RUHR

ECONOMIC PAPERS

Miriam Krieger
Thomas Mayrhofer

Patient Preferences and Treatment Thresholds under Diagnostic Risk

An Economic Laboratory Experiment

#321

Imprint

Ruhr Economic Papers

Published by

Ruhr-Universität Bochum (RUB), Department of Economics
Universitätsstr. 150, 44801 Bochum, Germany

Technische Universität Dortmund, Department of Economic and Social Sciences
Vogelpothsweg 87, 44227 Dortmund, Germany

Universität Duisburg-Essen, Department of Economics
Universitätsstr. 12, 45117 Essen, Germany

Rheinisch-Westfälisches Institut für Wirtschaftsforschung (RWI)
Hohenzollernstr. 1-3, 45128 Essen, Germany

Editors

Prof. Dr. Thomas K. Bauer
RUB, Department of Economics, Empirical Economics
Phone: +49 (0) 234/3 22 83 41, e-mail: thomas.bauer@rub.de

Prof. Dr. Wolfgang Leininger
Technische Universität Dortmund, Department of Economic and Social Sciences
Economics – Microeconomics
Phone: +49 (0) 231/7 55-3297, email: W.Leininger@wiso.uni-dortmund.de

Prof. Dr. Volker Clausen
University of Duisburg-Essen, Department of Economics
International Economics
Phone: +49 (0) 201/1 83-3655, e-mail: vclausen@vwl.uni-due.de

Prof. Dr. Christoph M. Schmidt
RWI, Phone: +49 (0) 201/81 49-227, e-mail: christoph.schmidt@rwi-essen.de

Editorial Office

Joachim Schmidt
RWI, Phone: +49 (0) 201/81 49-292, e-mail: joachim.schmidt@rwi-essen.de

Ruhr Economic Papers #321

Responsible Editor: Volker Clausen

ISSN 1864-4872 (online) – ISBN 978-3-86788-370-2

Ruhr Economic Papers #321

Miriam Krieger and Thomas Mayrhofer

Patient Preferences and Treatment Thresholds under Diagnostic Risk

An Economic Laboratory Experiment

UNIVERSITÄT DUISBURG ESSEN

Bibliografische Informationen der Deutschen Nationalbibliothek

Die Deutsche Bibliothek verzeichnet diese Publikation in der deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über: *http://dnb.d-nb.de* abrufbar.

http://dx.doi.org/10.4419/86788370
ISSN 1864-4872 (online)
ISBN 978-3-86788-370-2

Miriam Krieger and Thomas Mayrhofer[1]

Patient Preferences and Treatment Thresholds under Diagnostic Risk – An Economic Laboratory Experiment

Abstract

We study risk aversion and prudence in medical treatment decisions. In a laboratory experiment, we investigate the frequency and intensity of second- and third-order risk preferences, as well as the effect of the medical decision context. Risk preferences are assessed through treatment thresholds (the indifference point between not treating and treating). Under diagnostic risk, medical decision theory predicts lower thresholds for risk-averse than risk-neutral decision makers. Given a comorbidity risk, prudent individuals have an even lower threshold. Our results demonstrate risk-averse and prudent behavior in medical decisions, which reduce the (average) treatment threshold by 41% relative to risk neutrality (from 50.0% to 29.3% prevalence rate). Risk aversion accounts for 3/4 of this effect, prudence for 1/4. The medical decision framing does not affect risk aversion, but is associated with more and stronger prudent behavior. These findings have consequences for treatment thresholds, diagnostics, and QALYs, and thus for clinical guidelines.

JEL Classification: I10, C91, D81

Keywords: Medical decision making; treatment thresholds; risk aversion; prudence; laboratory experiment

March 2012

1 *Both University of Duisburg-Essen. – We are grateful to Stefan Felder, Timo Heinrich, and Hendrik Schmitz for their valuable comments that helped to improve the paper. We also thank Daniel Wiesen for his comments on the experimental design and Christoph Helbach for his help during the experimental sessions. Financial support from the German Research Foundation (DFG) through FOR 655 is gratefully acknowledged.* – *All correspondence to Thomas Mayrhofer, Faculty of Economics and Business Administration, University of Duisburg-Essen, Schuetzenbahn 70, 45117 Essen, Germany, E-Mail: thomas.mayrhofer@wiwinf.uni-due.de..*

1 Introduction

Following a general development in economics towards taking a more comprehensive view of risk preferences, the importance of higher-order risk preferences for health-related decisions have been the subject of recent theoretical work. We investigate the role of second- and third-order risk aversion in decisions over medical treatment in a laboratory experiment, thus contributing to establishing the empirical relevance of such preferences. This is particularly salient in the medical context, where decisions on treatment are subject to diagnostic uncertainty and are almost always characterized by background risks such as comorbidity or severity risks. If second- and third-order risk aversion prove to be relevant to such decisions – as we aim to show in this paper – then they should be taken into account in clinical guidelines and in decision technologies that are based on preference data. Such preferences would, for instance, affect test- and test-treatment thresholds and cutoff points for diagnostic tests (Felder and Mayrhofer 2011b) as well as the health outcome measure of quality adjusted life years (QALYs), which is widely used in the appraisal of health technologies, but often obtained by indirect methods that do not incorporate patient risk preferences (Richardson and Manca 2004).

We focus on the decision whether or not to undertake medical treatment, which is characterized by diagnostic risk.[1] Presented with a patient, a physician might assume a specific disease but can only estimate its probability (e.g. the prevalence rate of the disease). As a result, in deciding whether to treat the patient for this disease, the physician faces a trade-off between the utility a sick person gains from treatment and the utility a healthy person loses due to (unnecessary) treatment. This trade-off implies a prevalence threshold at which the utility gain and loss from treatment are equal. If the estimated probability of disease for a patient lies below this (treatment) threshold, it is beneficial not to treat him, while at probabilities above the threshold it is preferable to treat.[2]

The concept of the treatment threshold was first introduced by Pauker and Kassirer (1975). While their analysis is based in expected utility theory (EUT; von Neumann and Morgenstern 1947), they only use mortality and survival rates – and thus expected values – in their clinical examples. Not until Eeckhoudt et al. (1985) was EUT fully brought to bear on this problem. They show that when facing diagnostic risk the treatment threshold decreases for risk-averse individuals, who should therefore be treated earlier (i.e. at lower prevalence rates) than risk-neutral individuals.[3] Under diagnostic risk, the

[1] We do not differentiate between the physician and the patient as decision makers, assuming for our purposes that the physician is a perfect agent to the patient's principal.

[2] In practice, there is often a diagnostic test that can determine whether the specific disease is present. The decision is then for or against applying the test, which is also associated with harm to the patient. We discuss the more basic situation without the diagnostic test, but the same reasoning applies to the test decision.

[3] Eeckhoudt et al. (1985) also discuss how the optimal treatment intensity differs for risk-averse and risk-neutral decision-makers when outcome parameters vary. We assume here that the physician can only treat or not treat the patient, but not vary the intensity of the treatment.

decision to treat is a risk-reducing strategy (thus resembling an insurance against the most extreme health states), and risk-averse patients therefore prefer to be treated at lower prevalence rates than risk-neutral patients.

Eeckhoudt (2002) extended the model of Pauker and Kassirer to show that introducing an exogenous background risk, such as a comorbidity or severity risk, will lead to a further decrease in the treatment threshold (only) for decision makers who are both risk-averse and prudent. While a risk-averse individual is characterized under EUT as having a utility function with a negative second derivative, a prudent individual exhibits a positive third derivative (Kimball 1990). The comorbidity risk is present only in the sick state and is independent of the choice for or against treatment. A decision maker who is prudent will try to counteract this exogenous risk by reducing his endogenous risk, the diagnostic risk, and selecting treatment at lower prevalence rates (Eeckhoudt 2002; Felder and Mayrhofer 2011a, 2011b).

We investigate the empirical relevance of these theoretical results in a laboratory experiment, pursuing three research questions: Firstly, we study whether decisions over medical treatment in a laboratory experiment are subject to risk aversion and especially to prudence. Secondly, we establish the magnitude of the effect these preferences have on the treatment threshold. Finally, we investigate the role of second- and third-order risk preferences specifically in the medical context by comparing the effect of medical and neutral framing of the decision situation in the experiment.

There is a sizeable body of literature on the relevance of (second-order) risk aversion in medical decisions, going back to e.g. McNeal et al. (1978), Weinstein and Stason (1982), and Gafni and Torrance (1984). More recently, a number of studies have been published on the impact of higher-order risk preferences in the medical context. Eeckhoudt and Gollier (2005), for example, investigate the impact of prudence on optimal prevention levels and show that, contrary to intuition, prudent individuals exert less effort in prevention than risk-neutral individuals. Loosening restrictive assumptions on the distribution of risks, Courbage and Rey (2006) come to the similar conclusion that individuals will pursue more prevention, the more they fear sickness and the less prudent they are. Regarding the theory of medical decision making, Eeckhoudt (2002) shows that, in a situation of diagnostic risk, prudent individuals decrease their treatment threshold in the face of an exogenous comorbidity risk.[4] Felder and Mayrhofer (2011a, 2011b) extend this model and demonstrate that it also applies to test and test-treatment thresholds when a diagnostic test is introduced. Furthermore, they show that optimal cutoff points for tests should be lowered when the test properties (sensitivity and specificity) are endogenous and the decision maker is prudent.

[4] Bleichrodt et al. (2003) analyze the effect of comorbidity risk on treatment decisions in the context of *therapeutic risk* (uncertainty over the outcome of treatment). They assume that treatment only affects the quality aspect while comorbidity only impacts on duration. Since the prudence premium can be interpreted as the decision maker's risk premium for a longer life, they find comorbidities to have no effect on treatment decisions.

Prudence has also been investigated empirically by means of the experimental economic method.[5] Outside the medical context, the earliest study is by Tarazona-Gomez (2004), who compares the certainty equivalents elicited for different lotteries and finds weak evidence for prudence. Deck and Schlesinger (2010), Ebert and Wiesen (2011a, 2011b), and Noussair et al. (2011) all use the prudence lotteries proposed by Eeckhoudt and Schlesinger (2006) and confirm that individuals exhibit prudence as well as risk aversion. Furthermore, Ebert and Wiesen (2011b) report that the risk premium demanded in a situation involving third-order risk (prudence) is in fact higher than what is required in a situation involving second-order risk (risk aversion).

To our knowledge, there is only one empirical study that examines prudence in health decisions. Shafrin and Wiesen (2009) research the relationship between prudence and prevention preferences in an online survey. While 53.1% (15.0%) of their respondents are categorized as prudent (imprudent), they find only suggestive evidence for the expected negative impact of prudence on preventive efforts in a regression. Shafrin and Wiesen study two factors that are relevant to decisions in health care, but the questions used are not cast as medical decisions but as neutral lotteries over hypothetical monetary outcomes. The influence of framing such decisions medically – for which we provide evidence below – suggests that extrapolating from neutrally framed lotteries to medical decision behavior in practice might understate the influence of prudence preferences.

Our contributions to this literature are both methodological and substantive: We conduct the first incentivized experimental study of prudence preferences directly in a medical context. Our results provide important information that can contribute to improving cost-effectiveness research, health technology assessment, and clinical guidelines for diagnostics and treatment.

2 Theoretical Background

The theoretical background to our experiment is provided by Pauker and Kassirer (1975) and Eeckhoudt (2002, chapters 2-3). We briefly outline their results below.[6]

2.1 Risk Preferences and the Treatment Threshold

Assume that there are only two states of nature – healthy (h) and sick (s) – and that a decision maker can choose between treatment (+) and no treatment (–). The decision maker knows the outcomes of

[5] Other (non-experimental) empirical work deals with higher-order risk preferences for saving and consumption (the "precautionary savings" motive), e.g. Dynan (1993), Carrol (1994), and Carrol and Kimball (2008).
[6] See original publications for proofs, and see Felder and Mayrhofer (2011a) for a more exhaustive treatment.

treating and not treating, but not whether the patient is in fact healthy or sick. He only knows the patient's a priori probability of illness (p). Treatment is assumed to benefit the sick but harm the healthy. Overall, there are four possible health states H, ranked $H_s^- < H_s^+ < H_h^+ < H_h^-$. Figure 1 represents the treatment decision.

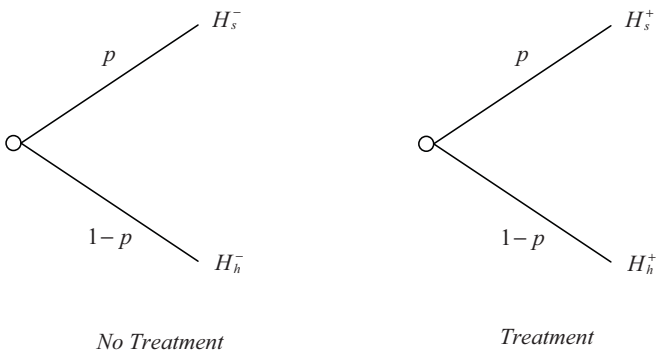


Figure 1: Simple Treatment Decision

Without further information (e.g. from a diagnostic test), the best decision can be derived from the treatment threshold, which is the probability of illness at which the decision maker is indifferent between treating and not treating. Applying EUT and solving for p leads to

$$
\begin{aligned}
EU^- = EU^+ \quad &\Leftrightarrow \quad pU\left(H_s^-\right)+(1-p)U\left(H_h^-\right) = pU\left(H_s^+\right)+(1-p)U\left(H_h^+\right) \\
&\Leftrightarrow \quad \tilde{p} = \frac{U\left(H_h^-\right)-U\left(H_h^+\right)}{\left[U\left(H_s^+\right)-U\left(H_s^-\right)\right]+\left[U\left(H_h^-\right)-U\left(H_h^+\right)\right]}
\end{aligned}
\qquad (1)
$$

where $\tilde{p}$ is the treatment threshold. Equation (1) depends on the utility gain from treatment for a sick patient, $U\left(H_s^+\right)-U\left(H_s^-\right)$, and the utility loss from treatment for a healthy patient, $U\left(H_h^-\right)-U\left(H_h^+\right)$. The treatment threshold thus indicates the trade-off between helping the sick and harming the healthy. If a patient's probability of illness lies below the threshold ($p < \tilde{p}$), then the best decision is not to treat; if the probability of illness lies above the threshold ($p > \tilde{p}$), then treatment is best.

For both risk-neutral and risk-averse decision makers the utility of health can be assumed to be increasing in health ($U'(H) > 0$). But while marginal utility is constant for a risk-neutral individual ($U''(H) = 0$), it is decreasing for a risk-averse individual ($U''(H) < 0$). Since utility is independent of

positive linear transformations in EUT, we can set $U_N\left(H_s^-\right)=U_A\left(H_s^-\right)$ and $U_N\left(H_h^-\right)=U_A\left(H_h^-\right)$, where *N* indicates risk-neutral and *A* risk-averse decision makers. Under diagnostic risk, treatment is a risk-reducing strategy since it leads to outcomes with a smaller spread: $\left(H_h^- - H_s^-\right)>\left(H_h^+ - H_s^+\right)$. Due to his concave utility function, a risk-averse decision maker will value the "inner" health states more than a risk-neutral one. This leads to $\left[U_A\left(H_s^+\right)-U_A\left(H_s^-\right)\right]>\left[U_N\left(H_s^+\right)-U_N\left(H_s^-\right)\right]$ and $\left[U_A\left(H_h^-\right)-U_A\left(H_h^+\right)\right]<\left[U_N\left(H_h^-\right)-U_N\left(H_h^+\right)\right]$. Rewriting and inserting into (1) leads to

$$(2)\qquad \tilde{p}_A=\frac{1}{1+\dfrac{U_A\left(H_s^+\right)-U_A\left(H_s^-\right)}{U_A\left(H_h^-\right)-U_A\left(H_h^+\right)}}<\frac{1}{1+\dfrac{U_N\left(H_s^+\right)-U_N\left(H_s^-\right)}{U_N\left(H_h^-\right)-U_N\left(H_h^+\right)}}=\tilde{p}_N,$$

showing that risk-averse individuals will treat at a lower probability of illness than risk-neutral individuals: $\tilde{p}_A<\tilde{p}_N$.

2.2 Risk Preferences, the Treatment Threshold, and Background Risk

We now assume – more realistically – that a sick patient may have additional health problems. We extend the model by a background risk $\tilde{\varepsilon}$ in the sick state that represents this comorbidity risk, or severity risk. Contrary to the diagnostic risk which can be reduced through treatment (i.e. is endogenous), $\tilde{\varepsilon}$ is independent of the treatment decision (i.e. exogenous). We assume that $\tilde{\varepsilon}$ is a binary random variable distributed with probability 0.5 and an expected value of zero; see Figure 2.

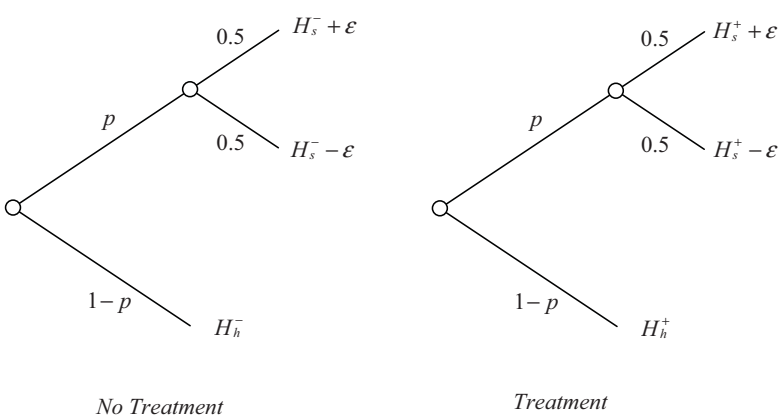


Figure 2: Treatment Decision with Background Risk

Introducing the background risk changes the expected utilities of the two actions such that

$$(3) \quad \tilde{\tilde{p}} = \frac{U\left(H_h^-\right) - U\left(H_h^+\right)}{\left[EU\left(H_s^+ + \tilde{\varepsilon}\right) - EU\left(H_s^- + \tilde{\varepsilon}\right)\right] + \left[U\left(H_h^-\right) - U\left(H_h^+\right)\right]},$$

where $\tilde{\tilde{p}}$ is the treatment threshold given a background risk. Equation (3) differs from (1) in the utility gained from treatment by the sick. To compare $\tilde{p}$ and $\tilde{\tilde{p}}$ it is therefore sufficient to compare $\left[U\left(H_s^+\right) - U\left(H_s^-\right)\right]$ and $\left[EU\left(H_s^+ + \tilde{\varepsilon}\right) - EU\left(H_s^- + \tilde{\varepsilon}\right)\right]$. For a risk-averse decision maker whose utility function has a positive third derivative, $U'''(H) > 0$, it can be shown[7] that $\left[EU\left(H_s^+ + \tilde{\varepsilon}\right) - EU\left(H_s^- + \tilde{\varepsilon}\right)\right] > \left[U\left(H_s^+\right) - U\left(H_s^-\right)\right]$, which leads to

$$(4) \quad \tilde{\tilde{p}} < \tilde{p}_A .$$

Introducing a background risk thus leads a risk-averse and prudent[8] decision maker to lower the treatment threshold even further.

3 Experimental Test of Risk Preferences

Our experiment consists of two tasks, one designed to elicit (second-order) risk preferences and one to test for prudence preferences.[9] In each task subjects face several decisions between two prospects, one riskier and one safer. Across these situations, the probabilities in the prospects change, while the outcomes remain constant in each task. The probability at which an individual is indifferent between the two options indicates his corresponding risk preference.

3.1 Risk preferences in the context of medical treatment decisions

These decisions are embedded in a medical context (see Appendix A for instructions). Subjects are told that they suffer from symptoms which are diagnosed by a doctor as being caused by a specific disease L (or M in task 2) with probability *p*. Meanwhile, the subject has a job opportunity for the next

[7] See Eeckhoudt (2002), p. 32-34.
[8] Second-order risk aversion is defined by the second derivative of the utility function, as prudence is defined by the third derivative. An individual can thus be prudent and risk-averse ($U''' > 0$ and $U'' < 0$) or prudent and risk-loving ($U''' > 0$ and $U'' > 0$), though not prudent and risk-neutral (if $U'' = 0$, then $U''' = 0$).
[9] The experiment also included a third task with a different focus, which is discussed in a separate paper.

10 days which pays on a daily basis, but only for days on which he works (not for days missed due to illness). Subjects can earn 100 Taler (the experimental currency, 35 Taler = 1 Euro) per day, up to 1000 Taler in total. There is a treatment T for the disease. T is time-consuming and thus harmful if the subject is healthy, but will lead to speedy recovery if the subject is ill. (We use "healthy" in reference to disease L, meaning that the subject's symptoms are caused by something else; accordingly "ill" means suffering specifically from L.) The subject must decide whether or not to undergo treatment. The options "treatment" and "no treatment" each resemble lotteries, given that he is uncertain about his health state. In task 2 of the experiment, a zero-mean background risk is added to the lotteries. Subjects are thus told that there is a genetic variant G which results in L progressing much more severely than usual. The probability of having G is always 50%. In both cases, "treatment" and "no treatment", recovery is slower if the subject has G than if he does not.

In order to control for the effect of the medical context we also administer a neutrally framed version of the experiment to some subjects (no subjects participated in both versions of the experiment). The prospects here are simply called lottery A and lottery B and the situations described in neutral language; otherwise the medically and neutrally framed versions are identical. For simplicity's sake we refer mainly to the medical framing in this paper.

Using the ability to earn income as a proxy for health and the loss of income due to absence from work as a proxy for sickness in our experiment is not unproblematic. Incentivizing decisions by non-hypothetical outcomes is one of the main tenets of the experimental method, but this poses a challenge when studying issues involving health: Conducting an experiment in which subjects experience outcomes in terms of different health states is ethically questionable at the least. Experimental studies therefore require a surrogate for health outcomes. Kroll et al. (2011) take an interesting approach, using thermal pain from holding one's hand into water of different low temperatures for various lengths of time to make the outcomes of a Holt and Laury (2002) type procedure non-hypothetical. Based on this cold pressor test, Kroll et al. estimate risk preferences over pain (a proxy for quality of life) and time as relevant to QALY models. While this is an innovative approach in experimental economics[10], it does not transfer easily to other situations in which risk preferences impact on decisions in health settings.

We follow a different path by framing a financially incentivized decision in a medical treatment context. This framing method has been shown to capture non-monetary aspects of decisions in the medical domain in other health economics experiments. Hennig-Schmidt et al. (2011) show that treatment decisions made by subjects in the role of physicians are not only determined by the incentives set by the remuneration system, but also substantially by the benefit that accrues to their patients. In this experiment, patients are not physically represented by other subjects, but their payoffs

[10] Pain has been used before as a measure of wellbeing in psychology experiments (e.g. Kahneman et al. 1993).

are donated entirely to a (real-life) charitable organization that cares for patients with eye disease. Ahlert et al. (2012) also show that framing an allocation task as a medical treatment decision strengthens the influence of professional norms on choices made by prospective physicians and economists.

3.2 Eliciting Second-Order and Third-Order Risk Preferences

The lottery choice tasks in our experiment are based on the procedure of Holt and Laury (2002). In each task, subjects first face 10 consecutive choices in which the probability of the sick state rises in 10% increments from 10% to 100%. Subjects are expected to choose the riskier prospect first and the safer prospect last, switching once in between. Subjects then face 4 further choices in which the probability of illness rises in 2% increments, refining the range between their last risky choice and their first safe choice. This narrows the elicited range within which a subject is indifferent between the two prospects to a 2%-point interval (which is the closest to an exact indifference point we find feasible in practice). After each set of 10 or 4 questions subjects are shown an overview of their decisions and given one chance to keep or retake them. Considering all 14 decisions, the probability at which a subject switches from "no treatment" to "treatment" represents his treatment threshold, or risk preference. In order to facilitate proper assessment of the probabilities, they are visualized using colored figures (in the medical framing) or urns with colored balls (in the neutral framing).

In the first task subjects face simple binary lottery pairs representing the decision between "no treatment" and "treatment" (see Figure 3). There is uncertainty regarding their health state. If they choose "treatment" and are healthy, they lose time in treatment but can still earn 600 Taler. If they are ill, they recover after some days and earn 500 Taler. If they choose "no treatment" and turn out to be healthy they can work the entire time, earning 1000 Taler. If they are ill, they recover slowly and can only work to earn 100 Taler. The treatment threshold in task 1 reveals second-order risk preferences. Switching at 50% or 52% indicates risk-neutral behavior[11], while a switch point below (above) 50% (52%) indicates risk-averse (risk-loving) behavior.

[11] Since a risk-neutral individual is indifferent between the alternatives at a prevalence of exactly 50%, he may not start treatment until he reaches the next probability of 52%.

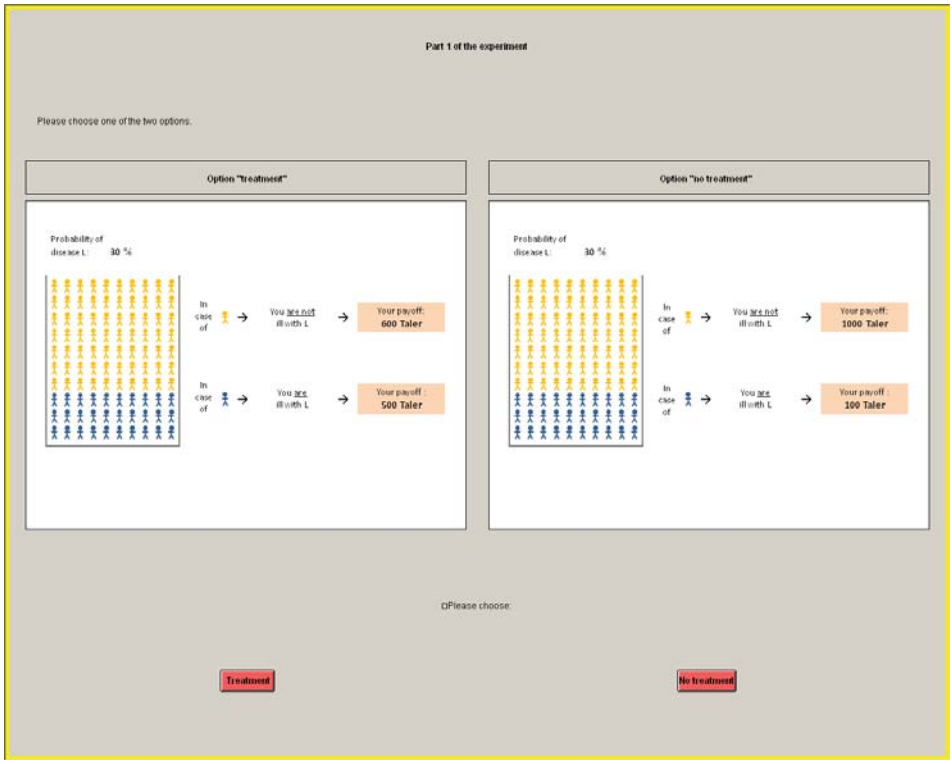


Figure 3: Screenshot of the treatment decision (medical framing) in task 1 at $p = 30\%$

In the second task a zero-mean background risk is added to the sick state of the same prospect (see Figure 4). This construct tests for prudence as defined by Eeckhoudt (2002): an aversion to a zero-mean exogenous risk added to the low-outcome state of a lottery. If a subject decides on "treatment" and is in fact healthy, he loses time in treatment but still earns 600 Taler. If he turns out to be ill and have the gene variation, recovery is slowed and he earns 400 Taler; without the gene variation he recovers sooner and earns 600 Taler. If the subject chooses "no treatment" and is healthy, he can work full time and earn 1000 Taler. If he is ill and has the gene variant, he becomes severely ill and does not recover in time to earn any money at all. If he is ill but does not have the genetic variant, he recovers in time to earn 200 Taler on the job.

Comparing the treatment thresholds in tasks 1 and 2 reveals a subject's prudence preferences: An individual that is neutral regarding prudence will not react to the introduction of the background risk, switching from "no treatment" to "treatment" at the same probability in both tasks. A prudent (imprudent) individual will have a lower (higher) switch point in task 2 than in task 1.

Figure 4: Screenshot of the treatment decision (medical framing) in task 2 at p = 30%

One limitation of this type of lottery choice task is that it operates within EUT, relying on the assumption that probabilities are objective items. Any subjective weighting of probabilities, as posited for instance by prospect theory (Kahneman and Tversky 1979), might confound the inferences drawn from this procedure. Alternative methods for experimental risk preference elicitation have been proposed (e.g. Wakker and Deneffe 1996; Ástebro et al. 2010; Abdellaoui et al. 2011; Ebert and Wiesen 2011a, 2011b), many of which vary the outcomes across decisions while keeping probabilities constant at 50%, a probability decision makers are assumed to treat objectively even outside of EUT. However, such procedures are not suitable for our research purpose, as we specifically aim to study treatment thresholds. (It is also difficult to imagine an experimental design in which outcomes that are defined over health states vary. If the potential subjective assessment of probabilities is problematic, the subjective judgment of health states should be even more so.) Moreover, as the theoretical basis for our experiment is firmly rooted in EUT, our empirical approach should not pose an additional

restriction. Interestingly, the potential practical application of our research may not even be affected by this issue: Research by John List (e.g. 2003, 2004) finds that EUT is better than alternative models at explaining the behavior of experienced decision makers – in our case physicians, trained professionals, making routine diagnostic and treatment decisions.

3.3 Experimental Procedure

Our experiment was conducted at the Essen Laboratory for Experimental Economics with help of the software z-tree (Fischbacher 2007). 152 subjects participated in eight 60-minute sessions between October and December 2011. 93 made their choices in the medical framing, 59 in the neutral setting. Subjects were provided with written instructions at the beginning of the session. The lottery choices were followed by a questionnaire on subject characteristics.

The experiment was incentivized by the random payment technique, where one situation is drawn at random and the lottery chosen by each subject in this situation executed to establish their payoffs. This procedure is widely used in economic experiments to avoid averaging and income effects in repeated decisions.[12] Subjects' payoffs were not determined until after the conclusion of the questionnaire. Average earnings from the experiment were € 18.40 (min.: € 3.00; max.: € 31.60), which included a show up-fee of € 3.00.

4 Results

4.1 Risk Preferences

Our analysis takes into account the choices made by 133 subjects in the experiment. We exclude 3 subjects who chose the dominated option "no treatment" at a prevalence rate of $p = 100\%$, assuming that they did not properly understand the task or the incentive mechanism of the experiment. We also exclude a further 16 subjects who switched back and forth between the options several times, which makes the comparison of switch points between the two tasks very difficult. 12 subjects (7.8%)[13] switched multiple times in the first task and 11 (7.2%) did in the second.

[12] While concerns have been raised that random payment dilutes the power of the monetary incentive, research addressing this question reports no adverse effect of random payment for non-complex choice tasks (e.g. Baltussen et al. 2010; Cubitt et al. 1998; Starmer & Sugden 1991). Laury (2006) specifically analyzes this problem and finds no difference in behavior between treatments where either all ten decisions are paid for or only one of them (although increasing the scale of payments overall does affect behavior).

[13] Holt and Laury (2002) report 6.6% and 13.3% multiple switchers in their treatments with comparable stakes.

Considering all 133 subjects (in both the medical and the neutral framing), we find evidence for risk-averse as well as prudent behavior. Recall that in the first task, we categorize individuals as risk-averse if they switch from "no treatment" to "treatment" at a prevalence rate below 50%, as risk-neutral if they switch at 50% or 52%, and as risk-loving if they switch at prevalence rates above 52%. We consider an individual prudent if they switch at a lower prevalence rate in the task 2 than in task 1; as prudent-neutral if their two switch points are equal, and as imprudent if they switch at a higher prevalence in task 2 than in task 1.

Table 1 – Subject classification by risk preference

	Prudent		Prudent-neutral		Imprudent		Total	
	%	subjects	%	subjects	%	subjects	%	subjects
Risk-averse	45%	60	17%	22	21%	28	83%	110
Risk-neutral	5%	7	3%	4	2%	2	10%	13
Risk-loving	6%	8	2%	2	0%	0	8%	10
Total	56%	75	21%	28	23%	30	100%	133

Table 1 summarizes the risk preferences observed in this experiment. We find risk-averse behavior for 110 subjects (83%), risk-neutral behavior for 13 subjects (10%), and risk-loving behavior for 10 subjects (8%). Overall, subjects reveal treatment thresholds between 2% and 86%. The average threshold in the risk aversion task is at 34.2% (median: 36%; see cumulative distribution in Figure 5). The difference between the average threshold and the risk-neutral position at 50% is highly significant ($p = 0.000$ in two-sided Wilcoxon signed-rank test[14]), clearly indicating the presence of risk aversion. On average, risk-averse individuals switch at a probability of 29.5%, while risk-loving individuals switch at 64.6%.

[14] Throughout this section one- and two-sided sign tests yield very similar results to Wilcoxon signed-rank tests and are not reported.

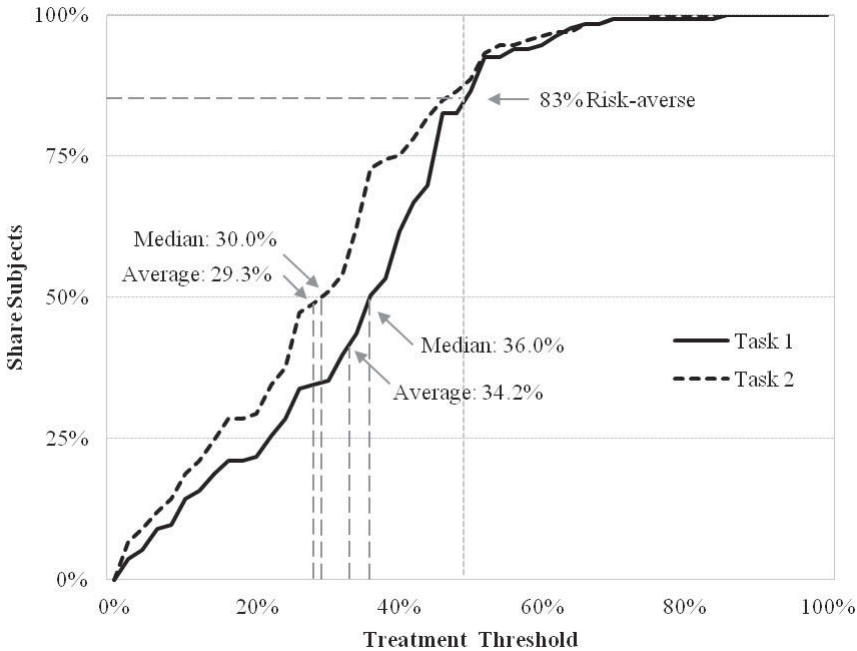


Figure 5: Cumulative distribution of treatment thresholds by task

Faced with the additional comorbidity risk in task 2, subjects lower their treatment threshold even further to 29.3% (median: 30%, see Figure 5). The shift downward from the task 1 threshold is highly significant and demonstrates prudence (p = 0.000 in two-sided Wilcoxon signed-ranks test). Comparing individual-level thresholds for tasks 1 and 2, we find 75 subjects (56%) to be prudent, 28 subjects (21%) to be prudent-neutral, and 30 subjects (23%) to be imprudent. Prudent individuals lower their treatment threshold by 12.0 %-points on average, while imprudent individuals raise theirs by 8.1 %-points. The share of prudent individuals in this experiment is just slightly lower than the 61% prudent individuals found by Noussair et al. (2011) and the 65% and 61% of prudent choices respectively reported by Ebert and Wiesen (2011a) and Deck and Schlesinger (2010).[15]

Figure 6 shows the treatment thresholds for all subjects, for task 1 on the x-axis and for task 2 on the y-axis. A switch point left of 50% on the x-axis indicates risk aversion, to the right of 52% indicates risk-loving. The diagonal marks the watershed between prudent (below the line) and imprudent (above the line) subjects.

[15] Note that only 8% (0%) of the subjects in the Ebert and Wiesen (2011a) study and 14% (2%) of those in Deck and Schlesinger (2010) make all choices prudently (imprudently).

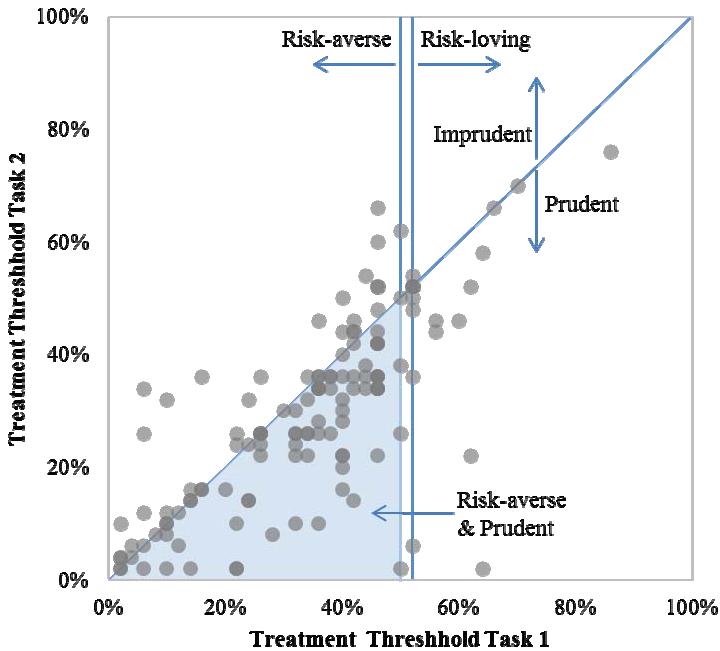


Figure 6: Treatment thresholds for all subjects in both tasks (dark markers for multiple observations)

Considering both preferences together (see Table 1), we find that 45% of our subjects are risk-averse and prudent. A further 17% are risk-averse and prudent-neutral. A Spearman rank correlation test rejects the null hypothesis that the distributions of second- and third-order risk preferences are correlated ($p = 0.156$), and that risk aversion and prudence are correlated ($p = 0.352$).

Result 1: We find risk-averse behavior among 83% of the subjects and prudent behavior among 56%.

The comparison of treatment thresholds to the risk-neutral point also allows us to infer the intensity of (second-order) risk and prudence preferences. In the risk aversion task subjects reveal an average treatment threshold that is 15.8 %-points below the risk-neutral threshold of 50.0%. Facing the background risk in task 2, they decrease their switch point by an additional 4.9 %-points. In relative terms, subjects lower their treatment threshold from the risk-neutral position by almost 31.6% in task 1

by another 14.3% in task 2. Risk-averse and prudent behavior is thus not only present, but also has a sizeable impact on the decision over medical treatment.

Result 2: We find risk preferences to have a considerable impact on the (average) treatment threshold, leading to its reduction by 41% relative to the risk-neutral position (from 50.0% to 29.3%). Risk aversion accounts for three quarters of this effect, prudence for one quarter.

4.2 Impact of Framing as Medical Treatment Decision

We find mixed evidence for the impact of framing the decision task medically or neutrally on treatment thresholds. Our subjects' second-order risk preferences are robust to differences in framing. In the first task, subjects in the medical framing switched on average at a prevalence of 35.7% (median: 40%; see Figure 7a), slightly later than the average switch point in the neutral setting at 32.1% (median: 34%). However, this difference in switch points is not statistically significant ($p = 0.215$ in Mann-Whitney U test[16]). Risk aversion is present in both framings, as both switch points differ highly significantly from the risk-neutral 50% ($p = 0.000$ in two-sided Wilcoxon signed-ranks tests). In the medical framing, 63 subjects (80.8%) are risk-averse, 6 (7.7%) are risk-neutral, and 9 (11.5%) are risk-loving. In the neutral framing, 47 subjects (85.5%) are risk-averse, 7 (12.7%) are risk-neutral, and 1 is risk-loving (1.8%). The distribution of classifications, does not differ significantly between the framing groups ($p = 0.370$ in Mann-Whitney U test).

Prudence preferences, on the other hand, appear to be reinforced by framing the decision in medical terms. Figure 7b highlights the shift in thresholds for both groups. Subjects in the medical framing switched to "treatment" at an average prevalence of 28.5% (median 29%) in the second task, 7.3 %-points earlier than in the first task (difference in medians: 11%). This shift is statistically significant ($p = 0.000$ in two-sided Wilcoxon signed-rank test). Of the 78 subjects here, 51 (65.4%) are classified as prudent, 11 (14.1%) as prudent-neutral, and 16 (20.5%) as imprudent. In the neutral framing, prudence preferences do not emerge quite as clearly: Subjects here have an average treatment threshold of 30.4% (median 32%), which is only 1.6 %-points lower than in the first task (difference in medians: 2%). This shift is weakly statistically significant ($p = 0.072$ in a one-sided sign test). Among the 55 subjects in the neutral framing of our experiment, 24 (43.6%) are prudent, 17 (30.9%) are prudent-neutral, and 14 (25.5%) are imprudent. Comparing the classification of subjects along prudence preferences in the medically and neutrally framed subsamples, the impact of the framing is

[16] Student's t-tests yield the same results as Mann-Whitney-U tests throughout this section.

significant (p = 0.037 when considering all three categories, p = 0.013 when considering only classification as "prudent" in Mann-Whitney U tests).

Based on the strength of the prudence preferences, i.e. the downward shift in treatment thresholds from task 1, the impact of the framing is even highly significant when including all subjects (p = 0.003 in Mann-Whitney U test). Considering only prudent individuals, subjects in the neutral framework shift their threshold by 8.5% while subjects in the medical framework do so by 13.7%, the difference being highly significant (p = 0.009 in Mann-Whitney U test) This means we find not only more frequent but also more intense prudent behavior when the decision is framed medically. This supports our claim that our experimental design succeeds in capturing an aspect of decision-making that goes beyond the pure monetary consequence of the decision.

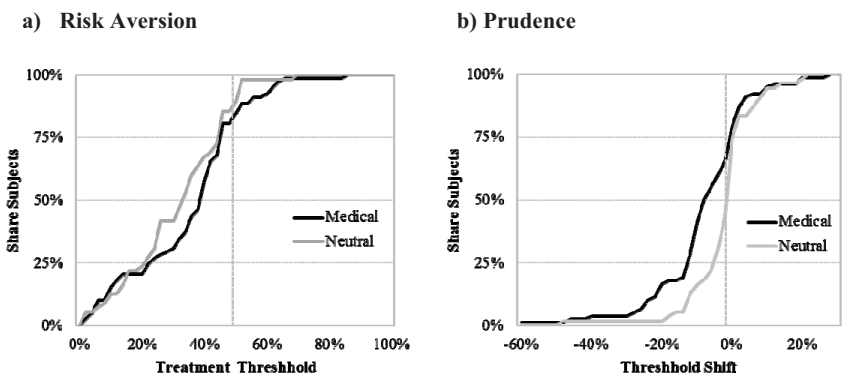


Figure 7: Cumulative distribution of treatment thresholds – medical vs. neutral framing

Result 3: We find that medical framing of the decision context does not affect second-order risk preferences, but is associated with more prudent behavior in terms of both frequency and intensity.

We also estimate OLS regressions to control for further influences on second- and third-order risk preferences (see Appendix B for results). The OLS regression confirms our finding that the medical framing is associated with the strength of prudence, i.e. the downward shift in treatment thresholds. These results are robust to variations in the order in which the tasks are completed. Demographic variables such as age and gender also provide no explanatory power. Only height is significantly associated with less risk aversion and paternal educational attainment with less risk aversion and more prudence, in line with literature on the demographic determinants of risk aversion (e.g. Dohmen et al,

2011). Finally, subjects' degree programs – specifically being a medical student – have no impact on treatment decisions in our experiment.

5 Conclusion

In our laboratory experiment on medical treatment decisions, we find evidence for risk-averse behavior among 83% of the subjects and for prudent behavior among 56%. These risk preferences have a considerable impact on the (average) treatment threshold, leading to its reduction by 41% overall relative to the risk-neutral position. Three quarters of this effect (i.e. a shift in the average treatment threshold from 50.0% to 34.2%) are due to second-order risk aversion and one quarter (i.e. a further shift to 29.3%) to prudential behavior. We find that the medical framing of the decision does not affect second-order risk preferences, but is significantly associated with more prudent behavior in terms of both frequency and intensity.

Our findings are not only relevant directly to test- and test-treatment-thresholds. They also impact diagnostic test technology when test characteristics sensitivity and specificity are endogenous and therefore a cutoff point has to be chosen which categorizes the test outcome as positive or negative (e.g. a certain concentration of a substance). Felder and Mayrhofer (2011a, 2011b) show that optimal cutoff points depend not only on the prevalence of the illness, but also on the risk preferences of the decision maker: Risk aversion lowers the optimal cutoff point and, given a comorbidity risk, prudence reduces it even further. Our results thus imply that risk preferences are also empirically relevant to diagnostic tests.

Furthermore, our findings have consequences for the construct of quality-adjusted life years (QALYs), which is widely used and sometimes even required (e.g. NICE 2004) to measure health outcomes for the appraisal of health technologies. Currently, QALYs often do not take risk preferences into account. Some of the methods used to obtain QALYs, such as the standard gamble, are grounded in EUT and thus in principle capable of incorporating at least second-order risk preferences. But although many health services researchers consider the standard gamble to be the "gold standard" (Gafni 1994), QALYs are in fact often obtained by means of questionnaires such as the EQ-5D and HUI (Richardson and Manca 2004). Recommended by the UK National Institute for Clinical Effectiveness, the EQ-5D is used to compute "utility weighted" QALYs (NICE 2004). However, this is not "utility" in the von Neumann and Morgenstern (1947) sense and does not reflect patient risk preferences. Given our results on the presence of second- and third-order risk preferences and their impact on medical treatment thresholds, it seems expedient to consider such preferences when assessing QALYs.

Test- and test-treatment thresholds as well as QALYs feature in clinical guidelines for diagnostic and treatment. Our findings suggest that these guidelines be revisited and possibly adjusted for second-order risk aversion as well as the impact of comorbidity risks given third-order risk aversion. The effect of lowering test-thresholds and optimal cutoff points for diagnostic tests, for example, might lead to a more favorable evaluation of screening programs (i.e. diagnostics in low prevalence settings).

References

Abdellaoui, M., Driouchi, A. & L'Haridon, O. (2011): Risk Aversion Elicitation: Reconciling Tractability and Bias Minimization, *Theory and Decision*, 71: 63-80.

Åstebro, T., Matab, J. & Santos-Pinto, L. (2010): *Preference for Skew in Lotteries: Evidence from the Laboratory*, Working Paper, HEC Paris.

Baltussen, G., Post, T., van den Assem, M. & Wakker, P.P. (2010): *Random Incentive Systems in a Dynamic Choice Experiment*, Working Paper, Erasmus University of Rotterdam.

Bleichrodt, H., Crainich, D. & Eeckhoudt, L. (2003): The effect of comorbidities on treatment decisions, *Journal of Health Economics*, 22, 805-820.

Carrol, C.D. (1994): How does Future Income Affect Current Consumption? *The Quarterly Journal of Economics*, 109(1): 111-147.

Carrol, C.D. & Kimball, M.S. (2008): Precautionary Saving and Precautionary Wealth, in: Durlauf, S. & Blume, L. (eds.): *The New Palgrave Dictionary of Economics*, Second Edition, MacMillan – London.

Courbage, C. & Rey, B. (2006): Prudence and optimal prevention for health risks, *Health Economics*, 15: 1323–1327.

Cubitt, R., Starmer, R. & Sugden, R. (1998): On the validity of the random lottery incentive system, *Experimental Economics*, 1(2): 115-131.

Deck, C. & Schlesinger, H. (2010): Exploring Higher Order Risk Effects, *Review of Economic Studies*, 77(4): 1403-1420.

Dohmen, T., Falk, A., Huffman, D., Sunde, U., Schupp, J & Wagner, G. (2011): Individual Risk Attitudes: Measurement, Determinants, and Behavioral Consequences. *Journal of the European Economic Association*, 9(39): 522-550.

Dynan, K.E. (1993): How Prudent are Consumers? *Journal of Political Economy*, 101(6): 1104-1113.

Ebert, S. & Wiesen, D. (2011a): Testing for Prudence and Skewness Seeking, *Management Science*, 57(7): 1334-1349.

Ebert, S. & Wiesen, D. (2011b): *Joint Measurement of Risk Aversion, Prudence and Temperance*, Working paper, University of Bonn.

Eeckhoudt, L. (2002): *Risk and Medical Decision Making*, Kluwer – Boston.

Eeckhoudt, L. & Gollier, C. (2005): The impact of Prudence on optimal Prevention, *Economic Theory*, 26(4): 989-994.

Eeckhoudt, L., Lebrun, T. & Sailly, J.C. (1985): Risk-Aversion and Physicians' Medical Decision Making, *Journal of Health Economics*, 4: 273-281.

Eeckhoudt, L. & Schlesinger, H. (2006): Putting risk in its proper place, *American Economic Review*, 96(1): 280-289.

Felder, S. & Mayrhofer, T. (2011a): *Medical Decision Making – A Health Economic Primer*, Springer – Berlin.

Felder, S. & Mayrhofer, T. (2011b): *Higher-Order Risk Preferences: Consequences for Test and Treatment Thresholds and Optimal Cutoffs,* Ruhr Economic Papers No. 287.

Fischbacher, U. (2007): Z-tree: Zurich toolbox for readymade economic experiments, *Experimental Economics*, 10(2): 171-178.

Gafni, A. (1994): The Standard Gamble Method: What Is Being Measured and How It Is Interpreted, *HSR: Health Services Research*, 29(2): 207-224.

Gafni, A. & Torrance, G.W. (1984): Risk Attitude and Time Preference in Health, *Management Science*, 30(4): 440-451.

Hennig-Schmidt, H., Selten, R. & Wiesen, D. (2011): How payment systems affect physicians' provision behavior – An experimental investigation, *Journal of Health Economics*, 30(4): 637-646.

Holt, C.A. & Laury, S.K. (2002): Risk Aversion and Incentive Effects, *American Economic Review*, 92(5): 1644-1655.

Kahneman, D., Frederickson, B.L., Schreiber, C.A. & Redelmeier, D.A. (1993): When More Pain is Preferred to Less: Adding a Better End, *Psychological Science*, 4(6): 401-405.

Kahneman, D. & Tversky, A. (1979): Prospect theory: An analysis of decision under risk, *Econometrica*, 47(2): 263-291.

Kimball, M.S. (1990): Precautionary Saving in the Small and in the Large, *Econometrica*, 58(1): 53-73.

Kroll, E.B., Trarbach, J.N. & Vogt, B. (2011): *Determining risk preferences for pain*, Working Paper, University of Magdeburg.

Laury, S.K. (2006): *Pay One or Pay All: Random Selection of One Choice for Payment*, Working Paper, Georgia State University.

List, J. (2004): Neoclassical Theory versus Prospect Theory: Evidence from the Market Place, *Econometrica*, 72(2): 615-625.

List, J. (2003): Does Market Experience Eliminate Market Anomalies?, *Quarterly Journal of Economics*, 118(1): 41-71.

McNeil, B., Weichselbaum, R., & Pauker, S. (1978): Fallacy of the five-year survival in lung cancer, *New England Journal of Medicine*, 299:1397-1401.

National Institute for Clinical Excellence [NICE] (2004): *Guide to the Methods of Technology Appraisal*, NICE – London.

von Neumann, J. & Morgenstern, O. (1947): *Theory of Games and Economic Behavior*. Second Edition, Princeton University Press – Princeton.

Noussair, C.N., Trautmann, S.T. & van de Kuilen, G. (2011): *Higher Order Risk Attitudes Demographics, and Financial Decisions*, Working Paper, Tilburg University.

Pauker, S.G. & Kassirer, J.P. (1975): Therapeutic Decision Making: A Cost Benefit Analysis. *The New England Journal of Medicine*, 293(5): 229-234.

Richardson, G. & Manca, A. (2004): Calculation of quality adjusted life years in the published literature: a review of methodology and transparency, *Health Economics*, 13(12): 1203-1210.

Shafrin, J. & Wiesen, D. (2009): *Why Aristotle didn't Get his Flu Shot: The Impact of Prudence on Prevention*, mimeo.

Starmer, C. & Sugden, R. (1991): Does the Random-Lottery Incentive System Elicit True Preferences? An Experimental Investigation, *American Economic Review*, 81(4): 971-978.

Tarazona-Gomez, M. (2004): *Are Individuals Prudent? An Experimental Approach Using Lottery Choices*, Working Paper, Laboratoire d'Economie des Ressources Naturelles (LERNA).

Wakker, P. & Deneffe, D. (1996): Eliciting von Neumann-Morgenstern Utilities When Probabilities Are Distorted or Unknown, *Management Science*, 42(8): 1131-1150.

Weinstein, M.C. & Stason, W.B. (1982): Cost-effectiveness of coronary artery bypass surgery, *Circulation*, 5(2): 56-66.

Appendix

A. Instructions

Sample instructions for the medically framed version of the experiment are shown below. The section referring to task number three, which is discussed in a separate paper, is omitted.

GENERAL INFORMATION

You are participating in an economic experiment on decision behavior. You can earn money in this experiment. Your payoff depends on the decisions you make and on random draws. There are no right or wrong decisions, please choose the alternative you consider best for yourself in every scenario.

You will make your decisions using this computer. You may not communicate with the other participants during the experiment. If you disregard this rule, you may be excluded from the experiment and forfeit your payoff. We guarantee that your choices will remain unknown to the other participants throughout the experiment.

All amounts of money are stated in Taler. 35 Taler equals 2 Euro (or 1 Taler = approx. 3 Cent). After the experiment has been completed you will be paid your earning in cash in Euro. Payoffs are explained in detail below. The entire experiment will take no more than 90 minutes.

YOUR CHOICES

In this experiment we ask you to make several decisions over medical treatment. There are three tasks.

Task 1

In this part of the experiment you have a job opportunity in each scenario, in which you can earn 100 Taler a day for up to 10 days, 1000 Taler in total. For each day you do not work, you earn nothing.

You are experiencing symptoms of illness that lead you to consult a doctor. He informs you that the symptoms are caused by **disease L** with a certain probability. This means that a known share of people with symptoms such as yours suffer from disease L; but whether you are among this share (or, alternatively, your symptoms are caused by something else) is unknown.

Example: Probability of illness of 30%

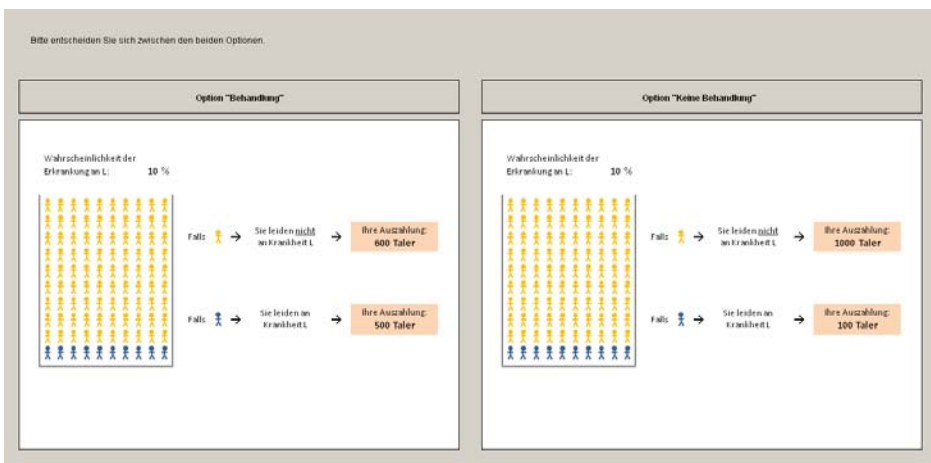
Assume that 100 people suffer of symptoms such as yours. If the probability of disease D is 30%, then for 30 of the 100 people the symptoms are due to L (dark figures in the image). The other 70 people's (light figures) symptoms have a different cause.

A treatment is available for disease L, but it is very time-consuming.

In every scenario you make a decision for or against this treatment.

If you decide to undergo treatment, you will miss some days at work as a result (no matter whether you are sick with L or not). If you are in fact suffering from L, you will recover quickly after treatment. Without treatment, you will miss several more days of work due to the illness. If you decide against treatment, you will also miss several days of work. Your payoff depends on how many days you are able to work, as you are paid 100 Taler for each day you are present.

The following picture illustrates the scenario for your decisions in this section of the experiment, using a probability of illness of 10% as an example:



If you choose treatment...

...and do not suffer from L (yellow figure), you earn 600 Taler (for 6 days of work).

...and do suffer from L (blue figure), you earn 500 Taler (5 days of work).

If you forgo treatment...

...and do not suffer from L (yellow figure), you earn 1000 Taler (for 10 days of work).

...and do suffer from L (blue figure), you earn 100 Taler (1 day of work).

Task 2

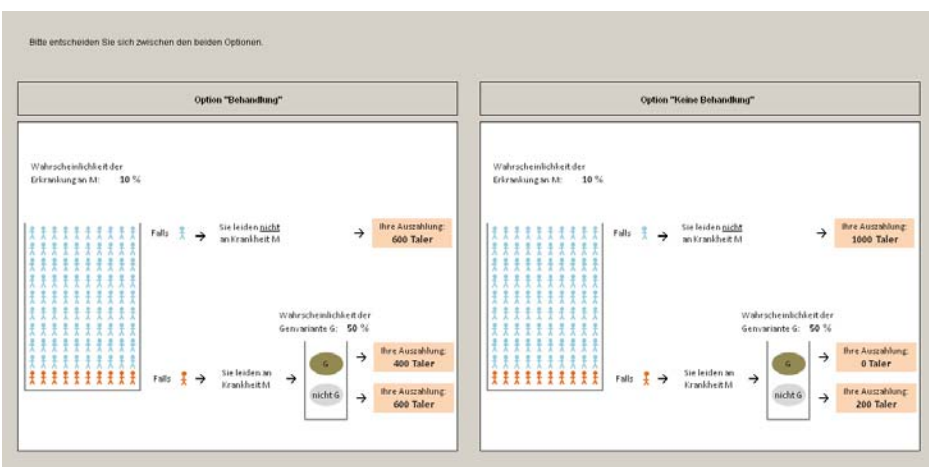
In this part of the experiment you also have a job opportunity in each scenario, in which you can earn 100 Taler a day for up to 10 days, 1000 Taler in total. For each day you do not work, you earn nothing.

You are experiencing symptoms of illness that lead you to consult a doctor. He informs you that the symptoms are caused by **disease M** with a certain probability. As in task 1, it is known that a certain share of people with your symptoms suffers from disease M, but not whether you are among this share.

Furthermore, half of the people suffering from M have a genetic variation G that causes them to experience a very severe course of the disease. The probability that you have gene variation G in addition to disease M is always 50%. Again, it is not known whether you yourself carry G or not.

A treatment is available for disease M as well, and it is also very time-consuming. The consequence of your decision for or against treatment – so how much money you are able to earn – now depends on whether or not you are suffering from M and whether or not you carry G.

The following picture illustrates the scenario for your decisions in this section of the experiment, using a probability of illness of 10% as an example:



If you choose treatment...

...and do not suffer from M (light blue figure), you earn 600 Taler (for 6 days of work).

...and do suffer from M and carry genetic variant G (brown figure and green ball), you earn 400 Taler (4 days of work).

...and do suffer from M and do not carry genetic variant G (brown figure and white ball), you earn 600 Taler (6 days of work).

If you forgo treatment...

...and do not suffer from M (light blue figure), you earn 1000 Taler (for 10 days of work).

...and do suffer from M and carry genetic variant G (brown figure and green ball), you earn 0 Taler (0 days of work).

...and do suffer from M and do not carry genetic variant G (brown figure and white ball), you earn 200 Taler (2 days of work).

Task 3 […]

Procedure during the experiment

All three sections of the experiment follow the same procedure: first you are asked to make decisions for 10 consecutive scenarios. The probabilities of the disease vary across these scenarios, from 10% in the first, to 20% in the second, to 30% in the third, and so on. In the last scenario the probability of the disease is 100%. (The consequences of your decision – so how many days you are able to work – remain unchanged throughout each section of the experiment.)

In every scenario you are asked to choose between "treatment" and "no treatment". Your decisions can follow three basic patterns:

1. You can select „treatment" at all probabilities of illness;
2. You can select „no treatment" at all probabilities of illness;
3. You can select „no treatment" at some probabilities and "treatment" at others.

If 1 is the case (you choose treatment in all 10 scenarios), you will proceed immediately to the next section of the experiment. If 2 or 3 are the case, you will be presented with 4 further (consecutive) scenarios. These contain the same consequences of your decision, and again vary in the probabilities. Here, too, you are asked to choose "treatment" or "no treatment" for each scenario. These four scenarios refine the previous 10 scenarios: If, for example, you chose "no treatment" up to and

including the probability of illness of 40% and "treatment" at 50% and above, then the next four scenarios will contain probabilities of illness between 40% and 50%.

Please consider each decision carefully. After completing each set of 10 or 4 scenarios, you will have the opportunity to see all your choices and, if necessary, repeat all 10 or 4 of them once.

After you have finished all three sections of the experiment, we kindly ask you to complete a brief questionnaire.

YOUR PAYOFF

Throughout the experiment you will make at least 30, at most 42 decisions. Your payoff will then be determined in the following steps:

1) First one out of all the first ten scenarios of all sections will be drawn at random (so 1 out of 30 scenarios), each with equal probability.

2) If this scenario is one in which you switched, e.g. from "no treatment" to "treatment", then one of the four scenarios following it will be selected in a second random draw (again, all four can be drawn with equal probability).

   Example for determining the scenario relevant to your payoff:

   Assume scenario 5 (with a probability of illness of 50%) is drawn by the first random generator.

   If you made the same choice in the fourth scenario (with probability 40%) and the fifth scenario (with probability 50%), so either "treatment" in both, or "no treatment" in both scenarios, then scenario 5 will determine your payoff.

   If you chose a different option for the first time in the scenario that was drawn, so in this example if you chose "no treatment" in scenario 4 and "treatment" in scenario 5, then the second random draw is carried out. Now one of the four follow-up scenarios with probabilities of illness of 42%, 44%, 46%, and 48% is randomly selected and will determine your payoff.

3) In the next step, your health state is determined for this scenario. Again, a random draw is carried out based on the probability of illness in the scenario.

   Example for determining health state:

   Assume the probability of illness in the payoff-relevant scenario is 44%.

   The random generator determines a number from 0 to 100. If this number is smaller than 44, you are ill with disease L, if it larger than 44 you are not.

4) Should the payoff-relevant scenario come from sections 2 or 3 of the experiment, then a final random draw will determine whether or not you carry the genetic variant G. Both cases can occur with equal probability.

5) Finally, your payoff consists of the Taler you earned in your job, which depends on your decision in the chosen scenario (steps 1 and perhaps 2), your health state (step 3), and, where relevant, your genetic variant (step 4).

At the end of the experiment, you will see a computer screen with buttons you can click to activate the random generators and determine your payoff.

Remember that your payoff is determined by <u>one</u> single decision you make. Each one of your decisions can thus affect your entire payoff.

In addition (and independent of your earned payoff) you will receive a lump sum participation fee of 3 Euro.

B. Additional Results

Table B1 – Controlling for Socio-Demographic Determinants and Order Effects

	OLS-Regression			
	Switch Point Risk Aversion Task		Difference in Switch Points Task 2 - 1	
Independent Variables	Coefficient	(Std. Err.)	Coefficient	(Std. Err.)
Constant	-87.30*	(46.93)	-18.82	(35.28)
Med Frame	2.26	(3.41)	-7.48***	(2.04)
Order of Task 1#	2.42	(3.53)	1.37	(3.19)
Order of Task 2#	1.75	(3.05)	1.91	(2.69)
Height	0.59**	(0.27)	0.08	(0.18)
Male	-2.68	(4.44)	-3.98	(3.86)
Age	0.10	(0.36)	0.10	(0.37)
Education Mother	0.56	(2.25)	1.01	(1.58)
Education Father	3.15*	(1.85)	-2.35*	(1.36)
Medical Student	0.57	(4.12)	3.01	(2.20)
Observations##	118		118	
R-squared	0.13		0.11	

*** 99% significance, ** 95% significance; * 90% significance (robust standard errors)

The third task in the experiment (which is not subject of this paper) serves as baseline.

This only includes individuals for whom all information is available.