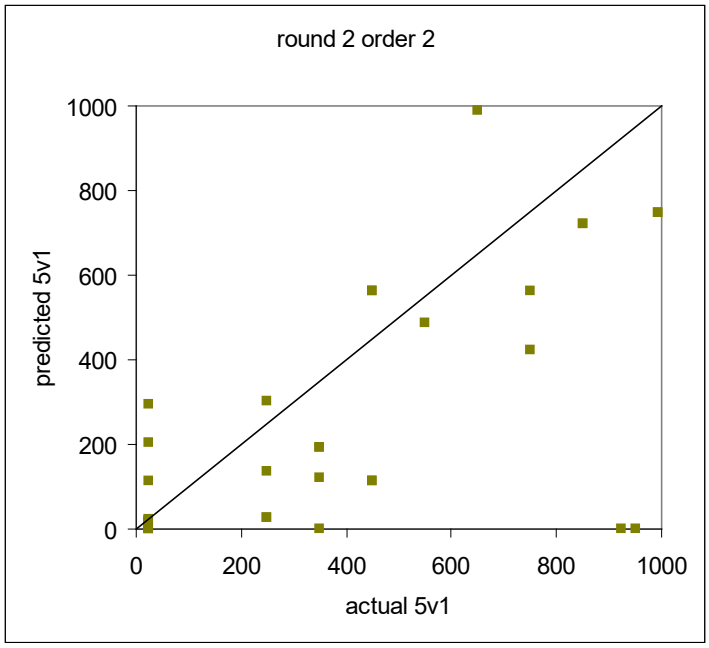


Table 1: Aggregate PTO results

Group	N	Actual MRS_{5v1}	Predicted MRS_{5v1} [a]	% actual > predicted	Actual / predicted [b]	β coefficient [c]
R1, O1	29	0.43	0.33	83	1.30	0.73*
R1, O2	31	0.31	0.34	52	0.91	0.92
R2, O1	34	0.37	0.25	85	1.48	0.64*
R2, O2	36	0.36	0.24	81	1.50	0.62*

[a] calculated as $MRS_{5v3} \times MRS_{3v1}$ at the individual level and then averaged
[b] calculated at the individual level and then averaged
[c] from the regression equation PRED = βOBS + e, through the origin, run on individual data –
the asterisks indicate that $p < 0.1$ for the 2-sided t-test that $\beta = 1.00$.

Table 2: Aggregate PTO results excluding inconsistent respondents

Group	Excluded	Actual MRS_{5v1}	Predicted MRS_{5v1} [a]	% actual > predicted	Actual / predicted [b]	β coefficient [c]
R1, O1	31%	0.33	0.33	75	1.00	0.92
R1, O2	10%	0.28	0.35	48	0.80	1.13
R2, O1	41%	0.27	0.28	75	0.96	1.00
R2, O2	28%	0.25	0.24	73	1.04	0.93

[a] calculated as $MRS_{5v3} \times MRS_{3v1}$ at the individual level and then averaged
[b] calculated at the individual level and then averaged
[c] from the regression equation PRED = βOBS + e, through the origin, run on individual data –
$p > 0.1$ for the 2-sided t-test that $\beta = 1.00$ for all β.