Deductibles and Health Care Utilization: An Experiment on the Role of Forward-Looking Behavior

Glenn W. Harrison*, Nadja Kairies-Schwarz† and Johann Han†‡

05 January 2023

Abstract

We investigate the effects of nonlinear deductible contracts on health utilization behavior by using a laboratory experiment in which we can control the likelihood of hitting the deductible. We also evaluate the effect of subjects receiving regular information updates on their remaining deductible. Our results show that varying the future price has a significant effect on health care consumption. At an individual level, we identify and richly characterize heterogeneity. We find fully forward-looking, fully myopic, as well as mixed types after controlling for risk preferences. We show that there is a substantial welfare loss due to a lack in forward-looking behavior. The distribution and drivers of the welfare loss are characterized, and differ sharply according to the model of risk preferences adopted for normative evaluation.

JEL classification: I13; C91; D6, D82, D9.

Keywords: Health insurance, nonlinear prices, forward-looking behavior, laboratory experiment, risk preferences, welfare evaluation.

---

* Department of Risk Management & Insurance and Center for the Economic Analysis of Risk, Robinson College of Business, Georgia State University, USA (Harrison). Harrison is also affiliated with the School of Economics, University of Cape Town. E-mail contact: gharrison@gsu.edu.

† Heinrich Heine University Düsseldorf, Institute for Health Services Research and Health Economics, Moorenstr.5, 40225, Düsseldorf, Germany & CINCH – National Health Economics Research Center, Essen; email: johann.han@uni-due.de, nadja.kairies-schwarz@hhu.de (corresponding author).

‡ We thank Daniel Avdic, Friedrich Breyer, Amitabh Chandra, Simon Decker, Ben Handel, Mathias Kifmann, Johanna Kokot, Jimmy Martinez Correa, Justin Sydnor, Christian Waibel, as well as participants at the dggö annual meeting of the German Health Economics Association 2017 in Basel, Young Economists Meeting in Brno, HESG Meeting in Aberdeen, 4th EuHEA PhD conference in Lausanne, the German Health Economics Association workshop on "Allocation and Distribution" in Hamburg, the 5th Workshop in Behavioral and Experimental Health Economics in Atlanta, the German Economic Association 2018 in Freiburg, the Socio-Scientific Committee of the German Economic Association 2019 in Essen, the CEAR/MRIC Behavioral Insurance Workshop 2022 in Munich. Financial support provided by the Bundesministerium für Bildung und Forschung (Federal Ministry of Education and Research) is gratefully acknowledged.

1. Introduction

In an effort to reduce health care spending, policy makers, health insurance companies, and employers have tried to incorporate some form of cost sharing in their policies. One popular way to do this is to include deductibles in health care plans. There is also an efficiency rationale for including deductibles in insurance contracts, that derives from a moral hazard problem when the insured has no incentive to engage in low-cost effort to mitigate risks. A deductible aligns the interests of the insured and the insurer, up to a point.[1] Such deductible plans affect health care prices over time since an individual will initially pay higher prices for the same flow of health services before reaching the deductible than after reaching the deductible.

How health insurance pricing influences consumers' medical spending has been actively studied with field data. The first major contribution was the RAND Health Insurance Experiment, followed by the Oregon Health Insurance Experiment.[2] Given the nonlinear pricing structure over time that is central to deductible plans, how do consumers react to implied future prices as well as to the current spot price? Recent studies have studied the effects of nonlinear contracts on utilization behavior either using firm or claim level data for health insurance plans, or data for Medicare Part D plans.[3]

To investigate the effect of dynamic incentives created by deductible plans on health care utilization behavior, one might keep the spot price constant while generating variation in the future price. For instance, Aron-Dine et al. (2015) use an empirical strategy that compares individuals who join the same deductible health plans at different times of the year. These individuals face the same spot price, when they join, but different future prices due to variations in the remaining time until the end of the year when the deductible resets. Using claim-level data from employer-provided health insurance in the U.S., they conclude that consumers indeed react to more than just the spot price. *Initial* health care utilization is higher for individuals who joined early in the year and thus face a low *future* price. However, their

[1] More complex insurance contracts might also use coinsurance where the insured pays a fraction of the costs above any deductible. We focus exclusively on deductibles, assuming full insurance beyond the payment of the deductible. Pauly (1968), Zeckhauser (1970), and Arrow (1971) are classic references on the mitigation of moral hazard in insurance contract design.

[2] See, for example, Keeler and Rolph (1988); Aron-Dine et al. (2013) for the RAND Health Insurance Experiment, and Finkelstein et al. (2012) for the Oregon Health Insurance Experiment.

[3] See, for example, Cardon and Hendel (2001), Kowalski (2015), Aron-Dine et al. (2015) or Brot-Goldberg et al. (2017) for health insurance plans, and Einav et al. (2015), Dalton et al. (2020) or Abaluck et al. (2018) for Medicare Part D plans. Earlier theoretical contributions addressing the issue include Keeler et al. (1977) and Ellis (1986).

approach depends on the assumption that reasons for joining in different months can be viewed as exogenous to health care utilization behavior.

Irrespective of the empirical strategy, each of these field studies investigating the effects of nonlinear contracts on utilization behavior face potential confounders, which make it difficult to manipulate the future price while holding the spot price constant. While seasonality might be relatively easy to account for, liquidity constraints, intertemporal substitution and comorbidities might substantially affect utilization and are difficult to control for in the field.

We complement the empirical field work on the effects of nonlinear deductible contracts on health utilization behavior by using a controlled laboratory experiment. Compared to the field, the laboratory allows one to perfectly control for a constant spot price while varying the future price, as well as for other confounding factors. In the experiment, subjects go through a cycle of periods and are insured by a health plan with a deductible. In each period they face probabilistic health events and have to choose between seeking treatment or not. Similar to Harrison and Ng (2016), Kairies-Schwarz et al. (2017) and Jaspersen et al. (2022), we also elicit risk preferences from each subject and infer risk preferences at the level of the individual. This allows to derive individually optimal treatment choices that we then compare to their actual decisions, allowing us to normatively evaluate the welfare effects of nonlinear deductible contracts.

We investigate various factors that might influence health care utilization in the context of dynamic incentives. First, we manipulate the channel through which the same future price is generated by exogenously varying the contract duration length or the deductible height. Second, we exogenously vary whether subjects receive regular information updates on their accumulated costs and remaining deductible. This information may be relevant in the context of episodic healthcare utilization over time. Health insurance plans with deductibles are complex, and it is not clear how well individuals understand their insurance policies.[4] Not understanding their health insurance plans implies that individuals may not to respond correctly to the incentives. One way to improve understanding of health care plans and thus health care utilization is to provide individuals with better information about their health plans, or simplify the decision process.[5]

[4] Some evidence suggests that many individuals do not completely understand them. The most immediate evidence is from the choice of dominated strategies, in particular when they are transparent, as in Bhargava et al. (2017), Biener and Zou (2022) or Samek and Sydnor (2020). For a broader discussion of insurance literacy see Harrison et al. (2022).

[5] Several studies suggest that providing individuals with information, or simplifying the decision process, can indeed affect decision outcomes. See Kling et al. (2012) for Medicare Part D plan choices, Hastings et al.

In line with the empirical evidence investigating the effects of nonlinear contracts on health utilization behavior, we find that individuals do respond to the dynamic incentives created by deductible plans. Our results show that the future price has a significant effect on spending behavior. The channel by which the future price is manipulated, whether the same future price is reached by changing the deductible or the number of periods, seems to be secondary. Controlling for individual risk preferences, we find fully forward-looking individuals as well as fully myopic individuals.

It is one thing to identify in detail the behavioral effects of deductibles, and related informational treatments, and another thing to show that these effects lead to a welfare gain for individuals. Perhaps some individuals made mistakes when processing the choice tasks before them. To evaluate the welfare gains or losses from changes in observed behavior, we also provide a structural evaluation of latent effects on expected consumer surplus. We show that individuals insured with an insurance contract with a deductible face substantial welfare losses due to a lack in forward-looking behavior, irrespective of whether we assume Expected Utility Theory or Rank Dependent Expected Utility risk preferences. However, the distribution of welfare losses differs for the two models of risk preferences. Under RDU we find a single mode for Efficiency under RDU at high levels close to 100%, and a slight tail of lower efficiency levels. For EUT we observe two modes: a significant number of subjects around 50 to 70% Efficiency, and then some subjects with between 0 and 25% Efficiency. These results point to more individuals making mistakes that were welfare costly under EUT, to the point where their Efficiency drops well below 50% in many cases. Finally, we show that the drivers of these welfare effects also differ for the two models of risk preferences. Under RDU, for example, we find a significant welfare reduction for women, while under EUT there is a significant increase in welfare for women.

In Section 2 we lay out the experimental design and procedures. In Section 3 we report our results before presenting conclusions in Section 4.

(2008) for low income families choosing schools with high test scores, Bhargava and Manoli (2015) for economically and socially disadvantaged families claiming eligible tax benefits, or Bhargava et al. (2017) and Samek and Sydnor (2020) for health plan choices with dominated options.

2. Experimental Design

2.1 Decision Situation

Basic Decision Scenario

We employed a laboratory experiment with a sequential design. In the first part, we elicited individual risk preferences. In the second part we analyzed health utilization behavior under dynamic incentives.[6]

The design of the risk preference elicitation in the first part was similar to Andersen et al. (2008) where subjects made decisions over a battery of binary choices over risky lotteries. Each subject made 20 decisions to identify risk preferences (See Appendix A.1).

In the second part of the experiment, we investigated subjects' health care utilization behavior given dynamic incentives. This part varied between treatment conditions as shown in Table 1.[7] In a periodic task, subjects went through a cycle of decision situations. Each period a subject received an income of 50 ECU (experimental currency unit)[8] and faced one of three possible events: (a) healthy, (b) sickness A or (c) sickness B. The health events were drawn from a probability distribution known to all participants. The 'healthy' event occurred with probability 0.6, and 'sickness A' and 'sickness B' each occurred with probability 0.2.

Conditional on the realization of the health event, subjects had to choose their action. If a subject was healthy, they did not face any costs, did not make any decision, kept the periodic income, and moved on to the next period. If a subject was sick, they had to decide whether to get treated or not. Costs for both decisions did not differ for sickness A (50 ECU each), but it was cheaper to leave the sickness untreated with sickness B (30 ECU) than getting treatment (again 50 ECU). Hence, in the absence of insurance, sickness A could be interpreted as the relatively more severe sickness. The optimal decision for the relatively less severe sickness B would be *no* treatment in a one shot game in which the decision to seek treatment occurs after realizing that it is sickness B.[9] There were no consequences of the decision on future health outcomes or probabilities, and subjects knew that.

---

[6] We used this order because we did not want income effects from the utilization part to affect the elicitation of preferences. Since the payments for the elicitation part were randomly determined after the whole experiment was concluded, this concern did not arise with the given order. We controlled for potential order effects by reversing the order of both parts in one condition.

[7] The general design was inspired by a dynamic model outlined in Aron-Dine et al. (2012) and is similar to Einav et al. (2015) in the Medicare D prescription drug context.

[8] The conversion rate at the time of the experiment was 1 ECU = 0.015 EUR.

[9] An obvious extension is to consider the compound lottery in which one sick or not, but 'treatment' is needed to identify, through medical tests and diagnosis, which sickness it is.

Table 1: Potential Health States, Costs of (Not) Treating, and Event Probabilities

Health State	Cost of Choosing Treatment	Cost of Not Choosing Treatment	Probability of Event
Healthy	0	0	0.6
Sickness A	50	50	0.2
Sickness B	50	30	0.2

Depending on the condition, subjects went through 52 or 26 decision periods. The order of health events was drawn prior to the experiment and was the same between all treatment conditions at least until the 26th period, since some sessions ended after that. Hence, *ex ante*, the spot price of care is constant across individuals. Table 2 displays the order of exogenous health events that subjects faced. In contrast to the field there is no need to control different histories of health events or "seasonality", where locally or temporarily concentrated health events could affect utilization behavior.

Table 2: Draw of Health Events for Every Period

Period	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**
Health	G	B	G	G	A	G	G	B	G	G	A	G	A
	14	**15**	**16**	**17**	**18**	**19**	**20**	**21**	**22**	**23**	**24**	**25**	**26**
	G	B	G	B	B	G	G	G	B	G	G	G	A
	27	**28**	**29**	**30**	**31**	**32**	**33**	**34**	**35**	**36**	**37**	**38**	**39**
	B	G	G	B	A	A	G	B	G	G	B	B	G
	40	**41**	**42**	**43**	**44**	**45**	**46**	**47**	**48**	**49**	**50**	**51**	**52**
	A	G	G	A	G	G	G	B	G	G	B	G	A

Notes. G = good health; A = sickness A; B = sickness B. Order of health events was drawn before the experiments and was identical for all sessions and conditions.

Insurance

Apart from receiving the periodic income of 50 ECU, subjects were told that they had insurance coverage for negative health events after they had incurred medical expenditures beyond a deductible. The deductible height varied by the experimental condition. If subjects decided to treat a sickness, they paid the cost themselves up to the remaining deductible. Medical spending beyond the deductible was free.

In our benchmark condition *LowPrice,* the deductible was set at 600 ECU. In this case, subjects would need to pay for 12 treatment decisions and would have free treatment thereafter. The costs of leaving the sickness untreated, however, did not affect the deductible, and subjects would always bear those costs out of their (cumulated) periodic income. In the case of sickness B, the treatment costs of 50 ECU would count towards the deductible, and the opportunity costs of non-treatment, 30 ECU, would not affect the deductible. Given the insurance setting, the optimal decision with sickness B becomes more interesting. After every period, subjects received information about their accumulated income and, depending on the condition, about their accumulated treatment costs and the remaining deductible.

The expected end-of-year price, or future price, p^f_{it} in our design plays a critical role in understanding the dynamic incentives central to our design. It lies between 0 and 1 and is defined as $p^f_{it} = 1 - Pr(h_{it})$, where $Pr(h_{it})$ is the probability of the event h that an individual i at period t will hit the deductible by the end of the experiment if the subject treats *all further sicknesses*, regardless of the severity. For a health insurance plan with no deductible the $Pr(h_{it})$ is naturally 1 and hence p^f_{it} is 0. The lower $Pr(h_{it})$ is the higher is p^f_{it}. When $Pr(h_{it})$ is 0 the future price p^f_{it} is 1. To have a better comparison to the constant treatment spot price of 50 ECU and the opportunity cost of non-treatment of 30 for sickness B in the experiment, we then derive the normalized future price by multiplying p^f_{it} by 50. Hence, the *normalized* future price is 50 when the $Pr(h_{it})$ is 0 and p^f_i is 1. The higher $Pr(h_{it})$ is, the lower are the p^f_i and the normalized future price.

The probability of hitting the deductible depends on the probability of falling sick, the height of the deductible, and the number of periods left in the game.[10] These parameters can be exogenously manipulated in the experiment between conditions to create variation in the future price. Within any exogenous condition, the future price is identical for each subject at the beginning. We vary the future price compared to *LowPrice* through two channels: by reducing the number of periods from 52 to 26 while keeping the deductible of 600 constant (*HighPricePeriod*), and by increasing the deductible to 1150 while keeping the number of periods at 52.

2.2 Experimental Conditions

We conducted seven experimental conditions shown in Table 3. Part A of Table 3 shows our main treatment conditions. The objective of the conditions *LowPrice*, *HighPricePeriods* and *HighPriceDed* was to investigate the relationship between future price and expenditure by keeping everything constant other than the future price. In our benchmark condition *LowPrice* the deductible was relatively low (600 ECU) over the duration of 52 periods. The likelihood

[10] Since we construct the future price under the assumption that all sicknesses are treated, it only matters if subjects are sick or healthy, not the severity of illness. Hence we can retrieve the future prices from the cumulative distribution function of the binomial distribution where the number of trials equals the number of periods and the number of successes is the number of treatments needed to hit the deductible, e.g. 12 for a deductible of 600. The probability of being sick, regardless of severity, is $p = 0.4$.

of spending beyond the deductible was therefore high, which implies a low future price of 0.003 translating into a normalized future price of 0.15. The normalized future price is hence very low compared to the opportunity cost of not treating sickness B of 30. In *HighPricePeriods* we increased the future price by decreasing the number of periods to 26, and in *HighPriceDed* we increased the future price by increasing the deductible to 1150 ECU. Both manipulations led to similar (normalized) future prices, 0.674 (33.7) and 0.687 (34.35) respectively. The normalized future prices are hence above the opportunity cost of not treating sickness B (30).

Part B of Table 3 shows our robustness checks. Conditions *LowPriceNoInfo* and *HighPriceNoInfo* aimed at investigating the role of information. Here subjects only learned their accumulated income, and not the accumulated treatment costs and remaining deductible, after each decision period. Condition *LowPriceReverse* was a control condition for order effects, in which we started with the health utilization decisions followed by the elicitation of risk preferences. Condition *LowPriceNeutral* was an additional control in which the decision situation was framed in a neutral indemnity insurance context, with no association to health.

Table 3: Experimental Conditions Overview

Condition	Spot Price	Deductible	# Periods	Normalized Future Price	*N*	Add. Info
Part A: Main Conditions						
LowPrice	50	600	52	0.15	48	Yes
HighPricePeriods	50	600	26	33.7	24	Yes
HighPriceDed	50	1150	52	34.35	24	Yes
Part B: Robustness Conditions						
LowPriceNoInfo	50	600	52	0.15	48	No
HighPriceNoInfo	50	1150	52	34.35	24	No
LowPriceReverse	50	600	52	0.15	20	Yes
LowPriceNeutral	50	600	52	0.15	47	Yes
Total					235	

2.3 Behavioral Expectations

To derive behavioral expectations, we first illustrate the relationship between the expected future price and the number of periods remaining for the respective conditions varying in the number of periods and height of the deductible. We then discuss the relationship between the expected future price and expected utilization behavior and transfer it to expectations based on our experimental parametrization.

Figure 1 is inspired by the intuitive model illustration in Aron-Dine et al. (2012) and adjusted to our parameters deductible height and number of periods), and definition of the future price. The upper panel in Figure 1 illustrates that the expected future price, and thus the probability of reaching the deductible, depends on the number of periods left as well as the height of the deductible. The circle on the bottom left marks our *LowPrice* condition where a subject has the full 52 periods to reach the deductible of 600. Here the probability of hitting the deductible is almost 100 percent, resulting in a future price of close to 0. Future prices of hypothetical later entry dates, implying fewer periods to utilize the deductible, into the same deductible plan of 600 are located along the dashed line from this circle. Keeping the deductible constant, the future price of utilization increases when the time to utilize the same deductible decreases. Our *HighPricePeriod* condition, where subjects have only 26 periods to reach the deductible of 600, also lies on this line and is marked with a diamond. Analogously, keeping the time to utilize constant at 52 periods, the future price increases with the height of the deductible. This is reflected in our *HighPriceDed* condition with a deductible of 1150, marked with a square. The connected line displays the future prices for different hypothetical entry dates into a 1150 deductible contract. The future price between the two *HighPrice* conditions is initially almost identical, which allows us to explore whether the channel that varies the future price affects utilization behavior differently. Our variation of the future price, while keeping the spot price constant, creates conditions akin to the "ideal experiment" proposed by Aron-Dine et al. (2015, p.726) in a simplified yet controlled setting. Hence, in the experiment we investigate the three specific contract combinations of deductible height and periods left to reach the deductible as depicted in Figure 1.

The lower part of Figure 1 illustrates the expected number of treatments for sickness A and sickness B, for the three actual contracts as well as other hypothetical contracts. Here we consider initial utilization, which we define as the number of treatments of sickness A and sickness B within the first 26 periods of the contract for two behavioral types (a fully forward-looking individual who takes all future periods into account and a myopic individual that does not look beyond the current period). For simplicity here, we assume risk neutrality. Taken together with the upper panel, the lower panel illustrates how the expected number of treatments relates to the future price.

The **completely myopic** individual only takes the costs of the current period into account. In the event of sickness B, this individual would never choose to treat for 50 ECU since the cost of non-treatment is only 30 ECU. In the event of sickness A, this individual would be indifferent between paying 50 for treating or incurring costs of 50 for non-treatment.[11] Thus, the expected number of doctor visits for a myopic individual over the initial 26 periods would be 5.2 (or lower).[12] Since a myopic individual does not consider the future price, this result holds regardless of the height of the deductible as well as the number of periods left to reach the deductible. Thus, the grey horizontal dash-dotted line illustrating the behavior of a completely myopic individual is horizontal. In total contrast, the black horizontal dash-dotted line on top, showing expected visits of 10.4, displays a hypothetical **fully-insured** individual who would treat every sickness since this individual does not have to worry about cost.

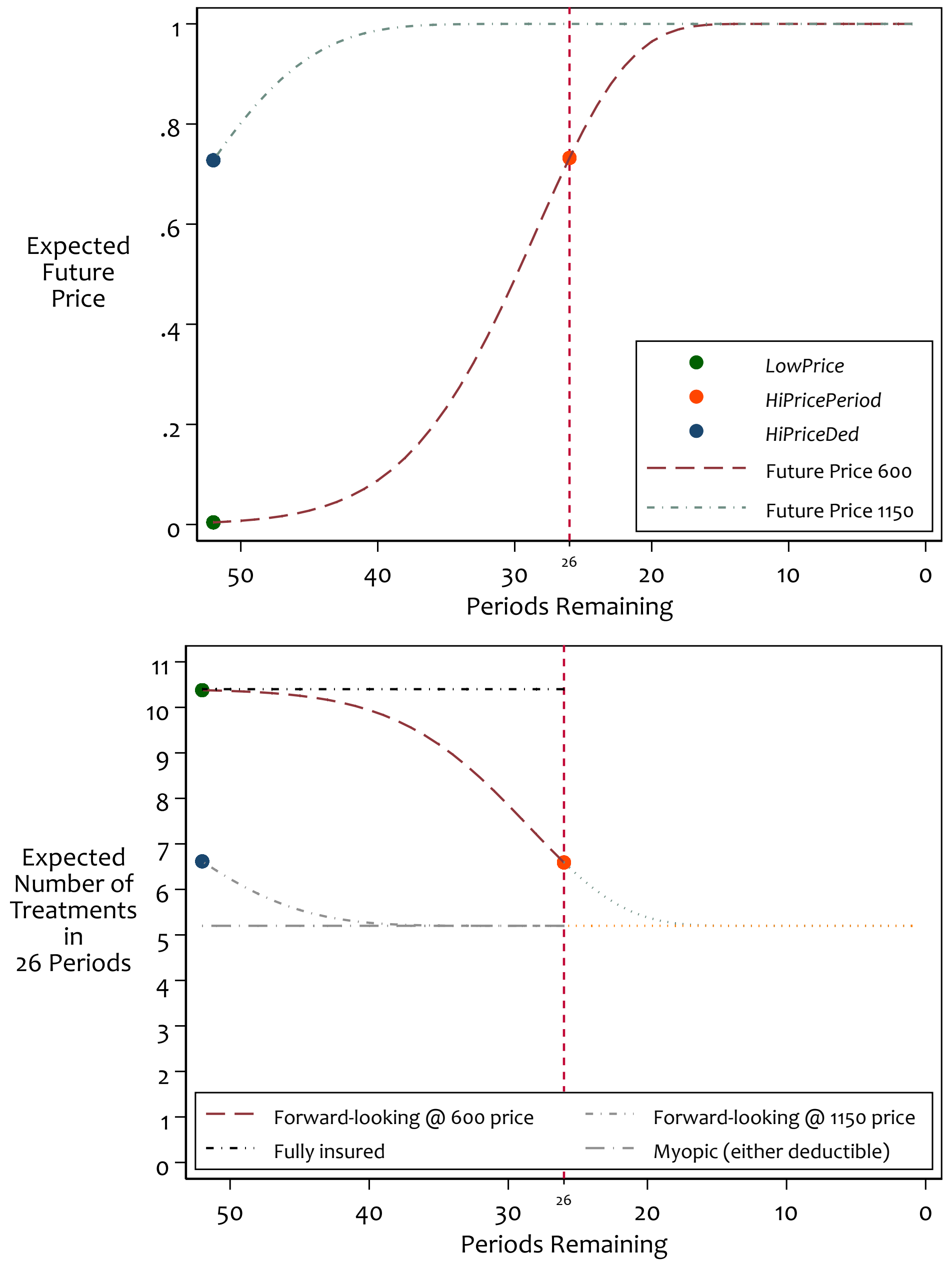
The **completely forward-looking** individual considers the future price. The expected initial utilization of this individual thus depends on the future price and differs between our *HighPrice* and *LowPrice* conditions. Its utilization behavior, given the *LowPrice* condition with a deductible of 600, is depicted by the black dashed line in Figure 1 that starts at the dot. Initial utilization of a completely forward-looking individual in these conditions is comparable to a fully insured individual, since the likelihood of using up the deductible is very high and the future price is close to 0 (as shown in the upper panel).[13] As the available time to utilize the deductible decreases, the future price increases and expected utilization decreases. Again, increasing the deductible, while holding the remaining periods constant as in our *HighPriceDed* contract (blue circle), increases the future price and reduces the expected initial utilization of the completely forward-looking individual.

[11] To calculate an expected number of visits, we assume a basic understanding of the dynamic incentives in the sense that it is reasonable to reduce the deductible. Hence, this individual would choose treat sickness A and visit the doctor in this case. If this individual had a preference for not treating A or would use a mixed strategy, the expected number of treatments would be even lower.

[12] Here, we multiply the probability of occurrence of sickness A with the number of number of the initial 26 periods, i.e. $Expected\#treatments_{myopic} = \Pr(SicknessA) \cdot 26 = 0.2 \cdot 26 = 5.2$

[13] For *LowPrice* in the beginning the expected number of treatments for a completely forward-looking individual is: $Expected\#treatments_{FL}^{LowPrice} = \Pr(SicknessA) \cdot 26 + \Pr(HitAB) \cdot \Pr(SicknessB) \cdot 26 = 0.2 \cdot 26 + 0.997 \cdot 0.2 \cdot 26 = 10.4$ Thus, we expect a forward looking individual to base his treatment decision for B on the initial probability to hit the deductible if it would treat both cases. $\Pr(HitAB) = 0.326$ in *HighPriceDed* and $\Pr(HitAB) = 0.3$ in *HighPricePeriod*.

Figure 1: Future Prices and Expected Initial Utilization



In order to derive behavioral expectations for our experimental conditions, we assume that individuals choose their individual treatment strategy based on the implied parameters before the utilization periods start and then stick to it. To build intuition, consider the two pure strategies of always treating only sickness A and always treating both sickness A and B.[14] In our *LowPrice* conditions, always treating sickness A and B can then be considered the forward-looking strategy. Given the actual draw of sicknesses within the experiment, this

[14] These are the dominant strategies when assuming risk neutrality and a preference for treating sickness A over not treating it, which signals a basic understanding of the dynamic structure.

strategy would yield 10 treatments over the first 26 periods. Treating sickness A only can be considered the myopic strategy in our *LowPrice* conditions resulting in 4 treatments. In the *HighPrice* conditions, both forward-looking and myopic types would choose the strategy to treat sickness A only[15], resulting in 4 treatments over the first 26 periods.[16] Hence we have these two hypotheses:

Hypothesis 1: Assuming risk neutrality and an ex ante choice of a pure decision strategy, forward-looking (myopic) individuals will choose the strategy to always treat sickness A and B (treat sickness A only) and will treat 10 (4) times in the first 26 periods in our *LowPrice* conditions. In the *HighPrice* conditions, both forward-looking and myopic individuals will choose to treat sickness A only and will treat 4 times in the first 26 periods.

Hypothesis 2: If the future price is relevant, we expect aggregate treatment rates to be higher in our *LowPrice* than in our *HighPrice* conditions. This difference should result from different treatment rates for sickness B.

2.3 Experimental Procedure

The experiments were conducted at the Essen Laboratory for Experimental Economics (elfe) in Essen in 2016 and 2017. Sessions lasted about 90 minutes and 235 students participated after being recruited by the online recruiting system ORSEE due to Greiner (2015).

Subjects were randomly assigned to their seats in the laboratory upon arrival. Before each part of the experiment, they were given the appropriate instructions and were given time to read them and ask questions. Any questions were answered in private by the same experimenter across all sessions. To ensure understanding of the decision task in each part, subjects had to answer a set of control questions, and the experiment did not start unless all subjects had answered the control questions correctly.

At the end of the experiment, two subjects per session were randomly selected and paid for one of their decisions in the first part.[17] The random selection took place at the end of the

[15] The forward-looking type recognizes that the probability to hit the deductible is smaller than 50 percent.

[16] The profits individuals make when playing either one of these two pure decision strategies given our draw of health events in the respective conditions can be found in Appendix A.2.

[17] The first part consisted of the 20 choices over risky lotteries referred to earlier, as well as 10 choices over time-dated amounts of money. These 10 choices are not relevant to the decisions within a 90-minute session, and are not discussed further. However, two subjects were randomly selected and were each paid in the first part for one of their 30 decisions. One individual was paid for one *ex post* randomly determined choice over risky lotteries and one individual for one *ex post* randomly determined choice over time-dated amounts of money. On average we had about 21 subjects within each session.

experiment to avoid income effects for the utilization behavior part of the experiment. In the utilization part, every subject was paid out, and the average payoff was 26.60 EUR.

The earnings were determined by the accumulated income from the utilization decision part of the experiment, which was the sum of periodic income after accounting for expenditures in sickness. Finally, subjects were asked to answer a short questionnaire with questions on demographics and questions related to their behavior in the previous decisions.

3. Results

3.1 Initial Aggregate Health Care Utilization Behavior

For the analysis of our results we focus predominantly on initial utilization behavior, the decisions to seek treatment or not, during the first 26 periods. First, the *ceteris paribus* assumption only holds over all conditions for this span since *HighPricePeriods* has only 26 periods. Second, we want to keep the spot price constant across conditions, and this is only the case before hitting the deductible: afterwards, the spot price would be 0. Given our random draw of health events, the earliest possible period to hit the deductible of 600 is period 30 if all 12 sicknesses are treated at this point. By period 26, 10 sickness events will have occurred, 4 sickness A events and 6 sickness B events (see Table 2).

Low Price vs. High Price

We begin by comparing the results of *LowPrice* and *HighPricePeriods*. This resembles the empirical strategy utilized by Aron-Dine et al. (2015) to compare individuals who join the same deductible plan at different times of the year. By reducing the number of periods from 52 to 26, while keeping everything else constant, subjects have less time to hit the deductible and face a higher future price. If subjects did not react to the future price, only the severe sickness A would be treated, for 4 treatment decisions overall, and we would not see a difference between the two conditions, since the order of health events is identical for the first 26 periods. If subjects behaved in a forward-looking manner, we would expect treatment for all health events in *LowPrice*, resulting in 10 treatment decisions by period 26.

Table 4.1 provides information on the average number of treatment decisions, regardless of the severity of illness, and the respective treatment rates for the severe sickness A and the mild sickness B for each condition. We see that the average number of treatment decisions by a subject is 8.15 (out of a possible 10) for *LowPrice,* while it is only 5.67 for *HighPricePeriods* over the first 26 periods. This difference is significant (p<0.001) according to a two-sided Mann Whitney U-test (MWU). Thus, subjects decide to treat significantly more when the future price is low. We can also infer that the difference stems from treatment decisions for the mild health events, sickness B. Although over 95 percent of severe sickness A cases are treated in both conditions, indicating that subjects recognize this as a dominant strategy, treatment rates for sickness B differ substantially across the two conditions. In particular, 71 percent of the mild cases B are treated when the future price is low, even though not treating would be cheaper in a one shot situation (30 ECU vs. 50 ECU). In contrast, only 31 percent are treated when the future price is higher, and the likelihood of spending beyond the deductible is low.

Table 4.1: Initial Behavior by Period 26 After 10 Sickness Cases

	Average Number of Treatment Choices	Treatment Rate for Sickness A	Treatment Rate for Sickness B
Part A: Main Conditions			
LowPrice	8.15	0.97	0.71
HighPricePeriods	5.67	0.96	0.31
HighPriceDed	5.04	0.94	0.22
Part B: Robustness Conditions			
LowPriceNoInfo	8.29	0.98	0.73
HighPriceDNoInfo	5.45	0.92	0.3
LowPriceReverse	8.1	0.9	0.75
LowPriceNeutral	8.34	0.96	0.75

Notes. Treatment rate indicates share of respective sickness cases treated by all subjects. See Table A.5 in Appendix A.6. for results after 52 periods.

Table 4.2: Initial Behavior by Period 26 After 10 Sickness Cases for Males and Females

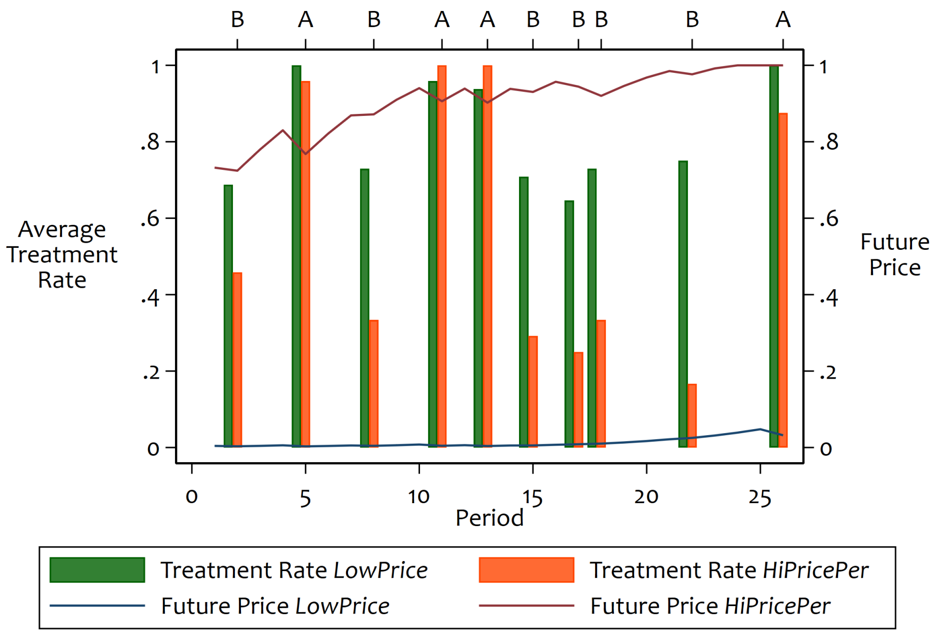
	Average number of treatment choices			Expected number of treatment choices		*p*-value (Males vs. Females)
	All	Males	Females	Forward-looking	Myopic	
Part A: Main Conditions						
LowPrice	8.15	8.77	7.61	10	4	0.14
HighPricePeriods	5.67	4.44	6.4	4	4	0.05
HighPriceDed	5.04	5.24	4.57	4	4	0.22
Part B: Robustness Conditions						
LowPriceNoInfo	8.29	8.87	7.76	10	4	0.07
HighPriceDNoInfo	5.45	6.5	4.71	4	4	0.27
LowPriceReverse	8.1	9.25	6.38	10	4	0.02
LowPriceNeutral	8.34	9.17	7.48	10	4	0.02

Notes. The table reports descriptive statistics on treatment choices for males and females (means); reported exact p-values are based on two-sided Mann-Whitney U tests.

Figure 2 illustrates this pattern over time. The green and orange bars reflect the treatments rates in *LowPrice* and *HighPricePeriods* by sickness periods, while the blue and the red line visualize the average respective future prices accounting for utilization behavior. It is apparent that subjects treat mild cases B less when the future price is high. Moreover, we observe a significant negative time trend for treating sickness B in *HighPricePeriods*.[18] A higher future price due to a higher deductible in *HighPriceDed* also leads to a significantly lower number of treatments of 5.04 compared to *LowPrice* (MWU, p<0.0001). This result indicates that subjects do not only react to the spot price but also show forward-looking behavior and anticipate that they will spend beyond the deductible in *LowPrice*.

[18] We run a random effects probit regression with the decision to seek treatment or not in *HighPricePeriod* as the dependent variable and the respective Periods of sickness B as independent variable. The coefficient is negative and significant on a level of 0.05. See Table A.3 in Appendix A.3.

Figure 2: Initial Utilization in *LowPrice* and *HighPricePeriods*



Result 1 corresponding to Hypothesis 2: A higher future price due to a lower number of periods in *HighPricePeriods* leads to a significantly lower number of treatments compared to *LowPrice*. The difference stems from different treatment rates for the mild sickness B. This indicates that subjects do not only react to the spot price but also show forward-looking behavior and anticipate that they will spend beyond the deductible in *LowPrice*.

Table 4.2 provides information on the average number of treatment choices by gender, as well as the expected number of treatment choices for a forward-looking or myopic individual for each condition. It shows that in all our *LowPrice* robustness conditions women treat sicknesses significantly less than men ($p \leq 0.07$). This result is in line with Hayen et al. (2021) who show that women react stronger to cost-sharing schemes than men. Our results further suggest that women are more myopic than men.

High Deductible vs. Fewer Periods

Do subjects react differently when manipulating the future price through either more decision periods or a higher deductible? We compare treatment behavior in *HighPricePeriods* with the behavior in *HighPriceDed*. If subjects reacted to the probability of hitting the deductible, the differences between both conditions should be marginal since the future price is almost identical. However, in *HighPricePeriods* the higher future price, or lower probability of hitting the deductible, is a product of halving the time to utilize the deductible, while in *HighPriceDed* the price was manipulated by increasing the deductible from 600 to 1150. We find that the average number of treatments is 5.67 in *HighPricePeriods* and 5.04 in *HighPriceDed* and the difference is not significant (p=0.38). From Table 4.1 we again see that for both conditions the treatment rate in sickness A is close to 100 percent. The treatment rate for sickness B is similar between treatments: 31 percent in *HighPricePeriods* and 22 percent in *HighPriceDed*.

Figure 3: Initial Utilization in *HighPricePeriods* and *HighPriceDed*

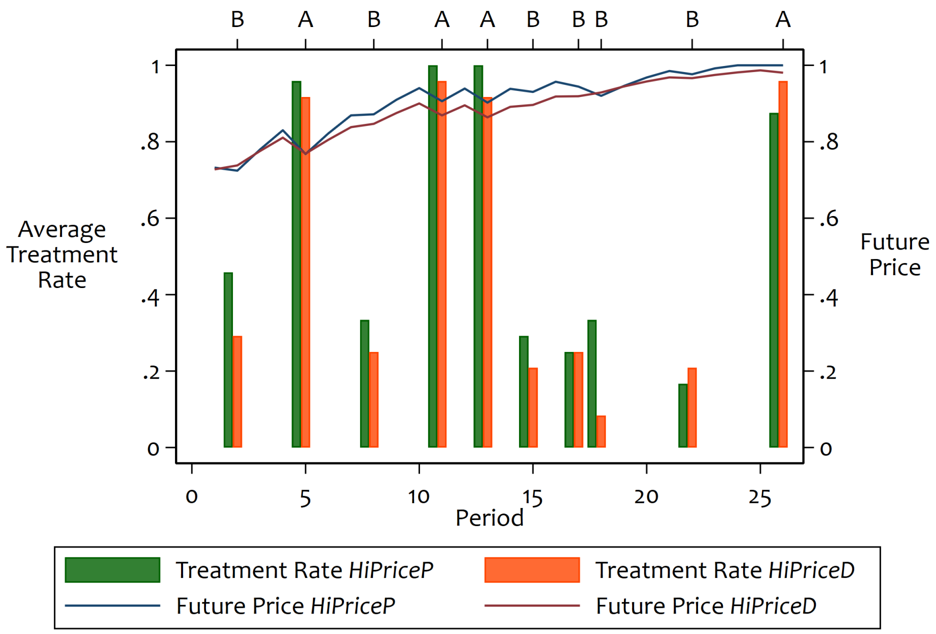


Figure 3 illustrates the dynamic relationship between treatment decisions and future price in these two conditions. The blue and red lines in Figure 3 mark the average future prices for the two conditions over time, accounting for past utilization behavior in every period. Since the future price is almost identical initially, the lines effectively have an identical starting point. Over time, they stay relatively close together while approaching the price of 1, the point where it is impossible to hit the deductible, at period 26. The green and orange bars in Figure 3 represent the respective treatment rates in *HighPricePeriods* and *HighPriceDed*. The difference in behavior between the two possible health events is clearly visible, and for both *HighPrice* conditions the majority opts against treating when the mild sickness B occurs. This result further supports the conclusion that subjects react to changes in the future price. Also, subjects seem to display a good understanding of the future price in both *HighPrice* treatments. Whether the likelihood of hitting the deductible is manipulated via number of periods or height of deductible appears to be secondary.[19]

Robustness Conditions: Information, Order and Framing Effects

To investigate the role of information on utilization behavior, we compare *LowPrice* and *LowPriceNoInfo*. In *LowPrice* subjects received information on their income, their cumulated treatment costs, and their remaining deductible after each period. In *LowPriceNoInfo* the information was reduced and the cumulated income update was the only feedback that subjects received. When we compare the initial utilization between both conditions, the information about the remaining deductible and cumulated treatment costs does not seem to lead to differences in utilization behavior (see Table 4.1). Utilization is almost identical: 8.15 in *LowPrice* compared to 8.29 in *LowPriceNoInfo,* and the difference is not significant (p=0.86). Similarly, not giving information on the remaining deductible does not significantly affect the utilization when we compare *HighPriceDed* with *HighPriceDNoInfo* (p=0.70).

Finally, we control for the order and framing. Reversing the order of part 1 and part 2 and eliciting preferences after the insurance part, does not have a significant effect. Utilization is 8.10 in *LowPriceReverse* and not significantly different from the benchmark condition (p=0.89). In *LowPriceNeutral* we changed the wording of the instructions and removed any association to health, and framed the task as general indemnity insurance. The average utilization is 8.34 and not significantly different from *LowPrice* (p=0.71).

3.2 Heterogeneity in Utilization Behavior

Our controlled experimental design allows for classifying individuals based on their observed individual utilization behavior. However, this is only possible in the conditions with a low future price since, assuming risk neutrality, the behavioral prediction is identical for both types given a high future price, not to treat sickness B. In the LowPrice conditions, however, decisions during sickness B provide a way to classify individuals.[20] To begin with, we classify individuals assuming risk-neutrality and expected costs. In the following, we will then account for risk preferences and derive individually optimal treatment paths and strategies.

[19] In a natural setting time preferences would play a role as well. In our experimental design, given that the time horizon was about 1 hour, it is plausible to assume that there is no time preference within the session.

[20] We aggregate the *LowPrice* treatments here. As shown above, we do not find significant differences in aggregate behavior treatment behavior between these conditions. For details, see Appendix A.4 Table A.4.

Individual Utilization Behavior and Risk Neutrality

Assuming risk-neutrality and expected costs, only forward-looking individuals would treat the mild cases sickness B before hitting the deductible. From Table 5 we can infer that about 15 percent of subjects in the *LowPrice* conditions never treat the mild cases B and that around 60 percent always treat sickness B.[21] Hence, in a scenario with a low future price, forward-looking behavior appears to be prevalent.

Result 2 corresponding to Hypothesis 1**:** We find individuals who consistently never treat *and* individuals who consistently always treat sickness B in our *LowPrice* conditions. Hence, our results are consistent with some subjects looking towards the future price and taking that into account, and some subjects not only looking at the spot price.

Table 5: Classification of Treatment Behavior in *LowPrice* Conditions by Period 26

	Never Treat B	Mixed Strategy	Always treat B	Total
LowPrice (pooled)	24 (14.7%)	43 (26.4%)	96 (58.9%)	163

However, this analysis does not account for individual risk preferences and thus we cannot disentangle whether an individual treats sickness B because the individual is forward-looking or because the individual is risk averse. Moreover, an optimal treatment path for an individual within a period depends on the previous events that have occurred as well as the individual's previous treatment decisions. For instance, for an individual who has previously deviated from their originally optimal path by not treating B in several sickness periods, it may at some later point become optimal not to treat sickness B anymore. We now derive such individually optimal strategies that account for individual risk preferences as well previous treatment behavior.[22]

[21] Thus, at least 70 percent of the individuals stick to one strategy: they always or never treat sickness B in one of the treatments with a low price. Taking a closer look at the average treatment rates of the subjects who do not stick to one of the two strategies, we do not find a clear pattern in their behavior over time. Individual decision patterns of these subjects show that few individuals decide against treating sickness B at its first occurrence, and always treat later, which could be attributed to learning. However, for the majority this is not a plausible behavioral explanation.

[22] For the horizon we are considering within the experiment we assume that individuals have no discounting at all. Accounting for risk preferences should therefore be sufficient to derive optimal strategies. Yet, a natural expansion of the experimental design would be to extend the horizon of the experiment and hence also account for individual time preferences.

Individual Utilization Behavior and Risk Preferences

We estimate risk preferences at the individual subject level, to be able to make normative inferences that respect the heterogeneity of risk preferences that we expect a priori for individuals. Following Gao et al. (2022), we adopt a Bayesian approach to this estimation task, specifically a Bayesian Hierarchical Model (BHM). Our subjects made 20 binary choices over lotteries with objective risks, and that number of observations would not be able to reliably generate individual estimates of risk preferences using the classical Maximum Likelihood methods used by Harrison and Ng (2016) for their normative evaluation of insurance choices. A BHM addresses this issue[23] by pooling the behavior over all subjects to estimate hyper-parameters for a model of the risk preferences of a single "representative agent" that can then be used as informative priors for the estimation of risk preferences at the individual level. In this manner the larger data set of the subjects facing the same binary choices can be used to generate more plausible estimates for one individual than if each subject was estimated in isolation. The informative prior employed here is also referred to as a "shrinkage prior" since it effectively shrinks extreme estimates towards the pooled risk estimates. The extent of the shrinkage towards the pooled estimates depends on how informative the 20 observations are for each individual. For some individuals the prior will have little effect on the estimates, since their 20 observations are relatively informative about their risk preferences. But for other individuals, with more noisy behavior, the informative prior will play a more important role. In this manner the BHM is said to naturally "regularize" the estimates for each individual.

We estimate risk preferences for each individual assuming Expected Utility Theory (EUT) or the Rank-Dependent Utility (RDU) model of Quiggin (1982).[24] Since EUT is nested within RDU, we could just estimate risk preferences assuming RDU. But for many policy-makers and economists, EUT is more normatively attractive than RDU for policy evaluation. We remain agnostic on that issue, for reasons explained in Harrison and Ross (2018; p.49ff.),

[23] Indeed, Gao et al. (2022; §3.1) evaluate exactly this issue, by comparing inferences about welfare when each subject made 80 binary choices with inferences about welfare when just 20 choices per subject were selected at random. They find an acceptably high correlation of inferences about individual welfare, precisely to guide experimental designs in answering the question of "how many binary choces are needed" to generate reliable welfare evaluations. Obvioulsy, more (informative) choices are always better than fewer, but in practice it is valuable to have guidance on the number of choices that are likely to be sufficient for reliable welfare evaluation.

[24] We could also estimate a Dual Theory (DT) model in which the utility function is assumed to be linear, and the risk premium for an individual is generated entirely by their estimated probability weighting function. DT is also nested in RDU. However, we have *never* found systematic evidence that any noticeable fraction of individual subjects exhibit DT behavior. And DT is not regarded as normatively attractive by anyone, as far as we are aware.

but prefer here to be able to evaluate welfare using EUT and RDU to be able to inform policy with either, and to see if the use of either model makes much difference to normative conclusions. For both EUT and RDU we assume a Constant Relative Risk Aversion utility function, and for RDU we assume a flexible two-parameter probability weighting function due to Prelec (1998). We evaluate behavior using the mean of the posterior distributions of each individual for their risk preferences parameters.

Evaluation of Treatment Choices

In Section 3.2., we classified utilization behavior based on the observed treatment choices in our *LowPrice* conditions assuming risk-neutrality and expected costs. In particular, we identified individuals who consistently treated sickness B over the first 26 periods and those who consistently did not treat the latter. However, this classification relies on risk neutrality and expected costs and thus does not consider individual risk preferences. Having characterized subjects by their elicited risk preferences, we can now use their individual risk parameters to derive optimal treatment choices. This allows us to reconcile actual choices made with these optimal choices. We can then identify myopic behavior or decision errors that cannot be explained by individual risk preferences and thus also forward-looking behavior.

For evaluating the series of treatment choices, we calculate the prospective expected costs and expected utilities according to RDU or EUT for each individual in each respective period for different treatment strategies. We evaluate the two strategies (i) either always treat both sicknesses A and B or (ii) only treat sickness A[25] for the remaining duration as potential optimal strategies. For each period and individual, we calculate the prospective expected utility under RDU and EUT if the individual would stick to one of these two strategies until the end. While in period 0, i.e. before making any treatment choices, a subject facing a deductible of 600 knows for sure that he will have to pay out of pocket for the first 12 sicknesses. The probability to pay out of pocket in later stages of the experiment depends on the previous treatment strategy of the subject as this affects the probability to hit the deductible. We use these probabilities to hit the deductible as well as the treatment cost of 50 or the opportunity cost of 30, to calculate the prospective expected costs and individual utilities for each prospective period and sum it up.[26] This gives us the cumulated expected

[25] Treat only the mild sickness B and treat nothing are other pure strategies but they are statistically dominated. For simplicity and as mentioned in Section 2.3., we do not look at mixed strategies.

[26] From the perspective of period 0 and with $P(OOP)_i^j$ being the probability to treat out of pocket in period j given strategy i, the expected costs in period 1 would be $\mathrm{P}(OOP)_{AB}^1 \cdot P(A \cup B) \cdot C(Treat)$ for strategy "treat

utility of the two strategies for each subject in period 0. We then repeat this exercise in every period, taking the actual utilization up to this point into account. This leave us with the expected utilities for both strategies for every subject in every period of the game which allows us to evaluate the optimal treatment choices and thus an optimal treatment path for each subject. We do this under the assumption of an RDU as well an EUT model or risk preferences.

Given actual treatment choices, we can then determine whether an individual deviates from the optimal choice in the respective period. We assume that individuals deviate from their RDU or EUT optimal path in case they don't treat sickness A in the respective periods, or do not treat sickness B although this would be optimal for them in the respective period, or do treat sickness B although this would not be optimal for them in the respective period. If someone is an optimal RDU or EUT decision maker it hence does not follow that they never or always treat B in every period. In fact, the optimal path for an individual might be a mixed strategy. The total number of these deviations serves as a measure of decision error that we can relate to our previous classification in the *LowPrice* conditions that was based on whether the mild sickness B was treated in the initial 26 periods. When a deviation from an optimal choice and the decision not to treat the mild sickness B overlap, it is very likely due to a lack in forward-looking behavior and not due to risk preferences. If the choice not to treat the mild sickness B, however, is in line with the preferences (no deviation from the optimal choice), we cannot rule out that the individual considered the future price and thus cannot label it as myopic.

For this, we now consider the first 26 periods and the mild sicknesses B. Given our ex-ante random draw, every subject faced 6 cases of sickness B. From Table 6, we can see the number of individuals for whom the previous classification based on risk neutrality and expected costs resembles their optimal decision based on their risk preferences in each category as well as the average number of deviations from the optimal choice. Among the subjects who never treat the mild sickness B, 21 (8) out of 24 individuals do so in line with their preferences under RDU (EUT). For them not treating sickness B is not a behavioral bias or lack in forward-looking behavior. On the other hand, 26 (88) out of 96 subjects behave optimally by treating both sickness cases under RDU (EUT). Inconsistent play, i.e. switching

both A and B" (abbreviated as AB) and $P(OOP)_A^1 \cdot P(A) \cdot C(Treat) + \mathrm{P(B)} \cdot \mathrm{C}(NotTreat)$ for the strategy to treat "A only" (A), where C are the costs for treating ($Treat$) or not treating ($NotTreat$) the respective sickness. To begin we sum up the expected cost for every upcoming period and yield the cumulated utility for these two strategies from the perspective of period 0. Then we continue with the perspective of period 1, and do the same calculations. To get the expected utilities we just replace the cost function with utility functions.

between treating and not treating B, is optimal for 1 (17) of the 43 subjects under the RDU (EUT) model. From the average number of deviations, we can also see that those who never treat sickness B are much less (more) prone to deviate and make behavioral mistakes under RDU (EUT). Their average number of deviations is 0.67 (2.71) while for those who always treat sickness B it is 3.75 (0.21) under RDU (EUT). (See A.5 Figures A.3.1 and A.3.2 for histograms with distributions for RDU and EUT). Assuming an RDU (EUT) model of risk preferences, a classification of observed treatment behavior based on risk neutrality seems to work quite well for the strategy to never treat sickness B (to always treat sickness B) but is suboptimal for the strategy to always treat sickness B (never treat sickness B).

Table 6 Number of Optimal Decision Makers and Average Number of Deviations at Period 26 (*LowPrice* Conditions Pooled)

	Never Treat B	Mixed Strategy	Always Treat B
Optimal Decision Makers RDU	21 out of 24	1 out of 43	26 out of 96
Avg. # of deviations RDU	0.67 (sd 1.81)	2.79 (sd 1.34)	3.75 (sd 2.69)
Optimal Decision Makers EUT	8 out of 24	17 out of 43	88 out of 96
Avg. # of deviations EUT	2.71 (sd 2.48)	1.74 (sd 1.72)	0.21 (sd 0.75)

Accounting for risk preferences, the optimal treatment paths now allow for classifying individuals by their degree of forward-looking behavior. For this, we assume that individuals that always deviate from their optimal path in periods with sickness B are fully myopic, while those who never deviate from their optimal path in these periods are fully forward-looking. Those individuals who deviate from their optimal path several times but not always are mixed types in the sense that they display some degree of forward-looking behavior. Table 7 displays the number (percentages) of the respective behavioral types in our *LowPrice* conditions. These results show that while under an RDU model of risk preferences we find a roughly equal distribution of myopic, mixed and forward-looking types, we find about two thirds of forward-looking types, one third of mixed types and only very few myopic types under the EUT model of risk preferences. These results suggest that the distribution of types differs substantially depending on which model of risk preferences one assumes to be the normative metric.

Table 7 also shows that the distribution of types differs by gender. Under an RDU model of risk preferences we find 41% (24%) myopic, 27% (49%) mixed and 32% (27%) forward-looking types for males (females). Under an EUT model of risk preferences we find 1% (5%)

myopic, 17% (38%) mixed and 82% (57%) forward-looking types for males (females). Hence, under an RDU model of risk preferences there is a tendency towards more male myopic types, while under an EUT model of risk preferences we find less forward-looking types. This can be explained by a larger shares of female mixed types under both models of risk preferences.

Table 7 Classification of Behavioral Type by Number of Deviations at Period 26 and Gender (LowPrice Conditions Pooled)

	Myopic	Mixed	Forward-looking	Total
All				
Total RDU	53 (33%)	62 (38%)	48 (29 %)	163
Total EUT	5 (3 %)	45 (28%)	113 (69 %)	163
Male				
Total RDU	33 (41%)	22 (27%)	26 (32%)	81
Total EUT	1 (1%)	14 (17 %)	66 (82%)	81
Female				
Total RDU	20 (24%)	40 (49%)	22 (27%)	82
Total EUT	4 (5%)	31 (38%)	47 (57%)	82

Result 3: Accounting for individual risk preferences, we identify consistent myopic and forward-looking behavior. The distribution of types differs depending on which model of risk preferences one assumes to be the normative metric: there are substantially more (less) forward-looking (myopic) types under the EUT model of risk preferences than under the RDU model of risk preferences.

We further investigate determinants for not treating according to individual RDU or EUT preferences. We run a random effect probit to describe the determinants of deviating from the optimal choice over the first 26 periods in all *LowPrice* conditions. Under RDU we find treatment effects. Giving individuals information on the remaining height of the deductible after each round significantly reduces the likelihood to deviate from the optimal path. A neutral frame, on the other hand, significantly increases the likelihood of deviating from the optimal path. Under EUT we find a strong gender effect: women have a higher likelihood for deviating from their optimal choice. Hence women appear to make more behavioral mistakes in the sense that they do not act in accordance to their risk preferences under EUT.

Table 8 Deviations from Optimal Paths and Characteristics at period 26

	RDU Deviate	EUT Deviate
Female	-0.022	0.740***
	(0.089)	(0.209)
Information	-0.267**	0.193
	(0.108)	(0.241)
Private Health Insurance	-0.077	0.324
	(0.119)	(0.254)
Neutral Framing	0.343***	0.029
	(0.107)	(0.234)
N	4238	4238

Notes. Based on 26 decisions made by 163 subjects. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Welfare effects

In a next step we undertake an evaluation of utilization choices in terms of their welfare effects, using the notion of the Expected Consumer Surplus (ECS) of observed utilization choice. The logic is conventional in terms of welfare economics, with one extension. Each utilization choice offers the subject a well-defined lottery. How the subject trades off the final outcomes of that lottery is determined by her utility function parameters, and how the subject weights the utility of different payoffs is determined by the probabilities of those payoffs and her probability weighting parameters. Assume for the moment that she is an EUT decision-maker, solely to ease the basic exposition. Then if we know her utility function parameters we can infer the EUT that she attaches to each possible choice. We can also then immediately calculate the Certainty Equivalent (CE) of that lottery to her, by solving for CE in the equation $U(CE) = EUT$. To take simple case, if we had used the power utility function $U(x) = x^r$, then the CE is equal to $EUT^{(1/r)}$. Even if closed-form expressions for the CE do not exist, it is a simple numerical matter to find the scalar CE that solves this equation. If the decision maker chooses one utilization lottery A over another utilization lottery B, we just evaluate EUT^A and

EUT^B, then evaluate CE^A and CE^B, and the ECS is simply CE^A - CE^B. In the familiar language from welfare economics, the ECS is just the certain amount of money that the individual would require to be just willing to give up her preferred utilization choice for the alternative.

All of this is familiar welfare economics, assuming we know the risk preferences of the individual or, as in our case, we can predict those risk preferences from a pooled model with demographic characteristics that differentiate each individual. We can extend it immediately to the case of an individual with RDU risk preferences, but the welfare-theoretic logic is identical and standard.

Some economists insist that only EUT risk preferences are normatively attractive, and it is a simple matter to substitute EUT parameters for that individual. We disagree with this assumption about EUT being normatively attractive, for reasons discussed in Harrison and Ng (2016), but that is a debate for another time, and for now we can consider the effects on our conclusions of also assuming that EUT risk preferences are the appropriate normative metric.

However, what is novel here is that we have a measure of the risk preferences of the individual that is independent of the observed utilization choice. For normative evaluation of the utilization choice, we must in fact have some independent measure of risk preferences. The reason is that if we inferred risk preferences from observed utilization choices, we would always infer, in expectation, that the ECS from the observed utilization choice was zero or positive by direct revealed preference. In the example above, we were careful to say that we observe the subject choosing utilization lottery A over B. We did not say that the ECS of that decision was positive. In fact, and this is the normative point of behavioral welfare evaluation, we might have observed the subject making a poor decision and generating a negative ECS. This approach to the normative evaluation of lotteries was developed by Harrison and Ng (2016), and reviewed methodologically by Harrison (2019).

Once we have evaluated the ECS for each individual and choice, we can collate these effects over all of the choices made by an individual. This measures Efficiency, in the language of experimental economists since Plott and Smith (1978): how much of the potential ECS did the subject actually make from her observed choices. The usual measures of Efficiency were developed for non-stochastic settings, in terms of the ability of the subject to "extract surplus" from the experimenter by appropriate choices. We simply utilize ECS and ask about the ability of the subject to extract surplus *in expectation*.

Figure 4: Total Foregone Welfare

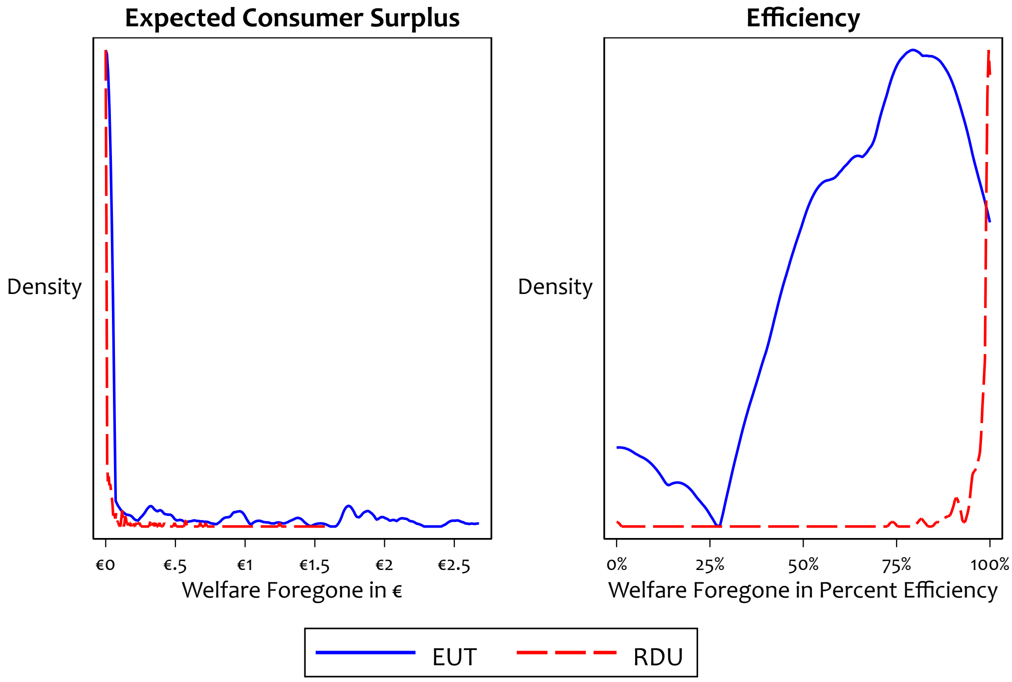


Figure 4 displays results for ECS in the left panel, and results for Efficiency in the right panel. We present results assuming RDU risk preferences in red, and assuming EUT risk preferences in blue. Under RDU we see a large number of small welfare losses, yet all losses are below the level of €2. On the other hand, under EUT we find many tiny ECS welfare losses, but a long tail of larger losses. These findings for ECS translate into a single mode for Efficiency under RDU at high levels close to 100%, and a slight tail of lower efficiency levels. For EUT we observe two modes: a significant number of subjects around 50 to 70% Efficiency, and then some subjects with between 0 and 25% Efficiency. These results point to more individuals making mistakes that were welfare costly under EUT, to the point where their Efficiency drops well below 50% in many cases. Under EUT we find that more than two thirds of the subjects are classified forward-looking. Hence, these costly mistakes appear to originate from a few subjects making very large mistakes.

We can explore these welfare results by examining the "marginal effects" of treatments or demographics on welfare distributions. In this case we consider the range of treatments and demographics shown in Figure 5 for RDU and Figure 6 for EUT. The diamond symbol is the point estimate of the marginal effect, and the bars either side of that symbol show the 95% confidence interval.

For both EUT and RDU we find a significant effect of the low probability of reaching a deductible leading to a welfare loss. Under RDU this effect is positive, but under EUT it is negative. The loss is particularly strong under EUT, and appears to be the driving factor behind the large tail of losses in ECS noted earlier under EUT for Figure 4. We also identify welfare effects that are specific with respect to the variations in our robustness conditions. Under RDU we find that there is a tendency for extra information on the remaining height of the deductible to lower welfare. Under EUT we find only a tiny positive effect of extra information. For the abstract insurance context, we find a lower welfare under RDU and only a tiny positive effect on welfare under EUT.

Figures 5 and 6 also show that there are welfare effects which depend on individual subject characteristics. Under RDU we find a significant welfare reduction for women. In contrast, under EUT there is a significant increase in welfare for women. Studying in the field of economics and having a private health insurance increase welfare under both RDU and EUT. The effect of statutory health insurance on welfare differs depending on the underlying model of risk preferences. Under RDU having statutory health insurance has a significant negative effect on welfare while it has a significant positive effect under EUT.

Result 4: There is a substantial welfare loss due to a lack in forward-looking behavior irrespective of whether we assume EUT or RDU risk preferences. The distribution and drivers of the welfare losses differ for the two models of risk preferences.

Figure 5: Marginal Effects on Efficiency under RDU

Marginal average effects from a Fractional regression model

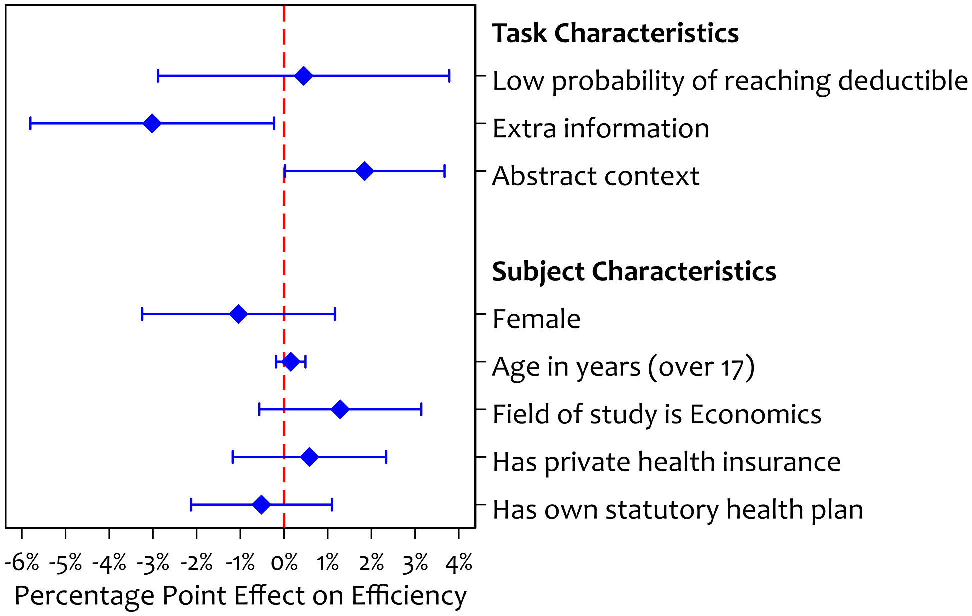
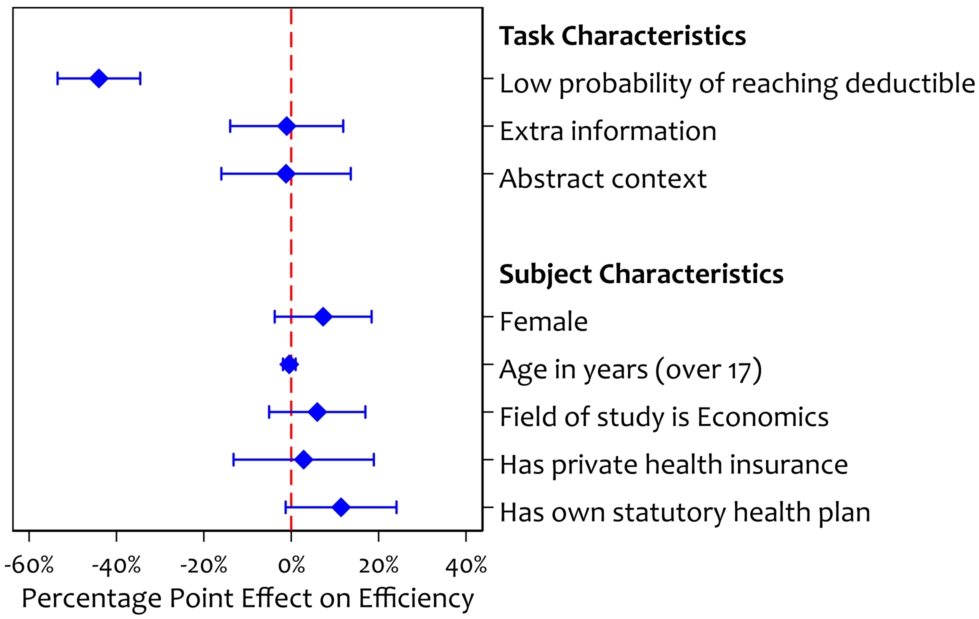


Figure 6: Marginal Effects on Efficiency under EUT

Marginal average effects from a Fractional regression model



4. Conclusions

We complement the empirical evidence on the effects of nonlinear deductible contracts on health care utilization behavior by using a controlled laboratory experiment. Compared to the field, the laboratory environment allowed us to control for keeping the spot price constant while varying the future price, as well as for other confounding factors such as seasonality, liquidity, or comorbidities. Subjects in the experiment went through a cycle of periods and were insured by a health plan with a deductible. In each period, they faced probabilistic health events and had to choose between treating or not. We also elicited individual risk preferences for each subject, allowing for welfare estimation of observed health care utilization behavior.

Consistent with recent empirical results for nonlinear health insurance plans with deductibles, and Medicare Part D plans, we find that subjects respond to the embedded dynamic incentives in aggregate and do not only react to the spot price. We also find a tendency for women to treat significantly less sicknesses under a deductible plan with a low future price. This is consistent with Hayen et al. (2021) who show that women react more strongly to cost-sharing than men. Whether the future price is manipulated through more decision periods or a higher deductible does not significantly affect utilization behavior as long as the likelihood of hitting the deductible is the same.

Aron-Dine et al. (2015) suggest that people understand the dynamic incentives of health insurance contracts with deductibles to some degree. This implies that both the spot price of insurance as well as the future price of insurance should be relevant to determine the price elasticities of demand for medical services. We contribute to the characterization of this heterogeneity. Specifically, our results suggest that, assuming an RDU (EUT) model of risk preferences, 33% (3%) of individuals are myopic and only take the spot price into account, 38% (28%) are mixed in their choice behavior and 29% (69%) are forward-looking and take the future price of insurance into account.

Our results further show that the drivers of welfare effects also crucially depend on the model of risk preferences one assumes to be the normative metric. The results from our treatment of giving individuals additional information on the remaining height of the deductible in the robustness conditions also provides some insights to policy makers. Assuming an RDU model of risk preferences we show that regularly giving individuals information on the remaining height of the deductible may actually decrease efficiency, while assuming EUT we find a tiny positive effect. Previous studies suggest that providing individuals with information, or simplifying the decision process, can indeed affect decision outcomes. We add to this literature and show that the direction of the effects of such policy measures can depend on the underlying model of risk preferences.

A further driver of welfare that depends on the underlying model of risk preferences is gender. Under a RDU model of risk preferences, we find a significant welfare *reduction* for women, whereas there is a significant welfare *increase* for women under an EUT model of risk preferences. We add to the literature on gender differences in health care consumption given a deductible. The tendency we find for women to treat less sicknesses under a deductible, which is consistent with the observations in the field, may translate either into a positive *or* negative welfare impact depending on the underlying model of risk preferences used for normative evaluation.

Future research could investigate an extension of the horizon of the experiment and hence also account for individual time preferences, or consider subjective beliefs about loss probabilities.

References

Abaluck, J., Gruber, J., & Swanson, A. (2018). Prescription drug use under Medicare Part D: A linear model of nonlinear budget sets. *Journal of Public Economics*, 164, 106-138.

Andersen, S., Harrison, G. W., Lau, M. I., & Rutström, E. E. (2006). Elicitation using multiple price list formats. *Experimental Economics*, 9(4), 383-405.

Andersen, S., Harrison, G. W., Lau, M. I., & Rutström, E. E. (2008). Eliciting risk and time preferences. *Econometrica*, 76(3), 583–618.

Aron-Dine, A., Einav, L., & Finkelstein, A. (2013). The RAND Health Insurance Experiment, three decades later. *Journal of Economic Perspectives*, 27(1),197–222.

Aron-Dine, A., Einav, L., Finkelstein, A., & Cullen, M. R. (2012). Moral hazard in health insurance: How important is forward looking behavior? *Working Paper 17802*. National Bureau of Economic Research.

Aron-Dine, A., Einav, L., Finkelstein, A., & Cullen, M. (2015). Moral hazard in health insurance: do dynamic incentives matter? *Review of Economics and Statistics*, 97(4), 725–741.

Arrow, K. (1971). *Essays in the Theory of Risk Bearing.* Amsterdam: North-Holland Publishing Company.

Bhargava, S., & Manoli, D. (2015). Psychological frictions and the incomplete take-up of social benefits: Evidence from an IRS field experiment. *American Economic Review*, 105(11), 3489-3529.

Bhargava, S., Loewenstein, G., & Sydnor, J. (2017). Choose to lose: Health plan choices from a menu with dominated option. *Quarterly Journal of Economics*, 132(3), 1319-1372.

Biener, C., & Zou, L. (2022). The Value of Choice in Mandatory Health Insurance. *Working Papers on Finance 2021-04*, University of St. Gallen, School of Finance.

Brot-Goldberg, Z. C., Chandra, A., Handel, B. R., & Kolstad, J. T. (2017). What does a deductible do? The impact of cost-sharing on health care prices, quantities, and spending dynamics. *Quarterly Journal of Economics*, 132(3), 1261-1318.

Cardon, J. H. & Hendel, I. (2001). Asymmetric information in health insurance: evidence from the National Medical Expenditure Survey. *RAND Journal of Economics*, 32(3), 408–427.

Dalton, C. M., Gowrisankaran, G., & Town, R. J. (2020). Salience, myopia, and complex dynamic incentives: Evidence from Medicare Part D. *Review of Economic Studies*, 87(2), 822-869.

Einav, L., Finkelstein, A., & Schrimpf, P. (2015). The response of drug expenditure to nonlinear contract design: Evidence from Medicare Part D. *Quarterly Journal of Economics*, 130(2), 841–899.

Ellis, R. P. (1986). Rational behavior in the presence of coverage ceilings and deductibles. *RAND Journal of Economics*, 17(2): 158–175.

Finkelstein, A., Taubman, S., Wright, B., Bernstein, M., Gruber, J., Newhouse, J.P., Allen, H., Baicker, K. & Oregon Health Study Group (2012). The Oregon Health Insurance Experiment: evidence from the first year. *Quarterly Journal of Economics*, 127(3), 1057-1106.

Gao, X. S., Harrison, G. W., & Tchernis, R. (2022). Behavioral welfare economics and risk preferences: A Bayesian approach. *Experimental Economics*, https://doi.org/10.1007/s10683-022-09751-0.

Greiner, B. (2015). Subject pool recruitment procedures: organizing experiments with ORSEE. *Journal of the Economic Science Association*, 1(1), 114-125.

Harrison, G. W. (2019). The behavioral welfare economics of insurance. *The Geneva Risk and Insurance Review*, 44(2), 137-175.

Harrison, G., Morsink, K., & Schneider, M. (2022). Literacy and the quality of index insurance decisions. *The Geneva Risk and Insurance Review*, 47(1), 66-97.

Harrison, G.W., & Ng, J. M. (2016). Evaluating the expected welfare gain from insurance. *Journal of Risk and Insurance*, 83(1), 91-120.

Harrison, G. W., & Ross, D. (2018). Varieties of paternalism and the heterogeneity of utility structures. *Journal of Economic Methodology*, 25(1), 42-67.

Harrison, G. W. , & Rutström, E. E. (2008). Risk aversion in the laboratory, in Cox, J.C. and Harrison, G.W., eds*., Risk Aversion in Experiments* (Research in Experimental Economics, Volume 12). Bingley, UK: Emerald Group Publishing.

Hastings, J. S., & Weinstein, J. M. (2008). Information, school choice, and academic achievement: Evidence from two experiments. *Quarterly Journal of Economics*, 123(4), 1373-1414.

Hayen, A. P., Klein, T. J., & Salm, M. (2021). Does the framing of patient cost-sharing incentives matter? The effects of deductibles vs. no-claim refunds. *Journal of Health Economics*, 80, 102520.

Holt, C. A. & Laury, S. K. (2002). Risk aversion and incentive effects. *American Economic Review*, 92(5), 1644–1655.

Jaspersen, J. G., Ragin, M. A., & Sydnor, J. R. (2022). Predicting insurance demand from risk attitudes. *Journal of Risk and Insurance*, 89(1), 63-96.

Kairies-Schwarz, N., Kokot, J., Vomhof, M., & Weßling, J. (2017). Health insurance choice and risk preferences under cumulative prospect theory - an experiment. *Journal of Economic Behavior & Organization*, 137, 374-397.

Keeler, E. B., Newhouse, J. P., & Phelps, C. E. (1977). Deductibles and the demand for medical care services: The theory of a consumer facing a variable price schedule under uncertainty. *Econometrica, 45,* 641–655.

Keeler, E. B. & Rolph, J. E. (1988). The demand for episodes of treatment in the health insurance experiment. *Journal of Health Economics*, 7(4), 337–367.

Kling, J. R., Mullainathan, S., Shafir, E., Vermeulen, L. C., & Wrobel, M. V. (2012). Comparison friction: Experimental evidence from Medicare drug plans. *Quarterly Journal of Economics*, 127(1), 199–235.

Kowalski, A. E. (2015). Estimating the tradeoff between risk protection and moral hazard with a nonlinear budget set model of health insurance. *International Journal of Industrial Organization*, 43, 122–135.

Loewenstein, G., Friedman, J.Y., McGill, B., Ahmad, S., Linck, S., Sinkula, S., Beshears, J., Choi, J.J., Kolstad, J., Laibson, D., Madrian, B.C., List, J.A., & Volpp, K.G. (2013). Consumers' misunderstanding of health insurance. *Journal of Health Economics*, 32(5), 850-862.

Pauly, M. V. (1968). The economics of moral hazard: comment. *The American Economic Review*, 58(3), 531-537.

Plott, C. R., & Smith, V. L. (1978). An experimental examination of two exchange institutions. *The Review of Economic Studies*, 45(1), 133-153.

Prelec, D. (1998). The Probability Weighting Function. *Econometrica,* 66(3), 497-528.

Quiggin, J. (1982). A Theory of Anticipated Utility. *Journal of Economic Behavior & Organization,* 3(4), 323–343.

Samek, A., & Sydnor, J. R. (2020). Impact of Consequence Information on Insurance Choice. *Working Paper 28003*. National Bureau of Economic Research.

Zeckhauser, R. (1970). Medical insurance: A case study of the tradeoff between risk spreading and appropriate incentives. *Journal of Economic Theory*, 2(1), 10-26.

A Appendix

A.1 Price Lists for Risk Preferences

The following tables display the tasks subjects had to complete for the elicitation of risk preferences. Subjects had to decide between Option A and Option B in every row. The (expected) payoff differences were not displayed to them. Each subject chose one line at which to switch from A to B: This is called a sequential Multiple Price List (sMPL) by Andersen et al. (2006).

At the end of the experiment one person per session was randomly selected to get paid for one random decision in the risk task. This procedure happened at the very end of the experiment to avoid income effects that could confound behavior in the second part.

Table A 1: Risk Lottery A

	Option A		Option B		Expected payoff difference
	20 EUR	16 EUR	38.50 EUR	1 EUR	
1	10%	90%	10%	90%	11.65 EUR
2	20%	80%	20%	80%	8.30 EUR
3	30%	70%	30%	70%	4.95 EUR
4	40%	60%	40%	60%	1.60 EUR
5	50%	50%	50%	50%	-1.75 EUR
6	60%	40%	60%	40%	-5.10 EUR
7	70%	30%	70%	30%	-8.45 EUR
8	80%	20%	80%	20%	-11.80 EUR
9	90%	10%	90%	10%	-15.15 EUR
10	100%	0%	100%	0%	-18.50 EUR

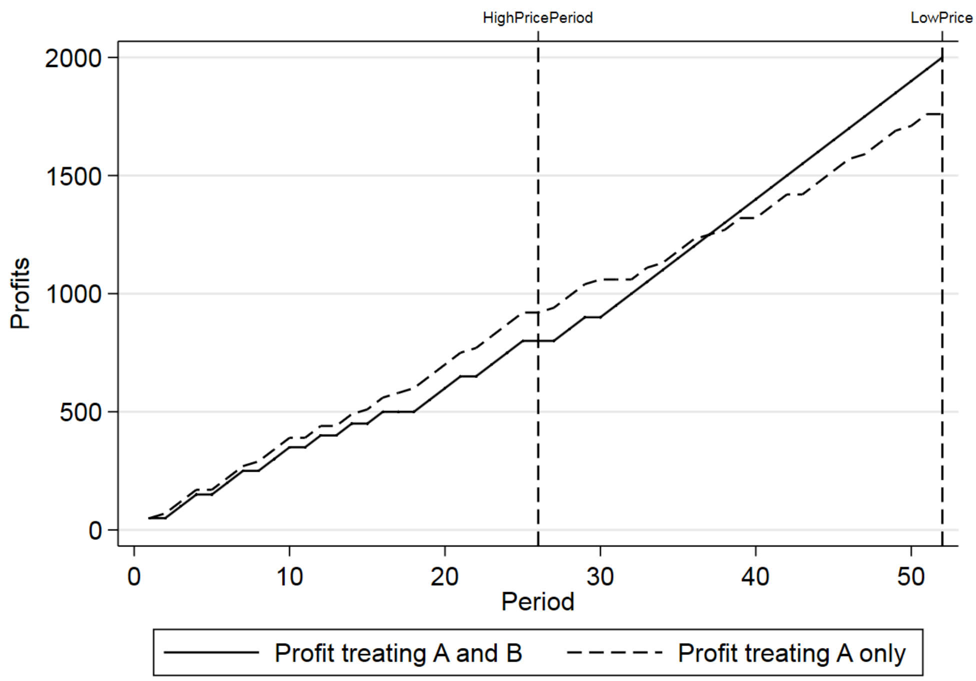
Table A.2: Risk Lottery B

	Option A		Option B		Expected payoff difference
	22.50 EUR	15 EUR	40 EUR	5 EUR	
1	10%	90%	10%	90%	7.25 EUR
2	20%	80%	20%	80%	4.50 EUR
3	30%	70%	30%	70%	1.75 EUR
4	40%	60%	40%	60%	-1.00 EUR
5	50%	50%	50%	50%	-3.75 EUR
6	60%	40%	60%	40%	-6.50 EUR
7	70%	30%	70%	30%	-9.25 EUR
8	80%	20%	80%	20%	-12.00 EUR
9	90%	10%	90%	10%	-14.75 EUR
10	100%	0%	100%	0%	-17.50 EUR

A.2 Profits From Pure Decision Strategies

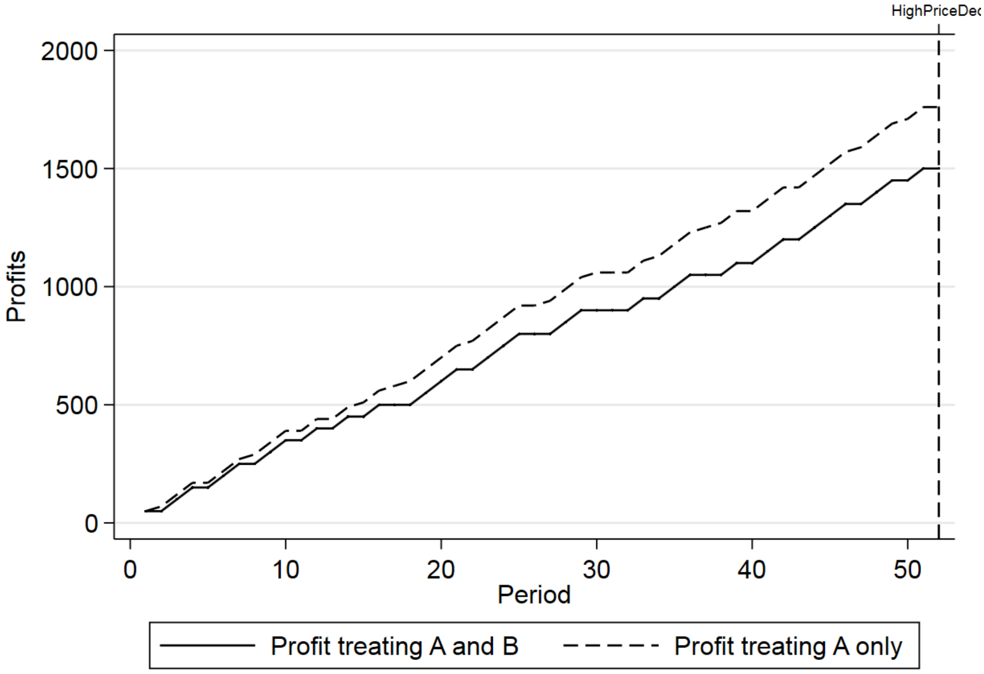
To substantiate our predictions we consider the profits generated by these two pure strategies for our *LowPrice* and *HighPrice* conditions when the actual draw of sicknesses within the experiment is taken into account, see Figures A.1 and A.2. Figure A.1 displays the profit[1] outcomes for a deductible of 600 and thus our *LowPrice* and *HighPricePeriod* conditions. In the *LowPrice* conditions, the strategy to always treat sickness A and B leads to the deductible being hit at period 30. Subsequent treatment is then free of cost. Always treating Sickness A and B, leads to higher profits than only treating sickness A, and profits thereafter remain higher until period 52. In the *HighPricePeriod* condition, the experiment ends after 26 periods. At that point, the profit lines have not crossed yet and treating sickness A only is more profitable. Similarly, in Figure A.2 we observe that, given a high future price due to a high deductible, *HighPriceDed*, always treating sickness A and B is less profitable than treating the A only.

Figure A.1: Profit Outcomes for *LowPrice* and *HighPricePeriod* by Treatment Strategy



[1] Profits in Experimental Currency Unit (ECU) with 1 ECU = 0.015 EUR.

Figure A.2: Simulated outcomes *HighPriceDed*



A.3 Time Trend in HighPricePeriods

Table A.3: Probability to Treat Sickness B in *HighPricePeriods*

	-1
Period	-0.0711^{***}
	-0.0249
_cons	-0.115
	-0.53
/	
lnsig2u	1.275^{**}
	-0.58
N	144
rho	0.782
sigma_u	1.892

Notes. Random effect probit regression with participant's decision to seek treatment in the six cases of sickness B during 26 periods is dependent variable. Standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

A.4 Heterogeneity in Utilization Behavior

Table A.4: Classification of Treatment Behavior by Period 26 per Condition

	Never treat B	Mixed Strategy	Always treat B	N
LowPrice	9 (18.75%)	11 (22.92%)	28 (58.33%)	48
LowPriceNoInfo	8 (16.67%)	12 (25.00%)	28 (58.33%)	48
LowPiceRev	3 (15.00%)	5 (25.00%)	12 (60.00%)	20
LowPriceNeutral	4 (8.51%)	15 (31.91%)	28 (59.57%)	47
HighPricePer	10 (41.67%)	11 (45.83%)	3 (12.50%)	24
HighPriceDed	13 (54.17%)	10 (41.67%)	1 (4.17%)	24
HighPriceDNoInfo	11 (45.83%)	8 (33.33%)	5 (20.83%)	24

A.5 Classification Behavioral Types

Figure A.3.1 Behavioral Types under RDU

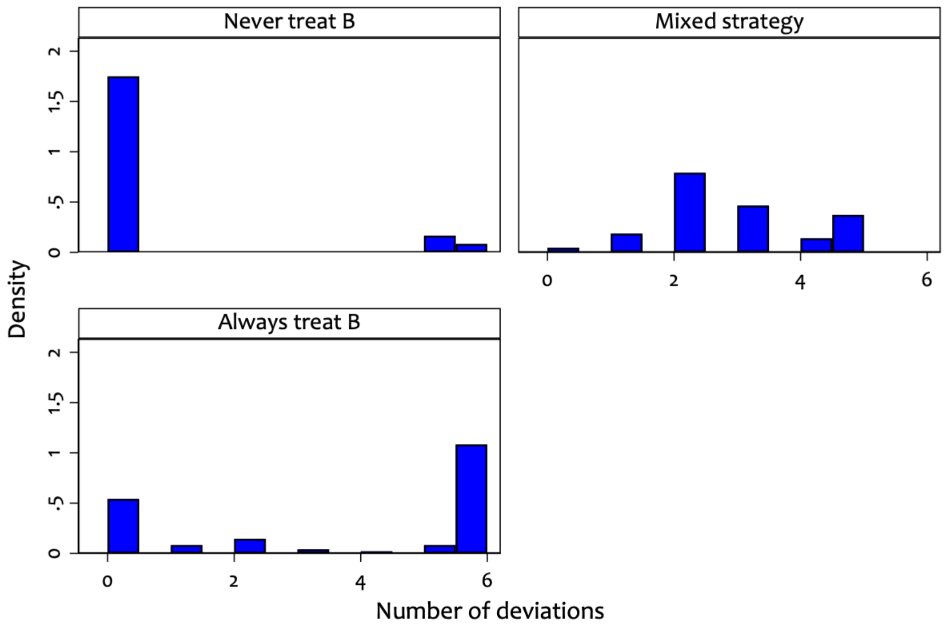
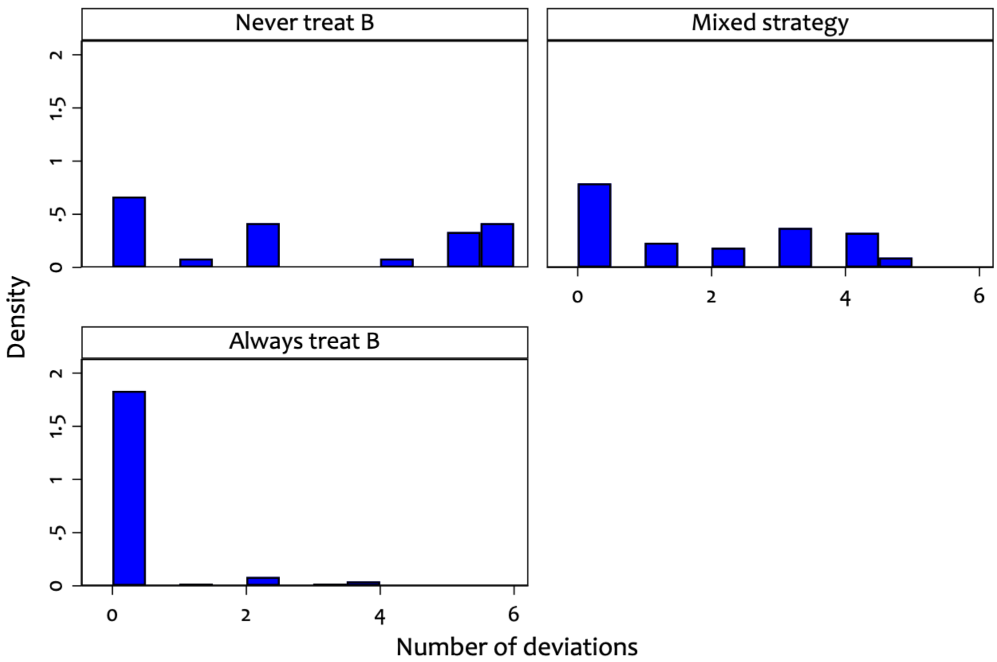


Figure A.3.2 Behavioral Types under EUT



A.6 After 52 Periods

Table A.5: Behavior by Period 52 (After 22 Sickness Cases; 9x A; 13xB)

	Avgerage number of treatment choices	Treatment rate for Sickness A	Treatment rate for Sickness B
LowPrice	18.38 -5.11	0.98	0.73
HighPriceDed	10.46 -3.6	0.94	0.15
LowPriceNoInfo	18.54 -5.28	0.98	0.75
HighPriceNoInfo	11.33 -5.09	0.93	0.23
LowPriceReverse	17.7 -7.08	0.88	0.75
LowPriceNeutral	18.68 -5.51	0.96	0.78

Notes. Treatment rate indicates share of respective sickness cases treated by all subjects.

A.7 Formal Description of Estimation of Individual Risk Preference Parameters

Assume that the utility of income from an experimental lottery choice task is defined by the following constant relative risk aversion (CRRA) specification:

$$U(x) = x^{(1-r)}/(1-r), \tag{1}$$

where x is the lottery prize and r represents a coefficient that indicates the level of constant relative risk aversion. With this specification $r = 0$ describes risk-neutrality, $r < 0$ corresponds to risk-loving preferences, and $r > 0$ corresponds to risk-averse preferences. Given an estimate of r one can calculate the expected utility of a typical lottery i. If lottery i has j possible outcomes, the EU of the lottery is given by

$$EU_i = \sum_j p(x_j)\, U(x_j). \tag{2}$$

Then, for each decision pair an index is calculated that indicates the difference in the expected utility of both lotteries in a decision pair. Formally,

$$\Delta EU = EU_L - EU_R, \tag{3}$$

where EU_L is the "left" lottery and EU_R is the "right" lottery in a decision pair as presented to subjects. The function that links the latent index in (3) to observed choice behavior is the cumulative density function (cdf) of the univariate normal distribution $\Phi(\cdot)$, resulting in a probit model. The probability of choosing the "right" lottery can be written

$$prob(choose\ R) = \Phi(\Delta EU). \tag{4}$$

Thus the latent index in (3) is linked to the observed choices by making the assumption that lottery R is chosen, when the $\Delta EU > 0.5$.

This basic approach can be extended in several ways. An important addition is accounting for behavioral errors. The structural probit model cannot predict individual decision making with certainty. Decision makers may deviate from their true underlying preferences for a variety of reasons. Behavioral error specifications can account for various error sources, ranging from random deviations due to attention lapses to systematic violations related to the psychology of perception and judgment. A particularly influential behavioral error specification is due to Fechner (1860). Its application to the evaluation of risky prospects was popularized by Hey and Orme (1994). The inclusion of the Fechner error specification expands the latent index in (3) to

$$\Delta EU = (EU_L - EU_R)/\mu, \tag{5}$$

where the new parameter μ allows the otherwise deterministic EUT model to account for deviations from the underlying preference structure.

Wilcox (2008, 2011) suggests an additional characterization of behavioral errors, called "contextual utilty." The intuition behind contextual utility originates from psychological experiments on signal detection and stimulus discrimination. These studies discovered that errors became more likely as the range of possible stimuli increase. Contextual utility respects this observation, by assuming that evaluative errors increase with the perceived range of outcomes. Econometrically, this implies that the standard deviation of the behavioral error is proportional to the range of utilities of the outcomes in a lottery pair. The contextual error specification is given by

$$\Delta EU = (EU_L - EU_R/\nu)\,/\mu, \tag{6}$$

where the new parameter ν is defined as the maximum utility over all outcomes minus the minimum utility over all outcomes in the lottery pair, i.e., over the context of that pair. This specification has a normalizing effect on the latent index, which then remains in the unit interval.[1]

Once the parameters of interest are defined, structural estimation can be undertaken using the procedures explained by Andersen, Harrison, Lau and Rutström (2008) and Harrison and Rutström (2008).

People may not necessarily behave as if given probabilities affect their lottery evaluations with objective values. Instead, they may distort these probabilities in their perception – a process that can be described by attaching subjective weights to probabilities. The Rank Dependent Utility (RDU) model, due to Quiggin (1982) derives probability weights from the entire distribution over ranked outcomes, not from individual probabilities and avoids any theoretical violations of first-order stochastic dominance. The resulting decision weights reflect subjective distortions of objective probabilities.

The RDU model nests the EUT model, and requires the introduction of a probability weighting function. A variety of weighting functions have been proposed in the literature, primarily by Quiggin (1982), Tversky and Kahneman (1992) and Prelec (1998). Prelec (1998) contributes a flexible two parameter specification of probability weighting:

$$\omega(p) = \exp[-\eta(-\ln p)^{\phi}], \tag{8}$$

[1] The contextual error specification is particularly parsimonious, since the parameter ν is defined by data, so that no additional parameter estimation is required. The specification also allows for inferences regarding "stochastically more risk averse" relationships. The latter refers to a stochastic notion of the familiar Arrow-Pratt metric of risk aversion. A stochastically risk averse subject is "on average" risk averse, but the metric is flexible enough to deal with choices that deviate from the subject's general risk aversion. With the latent index remaining within the bounds of the unit interval, one can compare the stochastic risk aversion of subjects who choose in dramatically different decision contexts (i.e., who face lotteries with very different prizes).

with $\eta > 0$ and $\phi > 0$. This weighting function is derived from several axioms that reflect apparent regularities of probability weighting, and requires the estimation of two additional parameters η and ϕ.

Another special case of RDU, due to Yaari (1987) and known as Dual Theory (DT), assumes that all of the risk premium is due to probability weighting, and that there is a linear utility function such that $r = 0$ and hence that $U'' = 0$. In this case a risk premium is generated entirely by "pessimistic probability weighting" with respect to better-ranked outcomes.

References

Fechner, G.T. (1860). *Elemente der Psychophysik*. Leipzig, Germany: Breitkopf and Härtel.

Hey, J. D. & Orme, C. (1994). Investigating Generalizations of Expected Utility Theory using Experimental Data. *Econometrica*, 62(6), 1291–1326.

Wilcox, N. T. (2008). Stochastic Models for Binary Discrete Choice under Risk: A Critical Primer and Econometric Comparison. In: J. Cox and G. W. Harrison (Eds), *Risk Aversion in Experiments* (12). Bingley, UK: Emerald, Research in Experimental Economics.

Wilcox, N. T. (2011). 'Stochastically More Risk Averse': A Contextual Theory of Stochastic Discrete Choice Under Risk. *Journal of Econometrics*, 162, 89–104.