Health Technology Assessment 2000; Vol. 4: No. 1

Methodology

The estimation of marginal time preference in a UK-wide sample (TEMPUS) project

JA Cairns
MM van der Pol

Health Technology Assessment
NHS R&D HTA Programme

Standing Group on Health Technology

Details of the membership of the HTA panels, the NCCHTA Advisory Group and the HTA Commissioning Board are given at the end of this report.

How to obtain copies of this and other HTA Programme reports.
An electronic version of this publication, in Adobe Acrobat format, is available for downloading free of charge for personal use from the HTA website (http://www.hta.ac.uk). A fully searchable CD-ROM is also available (see below).

Printed copies of HTA monographs cost £20 each (post and packing free in the UK) to both public **and** private sector purchasers from our Despatch Agents.

Non-UK purchasers will have to pay a small fee for post and packing. For European countries the cost is £2 per monograph and for the rest of the world £3 per monograph.

You can order HTA monographs from our Despatch Agents:

– fax (with **credit card** or **official purchase order**)
– post (with **credit card** or **official purchase order** or **cheque**)
– phone during office hours (**credit card** only).

Additionally the HTA website allows you **either** to pay securely by credit card **or** to print out your order and then post or fax it.

Contact details are as follows:

HTA Despatch
c/o Direct Mail Works Ltd
4 Oakwood Business Centre
Downley, HAVANT PO9 2NP, UK

Email: orders@hta.ac.uk
Tel: 02392 492 000
Fax: 02392 478 555
Fax from outside the UK: +44 2392 478 555

NHS libraries can subscribe free of charge. Public libraries can subscribe at a very reduced cost of £100 for each volume (normally comprising 30–40 titles). The commercial subscription rate is £300 per volume. Please see our website for details. Subscriptions can only be purchased for the current or forthcoming volume.

Payment methods

Paying by cheque
If you pay by cheque, the cheque must be in **pounds sterling**, made payable to *Direct Mail Works Ltd* and drawn on a bank with a UK address.

Paying by credit card
The following cards are accepted by phone, fax, post or via the website ordering pages: Delta, Eurocard, Mastercard, Solo, Switch and Visa. We advise against sending credit card details in a plain email.

Paying by official purchase order
You can post or fax these, but they must be from public bodies (i.e. NHS or universities) within the UK. We cannot at present accept purchase orders from commercial companies or from outside the UK.

How do I get a copy of *HTA on CD*?

Please use the form on the HTA website (www.hta.ac.uk/htacd.htm). Or contact Direct Mail Works (see contact details above) by email, post, fax or phone. *HTA on CD* is currently free of charge worldwide.

The website also provides information about the HTA Programme and lists the membership of the various committees.

The estimation of marginal time preference in a UK-wide sample (TEMPUS) project

JA Cairns*
MM van der Pol

Health Economics Research Unit, University of Aberdeen, UK

* Corresponding author

Competing interests: none declared

Published January 2000

This report should be referenced as follows:

Cairns JA, van der Pol MM. The estimation of marginal time preference in a UK-wide sample (TEMPUS) project. *Health Technol Assess* 2000;**4**(1).

Health Technology Assessment is indexed in *Index Medicus*/MEDLINE and *Excerpta Medica*/EMBASE. Copies of the Executive Summaries are available from the NCCHTA web site (see overleaf).

NHS R&D HTA Programme

The overall aim of the NHS R&D Health Technology Assessment (HTA) programme is to ensure that high-quality research information on the costs, effectiveness and broader impact of health technologies is produced in the most efficient way for those who use, manage and work in the NHS. Research is undertaken in those areas where the evidence will lead to the greatest benefits to patients, either through improved patient outcomes or the most efficient use of NHS resources.

The Standing Group on Health Technology advises on national priorities for health technology assessment. Six advisory panels assist the Standing Group in identifying and prioritising projects. These priorities are then considered by the HTA Commissioning Board supported by the National Coordinating Centre for HTA (NCCHTA).

This report is one of a series covering acute care, diagnostics and imaging, methodology, pharmaceuticals, population screening, and primary and community care. It was identified as a priority by the Methodology Group and funded as project number 94/35/01.

The views expressed in this publication are those of the authors and not necessarily those of the Standing Group, the Commissioning Board, the Panel members or the Department of Health. The editors wish to emphasise that funding and publication of this research by the NHS should not be taken as implicit support for the recommendations for policy contained herein. In particular, policy options in the area of screening will be considered by the National Screening Committee. This Committee, chaired by the Chief Medical Officer, will take into account the views expressed here, further available evidence and other relevant considerations.

Reviews in *Health Technology Assessment* are termed 'systematic' when the account of the search, appraisal and synthesis methods (to minimise biases and random errors) would, in theory, permit the replication of the review by others.

Criteria for inclusion in the HTA monograph series
Reports are published in the HTA monograph series if (1) they have resulted from work either prioritised by the Standing Group on Health Technology, or otherwise commissioned for the HTA Programme, and (2) they are of a sufficiently high scientific quality as assessed by the referees and editors.

Series Editors: Andrew Stevens, Ruairidh Milne and Ken Stein
Monograph Editorial Manager: Melanie Corris

The editors have tried to ensure the accuracy of this report but cannot accept responsibility for any errors or omissions. They would like to thank the referees for their constructive comments on the draft document.

ISSN 1366-5278

Published by Core Research, Alton, on behalf of the NCCHTA.
Printed on acid-free paper in the UK by The Basingstoke Press, Basingstoke.

Copies of this report can be obtained from:

The National Coordinating Centre for Health Technology Assessment,
Mailpoint 728, Boldrewood,
University of Southampton,
Southampton, SO16 7PX, UK.
Fax: +44 (0) 23 8059 5639 Email: hta@soton.ac.uk
http://www.hta.nhsweb.nhs.uk

Contents

List of abbreviations

CI	confidence interval*
df	degrees of freedom*
DU	discounted utility [model]
GLS	generalised least squares
LR	likelihood ratio*
OLS	ordinary least squares
QALY	quality-adjusted life-year
RIGLS	restrictive iterative generalised least squares
TEMPUS	The Estimation of Marginal Time Preference in a UK-wide Sample [project]
VAS	visual analogue scale

* Used only in figures and tables

Executive summary

Background

Generally, any individual would prefer to receive a benefit today rather than in the future and to incur a cost later rather than sooner. Economists call these time preferences. Such preferences are relevant in two ways in the context of health care. First, how individuals view future costs and benefits influences health-affecting behaviour like smoking, exercising and following dietary restrictions. Information on peoples' time preferences could help us to understand health-affecting behaviour and therefore be valuable with respect to the design of policies for the promotion of health. Second, because timing matters, and because different interventions have different time profiles of costs and benefits, methods are required to take into account the timing of costs and benefits when undertaking economic evaluation of healthcare interventions. This is achieved by discounting future costs and benefits to present values by attaching smaller weights to future events the further into the future they occur.

Objectives

1. To derive implied discount rates for future health benefits for a sample of the general public in the UK.
2. To establish whether individual inter-temporal preferences with respect to their own health differ from those with respect to the health of others.
3. To investigate the effect of different ways of asking questions on apparent inter-temporal preferences (specifically closed-ended and open-ended methods are compared).
4. To establish whether individuals value future health benefits in line with the traditional discounted utility model and to investigate, in addition, how well the hyperbolic discounting models explain individual responses.

Methods

Stated preference techniques comprising a series of health-related choices were used to elicit the time preferences of a random sample of adults. Two methods were used: an open-ended method and a discrete choice experiment (closed-ended method). Preferences were elicited for non-fatal changes in own health and others' health. Four different postal questionnaires were sent to a random sample of 5120 adults in England, Scotland and Wales. The data were analysed using a number of forms of regression analysis.

Results and conclusions

The median implied discount rates were 6.1% for own health and 6.2% for others' health using the open-ended method and, in the discrete choice experiment, 5.0%, 4.6%, 3.8% (5-, 8- and 13-year delay, respectively) for own health and 6.4%, 5.7%, 3.8% for others' health.

The results suggest that the implied discount rates for own and others' health are broadly similar. There are some differences but the similarities are much more striking, certainly in the case of the open-ended method.

The implied discount rates and the distribution of the implied discount were very similar for the open-ended method and the discrete choice experiment. The discrete choice experiment had a higher response rate and respondents considered that the discrete choices questions were easier to answer.

The results provide evidence against the discounted utility model. The key axiom of the discounted utility model, stationarity, was violated. The alternative, the hyperbolic discounting models, fitted the data better than the discounted utility model.

The implied discount rates elicited in this study should not be over-emphasised because of the unrepresentativeness of the study sample. However, it is notable how close the estimated median rates are to the rates advocated for use in economic evaluation in a range of countries (for example, 3% in the USA, 5% in Australia and Canada). The estimated implied discount rates in this study fall comfortably within the range of estimates from previous empirical studies.

Research recommendations

A single, albeit multifaceted, project such as TEMPUS adds significantly to our understanding but cannot by itself resolve the outstanding research issues, particularly as this is the first study in which a number of these issues have been addressed systematically. Three areas should be highlighted.

1. Continued refinement of the methods of eliciting time preferences is required. Relevant topics include the use of self-completed questionnaires versus interviews (face-to-face and telephonic) and the presence and impact of framing effects.
2. Further research is required on alternative models of time preference, in particular, models which allow for decreasing timing aversion. Also, the implications of using alternative models for policy making need to be investigated.
3. There is considerable scope for research to investigate the role played by time preference in explaining health-affecting behaviour. To what extent are individuals willing to incur short-term costs in order to secure longer-term benefits – for example, in the successful control of blood sugar levels by patients with diabetes or by participation in screening programmes?

Chapter 1

Introduction

Time preference and health

Generally, any individual would prefer to receive a benefit today rather than in the future and to incur a cost later rather than sooner. Economists call these time preferences. Such preferences are relevant in two ways in the context of health care. First, how individuals view future costs and benefits influences health-affecting behaviour like smoking, exercising and following dietary restrictions. Information on peoples' time preferences could help us to understand health-affecting behaviour and therefore be valuable with respect to the design of policies for the promotion of health. Second, because timing matters, and because different interventions have different time profiles of costs and benefits, methods are required to take into account the timing of costs and benefits when undertaking economic evaluation of healthcare interventions. This is achieved by discounting future costs and benefits to present values by attaching smaller and smaller weights to future events the further in the future they occur. These declining weights or discount factors are equal to $(1 + r)^{-t}$, where r is the discount rate and t the year in which the event occurs.

Rationale for discounting

It is conventional to explain the practice of discounting either in terms of social choices or in terms of individual preferences. As an economy it is possible to defer consumption and undertake investment so that a higher level of future consumption can be enjoyed. Thus the opportunity cost of current consumption is some higher level of future consumption and the discount factor is a formal recognition of this opportunity cost. At the individual level it is suggested that most people have time preferences, that is, they are not indifferent to the timing of future events. The reasons suggested for this include:

- individuals cannot be sure that they will be alive at any particular point in the future
- they expect to better off and will, as a result, attach less weight to further increments to their wealth or income
- what Pigou famously described as a defective telescopic faculty.

Time preference and health-affecting behaviour

Time preference influences health-affecting behaviour like smoking, exercising and following dietary restrictions. For instance, giving up smoking involves costs now in terms of foregone pleasure and the experience of withdrawal symptoms, while the benefits, such as improved quality of life and increased life expectancy, are largely in the future. Information on individuals' time preferences could help to understand health-affecting behaviour and therefore be valuable with respect to the design of policies for the promotion of health. For example, giving individuals with high rates of time preference (those who are more present-oriented) more information on the long-term consequences of their behaviour may have little impact on their short-term decision making. There have been a few empirical studies in which the interaction of time preference and health-affecting behaviour has been investigated (see chapter 2 for a brief overview of this literature).

Discounting in economic evaluations

Discounting practices often play a central role in determining the relative cost-effectiveness of different interventions. If evaluations are undertaken on an incorrect basis, the quality of decision making will suffer and health service efficiency will be reduced. Moreover, confusion or lack of agreement over standard discounting practice potentially undermines the credibility and value of economic evaluation. Different rates are applied in different jurisdictions but they are generally between 3% and 6%. The impact of discounting is especially marked when considering projects in which some of the effects are fairly far in the future. For example, the discount factor applied to costs arising in 20 years time would be 0.31 at the currently recommended UK discount rate ($r = 0.06$).

The appropriate relationship between individual time preferences and the social rate of discount has been the subject of debate for many years. This study does not address the normative question of the appropriate discount rate to use in economic evaluations but rather focuses on a number of specific questions concerning the nature of individual time preferences. However, information on these preferences can be seen, depending on one's viewpoint, as an input to discussions about the appropriate rate of social discount.

One normative view is that the rate that best represents peoples' preferences should be used to discount future health benefits. This viewpoint has resulted in a growing literature on the estimation of individuals' time preferences for health. A distinction can be drawn between those studies relating to life saving and those relating to non-fatal changes in health. These studies are reviewed in the next chapter under these two headings. It should be noted that time preferences measured empirically often include more than a pure time preference effect. Gafni and Torrance[1] hypothesised that an individual's preferences for health extended in time is the composite of a time preference effect, the effect of diminishing marginal utility and an uncertainty effect.

Research issues addressed by the TEMPUS project

Despite the fact that time preferences for health have been estimated previously current knowledge regarding the nature of these preferences is poor. Systematic investigation of the influence of type of choice on apparent time preferences, the characteristics of different methods of eliciting preferences, and the underlying models of time preferences for analysing responses, is required. The research issues explored in this, the TEMPUS (The Estimation of Marginal Time Preference in a UK-wide Sample) project, are as follows.

Different methods of eliciting time preferences

Little is known about the influence of different methods of eliciting preferences on estimates of implied discount rates. Identification of the direction and size of any bias introduced as a consequence of using a particular method is important because it can assist in the interpretation of the results of previous studies and the design of new studies. A major issue in the field of contingent valuation has been the use of closed-ended (Are you willing to pay £100?) versus open-ended questions (How much are you willing to pay?). In general, closed-ended methods generate a higher willingness-to-pay than those that are open-ended. This is largely the result of a difference in the incentives facing respondents to act strategically. Such differences are less likely to be present in a time preference context. In respect of the elicitation of time preferences, open-ended methods[2–10] have proved more popular to date than closed-ended[11–17] – at least, if those studies adopting a more indirect approach (via the valuation of different health outcomes[18–22]) are classified as open-ended. Comparison of the two types of study is difficult because of the many differences between individual studies. No studies have been published which were designed explicitly to compare different methods of eliciting time preferences.

Further issues relate to the use of self-completed questionnaires versus interviews (face-to-face and telephonic), and the presence and impact of framing effects. Both of these have yet to be systematically investigated. While there are undoubtedly lessons to be learned from work in other fields, such as contingent valuation, there is clearly considerable scope for designing time preference studies to address these issues directly.

Nature of the inter-temporal choice

An important question about which little is known is the extent and nature of differences between time preferences with respect to one's own health and to others' health. The distinction between own and others' health is important because time preferences relating to own health are likely to be more relevant in explaining health-affecting behaviour, whereas preferences relating to others' health are potentially more important if the focus is on the evaluation of publicly funded healthcare programmes. Previous studies have elicited time preferences either for own or others' health. Those studies relating to saving statistical lives clearly refer to the health of others.[2–5,11–15] Generally, questions with respect to non-fatal changes in health have been posed in terms of the respondents' own health.[6–8,19–22] Because of the many differences between the studies in terms of design and methods, no strong conclusions regarding differences between time preferences for own and others' health can be drawn. It is therefore important to design a study in which time preferences relating to own health are explicitly compared with time preferences relating to others' health.

Further issues concern the nature and extent of differences between time preferences for saving lives and those for non-fatal changes in health, and whether time preferences are to some extent health state specific. Loewenstein claims that "the dependence of discounting on the characteristics of awaited consumption means that discount rates estimated in specific contexts … cannot be generalised beyond the domain of behaviour in which they were derived".[23] Although this remark was made in the context of a comparison of the purchase of consumer durables and savings behaviour, it does raise the issue of whether implied discount rates are similar within the health domain. The health states examined to date have included an arthritic condition, a strangulated internal hernia, a depressive illness, inflammatory bowel disease, migraine and temporary blindness.

Underlying model of time preferences

As emphasised in the next chapter, economists have generally assumed a particular underlying model of time preference (the discounted utility (DU) model). It has been argued that the DU model is the only dynamically consistent model and, thus, the only appropriate normative model. This view has been challenged by some economists (see, for example, Ahlbrecht and Weber[24]). Moreover, it has become clear that the DU model performs poorly as a descriptive model. For example, a common finding is that implied discount rates are a decreasing function of the period of delay. In order to improve understanding of health-affecting behaviour, it is important to have a model which accurately describes individuals' time preferences and the DU model is therefore less appropriate. This necessitates investigation of alternative discounting models. Alternative models of time preference have been developed in the area of psychology called hyperbolic discounting models, and research has shown that these models have a better goodness of fit than the DU model. However, to date there have been few investigations of these alternative models in a health context.[4,5,12] Studies in humans have generally considered financial consequences and never involved future health effects. It is therefore important to investigate the descriptive properties of these models in the health context.

Further issues include whether time preferences for gains are different from time preferences for losses, and whether time preferences are a function of the magnitude of the consequence. These issues have been investigated for monetary consequences but not systematically for health consequences.

Chapter 2

The estimation of time preferences

Introduction

An overview of the largely recent but rapidly expanding literature on the estimation of time preferences for health is provided in this chapter. Specifically, the two broad approaches – revealed preference and stated preference – are outlined, the DU model and hyperbolic alternatives are introduced, consideration is given to whether time preferences for health are different, and the studies in which stated preference methods have been used to elicit time preferences for future health events are briefly reviewed. This brief literature review indicates that although time preference for health has been a vigorous field of enquiry, many issues remain to be addressed.

Revealed preference versus stated preference

Two broad approaches were used to estimate time preference rates – revealed preference and stated preference. In the former actual behaviour is observed, specifically inter-temporal decisions, whereas in the latter individuals are asked what they would do in particular hypothetical circumstances. Despite a predisposition in favour of revealed preference, economists have, in recent years, shown an increasing willingness to explore the stated preference approach. This is particularly true of health where, partly as a result of there not being markets for health, and frequently not even for health care, there are relatively few opportunities for individuals to reveal their time preferences. There are still concerns about the validity of the information generated and the ideal corroborating evidence remains observed behaviour.

A wide range of behaviours have been studied including the purchase of consumer durables,[25–28] educational investment decisions,[29] food consumption,[30] and labour market wage-risk choices.[31–33] These studies are, in general, based on larger sample sizes than those used in applications of the stated preference approach. Also, the estimation of discount rates is relatively indirect and quite complicated. This results partly from the difficulty of using data collected primarily for other purposes and the many more factors beyond the researchers' control (as compared with an experimental approach).

The stated preference approach has also been applied over a wide range of settings. These have included financial choices,[34–39] purchases of consumer durables,[40,41] saving lives, and non-fatal changes in health. Studies in the last two groups are, of course, of particular relevance.

The DU model

Underlying both these approaches has been a reliance on the DU model, which has dominated economic thought with respect to inter-temporal choice for over 50 years. The most important assumption made in the model is that individuals discount future events at a constant rate. The key axiom of the DU model is **stationarity** – the assumption that preference between two outcomes depends only on the absolute time interval separating them. However, in practice, preferences between two delayed outcomes often switch when both delays are incremented by a given constant amount. Loewenstein and Prelec[42] refer to this as the common difference effect. The common difference effect implies that discount rates should decrease as a function of the time delay over which they are estimated.

Although many studies provide evidence of discount rates decreasing over time, in only a few has the axiom of stationarity been explicitly tested. Two studies have directly tested for preference reversal or the common difference effect. Kirby and Herrnstein[43] offered participants choices between delayed monetary rewards or goods while manipulating the delays to those rewards. In all three of their experiments, the majority of participants reversed their preferences systematically as a function of delay. In the study by Green and colleagues,[44] participants made choices between hypothetical monetary rewards available after different delays. These authors found substantial evidence of preference reversal as equal increments were added to both delays.

Ahlbrecht and Weber[45] tested the stationarity axiom by looking for evidence of short-long term asymmetry in implied discount rates, derived from experiments in which participants had to match future risky outcomes and from experiments in which they had to choose between risky future outcomes. They found that the stationarity axiom was violated in the former case but not the latter.

To date, three studies have tested the stationarity axiom in a health context. Cairns and van der Pol[5] investigated preferences for saving future statistical lives with data elicited using open-ended methods. A broadly similar approach was adopted by Bleichrodt and Johannesson[46] but using dichotomous choice data derived from choices between different health profiles. Christensen-Szalanski[47] tested whether women's preferences for avoiding anaesthesia reverse during childbirth. In all of these studies the stationarity axiom was found to be violated.

Loewenstein and Prelec[42] identified three other inter-temporal choice anomalies that run counter to the predictions of the model. They described these as the absolute magnitude effect, the gain–loss asymmetry, and the delay–speed-up asymmetry. The DU model assumes that the discount rate applied will not be related to the magnitude of the event which is subject to discounting, or to whether the event represents a gain or a loss, or to whether it is being brought forward or delayed. Further anomalies, highlighted by Roelofsma,[48] are intransitive choice patterns, immediacy effects and what he described as "non-corresponding inverses in time-outcome value functions". This last anomaly is the situation in which the implied rate of discount differs depending on whether individuals are asked to adjust the delay between consequences or the magnitude of the consequences in order to be indifferent between two options.

These anomalies are explained by Loewenstein and Prelec[42] in general terms with reference to future consumption, and supported by evidence from monetary choices. However, there is no reason to suppose that they are any less in evidence when the outcomes are in terms of health. As shown above, there is already evidence that the axiom of stationarity is also violated in the case of health. Of course, evidence that the DU model is an unsatisfactory representation of individuals' time preferences with respect to health need not reduce the appeal of DU as a normative model. However, good representations of such preferences are required in order to provide greater insights into the determinants of individuals' health-affecting behaviour.

Hyperbolic models

Hyperbolic models (which do not require the assumption of stationarity) have been explored in several studies in the psychology literature on inter-temporal preferences for hypothetical monetary rewards.[49–54] In a number of studies hyperbolic models have been explicitly compared with the exponential model.[24,37,39,50] The authors of all these studies concluded that hyperbolic models fit the data better than the DU model.

Myerson and Green[37] fitted exponential and hyperbolic models to data from 12 undergraduates who had been asked to choose between pairs of hypothetical sums of money available after different delays. Hyperbola-like functions provided a better description of the data than exponential functions both for individuals and for the group as a whole.

Kirby and Marakovíc[39] reported two experiments, one using hypothetical monetary rewards and the other using real rewards in a simulated auction. The participants (22 college students) were offered five different rewards, each with six different delays. They had to indicate the smallest amount that they would accept immediately in exchange for the delayed rewards. Both hyperbolic and exponential models fitted the data very well but the hyperbolic function fitted better for all delayed rewards.

Albrecht and Weber,[24] using data from studies by Benzion and colleagues[36] and Shelley,[55] compared a hyperbolic model and the standard discounting model. They found the empirical data were more consistent with a hyperbolic rather than the standard discounting model.

These studies are of relatively limited relevance to health economics since the majority elicit inter-temporal preferences for monetary awards and none elicit inter-temporal preferences for health. Also, the econometric modelling, especially in the instances where discounting functions are fitted on an individual basis, is dubious owing to the small number of observations.

Cairns and van der Pol[4] analysed data on preferences for future financial and health benefits collected from 473 members of the general public. They compared the exponential model with two hyperbolic models: the Loewenstein and Prelec[42] model and a special case of this model introduced by Rachlin.[56] They found greater support for the hyperbolic models than for the exponential model. Cropper and colleagues[14] assumed an exponential and a hyperbolic discounting function when analysing their closed-ended data. The exponential function fitted the data better for delays equal to or shorter than 25 years while the hyperbolic function fitted the data better for delays greater than 25 years.

Is health different?

Before outlining the work done on the estimation of time preferences for health, it is worth addressing the basic question – is health different? There are at least two ways in which this question can be addressed – as a practical issue arising in economic evaluation and at a more methodological level.

The **practical** question is whether or not future health effects should be discounted at the same rate as future costs. This question has generated considerable interest in recent years and much of the empirical work that has been undertaken has aimed to inform this debate.

Parsonage and Neuberger[57] claimed that "… in practice, for most purposes, it is appropriate to use a zero discount rate for future health benefits". They argued that traditional sources of time preference are unimportant in the context of future health benefits. Whether or not monetary costs and non-monetary benefits should be discounted at the same rate can be viewed as largely an empirical question.[58] Is the sum of undiscounted health benefits a better or a poorer approximation to the true value of a stream of future benefits than the sum of the discounted health benefits? The answer would appear to depend on the shadow price of the health benefits and how it changes over time. The greater the increase in shadow price over time, the closer the undiscounted sum approximates to the true value. The smaller the increase in shadow price the closer the approximation provided by the discounted sum. A recent contribution to this debate by van Hout[59] made the point more formally: "… costs and health benefits need to be discounted using rates that are not necessarily equal; both should be based on their expected rates of growth and their elasticities regarding the social utility function". van Hout also introduced an inter-generational argument emphasising the need not only to reflect individual preferences about one's own health but also individuals' social values about future generations.

Currently the only jurisdiction in which there is formal guidance to encourage the differential discounting of monetary costs and non-monetary benefits is England and Wales, where 6% is recommended for financial values and 1 ½–2% for health effects quantified in physical units.[60] The advice is at odds with that given elsewhere, for example, in Canada[61] and the USA[62] (5% and 3%, respectively, applied to all costs and benefits).

The **methodological** issue relates to whether or not it is possible to identify individuals' rates of time preference with respect to future health effects and, in the context of this report, is rather more important. Gafni and Torrance[1] hinted at the potential difficulties in so doing and Gafni[63] developed the argument more fully. Gafni and Torrance related attitude towards risk in health to three distinct effects: a quantity effect, a time preference effect, and a gambling effect. It is interesting to note that Gafni and Torrance suggested that "time preference is measured by asking conventional time preference questions… but cast in the health, as opposed to financial domain". Also, they claimed that it was not necessary to speculate on the nature of time preference "…since it is empirically determinable". However, drawing on the study by Loewenstein and Prelec,[64] which highlighted the importance of another class of effects – sequence effects, Gafni[63] argued robustly that no measurement technique allows pure time preference to be distinguished, and the best that can be achieved is a measure of time preference for a given sequence of events. This may be true of preferences over one's own future health states. However, it is less clear that the sequence of events will be an important influence when considering preferences over life-saving profiles.

A further way in which some health benefits may be different arises when the change in health is measured in terms of quality-adjusted life-years (QALYs). Krahn and Gafni[65] argued that, because of the methods sometimes used to measure QALYs, time preferences may already have been taken into account. As a result further discounting of the QALYs might represent a form of double counting.

Using stated preference to estimate time preferences for health

As noted above, a distinction can be drawn between those studies relating to life saving and those relating to non-fatal changes in health. A brief overview of this literature is presented in *Table 1*.

TABLE 1 Summary of empirical time preference literature in health

Study	Median r	Mean r	Type	Delay	Sample
Cairns[6]	–	–0.001–0.030	Health states	10–28 years	29 (economics undergraduates)
Chapman[8]	0.200–0.350	0.300–0.500	Health states	1–12 years	148 (psychology undergraduates)
Chapman & Elstein[7]	0.360 & 1.000	0.640 & 1.240	Health states	1–12 years	104 (psychology undergraduates)
Chapman, *et al.*[10]	0.06–0.09[a]	0.40–0.83[a]	Health states	1–6 months	79 (patients) + 77 (college students)
Dolan & Gudex[22]	0.000	–0.029–0.014	Health states	9 years	39 (general public)
Lipscomb[19]	–	–	Health states	1–25 years	52 (undergraduate students)
MacKeigan, *et al.*[21]	–	–	Health states	1 week–1 year	108 (university staff, hospital volunteers)
Olsen[9]	0.02	0.10	Health states	5 & 20 years	90 (economics students) + 40 (doctors)
Redelmeier & Heller[20]	–	0.023–0.041	Health states	1 day–10 years	121 (medical students, house officers, physicians)
Cairns & van der Pol[4,5]	0.160–0.410	0.140–0.450	Lives	2–19 years	473 (general public)
Cairns[3]	0.160–0.380	0.140–0.370	Lives	4–19 years	223 (general public)
Cropper, *et al.*[12]	–	0.027–0.086	Lives	25–100 years	1600 (general public)
Cropper, *et al.*[13]	0.168–0.038	–	Lives	25–100 years	3200 (general public)
Enemark, *et al.*[18]	0.104	0.102	Lives	≈10 years	25 (vascular surgeons)
Horowitz & Carson[11]	0.045	–	Lives	5 years	75 (economics undergraduates)
Johannesson & Johansson[15]	0.080–0.250	–	Lives	20–100 years	850 (general public)
Olsen[2]	0.058–0.229	0.066–0.233	Lives and health states	4–19 years	250 (general public) + 77 (health planners)
Johannesson & Johansson[16]	–	0.013	Life-years	10–46 years	528 (general public)
Johannesson & Johansson[17]	–	0.010	Life-years	6–57 years	2577 (general public)

[a] *Monthly not annual discount rate*

Saving lives

The first studies of time preferences with respect to life saving[11–14] adopted fairly similar approaches. Horowitz and Carson[11] offered respondents dichotomous choices between a programme saving lives for the next 15 years and a programme saving lives for 10 years starting in 5 years' time. Cropper and colleagues[13] asked two linked dichotomous questions that offered a context-free choice between programmes saving lives now and at various times in the future. In both these studies, econometric methods were then used to identify median implied time preferences rates for the sample interviewed. A broadly similar approach was followed in a recent Swedish study[15] of intergenerational choices.

A novel approach adopted by Johannesson and Johansson[16,17] was to elicit time preferences via willingness-to-pay values. Individuals were asked whether they were willing to pay a specified amount for a 1-year increase in their life expectancy. The rate of time preference was estimated by calculating the ratio of the mean willingness-to-pay values for different ages.

Later studies by Olsen,[2] Cairns[3] and Cairns and van der Pol[4,5] used variants of an open-ended approach in which the data were collected with a self-completed questionnaire. Through this approach, the identification of rates of time preference can be made on an individual basis through simple manipulation of the responses. The price of this advantage may be increased problems with respect to framing effects and related biases.

A third approach followed by Enemark and colleagues[18] differed in that it was not concerned with choices between saving differing numbers of statistical lives at different times in the future. It related to the choice between watchful waiting and surgery for abdominal aortic aneurysms for a hypothetical cohort of patients. In a classification between life-saving and non-fatal changes in health, it is probably appropriate to put it with the life-saving studies but it is rather different from the other studies and, as such, may offer a new line of enquiry.

Non-fatal changes in health

The first significant study of time preferences for non-fatal changes in health was by Lipscomb.[19] His approach required participants to classify 96 scenarios on a 0–10 category scale. Regression analysis was then used to explain these scores using the characteristics of the health state. Lipscomb suggested that time preference "… be regarded as operationally equivalent to the (marginal) influence of the delay-of-onset variable on preference scores". Notwithstanding the practical difficulties of applying it (such as the demands on individual respondents), this approach represents an attractive means of determining the impact of delay on the valuation of an outcome.

In Cairns,[6] implied discount rates were calculated directly from the duration of ill health which rendered the respondent indifferent between being ill in the further future and in the near future. A broadly similar approach was followed by Chapman and Elstein[7] and by Chapman,[8] in which participants indicated what period of relief from ill health at a specified point in the future would make them indifferent between relief from ill health for a specified period now. Chapman and colleagues[10] introduced a slightly different approach in which respondents identified the frequency of symptoms which would render them indifferent between two ill health profiles.

Redelmeier and Heller[20] used the standard gamble and categorical scaling to elicit preferences for a number of different health states, one aspect of which was the timing of onset of ill health. MacKeigan and colleagues[21] also used a categorical rating scale to elicit preferences for different health states with different delays of onset. They examined the effect of duration of health gain and health loss, and delay before health change, on the valuation of health states but did not estimate the implied discount rates.

The time trade-off method of health state valuation was used in two studies to identify implied discount rates. Olsen[9] elicited time preferences through two time trade-off questions involving different durations of the same health state. The implied discount rate was that which ensured that the estimated health state value was the same for two durations. Dolan and Gudex[22] estimated time preference rates using time trade-off valuations for two health profiles which differed only in the timing of the period of ill health.

Time preference and health-affecting behaviour

One of the major reasons for being interested in time preferences is the potential role these preferences might play in terms of understanding individual health-affecting behaviour. An early start was made by Fuchs[35] in exploring the relationship between time preferences and health-affecting behaviour. Discount rates were estimated from answers to pair-wise choices offering the opportunity to delay the receipt of a money prize between 1 and 5 years in order to receive a larger prize. Five health behaviours were considered: smoking, weight, time since last dental check-up, frequency of exercise and wearing of seat belts. A significant positive association was found between time preference and smoking. However, there was little evidence of an association with the other health behaviours.

In subsequent years, research has focused on the estimation of time preferences for health events rather than on the influence of time preference on behaviour. However, in three further studies the relationship between money–time preferences and health-affecting behaviour has recently been examined. Vuchinich and Simpson[50] examined the time preferences of college students classified as light and heavy drinkers. Time preferences were elicited using questions involving the speeding-up of the receipt of money over periods ranging from 1 week to 25 years. They found that a hyperbolic model fitted the data better than the exponential model and that heavy drinkers exhibited higher hyperbolic discounting than light drinkers.

Bretteville-Jensen[66] estimated an annual and a weekly rate of time preference for injecting drug addicts, former drug users and non-users. Time preferences were elicited using questions concerning the speeding-up of the receipt of money. Active drug users were found to have higher discount rates than former users, who in turn had higher rates than non-users.

Chapman and Coups[67] investigated the time preferences of a workforce, some of whom had accepted and some declined the offer of a free influenza vaccination. They elicited time preferences with respect to both a monetary and a health choice. The former involved delaying the payment of a fine and the latter delaying a period of ill-health for 3 months. A striking feature of this study was the very large proportion of respondents with a zero rate of time preference (greater than 80%). Those expressing zero time preference for money were significantly more likely to have accepted a vaccination. No significant association was found in the case of health time preferences.

While in all of these studies some evidence of the expected association was found, in the cases of smoking, drinking and injecting behaviour it is not possible to determine whether the health-affecting behaviour influences time preferences or whether time preferences influence health-affecting behaviour. However, in the case of influenza vaccinations it seems implausible that causation could run from the vaccination decision to time preference.

Summary

This overview of the literature on the estimation of time preferences includes outlines of stated versus revealed preference methods and the DU and hyperbolic discounting models. The literature on the differences in time preferences for health and monetary consequences is reviewed. The stated preference approaches that have been used to estimate time preferences with respect to health are then introduced. As befits a new area of research, a range of approaches has been adopted and, as a result, comparisons between studies are hampered by differences in methods and in the periods of delay considered. Although many interesting questions are raised, there are, as yet, no adequate answers. The current study attempts to provide some of these.

Chapter 3

Study design

Introduction

The basic study design of the TEMPUS project is described in this chapter. The specific objectives of the project and the minimum requirements, given these objectives, are discussed. The questionnaires are introduced and the choices to be made when designing time preference questions are examined. A discussion of the health state used in this study is followed by the general questions included in all questionnaires. The method of data collection and the pilot studies are described.

Objectives of the TEMPUS project

1. To derive implied discount rates for future health benefits for a sample of the general public in the UK.
2. To establish whether individual inter-temporal preferences with respect to their own health differ from those with respect to the health of others.
3. To investigate the effect of different ways of asking questions on apparent inter-temporal preferences (specifically closed-ended and open-ended methods are compared).
4. To establish whether individuals value future health benefits in line with the traditional DU model and, in addition, to investigate how well hyperbolic discounting models explain individual responses.

Stated preference techniques comprising a series of health-related choices were used to elicit the time preferences of a random sample of adults. In order to meet the four objectives, the following were required:

- inter-temporal preferences from a UK-wide sample (objective 1)
- inter-temporal preferences for both own and others' health (objective 2)
- both an open-ended and a closed-ended method (objective 3); the closed-ended method used in this study was a discrete choice experiment
- several observations of implied discount rates per respondent over a wide range of delays (this is explained below) (objective 4); these can only be elicited using an open-ended method.

The four questionnaires

At first sight, the strongest design for investigating differences between preferences for own and others' health, and for investigating differences between an open- and a closed-ended method, would involve eliciting all types of preference from all subjects. This would allow comparisons on an individual level. However, this would clearly ask too much from the respondents. Also, it could induce strategic or biased responses. For example, if asked about own and others' health in the same questionnaire, some respondents might feel that they did not wish to be seen as treating others' health differently from their own. There are four different types of questionnaire:

(I) open-ended questions with respect to **own** health
(II) open-ended questions with respect to **others'** health
(III) discrete choices with respect to **own** health
(IV) discrete choices with respect to **others'** health.

Design choices

Several choices have to be made when designing either open-ended or closed-ended time preference questions. Most are relevant to both types of questionnaire but some are specific to one or other question type.

The following are potentially relevant to both types of question.

- ***Number of points in time***: respondents can be asked to consider two points in time or they can be presented with a profile. The standard approach has been the former, with few studies comparing profiles.[10,11,18]
- ***Base health state***: the base health state can be full health with respondents making choices with respect to the consumption of ill health,[6,20,21] or the base health state is ill health with respondents making choices with respect to the consumption of full health.[7,8,21]
- ***Number of states of ill health considered***: one state[6–8,21] or more than one state of ill health[19,20,22] can be considered.

- ***Single health state or mix of health states***: respondents could be invited to imagine being in only one health state in a year or experiencing several different health states in a year (all studies to date have used single health states).
- ***Time considered***: a limited period can be considered[6–8,20,21] (for instance, 5 years), or a scenario can describe remaining life.[18,22]

In open-ended questions, consideration must, in addition, be given to whether respondents are asked about: the **timing** of a given change in health; or the **magnitude** to be experienced at a certain point in time; or possibly the health-related **quality of life** to be experienced. In studies to date, individuals have been asked to specify the magnitude of the health benefit to be enjoyed at a particular point in the future (either in terms of lives saved,[2–5] duration of health state[6–8] or frequency of symptoms[10]). In no study have individuals been asked to specify timing or quality.

There are a number of criteria which are relevant when making these choices:

(i) how difficult the questions are to answer
(ii) the degree of realism
(iii) the ease of computation of discount factor
(iv) the degree of influence of factors other than time preference on response.

The first two are clearly relevant because of the impact that they might have on response rates and on the meaningfulness of responses. Other things being equal, the more straightforward the computation of implied discount factors is the better. Similarly, methods which are less likely to be influenced by factors other than time preferences are to be preferred.

The performance of the various options in respect to these four criteria are shown in appendix 1. There appears to be a trade-off between realism and how easy it is to answer, and between accuracy and how easy it is to answer. The specific approach adopted to elicit time preferences will have implications for the sample size required, the estimation methods to be used and the methods of data collection employed. The determination of the specific approaches to be adopted is clearly a matter of judgement. Most of the decisions made were in line with those generally adopted in the literature. Thus comparisons were between points in time rather than profiles, departures from full health rather than from ill health were considered, choices involved a single ill health state, and times were limited. In the open-ended questions, individuals were asked to choose durations of ill health.

Health state

The health state selected (which was used in both the open-ended method and the discrete choice experiment) is based on the EuroQol descriptive system (EQ5D). This has five dimensions: mobility, self-care, usual activities, pain/discomfort, anxiety/depression;[68] each dimension has three severity levels. A generic health state classification was chosen rather than a condition-specific one because the participants were members of the general public with widely differing experience and knowledge of specific health conditions. The use of particular conditions might influence the implied discount rates. Also, the more specific the health state, the harder it is to devise a plausible inter-temporal question. The EQ5D classification was selected because of its simplicity and its widespread use.

Individual's inter-temporal preferences are expected to differ according to the severity of the selected health state. Individuals might have a tendency to minimise duration for very serious health states and maximise delay for very minor health states. The aim was to select a health state that is not too severe but is generally regarded as serious enough. The following health state was selected:

- no problems in walking about
- no problems with self-care
- some problems with performing usual activities
- moderate pain or discomfort
- not anxious or depressed.

The tariff for this EQ5D state (11221) is 0.773, assuming a 1-month duration.[69] Individuals were asked to rate the health state using a visual analogue scale (VAS). This gave some insight into how serious they regarded the chosen health state. It also permitted some limited testing of the hypothesised relationship between inter-temporal preferences and perceived severity of the health state.

In the case of own health, individuals were asked to imagine being ill as described by the EQ5D health state. However, the case for others' health is more complicated. If respondents were asked to imagine a person of the same age and gender as themselves, there is a danger that they might answer the questions as if they related to their own health. This could be a problem even when nothing is specified about the individual. In order to reduce the likelihood of this happening, the question was formulated in terms of a group of middle-aged patients rather than in terms of an individual. Their approximate age was stated because time preference is expected to be a function of age.

General questions

Individuals were asked to indicate how they perceived the EQ5D state on a VAS (ranging from worst to best imaginable health). Questions about year of birth, gender, perception of current long-term health, number of cigarettes smoked per day, and education level attained were also included in all of the questionnaires. Individuals were also asked how they perceived the time preference questions on a five-point scale ranging from very difficult (1) to very easy (5).

Data collection

There are basically three methods of collecting data: personal (face-to-face) interviews; telephone interviews; and postal questionnaires. Some researchers consider that the face-to-face interview is a superior method, a view challenged by Sudman and Bradburn.[70] In their extensive review of the literature comparing response effects, they observed only small differences between methods of data collection.

The main advantage of the face-to-face interview is that the interviewer can explain the questions; this is especially useful for questions involving complex scenarios. It is also possible to obtain information on the reasons for an individual's responses, resulting in a richer data set. However, this method can introduce interviewer bias. The main disadvantage of face-to-face interviews is the high cost involved in conducting them, particularly when large numbers are required. This has led to the development of telephone interviews. However, this method is less appropriate when visual aids are required and when questions are complex. One advantage of postal questionnaires is that respondents can take as much time to reflect on the questions as they wish. However, their use can lead to sample non-response bias problems.

In this study, data were collected by postal questionnaire rather than by face-to-face or telephone interviews. No reminders were sent. The inter-temporal choices were viewed as too complex for the use of telephone interviews. While face-to-face interviews might have offered the opportunity to collect richer data per subject, the required sample sizes would have been prohibitively expensive to achieve by any method other than postal questionnaire.

Another issue that needs to be considered when collecting data is whether individuals should be given incentives in particular monetary incentives. It is sometimes argued that monetary incentives are necessary in order to obtain valid responses. However, research has shown that monetary incentives have no real impact on responses.[71,72] In a previous study by Cairns and van der Pol,[4] in which half of the participants were offered a choice of a charity to which £2.00 would be donated, the results indicated that incentives had no impact on either the response rate or the results. Given this, it was decided that individuals would receive no monetary incentives in this study.

Pilot studies

Two pilot studies were undertaken to examine whether the questionnaires were feasible and whether respondents understood the inter-temporal choices.

In the first pilot study, all four questionnaires were handed out to about 30 individuals working in the Department of Public Health at the University of Aberdeen. They were encouraged to make some general comments about the design of the questionnaire.

In a second pilot study, 180 members of the general public in Aberdeen were selected from the telephone book. Questionnaires I and IV were piloted, as well as a further questionnaire which contained open-ended questions relating to both own and others' health. Once a questionnaire was returned, respondents were interviewed by telephone. The issues included in the interview were the realism of the EQ5D health state, the clarity of the questions, whether they understood the questionnaire, and the reasons for their answers. From this pilot study it became clear that respondents found it very difficult to discriminate between the two types of questions (own and others' health); hence, it was decided that the two types of questions would be in separate questionnaires only. Questionnaires I and IV were redesigned after the pilot work. The questions were better explained in questionnaire I and questionnaire IV was changed substantially (see chapter 7); the latter was sent out to a further 60 members of the general public in Aberdeen.

The response rate was 35% for questionnaire I (21 out of 60 questionnaires), 28.3% for questionnaire IV (34 out of 120), and 18.3% (11 out of 60) for the questionnaire containing open-ended questions on both own and others' health.

Summary

Some basic design issues are discussed in this chapter. In order to meet the four specified objectives, four different questionnaires were required. The choices which have to be made when designing time preference questions are discussed. The generic health state used in this study was an EQ5D health state. Data collection was by postal questionnaires. Although this may lead to non-response bias, it was the most efficient way of obtaining the large sample sizes required for this study. The pilot studies undertaken in the TEMPUS study are described; the questionnaires were piloted both on university staff and members of the general public.

Structure of the remainder of this report

The data used in the TEMPUS project are described in chapter 4. The report then focuses on the open-ended method used. The questionnaire design and econometric methods are described in chapter 5, with the results presented in chapter 6. This is then repeated for the discrete choice experiment (chapters 7 and 8). Chapters 5–8 meet the first objective specified on page 11 – the derivation of implied discount rates. The remaining three objectives are dealt with in chapters 9–11. Finally, in chapter 12, some overall conclusions are drawn and areas for future research are identified.

Chapter 4

Data

Introduction

The data collected for the TEMPUS study are described in this chapter. The method used to select subjects is described, the response rates are reported, and the samples are described in terms of demographic and socio-economic characteristics.

Selection of subjects

In previous time preference studies the chosen participants have been university students, university staff, patients, health service employees or members of the general public (see *Table 1*). Ideally, a random sample of members of the general public would be selected as participants since this would allow generalisation of the results. However, as shown in previous studies,[3,4] the use of postal questionnaires to elicit time preferences from the general public results in low response rates (about 25%). The most likely reason for such low response rates is that the questions are complex. This is related in part to the inter-temporal nature of the questions – thinking about the future is not easy. The problem is exacerbated when considering future health events. In a financial context, many familiar instruments exist which facilitate trading consumption over time. It is much harder to produce credible questions or choices in a health context. A solution would be to select participants who are familiar with time preferences, such as economics students. However, this would not allow any generalisation of the results since such students are not at all representative of the general population. Also, young and healthy adults may fail to consider future consequences, as found by Green and colleagues.[49]

The selection of participants for the TEMPUS study was midway between the above selection methods. It was decided that participants would be members of the general public in six urban and rural areas in Scotland, England and Wales (Edinburgh, Caithness, Manchester, Norfolk, Cardiff and Pembrokeshire). However, to decrease the problem of low response rates, wards with a high percentage home ownership (using 1991 Census data) were selected from the electoral registers since, in previous studies,[3,4] it had been found that higher socio-economic status areas had a higher response rate. The disadvantage of this selection method is that the sample is not entirely representative of the UK population; however, it is much more so than a sample of economics students or university staff. Also, representativeness is not of primary importance since the aim of the study is to explore the nature of individuals' time preferences and not to identify the mean or median population discount rates. For instance, for the comparison of two methods of eliciting preferences, the only requirement is that the samples are similar in terms of their characteristics. This also holds for the comparison of time preferences for own and others' health.

About 20,000 names were obtained from the electoral registers for each of the six chosen areas. From these, 854 names were randomly selected from each area. Questionnaires I and II were each sent to 960 individuals (160 individuals per area) and questionnaires III and IV were each sent to 1600 individuals (266/267 individuals per area). The four questionnaires and the different versions of the questionnaires were randomly allocated across and within the six areas.

Response rates

A total of 159 usable responses (where the respondent answered at least some of the time preference questions) were received for questionnaire I, and 149 for questionnaire II. Five and eight questionnaires, respectively, were returned by the Post Office. The response rates were therefore 16.6% and 15.7%, respectively. For questionnaire III, 399 usable responses were received and for questionnaire IV, 388 responses. A total of 15 and 15 questionnaires, respectively, were returned by the Post Office. The response rates were thus 25.2% and 24.5%, respectively.

Descriptive statistics of samples

In *Table 2* some descriptive statistics for the sample are presented. The samples appear to be similar. A chi-squared test was used to test

***TABLE 2** Descriptive statistics*

Questionnaire		n I	n II	n III	n IV	% I	% II	% III	% IV	χ^2
Gender	Male	79	79	182	171	49.7	53.0	45.6	44.1	4.041 (p = 0.26)
	Female	80	70	217	215	50.3	46.0	54.4	55.4	
	Missing value	0	0	0	2	0.0	0.0	0.0	0.5	
Health	Good	129	103	296	267	81.1	69.1	74.2	68.8	10.717 (p = 0.10)
	Fair/poor	29	46	103	121	18.2	30.9	25.8	31.2	
	Missing value	1	0	0	0	0.6	0.0	0.0	0.0	
Smoke	No	136	127	345	327	85.5	85.2	86.5	84.3	4.714 (p = 0.58)
	Yes	22	22	54	61	13.8	14.8	13.5	15.7	
	Missing value	1	0	0	0	0.6	0.0	0.0	0.0	
Education	Secondary	48	43	108	129	30.2	28.9	27.1	33.2	5.435 (p = 0.49)
	Beyond secondary	107	105	285	254	67.3	70.5	71.4	65.5	
	Missing value	4	1	6	5	2.5	0.7	1.5	1.3	
		Median				**Range**				
Age (years)		48	54	448	49	19–90	19–84	18–95	17–94	2.918 (p = 0.40)
	Missing value	n = 0	n = 2	n = 1	n = 2					
VAS score		60	50	50	55	0–100	0–98	10–96	5–97	1.975 (p = 0.58)
	Missing value	n = 25	n = 16	n = 31	n = 49					

the null hypothesis that the distribution of respondents across categories for each of the individual's characteristics is independent of the type of questionnaire. With respect to age and rating of the health state on the VAS, a non-parametric test was used to compare the median of two or more independent samples. These tests were performed on all four samples and the results are shown in *Table 2*. The null hypothesis that the distribution of respondents across categories for each of the individuals' characteristics is independent of the type of questionnaire is accepted. There was also no statistically significant difference in the median age of respondents or in the rating of the health state.

Because the questionnaires were sent to a random sample of the general public (stratified by area), no information is available on non-respondents. The best that can be done is to compare the age, gender and education of the respondents with the 1991 Census data for the geographical areas. It should be noted that the questionnaires were sent out in 1998 while the Census data are from 1991. Also, information on education in the Census data is only available for a sample in smaller areas. A comparison of the sample with the population in terms of age, gender and education is shown in *Table 3*. The age and gender structure of the sample seems fairly representative for the population. However, the younger age group (< 30 years) tends to be under-represented. A higher percentage of respondents have university degrees than in the general population.

Summary

Data for the TEMPUS study were collected from members of the general public in six areas in Scotland, England and Wales. The samples are not representative of the UK population because areas with a higher percentage of home ownership were chosen in order to improve the response rate. The response rates were 16.6% for questionnaire I, 15.7% for questionnaire II, 25.2% for questionnaire III and 24.5% for questionnaire IV. The four samples were very similar in terms of age, gender, self-rated health, education, and smoking status.

***TABLE 3** Representativeness of the sample*

		Caithness		Edinburgh		Norfolk	
		% sample	**Census**	**% sample**	**Census**	**% sample**	**Census**
Age (years)	< 30	13.6	23.6	9.6	20.8	11.8	20.9
	30–44	28.4	25.9	31.3	25.0	24.1	27.8
	45–64	42.6	30.6	32.2	31.0	43.6	31.6
	> 64	14.8	19.9	26.5	23.2	20.0	19.8
Gender	Male	44.3	48.6	49.6	46.2	41.0	49.4
	Female	55.7	51.4	50.0	53.8	58.5	50.6
Education	University	17.6	5.8	27.4	10.4	21.5	7.2
		Manchester		**Cardiff**		**Pembrokeshire**	
		% sample	**Census**	**% sample**	**Census**	**% sample**	**Census**
Age (years)	< 30	28.7	28.9	14.6	17.8	13.2	20.3
	30–44	26.8	27.4	22.1	24.4	25.2	25.3
	45–64	20.4	24.0	30.7	30.5	37.7	31.4
	> 64	23.6	19.7	32.2	27.3	23.2	23.0
Gender	Male	45.2	47.3	47.2	46.1	52.3	48.6
	Female	54.8	52.7	52.3	53.9	47.7	51.4
Education	University	49	16.9	34.7	16.6	25.8	7.4

Chapter 5
Open-ended method

Introduction

As explained in chapter 3, both an open- and a closed-ended method were used in the TEMPUS study to elicit time preferences. The study design for the open-ended method is described in this chapter (the study design for the closed-ended method is described in chapter 7), together with the methods for the basic analysis of the data collected using the open-ended approach. These involve estimating implied discount rates, examining negative and zero time preferences, and identifying the factors that influence individuals' implied discount rates.

Study design

As described in chapter 2, open-ended methods have been used in several studies to elicit time preferences for health. The method adopted in this study was broadly similar to that used by Cairns.[6] Each time preference question asked the respondent to imagine being ill at some point in the future (for x days in year t) and offered the opportunity for this spell of ill health to be delayed (to year s, where $s > t$). Individuals cannot generally delay their ill health so, to make the question more realistic, the respondent was told that the delay was the result of a treatment. However, the disadvantage of introducing a treatment is that it may result in a bias – some individuals may be averse to any treatment, which may influence their responses. To reduce this effect, the respondents were told that it was a one-off minor treatment. Individuals were asked to identify a maximum number of days of future ill health (y) at which it would still be worthwhile receiving this treatment (see appendix 2 for an example). If the number of days of ill health in year s was zero, everyone would probably choose the treatment but, as the number of days of ill-health in that year increases, individuals would at some point no longer prefer to be treated. The interest in this study is in the **maximum** number of days that would still make the treatment worthwhile. The majority of individuals are expected to have positive time preferences and, therefore, be willing to be ill for a longer period ($y > x$) in year s.

Investigation of different discounting models requires a relatively wide range of delays and several observations per respondent. In this study six observations were obtained from each individual. Two different years were chosen as the initial point at which ill health would be experienced if treatment was not received. Having more than two starting points might make the questions too difficult to answer, whereas having a common starting point for all six questions might result in respondents getting bored. The use of two starting points also facilitates the testing of the stationarity assumption (see chapter 11).

The periods of delay chosen ranged from 2 to 13 years from the starting point. The range of delays were selected on the pragmatic grounds that if a delay was too short, factors other than time preference would have a major influence on the responses and, if it were too long, some respondents might have difficulty imagining it (particularly older respondents). Each subject was asked six questions: three with a starting point 2 years in the future and three with a starting point 3 years in the future. Each questionnaire contained three different delays for each starting point: a short-term delay (2–5 years); a medium-term delay (6–9 years); and a long-term delay (10–13 years). Also, the difference between the short- and medium-term delay and between the medium- and long-term delay was not the same in any questionnaire (for each starting point). Four different versions of the questionnaire (A, B, C and D) with these properties are shown in *Table 4*. The purpose of this somewhat complex design is to collect data for a wide range of delays

TABLE 4 *The four versions of the questionnaire (A–D)*

Starting point	Years before delayed ill health			
	A	B	C	D
2 years	2	3	4	5
2 years	7	9	6	8
2 years	10	11	12	13
3 years	3	5	2	2
3 years	6	7	8	9
3 years	10	11	12	13

and to ensure that the questions do not appear to conform to any pattern, in case respondents think that they are expected to respond in a certain way.

Basic analysis

The basic data analysis of the open-ended method is as follows. First, the method of deriving implied discount rates is described. There is evidence of negative and zero time preference in most time preference for health studies.[73] Then the method of identifying and analysing negative and zero time preference in this study is explained, followed by a description of the regression analysis of the implied discount rates on individual characteristics. This provides some insight into which factors determine individuals' time preference.

Estimating implied discount rates

The derivation of implied discount rates is generally straightforward with data elicited with open-ended methods. The individual chooses a duration of ill health, y, such that if it were experienced s years in the future it would be equivalent to 20 days experienced 2 years in the future. Letting $b = 1/(1 + r)$, the respondent's choice implies that $20b^2 = yb^{s-2}$ and

$$r = \left(\frac{y}{20}\right)^{1/s-2} - 1 \qquad (1)$$

The implied discount rates are first represented graphically in a histogram. They are then described in terms of their mean, median and range. Apart from the mean and median, the 5% trimmed mean is estimated. The 5% trimmed mean disregards the smallest 5% and the largest 5% of all observations, and the estimate is therefore not influenced by extreme values.

The Lilliefors test is used to test the hypothesis that the data are from a normal distribution. Skewness and kurtosis of the distribution are also measured. A distribution that is not symmetric but has more observations toward one end of the distribution than the other is skewed. Kurtosis indicates the extent to which, for a given deviation, observations cluster around a central point.

Negative and zero time preference

Individuals may exhibit negative time preference in some instances in the expectation of future unpleasant consequences (dread).[23,74] This is likely to influence time preferences for health. Individuals may prefer to experience a spell of ill health sooner rather than later in order to eliminate dread. To test whether any of the respondents have negative and zero time preferences, the responses to the six questions are classified into six categories:

(i) at least four positive rates and no negative rates
(ii) at least four negative rates and no positive rates
(iii) at least four zero rates
(iv) three zero and three positive rates
(v) three zero and three negative rates
(vi) negative and positive rates.

Respondents were not categorised by their average discount rate because this could be potentially misleading. For example, respondents with five small positive discount rates and one large negative discount rate could have a negative mean discount rate, although they might more appropriately be identified as having positive time preference.

It is interesting to explore the factors that influence whether individuals have negative or zero time preferences. Because of the binary dependent variable (which takes a value of 0 if individuals have positive time preferences and 1 if individuals exhibit either negative or zero time preferences), probit regression analysis is used.

The probit model can be expressed as follows:[75]

$$y_i^* = \alpha + \beta \boldsymbol{x_i} + \varepsilon_i,\ \varepsilon \sim N[0, 1] \qquad (2)$$
$$(y_i = 1 \text{ if } y_i^* > 0;\ 0 \text{ otherwise})$$

where y_i is a binary variable, x_i is a vector of independent variables, and ε_i is a random error term.

The following independent variables are included: health state perception on the VAS (dummy variables: VAS score < 40 and VAS score > 70); age of the respondent (dummy variables: age 30–43 years, age 44–63 years, and age > 63 years); whether the respondent perceived the time preference questions as difficult (dummy variable: degree of difficult < 2.5); and whether or not they had been educated beyond secondary school level (dummy variable: education). The data available on the individuals' gender, smoking status and self-rated health are not included in the model because of lack of strong hypotheses. The choice of the specific age group dummies is, to some extent, arbitrary. In this study, the 25, 50, and 75 percentiles were used which resulted in sensible age groups. The choice of health state perception dummies is based on the mean, plus and minus 15. This gave reasonably sized groups of respondents.

Four specific hypotheses are of interest. First, it is hypothesised that individuals who perceive a health state as more severe are more likely to have negative time preferences because they are likely to experience more dread.[74] Second, individuals who perceive a health state as less severe are more likely to have zero time preferences because they are less likely to be concerned with the timing of a health state. Third, younger individuals are more likely to have negative discount rates because they expect to have more responsibilities in the future. Older individuals are less likely to have negative discount rates because of their reduced life expectancy. Fourth, negative and especially zero time preferences may indicate that individuals did not understand the task. This hypothesis is supported to some extent if individuals who perceive the questions as difficult (and are possibly less well-educated) are more likely to have negative or zero time preferences.

The interpretation of coefficients in probit models is less straightforward than in ordinary least squares (OLS) because of the non-linearity of the relationship between the probability that $y = 1$ and each of the independent variables. The most common method used to interpret the coefficients is by estimating a range of values of dP/dx_i corresponding to various values of x. For dummy variables a common method of interpretation is to estimate the change in probability when the dummy variable changes from zero to one.[76] This is estimated while holding all other independent variables at their mean values. In this analysis, and throughout this report, a 5% level of statistical significance is used. However, it should be recognised that some statistical associations might be explained on the basis of multiple comparisons. To assess the performance of the model, the McFadden R^2 is estimated. The McFadden R^2 equals the percentage decrease in the log-likelihood of the full model compared with a model with only the constant.[77]

Identifying factors that influence implied discount rates

The implied discount rates for own and others' health are examined to identify the factors on which they depend:

- the period in years for which the ill health is to be deferred (*delay*)
- whether the starting point was 3 years (dummy variable: starting point)
- the individual's characteristics: their age in years (dummy variables: age 30–43 years, age 44–63 years, and age > 63 years); their gender (dummy variable: female); whether or not they currently smoke cigarettes (dummy variable: smoker); whether they described their long-term health as fair or poor rather than good (dummy variable: health); whether or not they had been educated beyond secondary school level (dummy variable: education); whether or not they live in an urban area (dummy variable: urban); and their perception of the health state on the VAS (dummy variables: VAS score < 40; VAS score > 70).

The two data sets of own and others' health are aggregated and interaction terms of a dummy variable for others' health with the independent variables described above are included. Aggregation has the advantage of an increase in statistical power and it offers a direct test of differences in own and others' health (see chapter 9). The regression analysis was also run separately for own and others' health to ensure that the aggregation of the data is appropriate.

It is hypothesised, particularly with respect to own health, that time preferences could differ by age group. Older respondents may be more keenly aware of their own mortality and, as a result, discount future benefits more heavily, whereas younger respondents have a longer life expectancy and might, therefore, attach more importance to the distant future. Respondents with poor long-term health would be expected to place a relatively high value on the near future. Smoking and education status may also be associated with different attitudes towards the future, although the direction of causation is not clear. The DU model assumes that the rate of discount is independent of the magnitude of the future event and thus, if this assumption is valid, the individual's perception of the severity of the future health state should not be associated with the implied discount rate.

There is a clear multilevel structure because there are six observations per respondent. This clustering of observations has implications for the nature of the regression analysis. In multilevel data sets, observations are not independent of one another. The use of OLS can underestimate standard errors and thus overestimate the statistical significance of explanatory variables. Multilevel analysis takes the multilevel structure of the data into account by analysing variation that occurs at the higher level (i.e. variation among respondents) separately from variation at the level of the responses.[78] It has two advantages. First, it generates statistically efficient estimates of the regression coefficients. Second, by using the clustering information it provides correct standard errors, confidence intervals, and significance tests.

The simple two-level model in which only the intercept is included in the random part of the model can be expressed as:

$$y_{ij} = \alpha + \beta x_{ij} + \mu_j + \varepsilon_{ij} \qquad (3)$$

In this equation, μ_j and ε_{ij} are random quantities. The respondent-level random variable, μ_j, is the departure of the j-th respondent's actual intercept from the overall mean value α. Thus, this is a level 2 residual. The observation-level random variable ε_{ij} is the observation-level residual for the i-th observation in the j-th respondent and measures random variation across observations.

This model only measures variation in individuals' implied discount rates. It is also possible that the relationship between *delay* and implied discount rates varies across individuals. To allow for variation in this relationship a two-level model including a random coefficient is used:

$$y_{ij} = (\alpha + \beta_1 x_{ij} + \beta_2 delay_{ij}) + (\mu_j + v_j delay_{ij} + \varepsilon_{ij}) \qquad (4)$$

The respondent-level random variable v_j is the departure of the j-th respondent's actual slope from the overall mean value β_2.

Generalised Least Squares (GLS) estimation produces biased estimates of the random parameters since it takes no account of the sampling variation of the fixed parameters, particularly for data in which few level 1 units are nested within level 2 units. Therefore Restrictive Iterative Generalised Least Squares (RIGLS) estimation is used.[78] MLwiN software (© University of London) is used.[79] All hypothesised variables are included in the initial model. The technique of backward elimination is used in which the variable which contributed least to the full regression equation, in terms of its level of statistical significance, is eliminated. The model is rerun and, again, the least statistically significant variable is eliminated. This procedure continues until the only remaining variables are those which are statistically significant at the 5% level.[80] The Ramsay RESET test is used to test for functional misspecification.[81] To assess the performance of the model further, the McFadden R^2 is estimated.

Summary

The study design of the open-ended method and the methods for the basic analysis of the data from the open-ended method is described. Each respondent is presented with six open-ended questions which vary in terms of the starting point (2 or 3 years) and the delay (2–13 years). It is relatively straightforward to derive implied discount rates from the responses to the open-ended method, and it can then be established whether any individuals exhibit negative and zero time preference. The implied rates are regressed on the individuals' characteristics to identify the factors that influence the implied discount rates. The results of these analyses are presented in the next chapter.

Chapter 6

Results for open-ended questions

Introduction

The results for the basic analysis of the data from the open-ended method are reported here. The implied discount rates of the full sample are first examined by calculating various descriptive statistics and are then explored for negative and zero time preference. Finally, the regression results of the implied discount rates on the individuals' characteristics are reported.

Descriptive statistics

Ten respondents (nine in the case of own health and one in the case of others' health) considered that treatment was only worthwhile if it cured them completely. The proportion who only regarded treatment as worthwhile if it cured them completely was higher in the case of own health, which accords with intuition.

The 95% confidence interval for the difference in the proportion between the two samples was 0.0499 ± 0.0382. These individuals were excluded from the analysis because it could be argued that the responses are an expression of preferences regarding the treatment rather than an expression of inter-temporal preferences.

The histograms of the implied discount rates are shown in *Figure 1*. The distributions are peaked, as confirmed by the kurtosis statistic from *Table 5*. The distribution is negatively skewed, especially in the case of others' health. The Lilliefors test indicates that the observations are not normally distributed. The mean implied discount rate is 0.073 in the case of own health and 0.065 in the case of others' health. The median implied discount rates are 0.061 and 0.062, respectively. There is a wide range of implied discount rates – from –1.000 to 0.106 for own health and from –1.000 to 0.095 for others' health. However, the 5% trimmed mean is very similar to the mean indicating that the mean is not heavily influenced by the extreme values.

Negative and zero time preference

The distribution of respondents across the six categories is shown in *Table 6*. The majority (about two-thirds) of respondents gave responses consistent with a positive discount rate. In the case of own health, 6.7% of respondents had negative discount rates and 12.7% zero discount rates. In the case of others' health, 12.2% of respondents had negative discount rates and 7.4% zero discount rates. Thus this study adds to the body of evidence that significant numbers of respondents have negative or zero time preference for health.

The median and mean implied discount rates for each of the categories are shown in *Table 7*. The median is 0.096 for category (i) in the case of own and others' health. For category (ii), the median discount rate is –0.098 for own health and –0.117 for others' health.

The regression results presented in *Table 8* show which factors determine whether individuals have negative, zero or positive time preferences. The data for own and others' health are aggregated because there are relatively few individuals with negative and zero time preferences, especially when cases with missing values on one or more of the independent variables are excluded. Also, the age group dummies and the health state perception dummies have been redefined because of the limited number of individuals with negative and zero time preference. The results show that individuals who perceive the health state as more severe are more likely to have negative discount rates. This supports the first hypothesis and may indicate that these preferences are true preferences. No support is found for the second hypothesis. There is some support for the third hypothesis, since older individuals are more likely to have positive time preferences. There is mixed evidence for the fourth hypothesis: individuals with secondary school education only are more likely to express zero time preferences. However, the signs on the coefficients for the degree of difficulty of answering the questions indicate that individuals who perceived the questions as difficult are more likely to have positive time preferences.

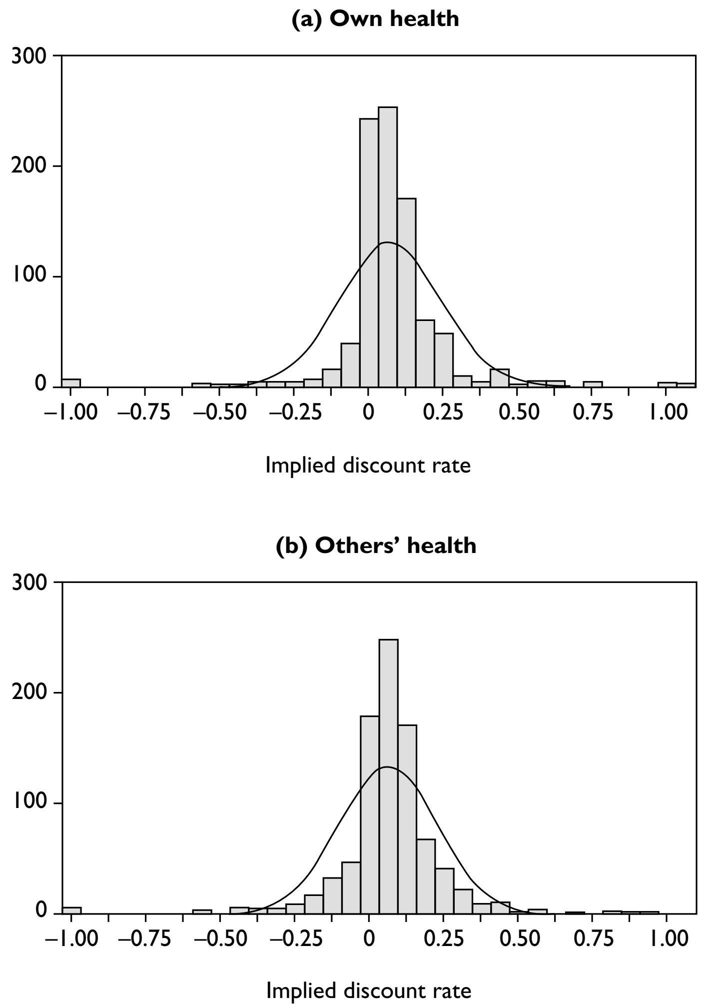


FIGURE 1 *Histograms of implied discount rates*
Own health: standard deviation 0.17; mean 0.07; n *= 897*
Others' health: standard deviation 0.16; mean 0.07; n *= 882*

Identifying factors that influence implied discount rates

The regression results presented in *Table 9* show that, for the full sample, five regressors are statistically significant. These are the intercept, the coefficients of *delay, starting point, health,* and the interaction term of *starting point* with the dummy variable for *others' health.* The longer the period of delay, the lower the implied discount rate. An inter-temporal choice with a starting point 3 years in the future generates a lower implied discount rate than a choice with a starting point 2 years in the future. Individuals in fair or poor health tend to have higher implied discount rates. These variables all have the expected sign. The dummy variable for others' health is not statistically significant, which indicates that when holding the other variables constant the implied discount rates for own and others' health are similar. Only one interaction term with the dummy variable of others' health is statistically significant, namely the interaction term with starting point. The sign of this interaction term indicates that, in the case of others' health, starting point has less of an impact on implied discount rates.

The significant *p*-values in the random part of the model indicate that multilevel analysis should be used. The size of the level 1 variance compared with total variance at level 2 suggests that the majority of variation is across individuals. Individuals vary greatly with respect to their time preferences but vary much

TABLE 5 *Descriptive statistics of implied discount rates*

	Own health	Other's health
Mean	0.073	0.065
95% CI	0.062–0.085	0.054–0.076
5% trimmed mean	0.071	0.068
Median	0.061	0.062
Variance	0.029	0.027
Standard deviation	0.170	0.165
Minimum	–1.000	–1.000
Maximum	1.062	0.955
Interquartile range	0.119	0.128
Skewness (standard error)	–0.411 (0.082)	–1.062 (0.082)
Kurtosis (standard error)	15.981 (0.163)	12.428 (0.165)
Lilliefors (*p*-value)	0.216 (0.000)	0.185 (0.000)

TABLE 6 *Distribution of respondents across categories*

Categories	Own health	Other's health
(i) ≥ 4 positive discount rates	98	98
(ii) ≥ 4 negative discount rates	10	18
(iii) ≥ 4 zero discount rates	19	11
(iv) 3 positive and 3 zero discount rates	0	2
(v) 3 negative and 3 zero discount rates	4	3
(vi) Positive and negative discount rates	16	13
Missing value[a]	3	3
Total	**150**	**148**

[a] *These respondents have one or more missing values and cannot, therefore, be classified*

TABLE 7 *Median and mean implied discount rates*

Categories	Own health (mean)	*p*-value median test	Others' health (mean)
(i) ≥ 4 positive discount rates	0.096 (0.134)	0.861	0.096 (0.129)
(ii) ≥ 4 negative discount rates	–0.098 (–0.177)	0.198	–0.117 (–0.180)
(iii) ≥ 4 zero discount rates	0.000 (–0.003)	0.124	0.000 (–0.001)
(iv) 3 positive and 3 zero discount rates	–	–	0.011 (0.024)
(v) 3 negative and 3 zero discount rates	0.000 (–0.016)	0.533	–0.005 (–0.023)
(vi) Positive and negative discount rates	0.000 (0.006)	0.689	0.000 (0.004)
Total	**0.061 (0.073)**	**0.981**	**0.062 (0.065)**

TABLE 8 *Regression results for discount category*

	Y_1[a]		Y_2[a]	
	dF/dx[b]	*p*-value[c]	dF/dx[b]	*p*-value[c]
Age 35–51 years	–0.069	0.215	–0.051	0.339
Age > 51 years	–0.071	0.202	–0.184	0.002
Education	–0.019	0.703	–0.103	0.055
VAS score < 40	0.104	0.035	–	–
VAS score > 70	–	–	0.033	0.482
Degree of difficulty < 2.5	–0.083	0.067	–0.101	0.030
Number of observations		186		190
McFadden R^2		0.0781		0.1780

[a] *Y_1 = 0 if ≥ 4 responses imply a positive discount rate*
Y_1 = 1 if ≥ 4 responses imply a negative discount rate
Y_2 = 0 if ≥ 4 responses imply a positive discount rate
Y_2 = 1 if ≥ 4 responses imply a negative discount rate
[b] *Change in the probability for a discrete change in each of the dummy variables*
[c] *Standard errors were estimated using the Huber/White/sandwich estimator*

less in their responses to different periods of delay. The sign of the covariance, cov (μ_j, *delay*), suggests that respondents with high intercepts tend to have smaller negative coefficients on *delay* than those with lower intercepts. This implies that the discount rate falls more rapidly with respect to increasing delay for individuals with relatively high short-run discount rates compared with those with relatively low short-run discount rates.

The regression analysis can only be repeated for category (i) since the number of respondents in other categories is too small. The regression results for category (i) are shown in *Table 10*. Compared with the results for the full sample, only one more variable is statistically significant, namely the dummy variable for the age group 64 years and over. Older individuals tend to have higher implied discount rates.

The results of the RESET test show that all models are mis-specified. This is not completely surprising since evidence suggests that the DU model is a poor representation of individuals' time preferences (see chapter 11, in which alternative discounting models are also explored). The regression analyses were run separately for own and others' health. The results for the two data sets were very similar, indicating that aggregating the data is appropriate.

TABLE 9 *Regression results for implied discount rates for the whole sample*

	Full model		**Reduced model**	
	β	***p*-value**	**β**	***p*-value**
Fixed effects				
Intercept	0.100	0.041	0.132	0.000
Others' health	–0.022	0.749		
Delay	–0.010	0.000	–0.008	0.000
Delay x others' health	0.004	0.184		
Starting point	–0.038	0.000	–0.038	0.000
Starting point x others' health	0.020	0.012	0.020	0.010
Age 30–43 years	0.038	0.317		
Age 30–43 years x others' health	0.002	0.968		
Age 44–63 years	0.022	0.562		
Age 44–63 years x others' health	0.024	0.653		
Age > 63 years	0.050	0.234		
Age > 63 years x others' health	0.024	0.682		
Female	0.024	0.317		
Female x others' health	–0.057	0.093		
Health	0.030	0.332	0.037	0.049
Health x others' health	0.010	0.810		
Smoker	0.018	0.617		
Smoker x others' health	–0.035	0.478		
Education	0.005	0.849		
Education x others' health	0.010	0.787		
Urban	–0.004	0.865		
Urban x others' health	–0.010	0.976		
VAS score < 40	0.015	0.631		
VAS score < 40 x others' health	–0.017	0.697		
VAS score > 70	0.017	0.529		
VAS score > 70 x others' health	–0.027	0.465		
Random effects				
Level 1: $\sigma^2(\varepsilon_{ij})$	0.0058	0.000	0.0058	0.000
Level 2: $\sigma^2(\mu_j)$	0.0538	0.000	0.0550	0.000
σ^2(delay)	0.0004	0.000	0.0004	0.000
cov (μ_j, delay)	–0.0038	0.000	–0.0039	0.000
–2 x Log-likelihood	–2387.17		–2375.43	
McFadden R^2	0.280		0.274	
RESET (*p*-value)	0.000		0.000	
n	1504		1504	

TABLE 10 *Regression results for implied discount rates category (i)*

	Full model		Reduced model	
	β	***p*-value**	β	***p*-value**
Intercept	0.235	0.000	0.239	0.000
Others' health	–0.042	0.401		
Delay	–0.017	0.000	–0.014	0.000
Delay x others' health	0.005	0.095		
Starting point	–0.044	0.000	–0.043	0.000
Starting point x others' health	0.021	0.009	0.020	0.004
Age 30–43 years	0.010	0.704		
Age 30–43 years x others' health	–0.006	0.873		
Age 44–63 years	0.024	0.337		
Age 44–63 years x others' health	0.005	0.889		
Age > 63 years	0.043	0.112	0.0258	0.008
Age > 63 years x others' health	0.006	0.873		
Female	0.009	0.522		
Female x others' health	–0.026	0.171		
Health	0.037	0.029	0.020	0.042
Health x others' health	–0.028	0.222		
Smoker	–0.010	0.596		
Smoker x others' health	0.004	0.881		
Education	0.001	0.952		
Education x others' health	0.005	0.818		
Urban	0.005	0.704		
Urban x others' health	–0.005	0.795		
VAS score < 40	0.018	0.317		
VAS score < 40 x others' health	–0.005	0.841		
VAS score > 70	0.011	0.465		
VAS score > 70 x others' health	0.005	0.803		
Random effects				
Level 1: $\sigma^2(\varepsilon_{ij})$	0.0037	0.000	0.0038	0.000
Level 2: $\sigma^2(\mu_j)$	0.0330	0.000	0.0343	0.000
σ^2(delay)	0.0003	0.000	0.0003	0.000
cov (μ_j, delay)	–0.0028	0.000	–0.0029	0.000
–2 x Log-likelihood	–2218.01		–2203.47	
McFadden R^2	0.476		0.466	
RESET (*p*-value)	0.000		0.000	
n	996		996	

Summary

The results of the basic analysis of the data using the open-ended method are reported. The mean implied rate is 0.073 for own health and 0.065 for others' health. The median rates are 0.061 and 0.062, respectively. The distribution of implied discount rates is non-normal. There is some evidence of negative and zero time preference; however, the percentage of individuals with such time preferences is relatively small. Individuals who perceive the health state as more severe are more likely to have negative time preferences, while younger individuals and those with only secondary school education are more likely to express zero time preference. The regression analysis of the implied rates on individuals' characteristics does not provide much insight into which factors determine individuals' time preferences. The dummy variable for fair or poor self-rated health is the only statistically significant individual characteristic. When the regression analysis is repeated for individuals with positive time preferences, only one other dummy variable becomes statistically significant (age group). The variables, period of *delay* and *starting point* are statistically significant both for the full sample and for respondents with positive time preferences.

As can be seen in *Table 1*, previous estimates of time preference rates with respect to future health states range from close to zero[5,19] to large positive rates.[6,7,67] The rates found in this study are closest to the 0.023–0.041 range reported by Redelmeier and Heller.[20]

Chapter 7
Discrete choice experiment

Introduction

The study design of the closed-ended method is described here. All previous closed-ended studies have presented respondents with either a single discrete choice or a discrete choice with follow-up. The disadvantage of these methods is that less information is obtained on each individual. From an individual's response it can only be inferred whether her/his discount rate is higher or lower than the rate implied by the choice. This study uses an alternative closed-ended method – the discrete choice experiment – to elicit inter-temporal preferences for non-fatal changes in health. Individuals are presented with several discrete choices which vary in terms of the timing and the size of a consequence. The use of several discrete choices improves the efficiency of the estimates and allows testing of assumptions, such as transitivity and continuity of preferences. Discrete choice experiments are gaining increased popularity in health economics but have not previously been used to elicit inter-temporal preferences. This is also the first application of a closed-ended method to elicit inter-temporal preferences for **non-fatal** changes in health rather than **fatal** changes.

Study design

Several discrete choices are presented which vary in terms of the trade-off between the timing and the duration of the EQ5D health state. The choice of combination of timing and duration of the health state is, to some extent, arbitrary. The scenarios should be plausible and informative, in the sense that it is not always obvious which one will be selected. There should be a reasonable range of delays since individuals might, for instance, find it hard to differentiate between 10 or 11 years from now. Also, a delay of 30 years would not be realistic for a 70-year-old individual. Three 'timings' are chosen: 2, 7 and 15 years from now (yielding three different periods of delay: 5, 8 and 13 years). The number of days of ill health are calculated for the different delays, assuming that there will be 20 days in the health state 2 years from now and constant discount rates of 0.04 and 0.05. This yields five durations of ill-health: 20, 24, 26, 33 and 38 days.

These levels give rise to 15 possible scenarios ($3^1 \times 5^1$) which can be combined into 105 potential discrete choices (as shown in appendix 3). In 30 of these choices, one scenario dominates the other in that, while both refer to ill health in the same year in the future, one involves a shorter duration of ill health. Each of the remaining discrete choices (assuming an exponential model) implies a particular discount rate. For instance, if individuals are indifferent between 20 days of ill health in 2 years time (scenario A) and 26 days of ill health in 7 years time (scenario H), then their implied discount rate is 0.054 ($r = (26/20)^{1/5} - 1$). If scenario A is preferred, the individual's discount rate will be less than 0.054; if scenario H is preferred, it will be greater than 0.054.

In order to include all 15 scenarios, at least eight discrete choices are required. There is no specific guidance in the literature on how best to pair the scenarios into discrete choices. The scenarios could either be paired on a random basis or on a pragmatic basis. Selecting the discrete choices randomly could lead to a set of discrete choices which represent a small range of implied discount rates (for example, a range of 0.01–0.04 or mainly negative discount rates). The scenarios are therefore paired on a pragmatic basis using the following three criteria:

(i) inclusion of all 15 scenarios
(ii) a wide range of implied discount rates
(iii) one discrete choice representing a negative discount rate.

The last criterion was used because previous studies have shown that usually a small proportion of respondents express negative time preference.[73]

Since no single set of discrete choices could satisfy the three criteria, two different sets of eight discrete choices were selected. This also allowed a test of whether the selection of discrete choices had a systematic influence on the results. In this case, however, only four choices were different and the test was therefore limited. The selected discrete choices are shown in *Table 11*. The implied discount rate ranges from –0.03 to 0.14. It was expected that this range would capture most individuals' time preferences.

TABLE 11 Selected pairwise comparisons

Choices	Scenarios	Year Scenario A	Days Scenario A	Year Scenario B	Days Scenario B	Discount rate
Version 1						
1	C & I	2	26	7	33	0.05
2	D & O	2	33	15	38	0.01
3	A & J	2	20	7	38	0.14
4	F & O	7	20	15	38	0.08
5	E & I	2	38	7	33	–0.03
6	F & M	7	20	15	26	0.03
7	A & I	2	20	7	33	0.11
8	F & L	7	20	15	24	0.02
Version 2						
1	G & O	7	24	15	38	0.06
2	D & O	2	33	15	38	0.01
3	A & J	2	20	7	38	0.14
4	F & O	7	20	15	38	0.08
5	G & K	7	24	15	20	–0.02
6	C & N	2	26	15	33	0.03
7	A & I	2	20	7	33	0.11
8	B & H	2	24	7	26	0.02

The eight discrete choices were presented to the respondents in the following context. In the case of own health, respondents were asked to imagine that they would be ill at somepoint in the future as described by the EQ5D state. The respondents were then informed that two treatments are available whose effects vary with regard to the timing and the duration of the health state. They were then asked to indicate which treatment they preferred (see appendix 2 for an example). Another possible context for others' health was that the two options represent two different groups of people who are ill at different points in time for different durations. Respondents were then asked to choose which group should be treated, assuming that only one can be treated. This design proved unsuccessful in the second pilot study. Most respondents reasoned that it was best to treat the group that became ill first because resources may be available later to treat the other group as well. The choices were presented so that scenario A was not always the nearest scenario.

Model

The basis for this is that individuals' time preferences are revealed by how much longer they are willing to be ill in exchange for a delay in the onset of that ill health. A relationship must therefore be specified between the duration of ill health and the year in which the ill health occurs and utility. The simplest and most commonly used model is the linear additive model, which assumes that the overall utility derived from any combination of attributes is given by the sum of the separate part-values of the attributes. The utility of scenarios A and B is equal to:

$$U_A = \alpha + \beta_1 days_A + \beta_2 year_A + \varepsilon_1$$
$$U_B = \alpha + \beta_3 days_B + \beta_4 year_B + \varepsilon_2 \qquad (5)$$

where: U = utility; $days$ = number of days in the selected health state; $year$ = year in which the spell of ill health occurs; and ε = random error term. The signs of the coefficients of $days$ are hypothesised to be negative since the longer the spell of ill health, the lower the utility. The signs of the coefficients of $year$ are hypothesised to be positive for the majority of respondents since the further in the future a spell of ill health occurs, the higher the utility. This does not hold for respondents with negative discount rates. The difference in utility between scenario B and A can be expressed as:

$$U_B - U_A = \beta_5\,(days_B - days_A) + \beta_6 delay_5 + \beta_7 delay_8 + \beta_8 delay_{13} + \varepsilon_3 \qquad (6)$$

where the dummy variable $delay_n = 1$ if $year_B - year_A = n$, and 0 otherwise.

Because of the collinearity of the three $delay_n$ dummy variables, the model is re-formulated as follows:

$$U_B - U_A = \alpha + \beta_9(days_B - days_A) + \beta_{10} delay_8 + \beta_{11} delay_{13} + \varepsilon_4 \qquad (7)$$

This difference model assumes that the coefficients β_1 and β_3 are equal. To test this assumption, the two coefficients are estimated and a t-test used to assess whether any of the differences are statistically significant. An alternative way of modelling the difference in utility would be to include the variable $(year_B - year_A)$ instead of the $delay_n$ dummy variables. This would measure the relative importance of the average delay and the implied discount rate would then be based on the average delay. This would only be appropriate if individuals are constant timing averse. However, evidence in the time preference literature suggests that the greater the delay the lower the implied discount rate (decreasing timing aversion).

To capture this effect, $delay_n$ dummy variables are included. The implied discount rates for the 5-, 8- and 13-year delay can then be tested for similarity or for whether they decrease the delay increases. The models are estimated separately for own and others' health, since aggregation of the data would make the model quite complex. Also, the implied discount rates for own and others' health could be similar even if interaction terms with others' health are statistically significant. Thus, the inclusion of interaction terms is not useful for the purpose of examining the differences in time preferences between own and others' health.

The ratio of the coefficients of $(days_B - days_A)$ and the intercept and $delay_n$ represent the marginal rate of substitution between duration and year, in other words how much longer individuals are willing to be ill for a specific delay in the onset of that ill health. The implied discount rate, r, assuming an exponential model, is estimated as follows:

$$r_1 = \left(\frac{days_n - \dfrac{\alpha}{\beta_9}}{days_n}\right)^{1/5} - 1;\; r_2 = \left(\frac{days_n - \dfrac{(\alpha+\beta_{10})}{\beta_9}}{days_n}\right)^{1/8} - 1;$$
$$r_3 = \left(\frac{days_n - \dfrac{(\alpha+\beta_{11})}{\beta_9}}{days_n}\right)^{1/13} - 1 \qquad (8)$$

The implied discount rate is a function of the original duration $(days_n)$. In previous discrete choice studies,[11–17] scenario A, and thus $days_n$, was held constant. In this study, scenario A was varied and it is therefore less obvious which value $days_n$ should be. However, the most appropriate value appears to be the average number of days of ill health offered in the A scenarios (24.2 days). If individuals' preferences are characterised by decreasing timing aversion, it would be expected that r_1 is greater than r_2 is greater than r_3.

Econometric issues

Ordered probit is used to estimate the coefficients.[82] This technique is preferred to OLS because of the ordinal nature of the dependent variable $(U_A > U_B;\ U_A = U_B;\ U_A < U_B)$. The ordered probit model is based on the following specification:[75]

$$\begin{aligned} &y^*_{ij} = \beta_{xi} + v_{ij} \\ &\text{and: } y_{ij} = 0 \text{ if } y^*_{ij} \le \mu_0;\ y_{ij} = 1 \text{ if } \mu_0 < y^*_{ij} \le \mu_1; \\ &y_{ij} = 2 \text{ if } y^*_{ij} > \mu_1 \end{aligned} \qquad (9)$$

where v_{ij} = observation specific error term.

The STATA program (© Stata Corporation)[83] was used to estimate the ordered probit model. As with the open-ended method, multiple observations were obtained from each individual. It was not possible to use multilevel analysis since no software was readily available for ordered probit multilevel analysis. The STATA 'cluster' option was used instead; this specifies that the observations are independent across individuals but not necessarily within repeated observations of individuals. This option changes the estimated standard errors and variance–covariance matrix of the estimators but not the estimated coefficients.

Since the implied discount rate is estimated using marginal rates of substitution, the calculation of standard errors for the rates is less straightforward. To estimate standard errors for the implied discount rates, the non-parametric method of bootstrapping was used.[84] The bootstrap method estimates the sampling distribution through a large number of simulations of the original data. Confidence intervals can then be constructed using this empirical estimate of the sampling distribution. The advantage of the bootstrap is that it makes fewer distributional assumptions than the parametric approach.[85]

It is important to test for heteroscedasticity in qualitative dependent variable models because this mis-specification leads to inconsistent estimators.[86] Hence, a multiplicative heteroscedastic model was also estimated and the likelihood ratio statistic used to test the homoscedasticity assumption of the model.[75] The software package LIMDEP (© Ecometric Software Inc.)[87] was used to test for homoscedasticity.

As with the probit model, the interpretation of the coefficients in the ordered probit model is less straightforward. However, in this discrete choice experiment the interest was not in the coefficients of the specific independent variables but in the marginal rates of substitution between two or more variables. The coefficients do not have to be adjusted in order to estimate the marginal rates of substitution.

Segmented model

To examine whether inter-temporal preferences vary across certain groups of respondents, the model is segmented. Data are available on the following respondent characteristics: age; gender; self-rated current long-term health; smoking status; education; whether they live in an urban area; and perception of the health state measured on a VAS. There is some empirical evidence that individuals vary greatly with respect to their time preferences but vary much less in their responses to different periods of delay[4] (see also chapter 6). Therefore only the variable ($days_B - days_A$) is segmented. There are not enough individuals to include interaction terms with the $delay_n$ dummy variables as well. There were potentially ten relevant interaction terms for ($days_B - days_A$) with the individuals' characteristics: age in years (dummy variables: age 30–43, age 44–63, and age > 63); gender (dummy variable: female); whether or not they currently smoked cigarettes (dummy variable: smoker); whether they described their long-term health as fair or poor rather than good (dummy variable: health); whether or not they had been educated beyond secondary school level (dummy variable: education); their valuation of health state on a VAS (dummy variables: VAS score < 40, and VAS score > 70); the version of the questionnaire (dummy variable: version 2). All interaction terms were included in the model initially. Backward elimination was again used until all variables in the model have a t-statistic greater than 1.96.

Testing the assumption of continuity

In a number of studies using discrete choice experiments evidence has been found of some respondents choosing between scenarios on the basis of a single attribute.[88] The existence of dominant preferences specifically violates the axiom of continuity. In this study, individuals could express dominant preferences for:

(i) longest delay
(ii) shortest delay
(iii) smallest duration of ill health
(iv) longest duration of ill health (this last on being, of course, unlikely).

One difficulty with the third group is that this pattern of choice also reflects the pattern of choice of a constant discount rate between –0.03 and 0.01. So it is not possible to distinguish between respondents with dominant preferences and those who have a discount rate within this particular range.

There are a number of reasons for people expressing dominant preferences. First, although a relatively wide range of discount rates is offered (from 0.03 to 0.14), it is possible that a respondent's discount rate is outside this range. Second, respondents might have set a certain target for an attribute and will not trade the attribute until this target is reached. For instance, an individual might not be willing to be ill for more than 20 days and, thus, would never choose an option with a longer duration, no matter how long the period of delay. Third, respondents might not be willing to trade, whatever the levels of the attributes. Fourth, some individuals might find the questions difficult to answer and respond by using rules of thumb, which do not necessarily always reveal an accurate picture of their preferences.

If an individual expresses a dominant preference for either of the first two reasons, they would trade if presented with an appropriate choice. An advantage of an interview compared with a postal questionnaire is that the former offers the opportunity to vary the options in response to previous choices, thus reducing the extent of apparently dominant preferences.

The distinction between those who trade and those who have dominant preferences is important, since the estimated marginal rate of substitution may be misleading if it is averaged across all respondents. The regression analysis was, therefore, first performed for the sample excluding those with dominant preferences. Since it was not clear if individuals who always chose the scenario with the shortest duration of ill health had dominant preferences for shortest duration or had an implied discount rate between –0.01 and 0.03, the regression analysis was also performed including these individuals.

The regression analysis was also undertaken for the full sample. The impact of including all respondents with potentially dominant preferences will depend on the relative numbers of the different groups of dominant preferences. It was expected that the implied discount rate would be higher when respondents with a dominant preference for longest delay were included. The implied discount rate was expected to be lower when respondents with a dominant preference for shortest duration were included.

Since it is not clear whether including respondents with dominant preferences is valid, another approach was also used. Another way of taking dominant preferences into account without including them in the regression analysis is to estimate a median discount rate. The implied discount rates are first estimated for the different groups of traders using the segmented model. The implied discount rates for the respondents with dominant preferences will be lower than –0.03 for those who have a dominant preference for shortest delay and higher than 0.14 for those with a dominant preference for longest delay. Assuming that there are no respondents who have a dominant preference for longest duration of ill health (this pattern of response does not imply a particular discount rate), the median implied discount rate can then be estimated for the overall sample. This approach assumes that dominant preferences are true preferences and that the responses can therefore be interpreted as implying a discount rate beyond the range offered. As noted above, it might be the case that some respondents use a rule of thumb when faced with discrete choices and their expressed preferences might not be their true preferences. Unfortunately, with the limited data available in this study, it was not possible to determine whether dominant preferences were true preferences.

In order to minimise dominant preferences, a relatively wide range of discount rates is included in the experiment. To test whether this range captures most individuals' time preferences, the questionnaire was piloted. The majority of respondents (71%) in the second pilot study traded between the attributes.

Testing the assumption of transitivity

Another assumption in discrete choice experiments is transitivity. This is to the assumption that if an individual prefers scenario A to scenario B and scenario B to scenario C, then the individual should prefer scenario A to scenario C. Because respondents were presented with several discrete choices, transitivity of choices could be tested. If, for instance, 20 days in 2 years time is preferred to 33 days in 7 years time, and 33 days in 7 years time is preferred to 38 days in 7 years time, then 20 days in 2 years time should be preferred to 38 days in 7 years time. The second discrete choice was not offered in this study but it is reasonable to assume that, other things being equal, respondents prefer the number of days of ill health to be low rather than high. The different transitivity tests that are possible with the two versions of the questionnaire are shown in appendix 4. This measure of transitivity is quite crude. It could, for instance, be the case that an 'intransitive' choice is the result of a mistake made by an individual. However, if respondents make several intransitive choices, the assumption of transitivity may be violated and this may give grounds for excluding these respondents from the analysis. The impact of intransitive respondents on the implied discount rates is examined by re-estimating the regression model without these individuals.

Summary

The study design of the closed-ended method is described. In the discrete choice experiment respondents are offered eight discrete choices. A utility function is then specified and regression analysis used to estimates this function. Two assumptions are tested, namely, transitivity and continuity or dominant preferences. The influence of including intransitive respondents and respondents with dominant preferences on the regression results are tested by repeating the regression analysis for different groups of respondents. The regression model is also segmented by individuals' characteristics; this determines which factors influence individuals' implied discount rates. The results of these analyses are reported in the next chapter.

Chapter 8

Results of the discrete choice experiment

Introduction

The result of the basic analysis of the data from the closed-ended method is reported here. The number of respondents with dominant preferences and the results of the transitivity tests are presented first. The regression results for the full sample and the sub-samples are then discussed.

A total of 14 respondents in the case of own health and seven in the case of others' health were indifferent with respect to all eight discrete choices. Since such a response indicates that the respondents did not understand the questions, these respondents have been excluded from the analysis.

Testing the assumption of continuity

Of the 385/381 respondents (own health/others' health), 182/185 (47.3%/48.6%) can be identified as traders. The scenario with the shortest duration of ill health was always chosen by 29/28 respondents, respectively. Dominant preferences for shortest delay were held by 92/90 respondents, respectively, and 82/78 respondents had dominant preferences for longest delay. As expected, none of the respondents had dominant preferences for longest duration of ill health.

To gain some insight into whether or not dominant preferences are genuine, probit regression was used to identify which characteristics determine whether individuals trade or do not trade (the dependent variable takes a value of 0 if the respondent trades and 1 if the respondent has a dominant preference). Individuals who always chose the scenario with the shortest duration of ill health were coded as having dominant preferences. The independent variables included in the model were: age (years) of the respondent (dummy variables: age 30–43, age 44–63, and age > 63); gender (dummy variable: female); whether or not they currently smoked cigarettes (dummy variable: smoker); whether they described their long-term health as fair or poor rather than good (dummy variable: health); whether or not they had been educated beyond secondary school level (dummy variable: education); health state valuation (dummy variables: VAS score < 40 and VAS score > 70); version of questionnaire (dummy variable: version 2); and degree of difficulty (dummy variables: difficulty < 2.5 and difficulty > 3.5).

The probit regression results are shown in *Table 12*. For both own and others' health, older respondents and respondents who perceived the questions as being relatively easy are more likely to have dominant preferences. The latter result is difficult to interpret. The questions are easy to answer for individuals with true dominant preferences. However, they might have used a rule of thumb which would also make the questions easier to answer. In the case of own health, the dummy variable for fair or poor health is also statistically significant indicating that respondents who rated their health as fair or poor were less likely to have dominant preferences. Individuals who always chose the scenario with the shortest duration of ill health were also considered to have dominant preferences. Coding these individuals as traders does not change the regression results.

Testing the assumption of transitivity

The results of the tests for transitivity are presented in *Table 13*. The number of respondents who can be identified as intransitive is relatively low. Only 35/43 (9.1%/11.3%) of the respondents made one or more intransitive choices and the majority of these (71.4%/83.7%) made only one intransitive choice. It could be the case that the respondents made an intransitive choice by mistake. Although this was a promising result, it should be noted that respondents with dominant preferences make transitive choices by definition. Excluding those with dominant preferences, the percentages of respondents making one or more intransitive choices increases to 19.2%/23.2%.

TABLE 12 *Probit regression results for traders versus non-traders*

	Own health[a]		Others' health[a]	
	dF/dx[b]	*p*-value[c]	dF/dx[b]	*p*-value[c]
Age 30–43 years	0.023	0.780	0.058	0.497
Age 44–63 years	0.221	0.004	0.068	0.406
Age > 63 years	0.257	0.002	0.180	0.041
Female	–0.046	0.436	–0.027	0.645
Fair/poor health	–0.150	0.029	0.093	0.165
Smoker	–0.007	0.936	0.014	0.868
Education	–0.004	0.952	0.098	0.123
VAS score < 40	–0.071	0.310	0.144	0.068
VAS score > 70	0.051	0.515	–0.026	0.735
Version 2	–0.032	0.572	–0.031	0.594
Degree of difficulty < 2.5	0.057	0.473	0.088	0.244
Degree of difficulty > 3.5	0.309	0.000	0.369	0.000
n	349		328	
McFadden R^2	0.103		0.095	

[a] *Dependent variable: 1 = non-traders; 0 = traders*
[b] *Change in the probability for a discrete change in each of the dummy variables*
[c] *Standard errors were estimated using the Huber/White/sandwich estimator*

TABLE 13 *Results of transitivity tests*

Test	Number of intransitive responses		Intransitive responses per respondent	Number of respondents			
	Own health	Others' health		Own health		Others' health	
1	3	3		version 1	version 2	version 1	version 2
2	4	7	0	169	177	169	166
3	5	5	1	13	12	12	24
4	8	4	2	7	2	3	1
5	10	3	3	1	0	1	0
6	6	9	Missing value	2	2	2[a]	3[a]
7	10	18	TOTAL	192	193	187	194

[a] *One respondent made one intransitive choice and had one or more missing values on other tests*

Regression results

The regression results for the basic model are shown in *Table 14*. The coefficients on ($days_B - days_A$) and the $delay_n$ variables are all statistically significant in all of the samples and have the hypothesised sign. The longer the duration of ill health is, the lower the utility. A delay in the period of ill health increased the individuals' utility. The null hypothesis that the coefficients are equal was not rejected for any of the samples. The results of the likelihood ratio test indicated that the models are homoscedastic.

The implied discount rates for the different samples are shown in *Table 15*. When all respondents with dominant preferences are excluded, the range of mean implied discount rates is 0.038–0.066, depending on delay and whether it relates to own health or others' health: the longer the period of delay, the lower the implied discount rate. The implied discount rates decrease when respondents with a dominant preference for duration of ill health are included.

Since the number of respondents with a dominant preference for longest delay (82/78) is lower than the total number of respondents with a dominant preference for shortest delay or shortest duration (121/118), it was expected that the implied discount rates for the full sample would be lower than for the sample

***TABLE 14** Regression results for the basic model*

	Own health		Others' health	
	β	***p*-value**	β	***p*-value**
Excluding respondents with dominant preferences				
Intercept	0.542	0.000	0.560	0.000
$Delay_8$	0.297	0.001	0.326	0.000
$Delay_{13}$	0.661	0.001	0.386	0.000
$Days_B - Days_A$	–0.071	0.000	–0.062	0.000
LR test (df)	1.940 (2)		0.266 (2)	
McFadden R^2	0.1277		0.0913	
n	1454 (182 individuals)		1469 (185 individuals)	
Including respondents with a dominant preference for shortest duration				
Intercept	0.447	0.000	0.466	0.000
$Delay_8$	0.293	0.000	0.305	0.000
$Delay_{13}$	0.450	0.000	0.233	0.004
$Days_B - Days_A$	–0.076	0.000	–0.066	0.000
LR test (df)	0.798 (2)		0.477 (2)	
McFadden R^2	0.1246		0.0936	
n	1685 (211 individuals)		1693 (213 individuals)	
Including all respondents with dominant preferences – full sample				
Intercept	0.200	0.000	0.195	0.000
$Delay_8$	0.143	0.000	0.160	0.000
$Delay_{13}$	0.231	0.000	0.144	0.003
$Days_B - Days_A$	–0.039	0.000	–0.034	0.000
LR test (df)	0.390 (2)		0.240 (2)	
McFadden R^2	0.0376		0.0298	
n	3071 (385 individuals)		3037 (381 individuals)	
Excluding both respondents with dominant preferences and intransitive respondents				
Intercept	0.569	0.000	0.657	0.000
$Delay_8$	0.413	0.000	0.423	0.000
$Delay_{13}$	0.822	0.000	0.454	0.000
$Days_B - Days_A$	–0.084	0.000	–0.077	0.000
LR test (df)	3.498 (2)		0.452 (2)	
McFadden R^2	0.1746		0.1331	
n	1175 (147 individuals)		1130 (142 individuals)	

excluding respondents with dominant preferences. The results presented in *Table 15* confirm this hypothesis.

The impact of including intransitive respondents is also shown in *Table 15*. The implied discount rates are slightly lower when excluding the intransitive respondents. Compared with the rates for own health, the rates for others' health tend to be slightly higher for the 5-year and 8-year delay but slightly lower for the 13-year delay. The statistical significance of these differences will be explored in chapter 9.

The regression results for the model are shown in *Table 16* and include the interaction terms for the sample excluding respondents with dominant preferences (and also excluding those who always choose the scenario with the shortest duration). In the case of own health, two age–duration interaction terms and the health–duration interaction term are statistically significant. In the case of others' health, only the health–duration interaction term is statistically significant. That the age of respondent has statistical significance for own health but not for others' health is highly plausible, in that when valuing others' future health, the respondent's age is not directly relevant.

The discount rates implied by these regression results are shown in *Table 17* for delays of 5, 8

TABLE 15 *Implied discount rates from the basic model*

Delay	Own health		Others' health		Difference (*p*-value)
	Implied rate	95% CI	Implied rate	95% CI	
Excluding respondents with dominant preferences					
5 years	0.056	0.046–0.065	0.066	0.055–0.076	–0.010 (0.145)
8 years	0.051	0.045–0.056	0.060	0.054–0.066	–0.009 (0.024)
13 years	0.041	0.036–0.046	0.038	0.032–0.044	0.003 (0.451)
Including respondents with a dominant preference for shortest duration					
5 years	0.045	0.035–0.054	0.053	0.042–0.062	–0.008 (0.218)
8 years	0.043	0.038–0.049	0.051	0.045–0.056	–0.008 (0.045)
13 years	0.031	0.027–0.035	0.028	0.023–0.033	0.003 (0.359)
Including all respondents with dominant preferences – full sample					
5 years	0.039	0.026–0.053	0.043	0.028–0.057	–0.004 (0.692)
8 years	0.040	0.032–0.047	0.046	0.037–0.053	–0.006 (0.235)
13 years	0.030	0.023–0.036	0.027	0.019–0.034	0.003 (0.523)
Excluding both respondents with dominant preferences and intransitive respondents					
5 years	0.050	0.041–0.059	0.062	0.053–0.072	–0.012 (0.080)
8 years	0.050	0.045–0.056	0.059	0.053–0.065	–0.009 (0.038)
13 years	0.041	0.036–0.046	0.037	0.031–0.042	0.004 (0.267)

TABLE 16 *Regression results for the segmented model*

Variables	Own health		Others' health	
	Full model β (*p*-value)	Reduced model β (*p*-value)	Full model β (*p*-value)	Reduced model β (*p*-value)
Intercept	0.559 (0.000)	0.558 (0.000)	0.596 (0.000)	0.596 (0.000)
$Delay_8$	0.307 (0.002)	0.306 (0.002)	0.325 (0.000)	0.322 (0.000)
$Delay_{13}$	0.683 (0.000)	0.679 (0.000)	0.425 (0.000)	0.419 (0.000)
$Days_B - Days_A$	–0.087 (0.000)	–0.083 (0.000)	–0.070 (0.000)	–0.068 (0.000)
$(Days_B - Days_A)$ x Age 30–43	0.018 (0.140)		0.018 (0.180)	
$(Days_B - Days_A)$ x Age 44–63	0.041 (0.001)	0.033 (0.002)	0.018 (0.173)	
$(Days_B - Days_A)$ x Age > 63	0.046 (0.000)	0.039 (0.000)	0.014 (0239)	
$(Days_B - Days_A)$ x Female	–0.002 (0.797)		–0.003 (0.778)	
$(Days_B - Days_A)$ x Health	–0.023 (0.014)	–0.024 (0.010)	0.018 (0.054)	0.019 (0.037)
$(Days_B - Days_A)$ x Smoker	–0.013 (0.324)		–0.001 (0.925)	
$(Days_B - Days_A)$ x Education	–0.001 (0.892)		–0.012 (0.200)	
$(Days_B - Days_A)$ x VAS score < 40	0.004 (0.666)		0.007 (0.576)	
$(Days_B - Days_A)$ x VAS score > 70	–0.006 (0.645)		0.004 (0.725)	
$(Days_B - Days_A)$ x Version 2	–0.002 (0.858)		–0.006 (0.497)	
McFadden R^2	0.1455	0.1438	0.1022	0.0973
n	1342 (168 individuals)		1260 (158 individuals)	

and 13 years (r_1, r_2 and r_3, respectively). Discount rates fall with increasing delay. In the case of own health, rates rise with increasing age. Respondents who rate their own health as fair or poor (rather than good) tend to have lower implied discount rates. In the case of others' health, respondents who rate their own health as fair or poor have higher implied discount rates.

Median discount rate

The implied discount rates for the subgroups identified in *Table 17* can be combined with those of the non-traders to yield median implied discount rates for the overall sample of 0.050, 0.046 and 0.038 (r_1, r_2 and r_3), for own health, and 0.064, 0.057 and 0.038, for others' health. These median rates are higher than the mean

TABLE 17 *Implied discount rates from segmented model*

		Discount rate			*n*
		r_1	r_2	r_3	
Own health					
Good health:	Age < 45 years	0.050	0.046	0.038	76
	Age 45–60 years	0.070	0.070	0.056	23
	Age > 60 years	0.089	0.078	0.061	19
Fair/poor health:	Age < 45 years	0.040	0.037	0.031	17
	Age 45–60 years	0.056	0.051	0.041	16
	Age > 60 years	0.061	0.055	0.044	17
Others' health					
Good health		0.064	0.057	0.038	115
Fair/poor health		0.085	0.074	0.049	43

implied discount rates estimated using the full samples.

Summary

The results of the basic analysis of the discrete choice experiment indicate that the number of respondents who made intransitive choices is relatively low. The existence of dominant preferences expressed by about 50% of respondents is more problematic. The regression results show that the implied discount rates vary, depending on the sample used. Excluding individuals with dominant preferences increased the estimated implied discount rates. Excluding intransitive respondents had a similar impact. As with the open-ended method, the implied discount rates do not seem to vary systematically with individuals' characteristics. Only age and self-rated health had an impact on the estimated implied discount rates.

The implied discount rates estimated with the discrete choice experiment are broadly comparable with other published estimates using a discrete choice approach. Horowitz and Carson[11] reported a median rate of 0.045. Cropper and colleagues[12] found mean values in the range 0.027–0.086 and Cropper and colleagues[14] reported median rates of 0.038–0.169. Johannesson and Johansson[16,17] reported mean rates of 0.013 and 0.010, respectively. Finally, the median rates estimated by Johannesson and Johansson[15] were from 0.080 to 0.250. In making comparisons, it is important to note that in the current study: the elicitation method used was different from that in previous studies; non-fatal changes in health were considered; and, generally, shorter periods of delay were involved than in earlier studies.

Chapter 9

Comparison of inter-temporal preferences for own and others' health

Introduction

The second objective of the TEMPUS project, focused on here, was to establish whether individual inter-temporal preferences relating to own health differ from those for others' health. It is not clear whether individuals apply a higher or a lower discount rate when considering their own rather than others' health. For example, individuals might discount future gains to others at a higher rate than gains to themselves because their altruism is weakened by delay. Alternatively, when making choices on behalf of others, they may demonstrate less impatience and a greater willingness to take a longer-term view. The former hypothesis derives some qualified support from the limited empirical literature. It is a question of some practical interest, since preferences relating to own health are likely to be of greatest relevance if the aim is to obtain a better understanding of individual health-affecting behaviour. Preferences relating to others' health are arguably more relevant when the aim is to inform discounting practice in economic evaluations. Although studies of both types of preference have been undertaken, none were designed expressly to compare these time preferences.

Previous studies are briefly reviewed here, followed by a discussion of the methods of comparison. Inter-temporal preferences were elicited using two different methods: an open-ended method and a discrete choice experiment. The methods of comparing time preferences for own and others' health are different depending on the elicitation method used and are therefore described separately. The results are presented and conclusions drawn.

Previous studies

The empirical literature comprises two types of study:

- those involving the saving of statistical lives
- those examining non-fatal changes to health and saving life-years.

Studies of preferences with respect to saving future statistical lives clearly refer to the health of others.[2–5,11–15] Studies of preferences with respect to non-fatal changes in health and saving life-years have generally been in terms of the respondents' own health,[6–8,20–22] although MacKeigan and colleagues,[21] while inviting respondents to consider their own health, asked them to imagine that they were 50 years old and married with children (none of whom lived with them). There have been three studies concerned with time preferences for the health of others. In the scenarios used by Lipscomb,[19] respondents were invited to consider "a person in your community who is now 25 years old". As part of his life-saving study, Olsen[2] repeated the life-saving questions but substituted "…a programme which improves the health of people in a chronic state of dysfunction and distress" for "…a programme which saves human lives". Finally, Enemark and colleagues[18] elicited the time preferences of vascular surgeons for their patients' future health.

None of these studies were designed to provide any insight into whether preferences for own and others' health differ. There does appear to be a broad pattern in the results, in that studies concerned with own health reported considerably lower estimated discount rates than studies concerned with others' health, although there are exceptions.[7] However, there are too many differences between the studies in respect of design and methods for any strong conclusions to be drawn.

Open-ended method

Inter-temporal preferences for own and others' health, when elicited using the open-ended method, were compared in three ways: in terms of the number of individuals exhibiting positive, negative and zero time preference; in terms of median and mean implied discount rates; and in terms of the factors that influence the inter-temporal preferences.

Respondents were classified into six categories (see chapter 5):

(i) at least four positive rates and no negative rates
(ii) at least four negative rates and no positive rates
(iii) at least four zero rates
(iv) three zero and three positive rates
(v) three zero and three negative rates
(vi) negative and positive rates.

The distribution of the respondents over these six categories was compared for own and others' health. A chi-squared test was used to test for statistical significance of possible differences in the distribution.

The mean and median implied discount rates were compared for the full sample and for each of the six categories. There were 14 different samples: seven for own health and seven for others' health. The one-sample Kolmogorov–Smirnov test was used to determine how well each of the 14 samples fitted the normal distribution. Depending on the results of this test, a non-parametric two-sample median test or a parametric two-sample t-test was used. A 5% significance level was used to identify whether the differences in implied discount rates between own and others' health were statistically significant.

How the implied discount rates were regressed on the period of delay, the starting point and the individuals' characteristics was described in chapter 5. Interaction terms of the dummy variable for others' health with the independent variables were also included in the regression model. The interaction terms tested directly for any difference in the relationship between the implied discount rate and the period of delay, starting point and individual characteristics for own and others' health. The dummy variable for others' health on its own tested for a difference in implied discount rates when all the independent variables included in the model were held constant.

Discrete choice experiment

Inter-temporal preferences for own and others' health, when elicited using the discrete choice experiment, were compared in four ways: in terms of the number of individuals expressing dominant preferences; in terms of mean implied discount rates; in terms of the factors that influence inter-temporal preferences; and in terms of the median implied discount rates.

As described earlier in chapter 7, individuals can express dominant preferences in discrete choice experiments. In the experiment performed in this study, individuals could have dominant preferences for shortest delay, longest delay or shortest duration of ill-health. The distribution of traders and individuals with dominant preferences were compared for own and others' health using a chi-squared test.

The regression analysis was performed for the full sample and three different sub-samples. The sub-samples excluded (some) individuals with dominant preferences and intransitive respondents. For each sub-sample, the estimated implied discount rates for own and others' health were compared. A t-test was used to test whether any of the differences in implied discount rates were statistically significant.

The segmented regression models indicate which individuals' characteristics influence inter-temporal preference. The segmented models for own and for others' health were compared to examine whether the same individuals' characteristics were statistically significant and whether the coefficients had the same sign.

The segmented models were used to estimate the medians for the full sample for own and for others' health. These medians for the full sample were compared for own and others' health. No test was available to test whether any differences in the medians were statistically significant.

Results

Open-ended method

The distribution of the respondents over the different categories was very similar for own and others' health. The null hypothesis that the distribution of respondents across categories is independent of whether the questions concerned own or others' health is accepted ($\chi^2 = 6.85$ with five degrees of freedom; p-value = 0.232).

The median and mean implied discount rates by category for own and others' health are shown in *Table 7*. The median rates with respect to own and others' health are similar for the full sample (0.061 and 0.062, respectively) and for all categories. The results of the Kolmogorov–Smirnov test casts doubt on the assumption of normality in ten of the 14 samples. For this reason the difference in medians was investigated. The p-values for the two-sample median test indicated that there were no statistically significant differences between own and others' health in any category.

The regression results in *Tables 9* and *10* show that for the full sample and for category (i) respondents the dummy variable for others' health is not statistically significant, which indicates that when holding the other variables constant the implied discount rates for own and others' health are similar. Only one interaction term with the dummy variable of others' health is statistically significant, namely, the interaction term with starting point. The sign of this interaction term indicates that, in the case of others' health, starting point has less of an impact on implied discount rates.

Discrete choice experiment

The distribution of the respondents over the different trading/non-trading categories is very similar for own and others' health. The null hypothesis that distribution of respondents across categories is independent of whether the questions concerned own or others' health is accepted ($\chi^2 = 0.19$ with three degrees of freedom; p-value = 0.979).

The implied discount rates for the full sample and the different sub-samples are reported in *Table 15*. The implied discount rates for others' health are slightly higher than those for own health in the cases of the 5-year and the 8-year delay. For the 13-year delay, the implied discount rates for others' health are slightly lower than those for own health. This pattern applies to the full sample and all sub-samples. The sixth column in *Table 15* shows the differences in implied discount rates and gives the p-value for the t-tests. The differences in implied discount rates between own and others' health are not statistically significant for the 5-year and the 13-year delay. The differences in implied discount rates for the 8-year delay are statistically significant for the sub-samples but not for the full sample.

The results for the segmented models shown in *Table 16* show that only self-rated health was statistically significant for both own and others' health. However, the coefficients of self-rated health have the opposite sign for own and others' health. In the case of own health, respondents in poor health have lower implied discount rates while, in the case of others' health, respondents in poor health have higher implied discount rates. In the case of own health, two age dummies are statistically significant. None of the age dummies are statistically significant in the regression model for others' health.

The estimated medians for the full sample are 0.050, 0.046 and 0.038 (r_1, r_2, r_3) for own health and 0.064, 0.057 and 0.038 for others' health. The median implied discount rate is higher for others' health for the 5-year and 8-year delay, and the same for the 13-year delay. Unfortunately, it was not possible to test whether the differences in medians have statistical significance.

Conclusion

Time preferences for own and others' health have been compared using an open-ended method and a closed-ended method (discrete choice experiment). The tentative conclusion to be drawn from the analysis of the data is that the time preferences for own and others' health are broadly similar. There are some differences but the similarities are much more striking than the differences, certainly for the open-ended method.

Open-ended method

The distribution of respondents across different categories and the median rates for the different categories of respondents were not significantly different. The regression analysis of the implied discount rates also revealed many similarities. There were some differences: the impact of starting point on implied discount rate; and the proportion of individuals who would only prefer the treatment if it cured them completely (see chapter 6). It is not completely surprising that starting point has more of an impact in relation to own health. Individuals are likely to have some idea of what they might be doing in the near future and becoming ill might interfere with this. Although this could affect their inter-temporal preferences relating to their own health, it might not affect discount rates relating to others' health. The larger proportion of individuals, in the case of own health, who considered a treatment worthwhile only if it cured them completely is possibly explained by individuals feeling more strongly about a treatment if it concerns themselves rather than others.

Discrete choice experiment

The distribution of the respondents over the trading/non-trading categories was not significantly different. The implied discount rates were only slightly higher for others' health in case of the 5-year and 8-year delay and slightly lower for the 13-year delay. For the full sample there were no statistically significant differences in implied discount rates for own and others' health. For the sub-samples, only the differences in implied discount rates for the 8-year delay were statistically significant. There were some differences in segmented models, with slightly more variables being statistically significant in the case of own health and one variable having the opposite effect. However, most individuals' characteristics did not have any impact on implied discount rates for both own and others' health. The median rates for the full sample, including the non-traders, were very similar for own and others' health.

This raises the question of whether clear differences are not observed because there is little difference between time preferences for own and others' health or because it results from differences in the way the inter-temporal choices were framed. It is very difficult, if not impossible, to ask the question about others' health in an entirely neutral fashion. Asking about a single unidentified stranger would appear to invite the respondent to imagine that he or she were that person – inviting them to consider an anonymous group possibly makes this less likely. However, the own health choices refer to a named individual (the respondent), whereas for others' health respondents are asked to consider the health of a group of unidentified individuals. This might be expected to drive apart the rates for own and others' health rather than encourage them to come together.

This is the first study in which inter-temporal preferences for own and others' health are directly compared. More research is required to assess whether these findings can be replicated. It is particularly important to investigate whether time preferences for own and others' health are also similar for different health states and life-saving consequences.

Chapter 10

Comparison of open-ended method and discrete choice experiment

Introduction

The third objective of the TEMPUS project was to investigate the effect of different ways of asking questions on apparent inter-temporal preferences. Both open-ended and closed-ended methods have been used before to elicit time preferences for health. It is important to know to what extent different methods of eliciting time preferences yield different results. The identification of the direction and size of any bias introduced by using a particular method can assist in the interpretation of previous studies and the design of new studies. Comparison of previous open-ended and closed-ended studies is difficult, because of the many differences between them (the periods of delay considered, the nature of the future health event, and whose preferences are being elicited). The current study is the first in which an open-ended and a closed-ended method for eliciting preferences relating to future health events are directly compared.

The methods previously used to elicit inter-temporal preferences are briefly reviewed, followed by a discussion of how the open-ended method and discrete choice experiment are compared in terms of implied discount rates and feasibility. The results are presented and conclusions drawn.

Previous studies

The impact of different methods for eliciting preferences has been emphasised in the contingent valuation literature (mainly in the area of environmental economics). Several studies have examined empirically whether different elicitation methods produce different willingness-to-pay values. A common finding has been that discrete choice methods produce higher estimates of willingness-to-pay than open-ended or payment card methods.[89–95]

These differences in results have often been explained in terms of differences in how individuals perceive the question, in particular in terms of differences in incentives for individuals to act strategically.[96] The expert panel formed under the auspices of the US National Oceanographic and Atmospheric Administration criticised open-ended or matching approaches as providing 'biased and erratic' results.[97] In the case of time preferences for health, there do not seem to be any obvious incentives to respond strategically in any of the elicitation methods; hence, it cannot be assumed that the findings in the contingent valuation literature also apply to time preferences.

Time preferences for health

Both open-ended and closed-ended (dichotomous choice question) methods have been used before to elicit time preferences for health. To obtain some insight into whether different elicitation methods yield different results in the case of time preferences for health, the median and mean implied discount rates of published studies can be compared (see *Table 1*). The implied rates of discount tended to be higher when open-ended methods were used. However, comparison of these studies is difficult because of the many differences between them. For instance, there is substantial evidence that implied discount rates are a function of the period of delay.[4] It is therefore difficult to compare studies which offer different periods of delay. No previous study has been specifically designed to compare open-ended and closed-ended methods.

Methods

Comparison of implied discount rates

Although comparing implied discount rates may appear relatively straightforward, there was a complication in this study. It was not clear whether individuals who do not trade in the discrete choice experiment should be included in the regression analysis since it was assumed that individuals do trade when calculating marginal rates of substitution. Therefore, individuals with dominant preferences for shortest and longest delay have been excluded when comparing mean implied discount rates. It did not seem meaningful to compare the sub-sample of the discrete choice experiment with the full sample of the open-ended method, since this full sample would include individuals with a mean implied discount rate greater than 0.14 or less than –0.03. These individuals would potentially have been non-traders had they been presented with the discrete choice experiment. They were therefore excluded to make the samples more comparable. For the full samples it was only possible to compare a rough distribution of implied discount rates.

Comparison of sub-samples

The mean implied discount rates were compared for the sub-samples in the discrete choice experiment and the open-ended method. It was expected that the discount rate would be a function of the period of delay. Since the periods of delay offered differ between the discrete choice experiment and the open-ended method, some allowance for delay is required. In the discrete choice experiment an implied discount rate was estimated for 5-, 8- and 13-year delays, while the period of delay in the open-ended method ranged from 2 to 13 years. This wider range of delays was chosen to facilitate other research involving the comparison of alternative discount functions. In order to estimate the mean implied discount rate for 5-, 8- and 13-year delays from the open-ended data, the implied discount rates were regressed on the period of delay. Multilevel regression was again used.[78] Mean implied discount rates for the appropriate periods of delay were predicted using the regression model. The statistical significance of the difference in implied discount rate for each delay was assessed using a t-test.

Comparison of full samples

For the full sample, only the distribution of mean implied discount rates can be compared. For non-traders in the discrete choice experiment it was only known that their implied discount rate was less than –0.03 or greater than 0.14. The distributions of mean implied discount rates from the open-ended and closed-ended approach were therefore compared using three categories: less than –0.03, –0.03 to 0.14, and greater than 0.14. A chi-squared test was used to test the hypothesis that the distribution of respondents over the three categories is independent of the elicitation method used.

Comparison of feasibility

Feasibility of an elicitation method refers to, among other things, whether the technique is acceptable and meaningful to respondents. In this study two aspects were assessed and compared across elicitation methods – response rate and the degree of difficulty in answering the questions as assessed by the respondents.

The response rate was calculated for each elicitation method by each area. The higher the response rate, the more desirable the technique is. This is an indication of how acceptable the technique is to the general public. Evidence from the contingent valuation area suggests that open-ended questions lead to more non-response.[98]

Respondents were asked to indicate how difficult they found it to answer the inter-temporal preference questions on a 5-point scale ranging from very difficult to very easy. The mean and median values for degree of difficulty were compared for the two elicitation methods. There were four different subgroups of respondents: two for the open-ended method and two for the discrete choice experiment. The one-sample Kolmogorov–Smirnov test was used to determine how well the values for each of the four subgroups of respondents fitted the normal distribution. Depending on the results of this test, a non-parametric two sample median test or a parametric two-sample t-test was used. A 5% significance level was used to identify whether the differences in degree of difficulty were statistically significant. The probit analysis described in chapter 8 indicated that traders and individuals with dominant preferences perceived the questions differently, with the latter group perceiving the questions to be easier. The analysis was therefore repeated for traders and individuals with dominant preferences separately.

It should be noted that the measure used in this study could be ambiguous. For instance, individuals who perceived the questions as easy may not have understood the task. So although one method may be perceived as easier, it does not ensure that the responses are more valid. However, even though this measure is far from perfect, it is thought that it does give some general indication of the acceptability of the technique.

Results

Comparison of implied discount rates

Comparison of sub-samples

Of the 150 (own health)/148 (others' health) respondents in the open-ended method, 14/20 respondents, respectively, had a mean implied discount rate of less than –0.03, and 32/37 respondents had a mean implied discount rate greater than 0.14. The regression results for the 104/91 respondents with a mean implied discount rate between –0.03 and 0.14 are presented in *Table 18*. As expected, the coefficient on period of delay is negative – in other words, the longer the delay, the lower the implied discount rate. This model is used to predict the implied discount rates for the different delays.

TABLE 18 *Regression results for the open-ended method*

	Own health β (*p*-value)	**Others' health β (*p*-value)**
Intercept	0.0852 (0.000)	0.0813 (0.000)
Delay	–0.0041 (0.000)	–0.0033 (0.000)
n	623	542
Log likelihood	–2087.62	–1455.02
McFadden R^2	0.0850	0.0140

TABLE 19 *Mean implied discount rates*[a]

Delay	**Open-ended**	**Discrete choice**	**Difference**	**t-test *p*-value**
Own health				
5 years	0.065	0.045	0.020	0.010
8 years	0.053	0.043	0.010	0.070
13 years	0.033	0.031	0.002	0.624
Others' health				
5 years	0.065	0.053	0.012	0.018
8 years	0.055	0.051	0.004	0.535
13 years	0.039	0.028	0.011	0.097

[a] *Respondents with mean implied discount rates < –0.03 or > 0.14 are excluded from the sample*

TABLE 20 *Distribution of respondents*

	Own health				**Others' health**			
	Open-ended		**Discrete choice**		**Open-ended**		**Discrete choice**	
	n	**%**	***n***	**%**	***n***	**%**	***n***	**%**
Mean rate < –0.03	14	9.3	93	24.2	20	13.5	91	23.9
Mean rate ≥ –0.03 and ≤ 0.14	104	69.3	209	54.3	91	61.5	209	54.9
Mean rate > 0.14	32	21.3	83	21.6	37	25.0	81	21.3
Total	150	100.0	385	100.0	148	100.0	381	100.0
χ^2 (*p*-value)	16.04 (0.000)				6.96 (0.031)			
χ^2 (*p*-value)	23.26 (0.000)							

The estimated mean implied discount rates for the different periods of delay are shown in *Table 19*. The mean implied discount rates for the open-ended method are predicted values from the regression model in *Table 18*. The implied discount rates for the discrete choice experiment are taken from *Table 15*. The sample is the full sample excluding respondents with dominant preferences for shortest or longest delay. The mean implied discount rates from the open-ended method tend to be slightly higher than those from the discrete choice experiment. This holds true for both own and others' health. The differences in the predicted mean values from the open-ended questions and the mean implied discount rate in the discrete choice experiment are only statistically significant for the 5-year delay. This is the case for both own and others' health.

Comparison of the full samples

The number of respondents in each of the three categories for the open-ended method and the discrete choice experiment are shown in *Table 20*. A lower percentage of respondents in the open-ended method have mean implied discount rates less than the –0.03 category. This is the case for both own and others' health. The chi-squared statistics reject the hypothesis that the distribution of respondents over the three categories is independent of the method used, whether the comparison is in respect of own health, others' health or is made across all four questionnaires.

Comparison of feasibility

In chapter 4, response rates of 16.6% and 15.7% were reported for questionnaires I and II, respectively, while the response rates for questionnaires III and IV were 25.2% and 24.5%, respectively. The response rate for the questionnaire containing the discrete choices is higher than that for the questionnaire containing the open-ended questions in all six geographic areas.

The results presented in *Table 21* show the mean and median values for degree of difficulty. For the full sample, the mean and median are higher for the discrete choice experiment than for the open-ended methods. This indicates that the discrete

TABLE 21 *Degree of difficulty*

	Discrete choice Mean (median)	χ^2 (*p*-value)	Open-ended Mean (median)
Own health			
Full sample	3.49 (3.50)	21.790 (0.000)	2.91 (3.00)
Traders	3.07 (3.00)	4.649 (0.031)	
Non-traders	3.86 (4.00)	41.600 (0.000)	
Others' health			
Full sample	3.32 (3.00)	15.591 (0.000)	2.68 (2.90)
Traders	2.88 (3.00)	1.394 (0.238)	
Non-traders	3.69 (4.00)	32.299 (0.000)	

choices are perceived to be easier than the open-ended questions. The Kolmogorov–Smirnov test indicated that none of the samples were normally distributed and a non-parametric test of differences in medians is therefore used. The difference in degree of difficulty between the discrete choice experiment and the open-ended method is statistically significant for both own and others' health. When splitting the sample for the discrete choice experiment into traders and individuals with dominant preferences, it can be seen that the latter group thought the questions were easier to answer. The difference in degree of difficulty between traders and the open-ended method is smaller and not statistically significant for others' health.

Conclusion

The effect of using an open-ended method versus a closed-ended method (discrete choice experiment) to elicit time preferences has been examined. The methods were compared in terms of the implied discount rates and in terms of feasibility. This first study, to compare directly an open-ended and a closed-ended method of eliciting time preferences for health, finds some support for the tendency observed in the literature for open-ended methods to elicit higher discount rates than closed-ended methods. The mean implied rates of discount (controlling for delay) are always higher in the case of open-ended questions. However, most of the differences are neither statistically significant nor large.

The rough distributions of mean implied discount rates for the full samples did show some differences between elicitation methods. More respondents had mean implied discount rates of less than –0.03 in the discrete choice experiment compared with the open-ended method. One possible reason may be that respondents used a rule of thumb when presented with the discrete choices because the task was too complex. These may not, therefore, be true preferences. However, if this argument holds, the proportion of respondents with a mean implied discount rate greater than 0.14 would be expected to be larger in the discrete choice experiment compared with the open-ended method. This was not found to be the case in this study. The results for the own health sample and the others' health sample are strikingly similar.

A higher response rate was achieved with the closed-ended method and respondents thought the discrete choices were easier to answer. This raises the question of whether the comparison of implied discount rates is valid when the two methods had a marked difference in response rates. However, comparison of respondent characteristics does not suggest that the additional responders in the context of the discrete choice experiment were systematically different.

The design of the open-ended questions places an upper bound on the implied discount rates. The maximum number of days for which an individual can be willing to be ill in the year further in the future is restricted to 365. This restricts the maximum implied discount rate, especially in the case of longer delays. For instance, in the most extreme case of a 13-year delay, the maximum discount rate is 21.36%. However, there is no evidence that this has influenced the results.

A limitation specific to the discrete choice experiment is that a large number of respondents were non-traders. It was not possible to estimate an implied discount rate for these respondents. In future studies, a wider range of discount rates should be offered to ensure that an individual's time preferences are captured. It is possible that some respondents will never be willing to trade. However, the number of non-traders is expected to decrease as larger trade-offs are offered.

Another limitation of this study is the difference in starting points between the two methods: 2 and 3 years in the open-ended method, 2 and 7 years in the discrete choice experiment. As was shown in chapter 6, starting point has an influence on the implied discount rates. The difference in starting points is therefore likely to introduce a bias in the comparison of implied rates.

In conclusion, the evidence of fairly limited differences between the two methods in terms of implied discount rates suggests that there are unlikely to be any major biases introduced by using one method rather than the other. As a result, the relative feasibility of the two methods is possibly more significant. While there is some evidence for regarding the discrete choice method as being the more feasible, the issue of dominant preferences requires further investigation before such a conclusion is other than tentative.

Chapter 11

Investigation of different models of time preferences

Introduction

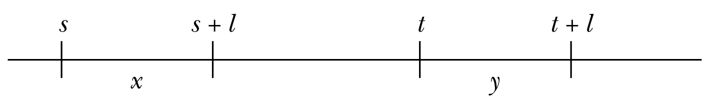
The fourth objective of the TEMPUS project – to establish whether or not individuals value future health benefits in line with the traditional DU model and to investigate, in addition, how well the hyperbolic discounting models explain individual responses – is focused on in this chapter. Although there is little empirical support for the DU model, economists have been generally reluctant to consider alternative discounting models. One reason for this is that only the DU model is considered to be dynamically consistent.[99] The key axiom required for dynamically consistent preferences is stationarity. Previous studies have shown that discount rates are not constant but decrease as a function of the period of delay over which they are estimated, indicating that this axiom is systematically violated.

Albrecht and Weber[24] argued that the dynamic consistency argument in favour of the DU model is not always convincing. Many decisions cannot be unmade in the future and, in these cases, stationarity is a stronger assumption than is required for rationality. It is then no longer obvious that the DU model is the only appropriate model and alternative discounting models, with more accurate descriptive properties, may be preferred instead. Also, where the primary motivation lies in improving the understanding of health-affecting behaviour, it is important to apply a discounting model which accurately describes individuals' behaviour. As highlighted in chapter 2, hyperbolic discounting models have been investigated by psychologists dissatisfied with the descriptive ability of the DU model. Their research suggests that hyperbolic models represent individuals' time preferences better than the DU model. However, the available evidence on hyperbolic discounting models cannot be generalised for the health domain, because some of the research has been conducted using animals and those experiments conducted with humans all relate to financial consequences. In only one study[4] have hyperbolic discounting models been investigated for health consequences.

Here, the DU model and some hyperbolic models are compared. The methods used to test the stationarity axiom and to compare the alternative models empirically are described and then applied.

Stationarity and the common difference effect

The key axiom of the DU model is stationarity. This is the assumption that preference between two outcomes, x and y, depends only on the absolute time interval separating them $(t - s)$. So individuals who prefer receiving £100 (x) after 1 month (s) to receiving £110 (y) after 2 months (t) should also prefer receiving £100 after 12 $(s + l)$ months to receiving £110 after 13 months $(t + l)$. However, in practice, preferences between two delayed outcomes often switch when both delays are incremented by a given constant amount (l). Loewenstein and Prelec[42] referred to this as the common difference effect. In this example, individuals would prefer to receive £110 after 13 months to £100 in 12 months if the common difference effect applies (see *Box* below).



time $0 < x < y, s < t$

Stationarity

$(x, s) > (y, t)$ and $(x, s + l) > (y, t + l)$

Common difference effect

$(x, s) > (y, t)$ and $(x, s + l) \leq (y, t + l)$

where > means preferred to.

The common difference effect implies that discount rates should decrease as a function of the time delay over which they are estimated. Decreasing timing aversion and the common difference effect have been observed in numerous studies both outside and inside the health field (see chapter 2 for a brief overview of the literature). In recent years, hyperbolic discounting models which allow for decreasing timing aversion have been explored in the area of economic psychology (see chapter 2).

Discounting models

Albrecht and Weber[24] provided a useful framework for comparing different discounting models. They defined decision weights (w_t) as follows:

$$w_t = 1/(1+r)^{\alpha(t)} \qquad (10)$$

where $\alpha(t)$ is a time perception function which indicates how fast time is perceived to pass in an individual's mind. A linear $\alpha(t)$ gives the standard discounting model:

$$w_t = 1/(1+r)^t \quad \text{and } \alpha t = t \qquad (11)$$

Concave $\alpha(t)$ yield hyperbolic discounting models. Three such discounting models are investigated here. The models are hyperbolic because the discount factor applied to each period declines as a hyperbolic function of time. Loewenstein and Prelec[42] proposed a general functional form which, using Albrecht and Weber's notation,[24] gives the following decision weights and time perception function:

$$w_t = 1/(1+gt)^{h/g} \quad \text{and } \alpha(t) = h\ln(1+gt)/g\ln(1+r) \qquad (12)$$

The parameter h measures the speed of an individual's time perception. The greater the value of h, the longer one period will be perceived to last. If $h = 0$, the periods are perceived as passing infinitely fast and the individual is timing indifferent. As h tends towards ∞, time is not perceived to pass at all. A value cannot be derived from future consequences and the discount factors for all periods of $t > 0$ are zero.[24] The parameter g determines how much the function departs from the traditional model. As g approaches zero, w_t approaches the exponential discount function.

Loewenstein and Prelec's two parameter model encompassed a number of hyperbolic models which have appeared in the literature. For example, setting $g = 1$ yields the model implied by Harvey:[100]

$$w_t = \frac{1}{(1+t)^h} \quad \text{and } \alpha(t) = h\,\frac{\ln(1+t)}{\ln(1+r)} \qquad (13)$$

Alternatively, setting $h/g = 1$ yields the model proposed by Mazur:[101]

$$w_t = 1/(1+gt) \quad \text{and } \alpha(t) = \ln(1+gt)/\ln(1+r) \qquad (14)$$

The one-parameter discounting models can also be reformulated in terms of their parameters:

DU model: $$r = \frac{1}{(w_t)^{1/t}} - 1 \qquad (15)$$

Harvey model: $$h = \frac{\log(1/w_t)}{\log(1+t)} \qquad (16)$$

Mazur model: $$g = \frac{1}{(w_t t)} - \frac{1}{t} \qquad (17)$$

When fitted to the same data, hyperbolic functions tend to be flatter at longer delays than DU models. The DU model and the one-parameter hyperbolic discounting models are shown in *Figure 2*. The discount factor for the 10-year delay is fixed for all three models at 0.614.

Methods

Only data from the open-ended methods were used to test stationarity and to model the discounting function. These data were more suitable because six implied discount rates per individual were available for a wide range of delays. The data from the closed-ended method provided implied discount rates for three delays only and for the full sample only and were thus less useful for modelling discounting functions.

Testing stationarity

As described in chapter 5, all participants were asked six open-ended questions with two starting points. In three questions, individuals were asked to imagine being ill in 2 years time and, in another three questions, they were asked to imagine being ill 3 years from now (so s is either 2 or 3 years). The period of delay offered ($t - s$) ranged from 2 to 13 years. These data permit a test of stationarity. Implied discount rates were estimated from the inter-temporal responses assuming the DU model.

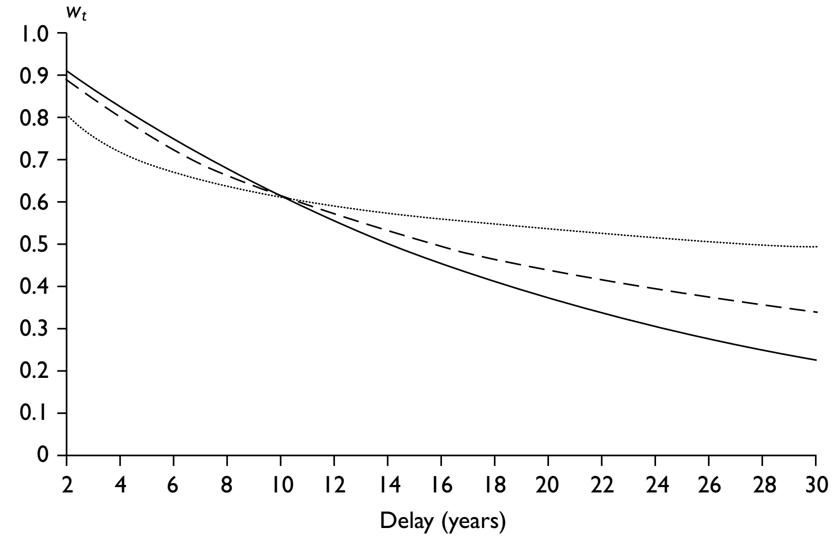


***FIGURE 2** Discounting models (w_{10} = 0.614) (——, DU model; r = 0.05; – – –, Mazur model, g = 0.06; ………, Harvey model, h = 0.20)*

If the preference condition of stationarity holds, individuals' implied discount rates should neither be a function of s nor of $(t-s)$; that is, $\delta r/\delta s = 0$ and $\delta r/\delta(t-s) = 0$. In other words, the importance of a fixed difference $(t-s)$ in the timing of the two benefits does not change as the timing of these benefits is moved into the future (or as s increases). If, on the other hand, individuals are decreasingly timing averse, less importance is attached to $(t-s)$ the further into the future the timing of the future benefits is moved and $\delta r/\delta s < 0$ and $\delta r/\delta(t-s) < 0$. Stationarity can thus be tested by regressing the implied discount rates on s and $(t-s)$ – as was done in chapters 5 and 6.

Selecting specific functional forms

Non-linear Least Squares Regression in the software package LIMDEP[87] (© Econometric Software Inc.) was used to select optimal values for r, g and h in the models. A non-linear regression model is one in which the first-order conditions for least squares estimation of the parameters are non-linear functions of the parameters.[75] Thus, non-linearity is defined in terms of the techniques needed to estimate the parameters rather than the shape of the regression function. Since there were only six observations per respondent, functional forms cannot be selected on an individual basis. Models were therefore selected on a group basis. One limitation of the non-linear regression model in the LIMDEP software is that it is not designed to take into account the multilevel structure of the data.

Although the majority of the respondents exhibited positive time preference, some respondents exhibited negative or zero time preference (see chapter 6). It is likely that alternative models are required for negative time preference, while zero time preference implies that individuals do not discount the future. The DU model and the three hyperbolic discounting models were therefore only fitted for respondents who had three or more positive discount rates and no negative discount rates (categories (i) and (iv) in chapter 5). The sample sizes were 98 respondents (62%) for own health and 100 (68%) for others' health.

It is hypothesised that the models vary according to individuals' characteristics. As the multilevel regression analysis in chapter 6 showed, individuals do vary greatly with respect to their time preferences. The following models which take into account individuals' characteristics were therefore estimated:

DU model

$$w_t = 1/(1 + r + \beta \boldsymbol{X}_i)^t \quad (18)$$

Harvey model

$$w_t = 1/(1+t)^{(h + (\beta \boldsymbol{X}_i)} \quad (19)$$

Mazur model

$$w_t = 1/(1 + (g + \beta \boldsymbol{X}_i)\, t) \quad (20)$$

Loewenstein and Prelec model

$$w_t = 1/(1+gt)^{(h + (\beta \boldsymbol{X}_i)/g} \quad (21)$$

where X_i is a vector of independent variables which include the age of the respondent (dummy variables: age 30–43, age 44–63, and age > 63); gender (dummy variable: female); whether or not they currently smoke cigarettes (dummy variable: smoker); whether they described their long-term health as fair or poor rather than good (dummy variable: health); whether or not they had been educated beyond secondary school level (dummy variable: education); perceived severity of health state measured on a VAS (dummy variables: VAS score < 40 and VAS score > 70); and starting point (dummy variable: starting point).

In principle, both parameters *g* and *h* could be allowed to vary across individuals' characteristics in the Loewenstein and Prelec model. However, the data sets were not rich enough to perform such an analysis. Hence only the parameter *h* was allowed to vary.

To test the goodness of fit of each equation, the proportion of variance accounted for by each equation (R^2) was estimated. It should be noted that R^2 is not bounded by [0, 1] because OLS was not used. R^2 was compared for the DU and hyperbolic models. The goodness of fit was expected to be rather low for any of the models because the form of the function describing aggregate data is not necessarily the same as the form of the function describing individual data. However, since only six observations were available for each individual, functions could only be fitted on a group basis.

Results

Testing the assumption of stationarity

The regression results for the full sample are shown in *Table 9* and the regression results for category (i) in *Table 10*. For the present purpose, the negative coefficients on the *delay* variable $(t - s)$ and the dummy variable for starting point (s) and their very large *t*-statistics are of greatest relevance. These findings hold true for the full sample as well as for the category (i) sample and for own health as well as others' health. This is strong evidence against stationarity. The negative coefficients indicate that there is evidence of decreasing timing aversion. The longer the period of delay is, the lower the implied discount rate. It is therefore likely that hyperbolic discounting models, which allow for decreasing timing aversion, will fit the data better.

The DU model versus hyperbolic discounting models

The non-linear regression results are shown in *Table 22*. Because of missing values, sample sizes are reduced to 480 and 515 observations, respectively. In the one-parameter models, the base case parameter values of *r*, *h*, and *g* are 0.0747, 0.2533, and 0.1125, respectively, for own health and 0.0569, 0.2171, and 0.0768, respectively, for others' health. In the case of own health, the parameter values vary with age, self-rated health and starting point. Older respondents and respondents who rated their health as fair or poor tend to have higher parameter values, while parameter values for starting points of 3 years tend to be lower. In the case of others' health, the parameter values vary with age, gender, education, health state valuation and starting point. Older respondents, respondents educated beyond secondary school level, males and respondents who scored the health state below 40 or above 70 tend to have higher parameter values. Again, parameter values for starting points of 3 years tend to be lower. The parameter values in the DU and Harvey models also vary with self-rated health. Individuals who rated their health as fair or poor tend to have higher parameter values.

The parameter values for the base case and the coefficients on the individuals' characteristics are very similar for own and others' health. The differences in those variables that are statistically significant in the non-linear regression analysis for both own and others' health are shown in *Table 23*. The results of the t-tests show that none of these differences are statistically significant.

Even though the base values for the parameters *r*, *h* and *g* and the coefficients are individually not significantly different for own and others' health, the predicted values for groups of respondents may differ. To test this, values of *r*, *h* and *g* are predicted for four groups of respondents in *Table 24*. None of the predictions for own and others' health are significantly different for any of the models.

The goodness of fit for all models is also shown in *Table 20*. For own health, the DU model has a very poor goodness of fit. All three hyperbolic models fit the data better than the DU model. However, the Mazur model does not seem to perform very well compared with the two other hyperbolic models. The Loewenstein and Prelec model has the best goodness of fit but, it should be noted, is a two-parameter model.

TABLE 22 *Non-linear regression results*

	DU model		Harvey model		Mazur model		Loewenstein & Prelec model	
	β	*p*-value	β	*p*-value	β	*p*-value	β	*p*-value
Own health								
Parameter (*r, h, g, g*)	0.0747	0.000	0.2533	0.000	0.1125	0.000	3.3253	0.064
Second parameter (*h*)	–	–	–	–	–	–	0.5436	0.012
Age 30–43 years	0.0166	0.061	0.0552	0.041	0.0253	0.061	0.1180	0.103
Age 44–63 years	0.0317	0.000	0.1007	0.000	0.0533	0.000	0.2140	0.026
Age > 63 years	0.0633	0.000	0.1912	0.000	0.1130	0.000	0.4045	0.013
Female	–0.0063	0.285	–0.0160	0.347	–0.0121	0.222	–0.0339	0.379
Health	0.0327	0.000	0.1011	0.000	0.0588	0.000	0.2150	0.021
Smoker	–0.0135	0.073	–0.0326	0.144	–0.0161	0.204	–0.0646	0.226
Education	–0.0035	0.617	0.0007	0.968	–0.0089	0.459	0.0038	0.928
VAS score < 40	–0.0029	0.719	–0.0076	0.734	–0.0052	0.689	–0.0167	0.726
VAS score > 70	–0.0051	0.424	–0.0059	0.741	–0.0081	0.441	–0.0078	0.841
Starting point(s)	–0.0227	0.000	–0.0724	0.000	–0.0407	0.000	–0.1537	0.020
n	480		480		480		480	
Adjusted R^2	0.0160		0.2185		0.1160		0.2302	
Others' health								
Parameter (*r, h, g, g*)	0.0562	0.000	0.2171	0.000	0.0768	0.000	0.8519	0.000
Second parameter (*h*)	–	–	–	–	–	–	0.1975	0.000
Age 30–43 years	0.0125	0.109	0.0340	0.171	0.0207	0.072	0.0314	0.174
Age 44–63 years	0.0328	0.000	0.1045	0.000	0.0536	0.000	0.0957	0.000
Age > 63 years	0.0643	0.000	0.1983	0.000	0.1177	0.000	0.1815	0.000
Female	–0.0226	0.000	–0.0711	0.000	–0.0419	0.000	–0.0651	0.000
Health	0.0130	0.034	0.0308	0.080	0.0230	0.033	0.0285	0.087
Smoker	–0.0023	0.704	–0.0036	0.841	0.0002	0.984	–0.0035	0.833
Education	0.0176	0.002	0.0484	0.003	0.0282	0.001	0.0446	0.007
VAS score < 40	0.0269	0.000	0.0814	0.000	0.0545	0.000	0.0743	0.001
VAS score > 70	0.0148	0.006	0.0445	0.005	0.0264	0.002	0.0407	0.011
Starting point(s)	–0.0203	0.000	–0.0613	0.000	–0.0333	0.000	–0.0561	0.000
n	515		515		515		515	
Adjusted R^2	0.2263		0.3670		0.3324		0.3674	

TABLE 23 *Difference between own and others' health*

	Difference in coefficients (*p*-value)			
	DU model	Harvey model	Mazur model	Loewenstein & Prelec model
Parameter (*r, h, g, g*)	0.0185 (0.215)	0.0363 (0.418)	0.0357 (0.134)	
Second parameter (*h*)	–	–	–	0.3461 (0.116)
Age 44–63 years	0.0011 (0.928)	0.0038 (0.912)	0.0003 (0.984)	0.1183 (0.238)
Age > 63 years	0.0010 (0.944)	0.0071 (0.865)	0.0047 (0.849)	0.2230 (0.184)
Starting point	0.0024 (0.741)	0.0111 (0.596)	0.0074 (0.549)	0.0976 (0.150)

The models for others' health all have a better goodness of fit than the models for own health. The DU model, again, has the poorest goodness of fit. Although the Harvey model again has better goodness of fit than the Mazur model, the difference is smaller than for own health. The Loewenstein and Prelec model does not have a much better fit than the Harvey model, even though the former is a two-parameter model.

Conclusion

One of the key assumptions of the DU model, namely stationarity, has been tested. The results show that this assumption does not hold and that there is evidence of decreasing timing aversion. Hyperbolic models which allow for decreasing timing aversion were investigated for respondents who exhibited positive time preference and were

TABLE 24 *Predicted parameter values for groups of respondents*

	DU model	Harvey model	Mazur model	Loewenstein & Prelec model
Age 25 years, male, in fair health, smoker, not educated beyond secondary school, health state 50				
Own health	0.094	0.322	0.155	0.694
Others' health	0.067	0.244	0.100	0.223
Age 45 years, male, in good health, non-smoker, university degree, health state 50				
Own health	0.088	0.309	0.129	0.665
Others' health	0.086	0.299	0.126	0.274
Age 65 years, female, in good health, non-smoker, not educated beyond secondary school, health state 60				
Own health	0.132	0.429	0.214	0.914
Others' health	0.098	0.344	0.153	0.314
Aged 50 years, male, in good health, non-smoker, educated beyond secondary school, health state 25				
Own health	0.100	0.347	0.152	0.745
Others' health	0.134	0.451	0.213	0.412

found to fit the data better than the DU model. The Loewenstein and Prelec model is the best performing hyperbolic model. This is not surprising since it was the only two-parameter model considered. The Mazur model has the poorest goodness of fit, especially for own health.

The models for own and others' health are very similar. The only differences are in the greater significance of the coefficient of self-rated health for own health than for others' health, and in the statistical significance of the coefficients of female, education and health state valuation for others' health and not for own health. There are also differences in the parameter values of the Loewenstein and Prelec model for own and others' health.

One limitation of the current study is that no models were fitted for respondents who exhibited zero or negative time preference (about one-third of the sample). The hyperbolic models have been developed assuming positive time preference. It is likely that alternative models are required to model negative time preference. This is an area for future research.

The overall finding, that hyperbolic models provide a better description of inter-temporal preferences than the DU model, is similar to those reported in other studies concerned with monetary consequences. This study is novel in that it is the first in which DU and hyperbolic models have been compared using data for non-fatal changes in health, and in which three different hyperbolic models are compared.

Although the evidence in favour of hyperbolic discounting models is growing, the economics profession is typically resistant to hyperbolic discounting. Camerer[102] suggested three reasons for this:

"(i) ignorance about the overwhelming empirical superiority and parsimony of hyperbolic discounting
(ii) confusion about the normative versus descriptive appeal of dynamic consistency
(iii) uncertainty about how to move away from the exponential model and still do analytical economics."

This study provides additional evidence for the descriptive superiority of hyperbolic discounting models. However, further research is required on both DU and hyperbolic models before most economists would embrace the radical step of discarding a familiar and successful model. In particular, research is required into the implications for policy making of the use of hyperbolic models.

Chapter 12
Conclusions

This project had four objectives:

(1) to derive implied discount rates for future health benefits for a sample of the general public in the UK
(2) to establish whether individual inter-temporal preferences with respect to their own health differ from those with respect to the health of others
(3) to investigate the effect of different ways of asking questions on apparent inter-temporal preferences (specifically a closed- and an open-ended method are compared)
(4) to establish whether individuals value future health benefits in line with the traditional DU model and to investigate, in addition, how well the hyperbolic discounting models explain individual responses.

The main findings of the TEMPUS study may be summarised as follows.

- The median implied discount rate ranges from 0.038 to 0.064 depending on the period of delay, the elicitation method used, and whether it concerns own health or others' health.
- The implied discount rates for own health and others' health are very similar.
- Implied discount rates tend to be higher when using the open-ended method. However, differences are small and, in most cases, not statistically significant.
- Individuals inter-temporal preferences are better represented by hyperbolic discounting models than by the DU model.

Basic study design and data

Stated preference methods were used to elicit individuals inter-temporal preferences. Four questionnaires were designed which differed in terms of the elicitation method used (open- or closed-ended) and in terms of whom it concerned (own health or others' health). Since the study was concerned with non-fatal changes in health, a health state had to be selected. A generic health state (an EQ5D health state) was used. For others' health it was stated that a group of middle-aged people would experience this health state. Data were collected through postal questionnaires. Although this may lead to non-responses, it was the most efficient way of obtaining the large sample sizes required for this study. The questionnaires were sent to 5120 members of the general public in six areas in Scotland, England and Wales. The response rates for the four questionnaires ranged from 15.7% to 25.2%. The four samples were very similar in terms of age, gender, self-rated health, education, and smoking status.

Open-ended method

In the open-ended method, each time preference question asked the respondent to imagine being ill at a point in the future and offered the opportunity for this spell of ill health to be delayed as result of a minor one-off treatment. Individuals were then asked to identify a maximum number of days of future ill health at which it would still be worthwhile receiving this treatment. Each respondent was presented with six time preference choices which varied in terms of the starting point of the ill health (2 or 3 years) and the delay (2–13 years).

The estimated mean implied discount rates were 0.073 in the case of own health and 0.065 in the case of others' health, with the median values being 0.061 and 0.062, respectively. Respondents were then classified as exhibiting positive, negative and zero time preferences. There is some evidence of negative and zero time preferences. However, the percentages of individuals with negative and zero time preferences were relatively small. Regression analyses were then performed to identify the factors that influence individuals' time preferences. This did not provide much insight into which individuals' characteristics determine time preferences. The dummy variable for fair or poor self-rated health was the only statistically significant individual characteristic. When the regression analysis was repeated for individuals with positive time preference, only one other dummy variable became statistically significant (the dummy variable for age group).

Discrete choice experiment (closed-ended method)

In the discrete choice experiment, individuals were asked to imagine that they would be ill at some point in the future, as described by the EQ5D health state. The respondents were then presented with two possible treatments which varied with the timing and duration of the health state. They were then asked to indicate which treatment they preferred. Each respondent was offered eight discrete choices which varied in terms of the timing and duration of the health state. To analyse data from discrete choice experiments a utility function has to be specified. Two assumptions were tested, namely transitivity and continuity or dominant preferences.

The number of respondents who made intransitive choices was relatively low. The existence of dominant preferences was more problematic. About 50% of the sample expressed dominant preferences. It is unclear whether individuals with dominant preferences should be included in the regression analysis. The implied discount rates ranged from 0.027 to 0.066, depending on the period of delay and whom it concerned (own health or others' health). Excluding individuals with dominant preferences increased the estimated implied discount rates. Excluding intransitive respondents had a similar impact. The regression model was also segmented by individuals' characteristics in order to examine which factors influence individuals' implied discount rates. As with the open-ended method, the implied discount rates did not seem to vary systematically with the individuals' characteristics. Only age and self-rated health had an impact on the estimated implied discount rates.

Inter-temporal preferences for own health and others' health

Inter-temporal preferences for own health and others' health were compared in terms of implied discount rates and other aspects which depended on the elicitation method used. In the open-ended method, the first test was for whether the percentages of respondents exhibiting positive, negative and zero time preferences were different for own health and others' health. The results showed that that none of the differences were statistically significant. The differences in the median implied discount rates for the full samples and for the sub-samples (respondents with positive, negative and zero time preference) were not statistically significant. The regression analyses of the implied discount rates with the individuals' characteristics also revealed many similarities. There were also some differences between own and others' health for the open-ended method: the impact of starting point on implied discount rate, and the proportion of individuals who would only prefer a treatment if it cured them completely.

In the discrete choice experiment, the difference in the relative number of respondents expressing dominant preferences was first tested and the results showed that they were not significantly different for own and others' health. The comparison of the implied discount rates showed that the rates were slightly higher for others' health for 5- and 8-year delays, and slightly lower for the 13-year delay. When using the full sample there were no statistically significant differences in implied discount rates between own health and others' health. For the sub-samples (which excluded intransitive respondents and/or individuals with dominant preferences), only the differences in implied discount rates for the 8-year delay were statistically significant. There were some differences in segmented models, with slightly more variables being statistically significant in the case of own health and one variable having opposite effects. However, most individual characteristics did not have any impact on implied discount rates for both own and others' health. The median rates for the full sample, including the non-traders, were very similar for own and others' health.

Thus, although there are some differences the similarities are much more striking. These results suggest that one of the several potential differences in the design and interpretation of preference elicitation studies – whether the future health consequences involve own health or others' health – is not important. However, more research is needed to confirm these findings, especially for different health states and for life-years.

Comparison of open-ended method and discrete choice experiment

The open-ended method and discrete choice experiment were compared in terms of the implied discount rates and in terms of feasibility. The mean implied discount rates for the sample, excluding individuals with dominant preferences, were compared with a sub-sample from the open-ended method for three different periods of delay. Some support was found for the tendency observed in the literature for open-ended methods to elicit higher discount rates than closed-ended methods. The mean implied rates of discount (controlling for delay) are always higher in the case of open-ended questions. However, most of the differences were neither statistically significant nor large.

The rough distributions of mean implied discount rates for the full samples were also compared. This did indicate some differences between elicitation methods. More respondents had mean implied discount rates smaller than –0.03 in the discrete choice experiment compared with the open-ended method. In terms of feasibility, a higher response rate was achieved with the closed-ended method. Respondents were asked to indicate how difficult the questions were on a 5-point scale (ranging from 1 (very difficult) to 5 (very easy)). Respondents thought the discrete choices were easier to answer.

Thus, on balance, the closed-ended method appears to be more attractive than the open-ended method. However, before the discrete choice method can be strongly recommended, some means must be found of reducing the extent of potentially dominant preferences. Also, there are some specific circumstances in which closed-ended methods are less satisfactory, for instance, when the purpose of the study is to examine alternative discounting models, since open-ended methods enable discount factors to be derived from each individual response.

Investigation of different models of time preferences

The key axiom of the DU model, the axiom of stationarity, was tested first. The results show that the assumption of stationarity did not hold and that there is evidence of decreasing timing aversion. The DU model and three hyperbolic discounting models (the Loewenstein and Prelec, Harvey, and Mazur models) were then fitted for respondents who exhibited positive time preferences using non-linear regression analysis. The hyperbolic models fitted the data better than the DU model. The Loewenstein and Prelec model was the best-performing hyperbolic model. This is not completely surprising because it was the only two-parameter model considered. The Mazur model had the poorest goodness of fit, especially in the case of own health.

The models for own health and others' health were very similar. The only differences were that the parameter values varied by self-rated health for own health and that the parameter values varied by gender, education, and health state valuation for others' health. Also, there were differences in the parameter values of the Loewenstein and Prelec model.

This study was concerned with the descriptive properties of the DU and hyperbolic discounting models. Hyperbolic are better descriptive models than the DU model. Hyperbolic models can possibly also be used as normative models. However, the implications for policy making of the use of hyperbolic models have yet to be systematically investigated.

Limitations

There are limitations inherent in the stated preference approach – specifically there are concerns over the validity of responses compared with values derived from observing behaviour. However, it is difficult to observe inter-temporal choices in a healthcare context and then to control for the large number of potential influences. Also, there is some evidence that hypothetical decisions are treated similarly to real decisions. Kirby and Marakovíc[39] found that implied discount rates for real monetary consequences were slightly higher than for hypothetical monetary awards but behaved similarly with respect to decreasing timing aversion and magnitude effect. The intertemporal choices in this study are not only hypothetical; it is also likely that the respondents were unfamiliar with the EQ5D state. However, a recent article by Chapman and colleagues[10] indicated that familiarity did not have an impact on implied discount rates.

The inter-temporal choices presented to respondents were complex. Thinking about the future is not easy, especially in the context of health. It could be argued that because of the complexity of the inter-temporal choices, face-to-face interviews are the only appropriate method of data collection. However, data in this study were collected by postal questionnaire. The use of postal questionnaires was driven largely by the desire to obtain reasonably large sample sizes. Such an approach offers less opportunity to establish whether or not the respondent understands the question; also the questions are predetermined, leaving no opportunity for them to be responsive to the previous answers. An interesting line of enquiry would be to assess the impact of using different methods of data collection on the results and to perform an economic evaluation to assess whether face-to-face interviews are worth the extra costs.

Comparability of results

The implied discount rates elicited in this study should not be over-emphasised because of the unrepresentativeness of the study sample. In particular, the respondents were markedly better educated than the population as a whole. This is not surprising, given the deliberate policy of sampling in more socio-economically advantaged areas. However, it is notable how close the estimated median rates are to the rates advocated for use in economic evaluation in a range of countries (for example, 3% in the USA, 5% in Australia and Canada). Also the degree of convergence between rates for own health and others' health and between the open-ended method and the discrete choice experiment is striking. The estimated implied discount rates in this study fall comfortably within the range of estimates from previous empirical studies.

Research recommendations

For future research, three of the themes addressed in this report require further investigation:

- the influence of different methods of eliciting preferences on estimates of implied discount rates
- the nature of the inter-temporal choice
- the underlying model of time preferences.

A single, albeit multifaceted, project such as TEMPUS adds significantly to our understanding but cannot, by itself, resolve the outstanding research issues, particularly given that this is the first study in which a number of these issues were addressed systematically. Three areas should be highlighted.

1. Continued refinement of the methods of eliciting time preferences is required. Relevant topics include the use of self-completed questionnaires versus interviews (face-to-face and telephonic) and the presence and impact of framing effects.
2. Further research is required on alternative models of time preference, in particular, models which allow for decreasing timing aversion. Also, the implications of using alternative models for policy making need to be investigated.
3. There is considerable scope for research to investigate the role played by time preference in explaining health affecting behaviour. To what extent are individuals willing to incur short-term costs in order to secure longer-term benefits – for example, the successful control of blood sugar levels by diabetic patients or participation in screening programmes?

Acknowledgements

The Health Economics Research Unit, University of Aberdeen, is funded by the Chief Scientist Office of the Scottish Executive Health Department. This study was commissioned by the NHS R&D Health Technology Assessment Programme (project 94/35/01). The views expressed in this report are those of the authors and not necessarily those of the Standing Group on Health Technology Assessment, the HTA Commissioning Board, the HTA panel members, the Department of Health or the Scottish Executive Health Department.

Our thanks are due to the referees for their perseverance in reading the report and for the quality of their comments.

References

1. Gafni A, Torrance GW. Risk attitude and time preference in health. *Manage Sci* 1984;**30**:440–51.
2. Olsen JA. Time preferences for health gains: an empirical investigation. *Health Econ* 1993;**2**:257–65.
3. Cairns JA. Valuing future benefits. *Health Econ* 1994;**3**:221–9.
4. Cairns JA, Pol van der MM. Saving future lives: a comparison of three discounting models. *Health Econ* 1997;**6**:341–50.
5. Cairns JA, van der Pol MM. Constant and decreasing timing aversion. *Soc Sci Med* 1997;**45**:1653–9.
6. Cairns JA. Health, wealth and time preference. *Project Apprais* 1992;**7**:31–40.
7. Chapman GB, Elstein AS. Valuing the future: temporal discounting of health and money. *Med Decis Making* 1995;**15**:373–86.
8. Chapman GB. Temporal discounting and utility for health and money. *J Exp Psychol Learn Mem Cogn* 1996;**22**:771–91.
9. Olsen JA. Persons vs. years: two ways of eliciting implicit weights. *Health Econ* 1994;**3**:39–46.
10. Chapman GB, Nelson R, Hier DB. Familiarity and time preferences: decision making about treatments for migraine headaches and Crohn's disease. *J Exp Psychol Appl* 1999;**5**:17–34.
11. Horowitz JH, Carson RT. Discounting statistical lives. *J Risk Uncertainty* 1990;**3**:403–13.
12. Cropper ML, Aydede SK, Portney PR. Discounting human lives. *Am J Agricultural Econ* 1991;**73**:1410–15.
13. Cropper ML, Aydede SK, Portney PR. Rates of time preference for saving lives. *Am Econ Rev* 1992;**82**:469–72.
14. Cropper ML, Aydede SK, Portney PR. Preferences for life saving programs: how the public discount time and age. *J Risk Uncertainty* 1994;**8**:243–65.
15. Johannesson M, Johansson P-O. The discounting of lives saved in future generations – some empirical results. *Health Econ* 1996;**5**:329–32.
16. Johannesson M, Johansson P-O. The value of life extension and the marginal rate of time preference: a pilot study. *Appl Econ Lett* 1997;**4**:53–5.
17. Johannesson M, Johansson P-O. Quality of life and the WTP for an increased life expectancy at an advanced age. *J Public Econ* 1997;**65**:219–28.
18. Enemark U, Lyttkens CH, Troëng T, Weibull H, Ranstam J. Implicit discount rates of vascular surgeons in the management of abdominal aortic aneurysms. *Med Decis Making* 1998;**18**:168–77.
19. Lipscomb J. Time preference for health in cost-effectiveness analysis. *Med Care* 1989;**27**:S233–53.
20. Redelmeier DA, Heller DN. Time preference in medical decision making and cost-effectiveness analysis. *Med Decis Making* 1993;**13**:212–17.
21. MacKeigan LD, Larson LN, Draugalis JR, Bootman JL, Burns LR. Time preference for health gains versus health losses. *PharmacoEcon* 1993;**3**:374–86.
22. Dolan P, Gudex C. Time preference, duration and health state valuations. *Health Econ* 1995;**4**:289–99.
23. Loewenstein G. Anticipation and the valuation of delayed consumption. *Economic J* 1987;**97**:666–84.
24. Ahlbrecht M, Weber M. Hyperbolic discounting models in prescriptive theory of intertemporal choice. *Z Wirt Sozialwiss* 1995;**115**:535–68.
25. Hausman JA. Individual discount rates and the purchase and utilisation of energy-using durables. *Bell J Econ* 1979;**10**:33–54.
26. Gately D. Individual discount rates and the purchase and utilisation of energy-using durables: comment. *Bell J Econ* 1980;**11**:373–4.
27. Dreyfus MK, Viscusi WK. Rates of time preference and consumer valuations of automobile safety and fuel efficiency. *J Law Econ* 1995;**38**:79–105.
28. Kooreman P. Individual discounting, energy conservation, and household demand for lighting. *Resource Energy Econ* 1996;**18**:103–14.
29. Lang K, Ruud PA. Returns to schooling, implicit discount rates and black–white wage differentials. *Rev Econ Stat* 1986;**69**:41–7.
30. Lawrance EC. Poverty and the rate of time preference: evidence from panel data. *J Political Econ* 1991;**99**:54–77.
31. Viscusi WK, Moore MJ. Rates of time preference and valuations of the duration of life. *J Public Econ* 1989;**38**:297–317.
32. Moore MJ, Viscusi WK. The quantity-adjusted value of life. *Econ Inquiry* 1988;**26**:369–88.
33. Moore MJ, Viscusi WK. Discounting environmental health risks: new evidence and policy implications. *J Environ Econ Manage* 1990;**18**:S51–62

34. Thaler R. Some empirical evidence of dynamic inconsistency. *Econ Lett* 1981;**8**:201–7.

35. Fuchs VR. Time preference and health: an exploratory study. In: Fuchs VR, editor. Economics aspects of health. Chicago: University of Chicago Press; 1982.

36. Benzion U, Rapoport A, Yagil J. Discount rates inferred from decisions: an experimental study. *Manage Sci* 1989;**35**:270–84.

37. Myerson J, Green L. Discounting of delayed rewards: models of individual choice. *J Exp Anal Behav* 1995;**64**:263–76.

38. Rachlin H, Raineri A. Cross D. Subjective probability and delay. *J Exp Anal Behav* 1991;**55**:233–44.

39. Kirby KN, Marakovíc NN. Modeling myopic decisions: evidence for hyperbolic delay discounting within subjects and amounts. *Org Behav Hum Decis Processes* 1995;**64**:22–30.

40. Houston DA. Implicit discount rates and the purchase of untried, energy-saving durable goods. *J Consumer Res* 1983;**10**:236–46.

41. Sultan F, Winer RS. Time preferences for products and attributes and the adoption of technology-driven consumer durable innovations. *J Econ Psychol* 1993;**14**:587–613.

42. Loewenstein G, Prelec D. Anomalies in inter-temporal choice: evidence and an interpretation. *Q J Econ* 1992;**107**:573–97.

43. Kirby KN, Herrnstein RJ. Preference reversals due to myopic discounting of delayed reward. *Psychol Sci* 1995;**6**:83–9.

44. Green L, Fristoe N, Myerson J. Temporal discounting and preference reversals in choice between delayed outcomes. *Psychonomic Bull Rev* 1994;**1**: 383–9.

45. Ahlbrecht M, Weber M. An empirical study of intertemporal decision making in the case of risk. *Manage Sci* 1995;**43**:813–26.

46. Bleichrodt H, Johannesson M. Time preference for health: a test of stationarity versus decreasing timing aversion. *J Math Psychol.* In press.

47. Christensen-Szalanski JJJ. Discount functions and the measurement of patients' values: women's decisions during childbirth. *Med Decis Making* 1984;**4**:47–58.

48. Roelofsma PHMP. Modelling intertemporal choices: an anomaly approach. *Acta Psychol (Amst)* 1996;**93**:5–22.

49. Green L, Fry AF, Myerson J. Discounting of delayed rewards: a life-span comparison. *Psychol Sci* 1994;**5**:33–6.

50. Vuchinich RE, Simpson CA. Hyperbolic temporal discounting in social drinkers and problem drinkers. *Exp Clin Psychopharmacol* 1998;**6**:292–305.

51. Green L, Myerson J. Exponential versus hyperbolic discounting of delayed outcomes: risk and waiting time. *Am Zoologist* 1996;**36**:496–505.

52. Green L, Myerson J, Lichtman D, Fry A. Temporal discounting in choice between delayed rewards: the role of age and income. *Psychol Aging* 1996;**11**:79–84.

53. Green L, Myerson J, McFadden E. Rate of temporal discounting decreases with amount of reward. *Memory Cogn* 1997;**25**:715–23.

54. Kirby KN, Marakovíc NN. Delay discounting probabilistic rewards: rates decrease as amounts increase. *Psychonomic Bull Rev* 1996;**3**:100–4.

55. Shelley MK. Outcome signs, question frames and discount rates. *Manage Sci* 1993;**39**:806–15.

56. Rachlin H. Judgement, decision and choice. New York: Freeman; 1989.

57. Parsonage M, Neuburger H. Discounting and QALYS. *Health Econ* 1992;**1**:71–6.

58. Cairns JA. Discounting and health benefits: another perspective. *Health Econ* 1992;**1**:76–9.

59. van Hout B. Discounting costs and effects. *Health Econ* 1998;**7**:581–94.

60. Department of Health. Policy appraisal and health. London: HMSO; 1995.

61. Torrance GW, Blaker D, Detsky A, Kennedy W, Schubert F, Menon D, *et al.* Canadian guidelines for economic evaluation of pharmaceuticals. *PharmacoEcon* 1996;**9**:535–59.

62. Gold MR, Siegel JE, Weinstein MC. Cost-effectiveness in health and medicine. Oxford: Oxford University Press; 1996.

63. Gafni A. Time in health: can we measure individuals "pure time preferences"? *Med Decis Making* 1995;**15**:31–7.

64. Loewenstein G, Prelec D. Preferences for sequences of outcomes. *Psychol Rev* 1993;**100**:91–108.

65. Krahn M, Gafni A. Discounting in the economic evaluation of health care interventions. *Med Care* 1993;**31**:403–18.

66. Bretteville-Jensen AL. Addiction and discounting. *J Health Econ* 1999;**18**:393–407.

67. Chapman GB, Coups EJ. Time preferences and preventive health behavior: acceptance of the influenza vaccine. *Med Decis Making* 1999;**19**:307–14.

68. EuroQol Group. EuroQol – a new facility for the measurement of health-related quality of life. *Health Policy* 1990;**16**:199–208.

69. Dolan P. Modelling valuations for health states: the effect of duration. *Health Policy* 1996;**38**:189–203.

70. Sudman S, Bradburn NM. Asking questions: a practical guide to questionnaire design. San Francisco: Jossey-Bass; 1989.

71. Beattie J, Loomes G. The impact of incentives upon risky choice experiments. *J Risk Uncertainty* 1997;**14**:155–68.

72. Stone DN, Zeibart DA. A model of financial incentive effects in decision making. *Org Behav Human Decis Processes* 1995;**61**:250–61.

73. van der Pol MM, Cairns JA. Negative and zero time preference for health. *Health Econ Lett* 1999;**3**:3–9.

74. Loewenstein G, Prelec D. Negative time preference. *Am Econ Rev* 1991;**81**:347–52.

75. Greene W. Econometric analysis. New York: Macmillan; 1993.

76. Long JS. Regression models for categorical and limited dependent variables. Thousand Oaks: SAGE Publications; 1997.

77. McFadden D. Conditional logit analysis of qualitative choice behavior. In: Zarembka P, editor. Frontiers of econometrics. New York: Academic Press, 1973;105–142.

78. Goldstein H. Multilevel statistical models. London: Kluwer Academic Publishing; 1995.

79. Goldstein H, Rasbash J, Plewis I, Draper D, Browne W, Yang M, *et al.* A user's guide to MLwiN (Multilevel Models project). London: Institute of Education, University of London; 1998.

80. Wetherill GB. Intermediate statistical models. London: Chapman and Hall; 1981.

81. Jones AM. An econometric investigation of low birth weight in the United States. *J Health Econ* 1991;**10**:81–99.

82. McKelvey R, Zaviona W. Statistical model for the analysis of ordinal level dependent variables. *J Math Psychol* 1975;**4**:103–20.

83. Statacorp. Stata statistical software: release 5.0. College Station, TX: Stata Corporation; 1997.

84. Efron B. Bootstrap methods: another look at the jackknife. *Ann Stat* 1979;**7**:1–26.

85. Briggs AH, Wonderling DE, Mooney CZ. Pulling cost-effectiveness analysis up by its bootstraps: a non-parametric approach to confidence interval estimation. *Health Econ* 1997;**6**:327–40.

86. Godfrey LG. Misspecification test in econometrics: the Lagrange multiplier principle and other approaches. Cambridge: Cambridge University Press; 1988.

87. Greene W. LIMDEP version 7.0. New York: Econometric Software Inc; 1998.

88. Bryan S, Buxton M, Sheldon R, Grant R. The use of magnetic resonance imaging for the investigation of knee injuries: an investigation of preferences. *Health Econ* 1998;**7**:595–603.

89. Brown TC, Champ PA, Bishop RC, McCollum DW. Which response format reveals the truth about donations to a public good? *Land Econ* 1996; **72**:152–66.

90. Boyle KJ, Johnson FR, McCollum DW, Desvousges WH, Dunford RW, Hudson SP. Valuing public goods: discrete versus continuous contingent valuation responses. *Land Econ* 1996;**72**:381–96.

91. Ready RC, Buzby JC, Hu D. Differences between continuous and discrete choice contingent valuation. *Land Econ* 1996;**72**:397–411.

92. Kriström B. Comparing continuous and discrete contingent valuation questions. *Environ Resource Econ* 1993;**3**:63–71.

93. Kealy MJ, Turner RW. A test of the equality of closed-ended and open-ended contingent valuations. *Am J Agric Econ* 1993;**75**:321–31.

94. Shechter MA. A comparative study of environmental amenity valuations. *Environ Resource Econ* 1991;**1**:129–55.

95. Seller C, Stoll JR, Chavas J-P. Valuation of empirical measures of welfare change: a comparison of non-market techniques. *Land Econ* 1985;**61**:156–75.

96. Carson RT. Contingent valuation: theoretical advances and empirical tests since the NOAA panel. *Am J Agric Econ* 1997;**79**:1501–7.

97. Arrow K, Solow R, Portney PR, Anderson GF, Dwyer SA, Faden RR. Report of the NOAA panel of contingent valuation. *Fed Register* 1993;**58**:4601–14.

98. Johannesson M, Jonsson B. Willingness to pay for antihypertensive therapy: results of a Swedish pilot study. *J Health Econ* 1991;**10**:461–74.

99. Strotz RH. Myopia and inconsistency in dynamic utility maximization. *Rev Econ Studies* 1956;**23**:165–80.

100. Harvey CM. Value functions for infinite-period planning. *Manage Sci* 1986;**32**:1123–39.

101. Mazur JE. An adjusting procedure for studying delayed reinforcement. In: Commons ML, Mazur JE, Nevin JA, Rachlin H, editors. Quantitative analyses of behavior: V. The effect of delay and intervening events on reinforcement value. Hillsdale, NY: Erlbaum; 1987.

102. Camerer C. Bounded rationality in individual decision making. *Exp Econ* 1999;**1**:163–83.

Appendix 1

Options and criteria

	Degree of difficulty to answer	Degree of realism	Computation of discount factor	Other factors than time preference influence response
Number of points in time Two points Profile	Having to consider only two points in time is easier.	A profile is more realistic whereby ill health is spread over more than one year instead of being concentrated in one particular year.	Discount factors are more easily computed when only two points in time are considered. Discount factors cannot, in some cases, be computed from profiles.	
Timing versus quantity Timing Quantity	It is easier to choose when something takes place than to choose how much health or ill health has to be consumed at a certain time.	Choosing how ill or how healthy is less realistic. Timing is more realistic (treatments can, for instance, postpone a certain disease taking place).	The choice of timing versus quantity does not have an influence on how difficult it may be, or whether it is possible, to calculate a discount factor.	In giving a choice of timing it is more likely that other factors will influence the response. This is certainly so in the short term (respondents might have things planned) and also in the long term.
Basic health state Full health Ill health	Since majority will be healthy it is easier to imagine full health.	Majority of people will be healthy so full health as basic health state appears more realistic; choosing when or how much to be ill is less realistic.	Choice of basic health state does not have an influence on how to compute the discount factor.	People are averse to ill health. When having to choose ill health, other factors might influence the response such as fear of being ill.
Number of ill health states One ill health state More than one	Considering two or more ill health states will make the question more complex.	It is more realistic to be in more than one ill health state when you have a spell of ill health (for instance, going from being ill to being less severely ill to being healthy).	If two or more ill health states are used, these have to be valued before a discount factor can be computed.	
Single or mix of health states in a year Single health state Mix of health states	Being in a single health state is easier because only one health state has to be considered for each year.	Being in one single health state during a whole year is less realistic unless it is a chronic condition.	It is probably easier to compute a discount factor when only one health state is used in a year.	
Period considered Limited period Remaining life	Respondents might have difficulty imagining their remaining life if they have to consider a long period.	Especially for older respondents, a long remaining life scenario is less realistic.	The choice between a limited period and remaining life does not influence whether a discount factor can be calculated but a limited period is less complex.	Age has more influence on the responses when considering remaining life. Older respondents might not be able or willing to consider a long remaining life.

Appendix 2

Examples of questions

Introductory question

Imagine the following state of ill health.

> You have some problems with performing your usual activities (e.g. work, study, housework, family or leisure activities) and you have moderate pain or discomfort. You have no problems in walking about, or with washing and dressing yourself and you are not anxious or depressed.

Please think of this state of ill health and indicate on the scale below how good or bad you think this ill health is.

Worst imaginable health 0 10 20 30 40 50 60 70 80 90 100 Best imaginable health

Open-ended method (Questionnaires I and II)

Imagine that you will be ill (as described in the box above) starting **2** years from now for **20** days. There is a minor, one-off, treatment available that will postpone this spell of ill health to a point further in the future. For instance, the treatment could have the following effects: your period of ill health would start **9** years from now instead of **2** years from now; and you would then be ill for **30** days instead of **20** days.

You might think this treatment is a good idea: the advantage of postponing the ill health outweighs the disadvantage of being ill for a longer period. Or you might think the treatment is not worthwhile: you do value the postponement but the advantage of this is outweighed by the disadvantage of being ill for a longer period; or you might simply prefer to be ill 2 years from now instead of 9 years from now.

> Imagine that you will be ill starting **2** years from now for **20** days and that treatment is available which will postpone this spell of ill health.

What is the maximum number of days of ill health that would still make the treatment worthwhile for you? For example, say that the treatment can postpone the period of ill health to **6** years in the future. If the number of days of ill health in that year were zero, probably everyone would choose the treatment. As the number of days of ill health in that year increases, individuals would at some point no longer prefer to be treated. What we are interested in is the **maximum number** of days of ill health at which you would still choose to be treated.

> If the ill health would then start 4 years from now, what is the maximum number of days of ill health that would still make the treatment worthwhile?

0 10 20 30 40 50 60 70 80 Days

more than 80 namely: ______ days

[For questionnaire II, substitute 'a group of middle-aged people' for 'you'.]

Discrete choice experiment (Questionnaires III and IV)

Imagine that you will be ill (as described in the box above). There are two **alternative** treatments (A and B) available. The effects of the treatments vary with regard to **when** the illness will occur and for **how long** (you cannot be cured completely). An example of how the effects can vary is shown below. With treatment A you will be ill starting **2** years from now for **20** days and with treatment B you will be ill starting **6** years from now for **48** days. Assuming that everything else about the treatments is the same (i.e. severity of the treatment, side-effects, costs), which treatment would you prefer?

• ***Example***

	When you are ill	For how long	Which treatment would you prefer?
Treatment A	In 2 years' time	20 days	
Treatment B	In 6 years' time	48 days	✔
		No preference	

Please consider each of the following choices **separately** and indicate which treatment you prefer by putting a tick (✔) in the appropriate box. Tick one box for each choice.

• ***Choice***

	When you are ill	For how long	Which treatment would you prefer?
Treatment A	In 2 years' time	20 days	
Treatment B	In 6 years' time	48 days	
		No preference	

[For questionnaire IV, substitute ' a group of middle-aged people' for 'you'.]

General questions (Questionnaires I, II, III and IV)

Finally, some general questions about yourself. All answers are **anonymous**. (Please tick appropriate boxes.)

1. Year of birth ☐

2. Male ☐ Female ☐

3. How would you describe your current long-term health?
   - Good ☐
   - Fair ☐
   - Poor ☐

4. How many cigarettes do you smoke, on average, per day?
   - None ☐
   - 1–5 ☐
   - 6+ ☐

5. What is the highest level of education you have completed, or are currently undergoing?
   - Secondary school ☐
   - Other professional or technical qualification after leaving school ☐
   - University degree ☐

Appendix 3

Implied discount rates for pair-wise comparisons

A	(2/20)														
B	(2/24)	Do													
C	(2/26)	Do	Do												
D	(2/33)	Do	Do	Do											
E	(2/38)	Do	Do	Do	Do										
F	(7/20)	**0.00**	–0.04	–0.05	–0.10	–0.12									
G	(7/24)	0.04	**0.00**	–0.02	–0.06	–0.09	Do								
H	(7/26)	0.05	0.02	**0.00**	–0.05	–0.07	Do	Do							
I	(7/33)	0.11	0.07	0.05	**0.00**	–0.03	Do	Do	Do						
J	(7/38)	0.14	0.10	0.08	0.03	**0.00**	Do	Do	Do	Do					
K	(15/20)	**0.00**	–0.01	–0.02	–0.04	–0.05	**0.00**	–0.02	–0.03	–0.06	–0.08				
L	(15/24)	0.01	**0.00**	–0.01	–0.02	–0.03	0.02	**0.00**	–0.01	–0.04	–0.06	Do			
M	(15/26)	0.02	0.01	**0.00**	–0.02	–0.03	0.03	0.01	**0.00**	–0.03	–0.05	Do	Do		
N	(15/33)	0.04	0.02	0.02	**0.00**	–0.01	0.06	0.04	0.03	**0.00**	–0.02	Do	Do	Do	
O	(15/38)	0.05	0.04	0.03	0.01	**0.00**	0.08	0.06	0.05	0.02	**0.00**	Do	Do	Do	Do
Scenario (years/days)		A	B	C	D	E	F	G	H	I	J	K	L	M	N

NB: 'Do' indicates dominated pairwise comparisons – the years in which the spell of ill health occurs is the same in the two scenarios and only the duration differs.

Appendix 4

Transitivity tests

Test	Choices	Years (scenario A and B in 1st choice)	Years (scenario A and B in 2nd choice)	Days (scenario A and B in 1st choice)[a]	Days (scenario A and B in 2nd choice)[a]	Intransitive if chooses:[b]
Version 1						
1	3 & 7	2–7	2–7	**20**–38	**20**–33	B A or I A
2	1 & 7	2–7	2–7	26–**33**	20–**33**	A B or I B
3	4 & 6	7–15	7–15	**20**–38	**20**–26	B A or I A
4	4 & 8	7–15	7–15	**20**–38	**20**–24	B A or I A
5	6 & 8	7–15	7–15	**20**–26	**20**–24	B A or I A
Version 2						
6	3 & 7	2–7	2–7	**20**–38	**20**–33	B A or I A
7	1 & 4	7–15	7–15	24–**38**	20–**38**	A B or I B

[a] *Numbers are set in bold type to highlight when they are identical for both choices*
[b] *I indicates no preference*

Health Technology Assessment panel membership

This report was identified as a priority by the Methodology Group.

Acute Sector Panel

* Previous Chair

continued

continued

continued

continued

Primary and Community Care Panel

Current members

Chair:
Dr John Tripp
Royal Devon & Exeter Healthcare NHS Trust

Mr Kevin Barton
East London & City Health Authority

Professor John Bond
University of Newcastle-upon-Tyne

Dr John Brazier
University of Sheffield

Ms Judith Brodie
Cancer BACUP

Mr Shaun Brogan
Ridgeway Primary Care Group, Aylesbury

Mr Joe Corkill
National Association for Patient Participation

Dr Nicky Cullum
University of York

Professor Pam Enderby
University of Sheffield

Dr Andrew Farmer
Institute of Health Sciences, Oxford

Dr Jim Ford
Department of Health

Professor Richard Hobbs
University of Birmingham

Professor Allen Hutchinson
University of Sheffield

Dr Aidan MacFarlane
Independent Consultant

Professor David Mant
Institute of Health Sciences, Oxford

Dr Chris McCall
General Practitioner, Dorset

Dr Robert Peveler
University of Southampton

Professor Jennie Popay
University of Salford

Dr Ken Stein
North & East Devon Health Authority

Past members

Professor Angela Coulter*
King's Fund, London

Professor Martin Roland*
University of Manchester

Dr Simon Allison
University of Nottingham

Professor Shah Ebrahim
Royal Free Hospital, London

Ms Cathy Gritzner
King's Fund, London

Professor Andrew Haines
RDRD, North Thames Regional Health Authority

Dr Nicholas Hicks
Oxfordshire Health Authority

Mr Edward Jones
Rochdale FHSA

Professor Roger Jones
Guy's, King's & St Thomas's School of Medicine & Dentistry, London

Mr Lionel Joyce
Chief Executive, Newcastle City Health NHS Trust

Professor Martin Knapp
London School of Economics & Political Science

Dr Phillip Leech
Department of Health

Professor Karen Luker
University of Liverpool

Dr Fiona Moss
Thames Postgraduate Medical & Dental Education

Professor Dianne Newham
King's College London

Professor Gillian Parker
University of Leicester

Dr Mary Renfrew
University of Oxford

Ms Hilary Scott
Tower Hamlets Healthcare NHS Trust, London

* Previous Chair

National Coordinating Centre for Health Technology Assessment, Advisory Group

Copies of this report can be obtained from:

The National Coordinating Centre for Health Technology Assessment,
Mailpoint 728, Boldrewood,
University of Southampton,
Southampton, SO16 7PX, UK.
Fax: +44 (0) 23 8059 5639 Email: hta@soton.ac.uk
http://www.hta.nhsweb.nhs.uk

ISSN 1366-5278