If nudge cannot be applied: a litmus test of the readers' stance on paternalism

Chen Li · Zhihua Li · Peter P. Wakker

Abstract A central question in many debates on paternalism is whether a decision analyst can ever go against the stated preference of a client, even if merely intending to improve the decisions for the client. Using four gedanken-experiments, this paper shows that this central question, so cleverly and aptly avoided by libertarian paternalism (nudge), cannot always be avoided. The four thought experiments, while purely hypothetical, serve to raise and specify the critical arguments in a maximally clear and pure manner. The first purpose of the paper is, accordingly, to provide a litmus test on the readers' stance on paternalism. We thus also survey and organize the various stances in the literature. The secondary purpose of this paper is to argue that paternalism cannot always be avoided and consumer sovereignty cannot always be respected. However, this argument will remain controversial.

Keywords Nudge · Libertarian paternalism · Debiasing · Standard gamble · Choice inconsistency

1 Introduction

Descriptive models in behavioral economics were devised to capture empirical deviations from classical decision principles. Studying these descriptive models helps us to better understand certain empirical patterns of inconsistencies and/or biases of human behavior in decision-making. Furthermore, taking the classical principles as normative, the empirical deviations from them, captured by descriptive models, leave space for

C. Li · Z. Li · P. P. Wakker (✉)
Erasmus School of Economics, Erasmus University Rotterdam, P.O. Box 1738,
3000 DR Rotterdam, The Netherlands
e-mail: Wakker@ese.eur.nl
http://people.few.eur.nl/wakker

improving the empirically observed decisions. Paradoxically, the prescriptive[1] value of normative models depends on the existence of such deviations; as Raiffa (1961) put it: "We do not have to teach people what comes naturally."

Attempts at prescriptive improvements of decisions can lead to paternalism, and consequently ethical and moral objections can be raised. There have indeed been many debates about paternalism. A central issue is whether one (for instance, a decision analyst) can ever go against the stated preferences of the people affected. Critics of paternalism emphasize that one should never impose a choice on people against their own will.

A new impulse for decision theory came from libertarian paternalism, often referred to as nudge, following Thaler and Sunstein (2003). A clever way was found to improve empirical decisions using the techniques of behavioral economics while never going against the will of the people affected. People are nudged in the right direction of choice, while left with full freedom of choice. This approach can be used in situations where people have no clear preferences (incompleteness), and it has led to a variety of applications (Jones et al. 2011; Thaler and Sunstein 2008).

In practice, situations arise where people's preferences are explicitly stated while at variance with normative principles. Then the *central question*, critical to debates on paternalism, is whether a decision analyst can seek to improve decisions while going against the stated preferences of a client, or whether one should respect stated preferences and forego purported improvements. Libertarian paternalism avoids addressing this central question. We use four thought experiments to show that this question cannot always be avoided. In each of these four experiments, a dilemma is faced, and a paternalistic decision should be made or forgone, with no possibility of avoiding the dilemma comprised in the central question.

The exact details of our four thought experiments are unlikely to arise in practice, and they were not developed for this purpose. They constitute gedanken-experiments common in philosophy, serving to maximally clarify the relevant issues. These issues are central in many practical decisions, playing a role in all decisions that affect other people. The four thought experiments represent real decision situations where the dilemma in the central question cannot be avoided, and a decision has to be made, one way or another. This leads to our main purpose: we provide a litmus test on the readers' stance on paternalism, and a way to organize stances in the literature. Our secondary purpose is to convince the readers that paternalism itself is sometimes appropriate, and consumer sovereignty should not always be respected. However, this argument is controversial and no consensus is likely to come soon.

2 Views on paternalism

Paternalistic interventions often improve people's well-being. Taxation policies are used to regulate addictive behaviors such as smoking. The "no drink driving" regulation that helps reduce traffic accidents is another example. Yet many objections to

---

[1] Prescriptive applications use normative models to improve decisions but reckon with descriptive limitations.

paternalistic interventions can still be raised. One such objection is based on a desire to avoid power manipulations. Paternalistic interventions give decision analysts/policy makers the power to influence the decisions of others, and may be misused for self-interest rather than for the well-being of the people affected. Although this is a strong argument against paternalism, it will not be considered in this paper for reasons of simplicity. We focus on what we called the central question and assume that the only purpose of a decision analyst is to optimize others' well-being.

A second objection against paternalism takes issue with the assumption that particular decision principles can be qualified as normative. In reply, our analysis will only assume stochastic dominance and transitivity as normative principles, whose normative status is uncontroversial, although some doubts have still been expressed (Bell 1982; Loomes and Sugden 1982; Mandler 2005). We emphasize that our discussion needs no commitment to a normative status of expected utility or any other specific decision theory. Our examples also pose a dilemma for people who advocate using heuristics rather than optimization theories (Berg and Gigerenzer 2010).

A third objection extends the second one. The argument is that individual decision-makers' preferences reflect their true values which they know better than any outsider can. Therefore, it is best for individual decision-makers to decide for themselves what is best for them. This objection is sometimes supported by Hume's famous citation "reason is, and ought only to be the slave of the passions," where preferences are taken as passions that should never be overruled by paternalistic reasons (Hume 1740).[2]

A fourth objection is that people should always have the freedom to follow their own preferences, irrespective of whether these preferences are optimal in some sense or not (McQuillin and Sugden 2012, p. 556). In a similar spirit, it can be argued that people need not get what others think to be best; instead, they should get what they choose for themselves. In defense of this objection, some may advance an evolutionary interpretation, claiming that the best preferences will then survive in the long run. Others may base their arguments on direct ethical principles, assigning an intrinsic value to freedom of choice (Sugden 2004).

For later discussion, we group these four objections, while raised from different perspectives, under the name anti-paternalism. Other terms used in the literature for anti-paternalistic positions are the Humean view on preference, or consumer sovereignty.

Several moderate positions have been advanced. Cautious paternalism (O'Donoghue and Rabin 1999) and libertarian paternalism are similar, combined in soft paternalism (McQuillin and Sugden 2012, p. 560). They advocate interventions, such as smartly designed choice structures, education, and/or incentive schemes, which nudge people to better choices, while keeping the ultimate decision power with individuals. In this way, people without a stable and consistent preference will be influenced by the structure and therefore go for the choice that is most probably the best for them. On the other hand, those who choose differently from the paternalistic suggestion are more likely the ones who understand themselves better and opt for a genuinely different preference.

---

[2] Our interpretation of Hume's citation is different, and we think that it does not preclude paternalism. Broome (1993) and Sugden (1998, p. 48) present critical discussions.

Asymmetric paternalism (Camerer et al. 2003) gives a special importance to rational individuals. It purports that rational individuals, who can independently make the best decisions, should be harmed as little as possible. Therefore, in cases where a paternalistic decision needs to be made for a group of people, but a separation between rational and irrational individuals is impossible, priority should be given to ensure that no harm is done to rational individuals. Paternalistic interventions should be restrained even if, on average, they benefit the majority of people.

Moderate positions make paternalistic interventions less controversial as they take into account an individual's own will. However, their applicability is limited. In many situations, they cannot provide us with guidance, as our thought experiments will demonstrate.

The most far-reaching form of paternalism is optimal paternalism, advocated by Zamir (1998) and O'Donoghue and Rabin (2003). They argue against the overweighting of rational individuals in asymmetric paternalism, and favor occasional coercive paternalistic interventions. The latter can be justified in situations where the cost of allowing irrational individuals to make errors significantly exceeds the harm brought to rational individuals by paternalistic interventions. These authors recommend using cost benefit analyses to find optimal policies, thus taking a purely economic perspective.

3 The decision to be taken, and two quality-of-life measurements

In the thought experiments that we now consider, you can imagine yourself to be a decision analyst (or a doctor) who has to decide on a medical treatment for a client. Your client's status quo is that he suffers from impaired vision. There is only one treatment available that may cure him and bring back perfect vision, but there is a risk of failure, leading to complete blindness. You must choose between treatment and the status quo (no treatment) for your client.

Unfortunately, in addition to impaired vision, your client is now unconscious, which is only temporary, has no serious consequences, and does not affect vision or health. This means that you cannot communicate with him now. You cannot wait until the client regains consciousness, because the treatment works only if carried out immediately. Hence, arguments for patient autonomy play no role here. For simplicity, we assume that the treatment is cheap, simple, and has no side effects. We assume that medical arguments give no clear verdict. The subjective quality-of-life perception of your client regarding the vision levels is the deciding factor.

As explained before, we assume that your only interest is to choose what is best for your client. You do not have to account for your decision to any outside party, but you do account to your own conscience. Legal considerations play no role, but moral ones do. Note that both a choice of treatment and a choice of the status quo are your choice, and morally you are equally responsible for both.

Your decision is based on information about the quality of life evaluation of the status quo (called status quo quality from now on) as relative to the quality of treatment (called treatment quality from now on). To determine the former, we consider two quality of life measurement methods commonly used in the health domain. We assume
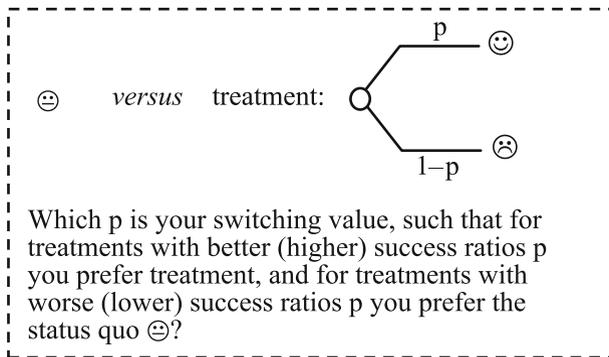
Fig. 1 Probability equivalent question

that, besides these two measurements, no other relevant information about the status quo quality is available. In particular, you cannot make your own assessment of this quality and must go entirely by the information from the quality of life measurements.

The first, traditional, measurement method is the *probability equivalent* (*PE*) method, often called the standard gamble method in the health domain (Drummond et al. 1997). In the measurement survey, participants are asked to choose between the status quo (indicated by a symbol ☺ in Fig. 1) and a hypothetical treatment. All treatments considered either result in complete recovery with perfect vision (indicated by a smiley symbol ☺) or the worst possible outcome of complete blindness (indicated by a ☹ symbol). The treatments differ only in their success ratio—the probability of complete recovery (p in Fig. 1). Participants are asked the question in Fig. 1.

The switching value p is an index of the quality of life of the status quo. We call p the *status quo quality*, abbreviated *sqq*. The higher sqq is, the better the status quo is judged to be by the client. Although this method is often analyzed assuming expected utility, we do not need this assumption in what follows. We will never assume more than transitivity and stochastic dominance.[3] Given the switching value sqq of the client, we have the indifference in Figure 2.

The second measurement method is the *certainty equivalent* (*CE*) method, which is indirect and more laborious. Each survey participant is presented with a rich set of vision levels varying from perfect vision to blindness. The status quo is contained in this set. Initially, some treatment with success ratio p_1 is chosen. Participants are asked the question in Fig. 3.

If the switching vision level given by the client is better than the status quo, then the success ratio is decreased, and if it is worse, then the success ratio is increased. The choice question is repeated with the new success ratio. This process goes on

---

[3] Our assumption of indifference at the switching value amounts to a kind of continuity, but is not essential to our arguments and is only made to simplify the presentation. Strict preferences at either side of the switching value suffice for our analysis. The fact that a choice must be made in each decision situation does not reflect an extra completeness assumption, but is intrinsic to the decision situation. This is discussed further in §5.3
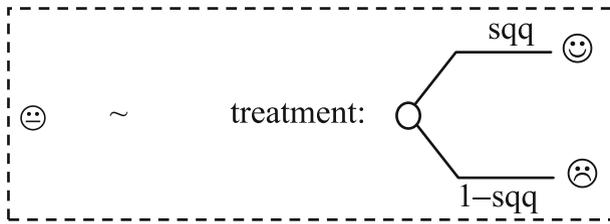
Fig. 2 Status quo quality (probability sqq)
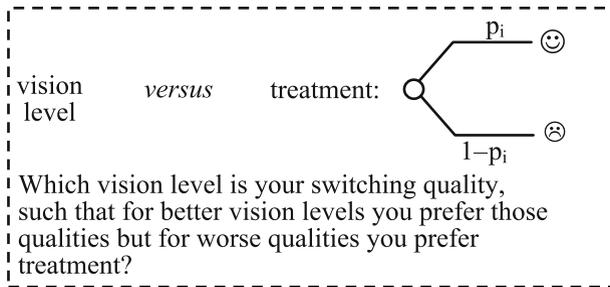
Fig. 3 Certainty equivalent question
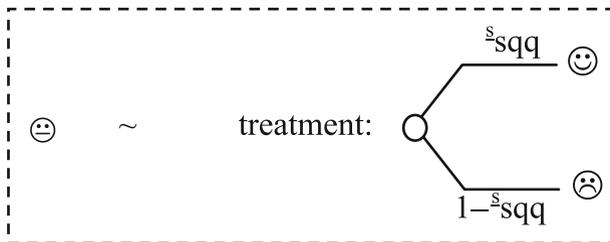
Fig. 4 Second status quo quality (probability ^{s}sqq)

until a success ratio p_i is found for which the switching vision level is the status quo. The resulting success ratio is again an index of quality of life of the status quo. We abbreviate it as the *second status quo quality*, or $^s sqq$. The higher ^{s}sqq is, the better the status quo is judged to be by the client. We now have the indifference in Figure 4.

It is natural to expect that sqq $=^s$sqq, which follows from the classical decision principles of transitivity and stochastic dominance. The equality is especially natural if the interviews were conducted carefully, with clients understanding the stakes well and (if possible …) being proficient in decision theory. It would also help if interviewers could interact with the subjects and influence answers, which is useful in prescriptive as opposed to descriptive applications (Keeney and Raiffa 1976, §5.8; Payne et al. 1999; Edwards and Elwyn 2006). However, we assume that the measurements were conducted as is common in most descriptive and applied studies, where there is only limited time for the measurements and explanations, where there can be no interaction to influence answers, and where interviewers themselves are often not well-versed in
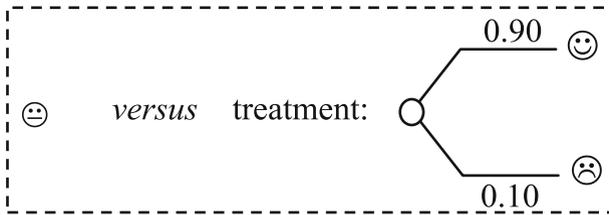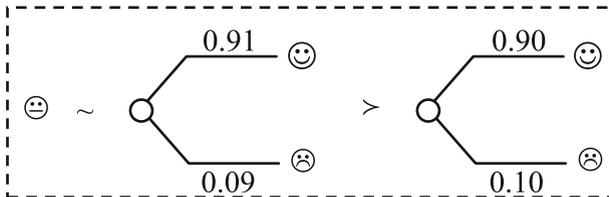
Fig. 5 The real treatment decision

Fig. 6 The analysis of Case 4.1S

decision theory. These limitations are unavoidable in large-scale quality-of-life measurements, and obviously preclude any interactions to discover or correct violations of decision principles. In this case, violations of the equality sqq = ^{s}sqq can result.

4 The four cases

This section presents the four thought experiments used to identify stances on paternalism. Henceforth, we will call them four cases. From now on, imagine that the only treatment available for your client has a success ratio 0.90. Because the success ratio is an index of the quality of the treatment, we call it the *treatment quality*, or *tq*, in analogy to sqq. The decision situation faced by you is depicted in Fig. 5.

In each of the following four cases, our question to the readers will be: given that the tq of the treatment is 0.90, would you choose treatment or the status quo for your client? The readers may want to make up their own mind immediately after reading each case, and before reading the pro and con arguments put forward by others.

CASE 4.1S [statistical information]. You do not know the sqq of your client. But you do know the sqq, based on the PE method, of 10,000 similar clients with the same status quo vision level as your client. The mean, median, and mode sqq over these 10,000 clients are all 0.91, slightly exceeding the tq of 0.90. □

In Case 4.1S, although the sqq of your client is unknown, you probably consider 0.91 to be the best estimate of your client's sqq, based on the information gathered from 10,000 similar clients. You will probably choose for the status quo, which follows naturally from the analysis in Fig. 6, which assumes only transitivity and stochastic dominance. [4] We expect this to be the answer of most readers (when first reading).

---

[4] The left indifference follows the PE measurement (Fig. 2) with sqq = 0.91. The strict preference follows from stochastic dominance. The preference of status quo over treatment then follows from transitivity.

CASE 4.2DS [double statistical information]. In this case, besides the sqqs of Case 4.1S, you also have information about the ^{s}sqqs of 10,000 similar clients. The mean, median, and mode of ^{s}sqq are, however, 0.85, which is considerably lower than the sqq of 0.91 found before, and the tq of 0.90.   □

In Case 4.2DS you face a large inconsistency in the data. At least one of the two measurements is incorrect and cannot reflect a genuine maximization of happiness. A natural first reaction is that the data are of poor quality, and that no decision should be based on such poor data. You would want to obtain better information. This could be done by interacting with 10,000 similar clients, or at least with some of them, to reconcile the inconsistencies. You could also communicate with the client or people close to him. However, in all of the four thought experiments, we assume that you have to make a decision now, with no other information available, nor any possibility to interact. Such situations are common in practice, especially in policy decisions. Facing the incoherence in revealed preferences, our thought experiment requires that you, as a decision analyst, devise a strategy to resolve it. Bear in mind that a refusal to take any action is not possible. Doing nothing means choosing the status quo.

As an sqq of 0.91 suggests a preference for the status quo, whereas an ^{s}sqq of 0.85 suggests the opposite, the two methods give contradictory suggestions. You may reason that, if the truth is in the middle between 0.91 and 0.85 in Case 4.2DS, then 0.88 is the best estimate of the status quo quality. It is below the treatment quality tq = 0.90, suggesting that treatment is better. This answer is enhanced by the observation that a wrong treatment decision is off by only 0.01 if the sqq of 0.91 reflects the true preference, whereas a wrong choice of the status quo is off by 0.05 if the ^{s}sqq of 0.85 reveals the truth. Hence, we expect most readers to choose treatment in Case 4.2DS.

CASE 4.3SI [statistical and individual information]. In this case, besides the sqqs and ^{s}sqqs of the previous cases, you also know the sqq of your client, and it is 0.91.
                                                                                                       □

Here we expect the readers to be divided. Although the information about 10,000 clients reveals large inconsistencies, the information about the preference of the client himself does not. His sqq exceeds the tq, suggesting a preference of the status quo over treatment. It is natural to think that the client's own stated preference should be more relevant than evaluations by other similar clients, further supporting your choice of the status quo. Yet an argument for treatment can also be supported. Although there is no direct evidence of inconsistent preference of your client, you can expect that he is probably like the 10,000 others, and is probably subject to the same inconsistencies. Taking possible inconsistencies into account, you can still adhere to the treatment decision of the previous Case 4.2DS, and also choose treatment in the present Case 4.3SI.

CASE 4.4I [individual information]. In this case, you do not have the information on similar clients. You only know the sqq of your client, which is 0.91.   □

Case 4.4I seems to be even clearer than Case 4.1S. Unlike Case 4.2DS and Case 4.3SI, there are no apparent inconsistencies in the data available on this case. Given that your client's sqq exceeds tq, we expect most readers to prefer the status quo here.

Table 1 The most common stances on paternalism

	4.1S (statistical); sqq = 0.91	4.2DS (double statistical); sqq = 0.91; $^{\underline{s}}$sqq = 0.85	4.3SI (statistical and individual); sqq = 0.91; $^{\underline{s}}$sqq = 0.85; individual sqq = 0.91	4.4I (individual); sqq = 0.91
1. Anti-paternalism	sq	tr	sq	sq
2. Weak anti-paternalism 1	sq	tr	tr	sq
3. Weak anti-paternalism 2	tr	tr	sq	sq
4. Weak paternalism	tr	tr	tr	sq
5. Strong paternalism	tr	tr	tr	tr

tr choose treatment, *sq* choose the status quo

5 Using the cases as a litmus test on paternalistic views

Table 1 lists the most common stances, ordered from least to most paternalistic, with the number of treatment (tr) choices increasing.

5.1 Invalid stances

Before discussing valid stances, we first discuss some invalid ones.

Refusal stance This stance entails the refusal to take any decision in Case 4.2DS. Strictly speaking, this stance is not possible in our thought experiments, but we know from experience that some readers will still want to consider it. The argument for refusal can be that no decision can be taken on such poor information, and one really has to get better information. Expressing doubts about decisions without committing to anyone may be possible in philosophical debates, but is not possible in our thought experiments, as it usually is not in practice. A purpose of our thought experiments is to show that taking a position on reconciling inconsistencies is unavoidable, contrary to what the refusal stance entails. People sympathetic to the refusal stance may include practitioners reluctant to involve in gedanken-experiments, and experimenters reluctant to consider hypothetical choice, even though the latter is essential in *prescriptive* applications (Keeney and Raiffa 1976, §1.4.3). We will not consider the refusal stance any further.

The ostrich stance [row 1 or 2] A widespread misunderstanding is to think that if your data do not show any violation of a model, then you may use that model for data analysis, even though other studies designed to test the model did find violations.[5]

---

[5] See Birnbaum (1992, p. 21, 2nd column, 2nd para), Cohen and Einav (2007, pp. 746–747), and Diamond (2008, p. 1860, 1st para). A related problem is that many medical applications use PE measurements as the

The data of Cases 4.1S and 4.4I do not show a violation of expected utility and, hence, according to the ostrich stance, one may use expected utility to analyze these cases. Then the status quo is chosen, as easily follows from transitivity and stochastic dominance (conditions verified by expected utility). Thus, weak anti-paternalism in row 2 may result, or, more likely, anti-paternalism in row 1.

Many applications of expected utility and, for instance, many analyses of PE and CE measurements, are based on this misunderstanding. We disagree with the ostrich stance. In the Cases 4.1S and 4.4I, no violations of expected utility showed up only because the data did not test for them. Assuming the absence of violations is like a doctor who declares a disease non-existent simply because he did not test for it. Inconsistencies are not the real problem, but they are a symptom of the real problem. The real problem concerns the biases in people's preferences. They will be present in Cases 4.1S and 4.4I as much as in the other two cases. Evidence for general violations of expected utility is reviewed by Camerer (1995) and Starmer (2000). Violations for the types of questions considered in our thought experiments are referenced later.

Non-statistical stance [row 1 or 3] Many people erroneously think that statistical information is not relevant to individual cases. Therefore, the data of the 10,000 similar subjects supposedly should play no role in the treatment decision of your client. This stance will lead to a choice of the status quo in Cases 4.3SI and 4.4I, only paying uncritical attention to the information coming from one individual client there. This stance can lead to anti-paternalism in row 1 or weak anti-paternalism in row 3. The non-statistical stance is surprisingly widespread (Steiner 1999). In the health domain, it is often based on the Hippocratic Oath that doctors take, swearing to do what is best for their patient (Fuchs 1974). Many doctors erroneously take this to imply that they can ignore not only the interests of patients other than their own, but also relevant *information* from patients other than their own. In an historical study, Murphy (1981) discusses the non-statistical stance in medicine in France in the 19th century. This stance also underlies the distinction often made between statistical lives and identified lives. Society commonly weighs identified lives higher than statistical lives (Schelling 1968), leading, for instance, to overspending on the treatment of rare diseases.

Some extreme advocates of the frequentist interpretation of probability, and critics of the concept of subjective probability, are also open to the non-statistical stance (Gigerenzer 1991, pp. 260–261, examples 1 and 2; Lopes 1981; Shackle 1949, p. 71). Such views also appear in discussions of the Monty Hall problem. For instance, Baumann (2005) wrote, supporting the non-statistical stance: "If the best argument so far for switching in an isolated individual case (not in a series of cases) fails, then one might wonder whether probabilistic arguments say anything at all about isolated individual cases."

Incoherence stance [row 1] Berg and Gigerenzer (2010, p. 148) argue that there is no irrationality in inconsistent preferences, writing "No studies we are aware of show that

---

Footnote 5 continued

gold standard, based only on the *normative* expected utility foundation, without concern about the many descriptive biases that have been documented (Drummond et al. 1997; Torrance and Feeney 1989, p. 560).

deviators from rational choice earn less money, live shorter lives, or are less happy." They argue for ecological rationality, with decision heuristics adapted to environments. Their ecological argument is not useful for the dilemma presented here and, in general, is tangential to the problems of inconsistency. Our gedanken-experiment requires a position to be taken regarding the inconsistency, and environments and ecologies provide no escape from the dilemma. Other arguments in favor of outcome-oriented ("correspondence") policies, with no direct concern for internal coherence, have been advanced (Friedman 1953; Hammond 2006; Smith 2008). These authors do not reject the usefulness of resolving inconsistencies explicitly as do Berg and Gigerenzer (2010).

5.2 Valid stances

This and the next subsection first present our viewpoint on the four cases. Then we survey several other stances in the literature, using the four cases to discriminate between them. Hence these two subsections are more detailed than the rest of this paper.

Strong paternalism [row 5; our stance] To position our upcoming discussion of various valid stances, we start with our own. We recommend treatment in all cases. In Case 4.2DS, we expect that virtually all readers will choose treatment. The information, poor as it is, directs to this decision, as discussed before. In Case 4.3SI, the individual information added suggests that the client is like the 10,000 others, making it likely that he is subject to the same inconsistencies as the others. Hence, we recommend treatment here as we do in Case 4.2DS.

In Cases 4.1S and 4.4I, our data contain no inconsistencies. Yet, following up on our criticism of the ostrich stance, many studies have demonstrated that discrepancies as found in the other two cases do occur. General discrepancies of such a fundamental nature, called preference reversals, were first demonstrated by Lichtenstein and Slovic (1971), and have been extensively confirmed since (reviewed by Slovic 1995). Comparisons of PE and CE measurements invariably found higher PE values, as in our example.[6] Such discrepancies show that at least one of the measurements concerned contains biases, and the million dollar question then is what these biases are and how to correct for them.

Behavioral economics can serve to provide diagnostic tools to identify the biases underlying the above discrepancy. Bleichrodt (2002) showed that most biases occur for the PE measurement, which generates serious biases upward.[7] This further supports our recommendation of deviating from the status quo decision suggested by the sqq = 0.91 observation, and of choosing treatment instead. Hence we also rec-

---

[6] See for instance Baron (1994, §17.1.4), Bleichrodt et al. (2001, p. 1505), Bleichrodt et al. (2007), Delquié (1993), Hershey and Schoemaker (1985), Johnson and Schkade (1989), Morrison (2000), Seidl and Traub (1999), Slovic et al. (1990, study 5), Stiggelbout et al. (1994, p. 87).

[7] Other studies confirming the poor performance of PE measurements include Chilton and Spencer (2001), Doctor et al. (2010), Dolan (2000, p. 1745), Hershey and Baron (1987, p. 208), Jansen et al. (1998), Officer and Halter (1968, p. 260), Revicki and Kaplan (1993), Richardson (1994), Stalmeier (2002), Stiggelbout and de Haes (2001, p. 224), and Torrance (1987).

ommend treatment in Cases 4.1S and 4.4I, even though they do not directly reveal inconsistencies.

We expect that most readers did not know about the literature just cited when they started reading this paper, and hence favored the status quo in Cases 4.1S and 4.4I. We hope that now, after learning about this literature, they side with us and favor treatment in all cases. We similarly hope that readers who knew about this literature beforehand, favored treatment in all cases from the beginning. At any rate, from now on, we assume the cited literature on biases to be known in the discussions that follow.

Weak paternalism [row 4] This stance may be taken by readers who find an argument based only on related literature too weak to overrule the stated preference of the affected client. Then the status quo may be chosen in Case 4.4I. On the other hand, paternalism is still considered necessary when no preference of the affected individual is clearly stated (Case 4.1S) or when there is clear evidence of inconsistencies in similar cases (Cases 4.2DS and 4.3SI).

Individual autonomy [rows 1 or 3] Other readers will take the dividing line between cases where the individual preference is clearly stated (Cases 4.3SI and 4.4I) and those where it is not (Cases 4.1S and 4.2DS). They will not want to overrule the preference of an individual client based on inconsistencies among other clients (Molenaar et al. 2004, p. 2129). It leads to the same decisions as the non-statistical stance but for more valid ethical reasons. Redelmeier and Tversky (1990, p. 1162) argued for the opposite: "Physicians and policy makers may wish to examine problems from both perspectives to ensure that treatment decisions are appropriate whether applied to one or to many patients."

Asymmetric paternalism [rows 1 or 3] Going against the stated preference of the client means harming him if he is fully rational (in the sense of satisfying the classical normative decision-theory principles). Asymmetric paternalism wants to avoid harming rational clients, and recommends against treatment in Cases 4.3SI and 4.4I. We acknowledge the harm of choosing treatment for a rational client in the two cases, but think that it is far more likely that the client violates classical decision principles and, in this sense, is irrational. Asymmetric paternalism leads to the same conclusions as the individual autonomy stance whenever the level of rationality of our clients is unknown.

Strict sampling [row 2] This stance concerns readers who find arguments based on the literature too weak. They are willing to overrule preferences that themselves contain inconsistencies as in Case 4.2DS, and they are even willing to do so based on similar observations as in Case 4.3SI. However, they find the inconsistencies in the literature, based on more remote samples and stimuli, too unconvincing.

Libertarian paternalism This stance may suggest, if possible, using only the CE measurements yielding $^{\underline{s}}$sqqs, and avoiding the PE measurements yielding sqqs, in the above cases. However, such an escape from the dilemma is not possible for us, with PE measurements already available. Libertarian paternalism seeks to avoid dilemmas and offers no guidance for cases as those considered here.

Cautious and soft paternalism These terms combine various positions that avoid coercion, and best correspond with the middle rows.

Anti-paternalism [row 1] Row 1 combines the two versions of weak anti-paternalism. This stance holds that preferences are only over-ruled if they are inconsistent within themselves, so that overruling one way or the other in Case 4.2DS cannot be avoided anyhow. It also implies that one then never relies on indirect suggestions of inconsistencies, and hence favors the status quo in the other three cases. In this spirit, Molenaar et al. (2004, p. 2129) wrote: "The use of a decision aid did not influence the kind of treatment selected. This is a desirable outcome as the aim of the decision aid is to assist patients in the decision-making process, and not to prescribe a course of action." Bernheim and Rangel (2009) introduced a revealed preference model where preference, in general, can depend on the context ("ancillaries"), and on no information other than directly revealed choices. This model accepts preferences as non-suspect only in the special case where they are not context-dependent; i.e., they are not part of an inconsistency. This model will formally recommend the status quo in Cases 4.1S and 4.4I, and probably also in Case 4.3SI if it does not consider outside information, leading to strong anti-paternalism.

Stances that avoid the responsibility of choosing treatment, for instance, in Cases 4.1S and 4.4I, could be considered to be forms of paternalism. Similarly, strict adherence to consumer sovereignty can be interpreted as paternalism. Own assessments of non-commitment are then given priority over assessments of the client's best interests. We will, however, adhere to the most common interpretations, where these stances are taken as what they aim to be: non-paternalistic.

5.3 Some related views on the (non-) existence of preference in the literature

We start with views that lead to strong paternalism. Throughout the history of decision theory, empirical studies have found inconsistencies in preferences that signal that biases are effective, and have argued for the desirability to debias (Edwards 1962; Loewenstein and Ubel 2008, §6.1.2). One of the most notorious cases concerns the discrepancy between willingness to pay and willingness to accept (Schmidt and Traub 2009). Many decision analysts recommended measuring utilities and other relevant quantities in several manners, and resolving inconsistencies using cross checks (Keeney and Raiffa 1976, §5.8.3; Kriegler et al. 2009, p. 5046; McCord and Neufville 1985; Weinstein et al. 1996, p. 1257).[8]

In one form of debiasing, one develops measurement tools that avoid biases.[9] Such debiasing would, like libertarian paternalism, suggest using CE measurements rather than the more biased PE measurements, if possible. In our cases, such avoidance is not possible since the biased PE measurements have already been implemented. A

---

[8] A remarkable study is Elstein et al. (1986), who found a case where more than half of the physicians' judgments deviated from the recommendation of a decision analysis.

[9] The numerous references include Baron et al. (2001), Huber et al. (2001), Lefebvre et al. (2011), McFadden (2006), Tversky and Kahneman (1981). Loewenstein and Ubel (2008, pp. 1806–1807) discuss the health domain.

second form of debiasing is still possible in our cases, because it is carried out ex post: If we have estimates of the biases, we can correct the measurements already obtained (Anderson and Hobbs 2002; Bleichrodt et al. 2001; Kahneman 2003, p. 1468; Viscusi 1995 last paragraph).

Avoiding and correcting biases is, of course, common in all empirical sciences. It becomes delicate in decision theory if biases are interpreted as human irrationalities. In this case, no clear objectively correct gold standard is available (Tversky and Kahneman 1981, p. 437, 3rd column), and ethical complications may arise. We expect that authors working on debiasing are close to our paternalistic stance in row 5. Many authors have argued, as we do, that behavioral economics gives tools to improve human decisions (Diamond 2008; Kahneman 2003; McFadden 1999; Oliver 2013), which also suggests that they oppose to anti-paternalism.

The constructive view of preference (Payne et al. 1999) entails that preferences obtained in measurements are constructed on the spot by participants, just so as to answer unfamiliar questions. These constructed preferences have little or even no relation to the underlying true preferences. Some constructivists will conclude that we should dig deeper to find true preferences (McFadden 1999; Slovic 1995; Tversky 1996). In our four cases, where we cannot dig deeper, we would have to accept that there are biases, and correct for them to our best ability, leading to our recommendations of treatment.

We now turn to a number of views of preference that lead to anti-paternalism. We start with constructivists who go in a direction opposite to the one considered above. They argue that true underlying preferences are a meaningless and non-existing concept (Güth 1995, p. 342; Starmer 1996; discussed by Camerer 1995, p. 673). As a result, they will probably not search for the implementation of true preferences, but prefer to minimize intervention in any form, leading to the anti-paternalistic stance.

Closely related to the second constructivist stance just discussed is the view that people may not be able to have any preference at all in some situations, leading to incompleteness of preference.[10] This may, for instance, happen if different values are considered to be incommensurable. Incompleteness has often been defended in descriptive applications. In the prescriptive context considered here, this position is invalid, amounting to the refusal stance. Valid prescriptive variations of this stance will probably lead to recommendations that coincide with anti-paternalism.

For some authors, the volatility of observed preferences has led to a position between the existence and non-existence of preference: they consider preferences to be fundamentally random (Hey 2005; Regenwetter et al. 2011; Wilcox 2008). We take their position as merely descriptive, but expect advocates of random preference not to seek for true preferences and, hence, to be sympathetic to anti-paternalism in our cases. Random preferences are untenable in prescriptive applications:

---

[10] See Baucells and Shapley (2008), Bewley (1986), Danan (2008), Dubra et al. (2004), Gilboa et al. (2008, pp. 179–180), Mandler (2005), and Nau (2006).

No-one would want a doctor, judge, or waiter in a restaurant to decide by a coin toss.[11]

Loewenstein and Ubel (2008) argue that many discussions of paternalistic interventions require speculations about genuine welfare. Examples are the speculations that people save too little for retirement, overeat, or lack discipline to quit smoking. The authors point out that such speculations, whether based on decision utility or experienced utility, are always problematic. In our gedanken-experiments, the choice inconsistency problem arises independently of any such speculation. Our recommendation of treatment in all four cases is based on the following two arguments. First, an underestimation of 0.90 relative to 0.91 is less serious than an overestimation of 0.90 relative to 0.86. Second, extensive literature suggests great overestimations of probabilities in PE questions. Neither of these two arguments entails direct speculations on genuine welfare.

There have been many other relevant debates in the literature on choice inconsistencies and ways to reconcile and improve them. A complete review of this literature is beyond the scope of this paper. We only briefly mention debates in intertemporal choice, where present and future selves can disagree, leading to intertemporal inconsistencies and self-control problems (Strotz 1956). Some have argued that even for addictions, no outsiders should intervene ("rational addiction"; Becker and Murphy 1988), which entails an anti-paternalistic stance. Others have favored interventions of benevolent social planners (Gruber and Köszegi 2001; Heil et al. 2003).

6 Conclusion

To discuss the central question concerning paternalism that plays a role in many practical situations, we used a hypothetical gedanken-experiment to maximally clarify the relevant issues (mainly our Case 4.4I). The gedanken-experiment provides a litmus test for the readers' stance on paternalism. Providing this test was the primary purpose of this paper. It shows that one cannot avoid taking a position (disproving many claims to the contrary in the literature). Our thought experiments involved only minimal rationality conditions, being transitivity and stochastic dominance. Thus, we separated debates about paternalism from debates about rationality of expected utility or other theories.

We further argued for deviating from what at first sight seems to be the true and consistent preference of a client in some situations (but on closer inspection need not be, as we argued). Although it is easy to cast doubt on our stance, as sometimes it can lead to wrong decisions, it is not easy to suggest better stances. Every other stance can also lead to wrong decisions, especially if ignoring a large literature on well-documented biases, or when serving only to avoid responsibility. The latter is impossible in our gedanken-experiment as it often is impossible in practice. For example, medical treatment decisions have to be taken one way or the other, and money spent on one treatment cannot be spent on another.

---

[11] Randomization in game theory only serves to be unpredictable by opponents.

Our analysis draws on the vast literature on biases in behavioral economics, and
further clarifies how behavioral economics can be of use in prescriptive decision-
making.

References

Anderson, R. M., & Hobbs, B. F. (2002). Using a Bayesian approach to quantify scale compatibility bias. *Management Science*, *48*, 1555–1568.

Baron, J. (1994). *Thinking and deciding* (2nd ed.) Cambridge: Cambridge University Press.

Baron, J., Zhijun, W., Brennan, D. J., Weeks, C., & Ubel, P.A. (2001). Analog scale magnitude estimation, and Person trade-off as measures of health utility: Biases and their correction. *Journal of Behavioral Decision Making*, *14*, 17–34.

Baucells, M., & Shapley, L. S. (2008). Multiperson utility. *Games and Economic Behavior*, *62*, 329–347.

Baumann, P. (2005). Three doors two players, and single-case probabilities. *American Philosophical Quarterly*, *42*, 71–79.

Becker, G. S., & Murphy, K. M. (1988). A theory of rational addiction. *Journal of Political Economy*, *96*, 675–700.

Bell, D. E. (1982). Regret in decision making under uncertainty. *Operations Research*, *30*, 961–981.

Berg, N., & Gigerenzer, G. (2010). As-if behavioral economics: Neoclassical economics in disguise? *History of Economic Ideas*, *18*, 133–166.

Bernheim, B. D., & Rangel, A. (2009). Beyond revealed preference: Choice-theoretic foundations for behavioral welfare economics. *Quarterly Journal of Economics*, *124*, 51–104.

Bewley, T. F. (1986). Knightian decision theory part I. Cowles Foundation discussion paper no. 807. *Decisions in Economics and Finance*, *25*, 79–110 (2002, reprinted).

Birnbaum, M. H. (1992). Should contextual effects in human judgment be avoided?, Book review of: Poulton, E. C. (1989). *Bias in quantifying judgments (Contemporary Psychology*, Vol. 37, pp. 21–23). Hillsdale, NJ: Erlbaum.

Bleichrodt, H. (2002). A new explanation for the difference between SG and TTO utilities. *Health Economics*, *11*, 447–456.

Bleichrodt, H., Abellán, J. M., Pinto, J. L., & Mendez-Martinez, I. (2007). Resolving inconsistencies in utility measurement under risk: Tests of generalizations of expected utility. *Management Science*, *53*, 469–482.

Bleichrodt, H., Pinto, J. L., & Wakker, P. P. (2001). Making descriptive use of prospect theory to improve the prescriptive use of expected utility. *Management Science*, *47*, 1498–1514.

Broome, J. R. (1993). Can a Humean be moderate? In R. G. Frey & C. W. Morris (Eds.), *Value, welfare and morality*. Cambridge: Cambridge University Press.

Camerer, C. F. (1995). Individual decision making. In J. H. Kagel & A. E. Roth (Eds.), *Handbook of experimental economics* (pp. 587–703). Princeton, NJ: Princeton University Press.

Camerer, C., Issacharoff, S., Loewenstein, G., O'Donoghue, T., & Rabin, M. (2003). Regulation for conservatives: Behavioral economics and the case for "asymmetric paternalism." *University of Pennsylvania Law Review*, *151*, 1211–1254.

Chilton, S., & Spencer, A. (2001). Empirical evidence of inconsistency in standard gamble choices under direct and indirect elicitation methods. *Swiss Journal of Economics*, *137*, 65–86.

Cohen, A., & Einav, L. (2007). Estimating risk preferences from deductible choice. *American Economic Review*, *97*, 745–788.

Danan, E. (2008). Randomization vs. selection: How to choose in the absence of preference? *Management Science*, *56*, 503–518.

Delquié, P. (1993). Inconsistent trade-offs between attributes: New evidence in preference assessment biases. *Management Science*, *39*, 1382–1395.

Diamond, P. A. (2008). Behavioral economics. *Journal of Public Economics*, *92*, 1858–1862.

Doctor, J. N., Bleichrodt, H., & Lin, J. H. (2010). Health utility bias: A meta-analytic evaluation. *Medical Decision Making*, *30*, 58–67.

Dolan, P. (2000). The measurement of health-related quality of life for use in resource allocation decisions in health care. In A. J. Culyer & J. P. Newhouse (Eds.), *Handbook of health economics* (pp. 1723–1760). Amsterdam: Elsevier.

Drummond, M. F., O'Brien, B. J., Stoddart, G. L., & Torrance, G. W. (1997). *Methods for the economic evaluation of health care programmes* (2nd ed.). Oxford: Oxford University Press.

Dubra, J., Maccheroni, F., & Ok, E. A. (2004). Expected utility without the completeness axiom. *Journal of Economic Theory*, *115*, 118–133.

Edwards, A., & Elwyn, G. (2006). Inside the black box of shared decision making: Distinguishing between the process of involvement and who makes the decision. *Health Expectations*, *9*, 307–320.

Edwards, W. (1962). Subjective probabilities inferred from decisions. *Psychological Review*, *69*, 109–135.

Elstein, A. S., Holzman, G. B., Ratvick, M. M., et al. (1986). Comparison of physicians' decisions regarding oestrogen replacement therapy for menopausal women and decisions derived from a decision analytic model. *American Journal of Medicine*, *80*, 246–258.

Friedman, M. (1953). *Methodology of positive economics*. Chicago: University of Chicago Press.

Fuchs, V. R. (1974). *Who shall live? Health, economics, and social choice*. New York: Basic Books.

Gigerenzer, G. (1991). From tools to theories: A heuristic of discovery in cognitive psychology. *Psychological Review*, *98*, 254–267.

Gilboa, I., Postlewaite, A. W., & Schmeidler, D. (2008). Probability and uncertainty in economic modeling. *Journal of Economic Perspectives*, *22*, 173–188.

Gruber, J., & Köszegi, B. (2001). Is addiction rational'? Theory and evidence. *Quarterly Journal of Economics*, *116*, 1261–1303.

Güth, W. (1995). On ultimatum bargaining experiments: A personal review. *Journal of Economic Behavior and Organization*, *27*, 329–344.

Hammond, K. R. (2006). *Beyond rationality*. New York: Oxford University Press.

Heil, S. H., Tidey, J. W., Holmes, H. W., & Higgins, S. T. (2003). A contingent payment model of smoking cessation: Effects of abstinence and withdrawal. *Nicotine and Tobacco Research*, *5*, 205–213.

Hershey, J., & Baron, J. (1987). Clinical reasoning and cognitive processes. *Medical Decision Making*, *7*, 203–211.

Hershey, J. C., & Schoemaker, P. J. H. (1985). Probability versus certainty equivalence methods in utility measurement: Are they equivalent? *Management Science*, *31*, 1213–1231.

Hey, J. D. (2005). Why we should not be silent about noise. *Experimental Economics*, *8*, 325–345.

Huber, J., Ariely, D., & Fischer, G. (2001). Expressing preference in a principal-agent task: A comparison of choice, rating, and matching. *Organizational Behavior and Human Decision Processes*, *87*, 66–90.

Hume, D. (1740). *A treatise of human nature*. Oxford: Clarendon Press.

Jansen, S. J. T., Stiggelbout, A. M., Wakker, P. P., Vliet Vlieland, Leer, J. W. H., Nooy, M. A., & Kievit, J. (1998). Patient utilities for cancer treatments: A study of the chained procedure for the standard gamble and time tradeoff. *Medical Decision Making*, *18*, 391–399.

Johnson, E. J., & Schkade, D. A. (1989). Bias in utility assessments: Further evidence and explanations. *Management Science*, *35*, 406–424.

Jones, R., Pykett, J., & Whitehead, M. (2011). Governing temptation: Changing behaviour in an age of libertarian paternalism. *Progress in Human Geography*, *35*, 483–501.

Kahneman, D. (2003). Maps of bounded rationality: Psychology for behavioral economics. *American Economic Review*, *93*, 1449–1475.

Keeney, R. L. & Howard, R. (1976). *Decisions with multiple objectives*. Wiley: New York (1993, 2nd ed., Cambridge University Press: Cambridge).

Kriegler, E., Hall, J. W., Held, H., Dawson, R., & Schellnhuber, J. (2009). Imprecise probability assessment of tipping points in the climate system. *Proceedings of the National Academy of Sciences*, *106*, 5041–5046.

Lefebvre, M., Vieider, F. M., & Villeval, M. C. (2011). The ratio bias phenomenon: Fact or artifact? *Theory and Decision*, *71*, 615–641.

Lichtenstein, S., & Slovic, P. (1971). Reversals of preference between bids and choices in gambling decisions. *Journal of Experimental Psychology*, *89*, 46–55.

Loewenstein, G. F., & Ubel, P. A. (2008). Hedonic adaptation and the role of decision and experience utility in public policy. *Journal of Public Economics*, *92*, 1795–1810.

Loomes, G., & Sugden, R. (1982). Regret theory: An alternative theory of rational choice under uncertainty. *Economic Journal*, *92*, 805–824.

Lopes, L. L. (1981). Decision making in the short run. *Journal of Experimental Psychology, Human Learning and Memory*, *7*, 377–385.

Mandler, M. (2005). Incomplete preferences and rational intransitivity of choice. *Games and Economic Behavior*, *50*, 255–277.

McCord, M. R., & de Neufville, R. (1985). Assessment response surface: Investigating utility dependence on probability. *Theory and Decision*, *18*, 263–285.

McFadden, D. L. (1999). Rationality for economists? *Journal of Risk and Uncertainty*, *19*, 73–105.

McFadden, D. L. (2006). Free markets and fettered consumers. *American Economic Review*, *96*, 5–29.

McQuillin, B., & Sugden, R. (2012). Reconciling normative and behavioural eonomics: The problems to be solved. *Social Choice and Welfare*, *38*, 553–567.

Molenaar, S., Oort, F. J., Sprangers, M. A. G., Rutgers, E. J., Mulder, J., & de Haes, H. C. J. M. (2004). Predictors of patients' choices for breast-conserving therapy or mastectomy: A prospective study. *British Journal of Cancer*, *90*, 2123–2130.

Morrison, G. C. (2000). The endowment effect and expected utility. *Scottish Journal of Political Economy*, *47*, 183–197.

Murphy, T. D. (1981). Medical knowledge and statistical methods in early nineteenth century France. *Medical History*, *25*, 301–319.

Nau, R. F. (2006). The shape of incomplete preferences. *Annals of Statistics*, *34*, 2430–2448.

O'Donoghue, T., & Rabin, M. (1999). Procrastination in preparing for retirement. In H. Aaron (Ed.), *Behavioral dimensions of retirement economics*. New York: The Brookings Institution.

O'Donoghue, T., & Rabin, M. (2003). Studying optimal paternalism, illustrated by a model of sin taxes. *American Economic Review, Papers and Proceedings*, *93*, 186–191.

Officer, R. R., & Halter, A. N. (1968). Utility analysis in a practical setting. *American Journal of Agricultural Economics*, *50*, 257–277.

Oliver, A. J. (2013). Should behavioural economic policy be anti-regulatory? *Health Economics*, *22*, 373–375.

Payne, J. W., Bettman, J. R., & Schkade, D. A. (1999). Measuring constructed preferences: Towards a building code. *Journal of Risk and Uncertainty*, *19*, 243–270.

Raiffa, H. (1961). Risk uncertainty and the Savage axioms: Comment. *Quarterly Journal of Economics*, *75*, 690–694.

Redelmeier, D. A., & Tversky, A. (1990). Discrepancy between medical decisions for individual patients and for groups. *New England Journal of Medicine*, *322*, 1162–1164.

Regenwetter, M., Dana, J., & Davis-Stober, C. P. (2011). Transitivity of preferences. *Psychological Review*, *118*, 42–56.

Revicki, D. A., & Kaplan, R. M. (1993). Relationship between psychometric and utility-based approaches to the measurement of health-related quality of life. *Quality Life Research*, *2*, 477–487.

Richardson, J. (1994). Cost utility analysis: What should be measured? *Social Science and Medicine*, *39*, 7–20.

Schelling, T. C. (1968). The life you save may be your own. In S. B. Chase Jr (Ed.), *Problems in public expenditure analysis* (pp. 127–162). Washington DC: Brookings Institution.

Schmidt, U., & Traub, S. (2009). An experimental investigation of the disparity between WTA and WTP for lotteries. *Theory and Decision*, *66*, 229–262.

Seidl, C., & Traub, S. (1999). Biases in the assessment of von Neumann–Morgenstern utility functions. *Journal of Economics Supplement*, *8*, 203–239.

Shackle, G. L. S. (1949). A non-additive measure of uncertainty. *Review of Economic Studies*, *17*, 70–74.

Slovic, P. (1995). The construction of preference. *American Psychologist*, *50*, 364–371.

Slovic, P., Griffin, D., & Tversky, A. (1990). Compatibility effects in judgment and choice. In R. M. Hogarth (Ed.), *Insight in decision making. A tribute to Hillel J. Einhorn* (pp. 5–27). Chicago: The University of Chicago Press.

Smith, V. L. (2008). *Rationality in economics: Constructivist and ecological forms*. Cambridge: Cambridge University Press.

Stalmeier, P. F. M. (2002). Discrepancies between chained and classic utilities induced by anchoring with occasional adjustments. *Medical Decision Making*, *22*, 53–64.

Starmer, C. (1996). Explaining risky choices without assuming preferences. *Social Choice and Welfare*, *13*, 201–213.

Starmer, C. (2000). Developments in non-expected utility theory: The hunt for a descriptive theory of choice under risk. *Journal of Economic Literature*, *38*, 332–382.

Steiner, J. F. (1999). Talking about treatment: The language of populations and the language of individuals. *Annals of Internal Medicine*, *130*, 618–622.

Stiggelbout, A. M., & de Haes, H. C. J. M. (2001). Patient preference for cancer therapy: An overview of measurement approaches. *Journal of Clinical Oncology*, *19*, 220–230.

Stiggelbout, A. M., Kiebert, G. M., Kievit, J., Leer, J.-W. H., Stoter, G., & de Haes, H. C. J. M. (1994). Utility assesment in cancer patients: Adjustment of time tradeoff scores for the utility of life years and comparison with standard gamble scores. *Medical Decision Making*, *14*, 82–90.

Strotz, R. H. (1956). Myopia and inconsistency in dynamic utility maximization. *Review of Economic Studies*, *23*(3), 165–180.

Sugden, R. (1998). Measuring opportunity: Toward a contractarian measure of individual interest. *Social Philosophy & Policy*, *15*, 34–60.

Sugden, R. (2004). The opportunity criterion: Consumer sovereignty without the assumption of coherent preferences. *American Economic Review*, *94*, 1014–1033.

Thaler, R. H., & Sunstein, C. R. (2003). Libertarian paternalism. *American Economic Review, Papers and Proceedings*, *93*, 175–179.

Thaler, R. H., & Sunstein, C. R. (2008). *Nudge: Improving decisions about health, wealth, and happiness*. New Haven: Yale University Press.

Torrance, G. W. (1987). Utility approach to measuring health-related quality of life. *Journal of Chronic Diseases*, *40*, 593–600.

Torrance, G. W., & Feeney, D. H. (1989). Utilities and quality-adjusted life years. *International Journal of Technology Assessment in Health Care*, *5*, 559–575.

Tversky, A. (1996). Rational theory and constructive choice. In K. J. Arrow, E. Colombatto, M. Perlman, & C. Schmidt (Eds.), *The rational foundations of economic behavior: Proceedings of the IEA conference held in Turin, Italy* (pp. 185–197). New York: St. Martins Press.

Tversky, A., & Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science*, *211*, 453–458.

Viscusi, W. K. (1995). Government action, biases in risk perception, and insurance decisions. *Geneva Papers in Risk and Insurance Theory*, *20*, 93–110.

Weinstein, M. C., Siegel, J. E., Gold, M. R., Kamlet, M. S., & Russell, L. B. (1996). Recommendations of the Panel of Cost-Effectiveness in Health and Medicine. *JAMA*, *276*, 1253–1258. for the Panel on Cost-Effectiveness in Health and Medicine.

Wilcox, N. T. (2008). Stochastic models for binary discrete choice under risk: A critical primer and econometric comparison. In J. C. Cox & G. W. Harrison (Eds.), *Risk aversion in experiments; research in experimental economics 12* (pp. 197–292). Bingley: Emerald.

Zamir, E. (1998). The efficiency of paternalism. *Virginia Law Review*, *84*, 229–286.