SUSTAINABLE
FINANCE
LAB

CHOICE COMPLEXITY, BENCHMARKS AND COSTLY INFORMATION

Sustainable Finance Lab working paper XXX

Job Harms

Stephanie Rosenkranz

Mark Sanders

22-03-2017

Abstract

In this study we investigate how two types of information interventions, providing a benchmark and providing costly information on option ranking, can improve decision-making in complex choices. In our experiment subjects made a series of incentivized choices between four hypothetical financial products with multiple cost components. In the benchmark treatments one product was revealed as the average for all cost components, either in relative or absolute terms. In the costly information treatment subjects were given the option to pay a flat fee in order to have two products revealed as being suboptimal. Our results indicate that benchmarks affect decision quality, but only when presented in relative terms. In addition, we find that the effect of relative benchmarks on decision-quality increases as options become more dissimilar in terms of the number of optimal and suboptimal features. This result suggests that benchmarks make these differences between products more salient. Furthermore, we find that decision-quality is improved by providing costly information, specifically for more similar options. Finally, we find that absolute – but not relative – benchmarks increase demand for costly information. In sum, these results suggest that relative benchmarks can improve decision-making in complex choice environments.

JEL classification: D03, C91, D14, G18

1 Introduction and background

Complex choices are omnipresent. Consumers are faced with more and more products that are complex in themselves and do no longer fit into the straightforward pick-and-pay experience. Their complex nature does not just come from the number of product variations, but their many attributes, usually determined by the underlying pricing models.[1] Mortgage products, financial retirement plans, health insurance policies, or mobile phone subscriptions would typically fall into this category.

Rational choice theory assumes that it is better to have more rather than fewer options. Empirical evidence, however, suggests that people indeed have difficulties in selecting a suitable product when choices are complex.[2] Recent studies found that increasing the size of choice sets reduces the likelihood of making a decision, the quality of that decision (e.g. Besedeš et al., 2012a, 2012b; Heiss et al., 2013; Iyengar et al., 2004), and the satisfaction with the decision (Iyengar & Lepper, 2000; Iyengar & Kamenica, 2006). Both decision quality and satisfaction are moreover negatively affected by the complexity of the choice (Greifeneder et al. 2010).

In this context, decisions concerning complex financial products are of specific interest, as mistakes in such decisions can be very costly for the individual as well as for society. Financial literacy seems to be key to financial well-being: financially literate individuals make fewer mistakes and are in better financial condition than financial illiterates (Lusardi and Mitchell, 2006).[3] However, many individuals generally score

---

[1] Choice complexity is defined as the amount of information a choice involves: a choice between objects with one or two important attributes is simple, whereas a choice between objects for which many attributes are important is complex (Dijksterhuis, et al., 2006).

[2] See Chernev et al. (2015) for an overview on empirical studies. In an experimental study, Besedeš et al. (2012a) find that specifically for older subjects the probability of a person selecting the optimal option declines in the number of options, and that older subjects rely more on suboptimal decision rules.

[3] Many people have limited capacity to interpret numerical information (Kirsch et al. 2002; Reyna and Brainerd, 2007) and low numeracy is shown to be associated with suboptimal financial decisions.

poorly on financial literacy, and it remains an open question what aids to offer these individuals to help them make better decisions. Where increasing financial literacy may be first best, it is also doubtful if this can ever materialize and be sufficient. Experimental research, however, has shown that in redesigning the choice architecture, we can significantly reduce decision complexity and improve decision making for given levels of financial literacy. The present paper provides one of the first investigations of the effectiveness of two specific decision aids.

The financial crisis sparked an intense policy debate on how to better enable consumers to make informed decisions in the domain of financial products.[4] Proposed interventions range from improving the comparability of options, e.g. by imposing disclosure requirements on issuers of financial products, or the obligation of issuers to offer a standardized product, to the provision and the incentivization of the use of financial guidance services and stricter monitoring of the advisory process (Chater et al., 2010, Barr et al., 2008).

Regarding the comparability of products, in 2016 the European Commission suggested in a Green Paper the creation of standardized pan-European personal pension and life insurance products. Since 2014 also in the Netherlands the potential effects of financial standard products on the choice process of consumers was repeatedly debated.[5] In the UK the government announced in July 2010 that it wanted to see a new range of simple financial products to help people take responsibility for their finances.[6] Regarding financial guidance, recently, the European Commission has renewed its commitment to

---

[4] While the crisis lead to several changes in financial regulation worldwide, (e.g. in Europe the "Markets in Financial Instruments Directive (MiFID 1 and 2)") aiming at making financial markets more efficient, resilient and transparent, and to strengthen the protection of investors, we here focus only on those initiatives that specifically aim at improving consumer decision making.

[5] See https://www.rijksoverheid.nl/documenten/kamerstukken/2015/12/15/kamerbrief-onderzoeksopzet-standaardproducten.

[6] See https://www.gov.uk/government/consultations/simple-financial-products.

empower consumers of retail financial services to make informed choices by ensuring the provision of advice.[7] Similarly, in the UK in 2008 the "Thorensen review" suggested a national approach to providing information in the form of generic financial advice as a key instrument to improve individual financial decision making.[8]

However, empirical evidence on the effectiveness of such decision aids on decision making in complex choice situations remains scarce. With the present paper we therefore focus on the two debated instruments, the provision of a benchmark product, and the provision of advice. To avoid strong assumptions about the relevant choice set and the nature of preferences we conduct incentivized laboratory experiments, which allows us to control costs and benefits of individual decisions as well as choice complexity, and at the same time isolate the effects of the instruments on the quality of the individual decision.

In our individual choice task subjects make a sequence of 10 choices among four different multi-attribute options, one of which being objectively optimal. Choices are incentivized and framed as a choice among financial products, and are made under a time constraint to mimic complexity.[9] *One treatment variable is the provision of a*

---

[7] See the recent "Study on access to comprehensive financial guidance for consumers" commissioned by the European Commission to identify best practice regarding both the provision and the incentivization of the use of financial guidance services, retrievable at http://ec.europa.eu/finance/finservices-retail/docs/fsug/papers/1611-study-financial-guidance_en.pdf.

[8] See http://webarchive.nationalarchives.gov.uk/+/http:/www.hm-treasury.gov.uk/media/8/3/thoresenreview_final.pdf.

[9] The ranking of options does not depend on subjects' risk preferences and but on the assumption that they are money maximizers. The full choice set is clearly defined, as is the value of each option. While the optimal option is always unique, its identity is concealed from subjects by manipulating the value of 5 attributes that need to be combined to calculate the value of each option. Complex decisions are typically made without a time constraint, but are characterized by a large number of alternatives, parameters, variables, and uncertainties, such that it is usually difficult to make the optimal decision. Setting a time constraint makes it difficult for subjects to make the calculations that are needed to find the optimal option with certainty. Payne,

benchmark. In the benchmark treatments we label one choice option as the "average product", for which we use two different frames: in one frame the attributes for each option are presented as relative deviations from the average product, of which all attributes are set to 100. In the other frame attributes are presented in absolute positive or negative deviations from the average product. In all treatments the values of the attributes are randomly varied for all financial products except the benchmark to be able to control for product similarity.

Providing a benchmark option with average values for each attribute, changes the presentation of decision-relevant information without altering the actual choice set. With this treatment we target the presentation of the choice attributes, by explicitly providing a benchmark as a reference for comparison.[10] Rational decision making in complex choices over options with multiple attributes requires a comprehensive evaluation strategy. Often a high value for one attribute can compensate for a low value for another, as for example in the case of a phone subscription that consists of various costs components such as the co-payment for the phone, fixed monthly costs, flexible monthly costs, etc. In the face of constraints to time and cognitive capacity, decision-makers might benefit from a benchmark as it makes it easier to compare available options. The rationale for this approach is that the decision-maker is steered towards using an elimination-by-aspects strategy. Someone using this strategy first decides what aspect is most important, establishes a cut-off level, and then eliminates all alternatives that do not meet this cut-off level. This process may be repeated, attribute by attribute, until either a choice is made or the set is sufficiently narrowed down to switch over to the evaluation of the remaining options. Alternatively, a decision-maker may be triggered to

---

Bettman and Johnson (1988) and Ben Zur and Breznitz (1981) show that decision-making under time pressure indeed leads to suboptimal decision making.

[10] Note that we do not claim that the benchmark product serves as a reference point in the sense of Kahneman and Tverksy (1982). Our benchmark product is part of the choice set (which does not need to be the case for a reference point) and does not necessarily determine clear domains of gains and losses.

substitute a comprehensive evaluation strategy by simpler alternative strategies (weighted additive, satisficing, lexicographic), or by (e.g. fast and frugal) heuristics relying on only a few comparisons.[11] This strategy may be particularly effective for subjects who are less financially literate, i.e. individuals with higher cognitive costs. In line with the notion that presentation of information influences how easily it can be processed (Bettman and Kakkar, 1977), we conjecture that the relative benchmark treatment will be more effective, as it enables the decision-maker to see directly how the different attributes of the various products compare to their respective market average, without having to compute these averages themselves. Finally, the effect of benchmarks on decision making may even be negative: Charter et al. (2010) argue that a significant group of consumers does not actively engage in product search but simply chooses a default if it is available.[12] It is thus not clear ex-ante whether benchmarks will improve decision-making in complex choices.

Besedeš et al. (2015) study two other forms of choice architectures, approaches that - in the spirit of "choice architecture" - assist decision makers without reducing the choice set (Thaler & Sunstein, 2003), which both reduce a large decision problem into a series of smaller ones. The authors show that a tournament-style choice architecture, in which a large choice set is broken down in several smaller choice sets from which

---

[11] Regarding strategies for choice with multiple attributes see e.g. Thaler et al. (2014) or Gigerenzer and Todd (1999), or see Payne, Bettman, and Johnson (1988) for an overview.

[12] Whether choosing the benchmark improves decision making depends on the specific situation. In our experiment such behavior will lead to suboptimal choices as the benchmark consists of the average values for all attributes. Given the cost-function, such an average product is rarely the optimal choice. Our product attributes are different types of costs and were randomly generated in every decision. The chance that the average over all randomly generated attributes is optimal approaches zero as the average is always higher than the minimum. We did not eliminate the possibility of attributes exactly compensating such that the average on all attributes could still be the best choice out of four. In our data we found that the benchmark option was never the cheapest, in 35% of the tasks it was the second cheapest option, in 64% of tasks it was the second most expensive option and in 1% of tasks it was the most expensive option (STATA command: tab Rank_aver)

respective optimal options are selected into a final choice set, reduces choice overload and thereby improves decision making. When offering subjects the choice between the different choice architectures, Besedeš et al. (2015) find that subjects' preferences for choice architectures are negatively correlated with performance, suggesting that providing choice over architectures might reduce the quality of decisions. Even when controlling for effort costs, the authors cannot exclude that this problematic result is related to self-sorting. Petes et al. (2009) show that difficult-to-evaluate attributes are more accounted for by decision makers when graphical decision aids make it easier for the decision maker to map these attributes on a good/bad scale. In a similar vein, Agnew and Szykman (2010) find that subjects with higher levels of financial knowledge are less likely to suffer from choice overload when information about various products is presented in a table format than when it is presented in a booklet format. Soll et al (2013) investigate the effect of the so-called "Credit Card Accountability Responsibility and Disclosure" (CARD) act, which forces providers to issue information to make the relationship between attributes (monthly repayments and total repayment periods of credit cards) easier to understand. In an online survey experiment the authors find that this information indeed improves decision quality (it reduced people's bias towards underestimating the duration of repayment periods).

We contribute to this literature on decision aids with our finding that benchmark products promote decision-quality, but only when attribute values of the products are expressed in relative terms. Furthermore, we find that this effect of a relative benchmark is greater for choices in which the options are more dissimilar in terms of the number of above- and below-average attributes. These results suggest that the relative benchmark treatment improved decision making by making the dissimilarity between products more salient and thereby enabling subjects to more easily select the optimal product. In addition we find evidence of a learning effect: the positive effects of the relative benchmark treatment are greatest in later rounds of the experiment.

Our second treatment variable is the provision of (costly) information. In a set of treatments we offer our subjects the possibility to buy information (which we label "advice") about the relative profitability of various options that can be chosen. When advice is bought, the computer truthfully marks two of the suboptimal options, the worst and randomly one of the other non-optimal options, in the set of four. This information is costly and can be ignored. When the decision is taken randomly, the two scenarios with and without costly information differ only in the variance not in the expected value of the consequence of the decision, with the variance when information is acquired being lower.

Providing the possibility to discard suboptimal options from the choice set reduces the cognitive burden of the decision maker because fewer options need to be evaluated. With this treatment we target the structure of the choice task as we vary the number of relevant alternatives. However, we present the decision maker with the possibility to receive additional information on suboptimal options at the expense of the benefits of selecting the optimal option. The effectiveness of this approach relies on the assumption that people who request the additional information are benefited by it, and those who do not are benefited by the original choice set.[13] A rational decision-maker following a comprehensive evaluation strategy should only make use of this option when the reduction in marginal cognitive costs (due to reduced difficulty of the decision task) overcompensates the reduction in marginal benefits. However, the more such a maximizing strategy proves impossible, e.g. for subjects who are less financially literate or who are susceptible to choice overload and thus have higher marginal cognitive costs, the more attractive becomes the information that implicitly reduces the choice set, as also the chance to choose the optimal option increases.

Our paper thus also contributes to the literature on the effects of advice on decision quality. In an experimental study Gino and Moore (2007) show that costly advice is

---

[13] Note that providing costless information regarding the ranking of options, i.e. without affecting consequences, would simplify the decision problem for all decision makers in a trivial way without allowing us to differentiate between decision makers applying different choice strategies.

overweighed in complex choice tasks but underweighted in simple choice tasks. In a hypothetical choice experiment Hung and Yoong (2010) compare the effect of unsolicited and solicited advice. They find that unsolicited advice does not affect investment behavior, but when advice is optional, individuals with low financial literacy are more likely to ask for it. The authors also find that notwithstanding this negative selection on ability, individuals who actively solicit advice indeed make better choices. Hackethal, Haliassos, and Jappelli (2012) find that self-selection largely explains their finding of better outcomes for advisees in the context of German Internet brokerage accounts. A robust finding in the literature is that individuals who receive advice by default tend to significantly discount it (Bonaccio and Dalal, 2006; Yaniv 2004a, 2004b; Yaniv and Kleinberger, 2000). While advice that is explicitly solicited is perceived as helpful, unsolicited advice is perceived as intrusive and might even lead to worse decisions (Deelstra, 2003; Goldsmith, 2000; Goldsmith and Fitch, 1997). Gino (2008) shows that individuals are more likely to use decision-related information when they pay for this information if compared to when they get it for free.

We find that offering costly information positively affects decision making. Not only does it result in more optimal decisions and fewer suboptimal decisions, it also leads to a higher payoff even net of the advice cost. Furthermore, we find that subjects benefit more from the option to buy advice when the options are more similar, hence when simplifying the choice set to two options and thus reducing the cognitive effort needed for optimal decision making has greater marginal effect.

We conduct the experiment with a 3 × 2 factorial design (relative, absolute, no benchmark × advice, no advice), allowing us to study possible interaction effects of benchmarks and advice. We are especially interested in understanding whether these two approaches indeed improve decision making in complex choices, and whether the instruments are complementary or rather substitutable. Thus, our study also contributes to the literature on demand for financial advice. Various studies show that demand for financial advice, such as consultation of a bank advisor, is positively correlated to individuals’ level of financial literacy, even when controlling for income and education levels (Calcagno and Monticone, 2015; Hackethal, Haliassos, and Jappelli 2012; Collins, 2012). One interpretation for this result is that individuals with higher financial literacy better understand the potential benefits of seeking costly advice, for example to avoid even higher costs resulting from poor financial decisions (Robb et al., 2012). Gino and Moore (2007) find that subjects are not more likely to seek even costless advice in a difficult version of the task compared to an easy version of the task.

We find no significant interaction effects between the benchmark and the option to buy advice. Moreover, benchmarks do not affect demand for costly information, regardless of whether the benchmark is presented in relative or absolute terms. The latter is relevant in the light of the debate on the role of the advisory process and its interaction with other instruments.

This paper proceeds as follows. Section 2 describes the experimental design, and procedures. Section 3 develops the hypotheses we test in our study and our empirical strategy. Our results are presented in Section 4 and we discuss our findings and conclude in Section 5.

2 Experimental design and procedures

2.1 Design

The experiment consisted of three parts. The first part of the experiment was an individual choice task, where subjects made choices among different options, which were framed as financial products. Subjects were presented a table with four different options, labelled product A, B, C and D, and were instructed and incentivized to select the option with the lowest total ‘costs’. To mimic decision making for complex financial products, the total costs of the products were not explicitly given, but the subjects were presented with five different cost elements, framed as costs and tax deduction, of each product.

Subjects were informed that the product had a maturity of one year (12 months) and that it was the subjects' task to determine which product has the lowest total costs. The optimal product could be calculated using the formula:

Total cost of product = start costs + (12 x monthly costs) + maturity costs + management fee (percentage of start costs) - tax deduction (percentage of monthly cost)

This formula was not given explicitly, but in the instructions all cost elements and their influence on the total costs were carefully explained (see the instructions in the online appendix). The values for the cost elements were randomly generated for three of the products while for the fourth product the cost elements were calculated as the average of the three randomly generated others. The intervals in which the cost elements were randomly varied were displayed on a whiteboard in the room. This design allows for an objective evaluation and ranking of options, independent of subjects' tastes and risk preferences as long as subjects are not satiated in money (see Besedeš et al., 2012a, 2012b). The products and an example of cost elements were presented as in Table 1.

Table 1: Example Payoff Matrix

	Product			
	A	B	C	D
Starting costs	87	92	103	94
Monthly costs	35	49	64	49
Maturity costs	72	91	2	52
Management fee (%)	15	31	16	21
Tax deduction (%)	11	10	10	10

The task of choosing the optimal product was repeated ten times, and each time the cost elements for three of the products were randomly determined. The position of the average product within the table was randomly assigned by the computer every round. The treatment to which subjects were assigned determined whether they were informed about the existence of this average product or not. Subjects had to perform the task of choosing the optimal product within 30-seconds, which were presented by a timer counting down in the upper right corner of the screen. This time limit was introduced to simulate the complexity of the financial product choice. If a subject did not choose a product within these 30 seconds, the computer automatically implemented the product with the highest total costs as the subject's choice.

As treatments, we varied the information regarding the benchmark option, as well as the costly information regarding the ranking of the options, in a 3 × 2 factorial design. All treatments were employed in a between-subject design. Regarding the benchmark we varied the information subjects had on the existence of the average product as well as the presentation of this average product. In the control treatment, subjects were not informed that the cost elements of one of the options was the respective average of the other three options. In our benchmark treatments, we provided subjects with a reference product by informing them about the fact that one product is an average product in all cost elements. This average product was indicated by a blue font and it was also explicitly stated that the respective product is an average product. In one set of treatments the average product was represented in absolute values (ABSOLUTE) while in another treatments (RELATIVE) the cost elements of all other products were presented in percentages relative to the average product. The average product thus had a value of 100 for every cost element, indicating that all other values are relative to it.

We varied the information regarding the ranking of the options by giving subjects in one set of treatments (ADVICE) the opportunity to have the worst option and another randomly chosen suboptimal option indicated by the computer before making their choice, against the payment of a fixed price. While this potentially left the subject with only two options, the advice could be ignored and the indicated suboptimal options could still be chosen. We framed this decision as 'buying advice' and the price for this information was set such that the decision was only profitable if indeed the optimal option was selected.

Payoffs are shown in Table 2. The subjects were rewarded based on how well they made their decisions. The options were ranked in order from optimal (cheapest) to worst (most expensive). Payoffs were directly linked to whether the optimal, second best, second worst or worst product was chosen. Subjects started with an initial endowment of €8 and the payment was added or deducted from this amount depending on the choices the subject made. Choosing the second best option after buying advice led to a payoff of zero, while choosing a worse product led to deductions from the endowment. Note that it was made clear that the worst option and one of the remaining suboptimal options would be indicated by the advice, such that it was not clear to subjects if the remaining worst option would be the second or third best option. For purely random choices the expected value of buying advice would then be 0.5*5+0,5*(0,5*-2.5+0,5*2.5) – 2.5 = 0, i.e. identical to the expected value of not buying advice.[14]

Table 2: Payoffs

Choice	No advice	Advice (cost=2.5)
Optimal	5	2.5
2nd best	2.5	0
3rd best	-2.5	-5
Worst	-5	-7.5

In the end, one of the situations was randomly drawn and the actual decision that was made in this round was realized to establish the actual payoff. Subjects earned on average €9.16 in this part of the experiment.

---

[14] Note that a risk-averse decision-maker employing such a random choice strategy should acquire information as this reduces the variance of the expected outcome.

In the second part of the experiment, to assess participants' risk aversion, a sequence of binary lottery decisions was administered, which was equivalent to the one introduced by Holt & Laury (2002). The task asked the subjects to choose between two lotteries in ten different cases. The lottery choice screen is shown in Appendix 2. The amounts that could be won did not change in the ten lotteries, only the probability of occurrence of each amount changed. One lottery paid either €3 or €0 and the other paid either €1.50 or €1. The switching point from one lottery to the other is the crucial point that reveals the individual's risk aversion. From Holt and Laury (2002) it is known that even when people switch back and forth between lotteries, the number of safe options gives a good indication of the subject's level of risk aversion. After the ten decisions were made, one of the ten lotteries was chosen randomly by the computer and subsequently the lottery was played to determine the payoff for the subject in this round. The payoff received in this part of the experiment was added to the amount the subject earned in the first part. Subjects earned on average €1.74 in the lottery part.

It is important to clarify some aspects of the experimental design. Our first remark concerns the fact that we deliberately abstracted from heterogeneity in subjects' tastes. Of course this limits the external validity of our findings. However, this allowed us to implement a benchmark with an objective value that is independent of subjects' tastes and therefore increase internal validity by abstracting from subjective beliefs. The same argument holds for the absence of risk in the choices in our treatments. The possibility to objectively rank options allows us to observe decision quality without relying on extra measures of risk preferences. Our second remark concerns the relation between the benchmarks as we implemented them in our experiment and the real-world examples discussed in the introduction. A benchmark that corresponds to a product that suits the preferences of an average consumer may be more realistic. However, in our simplified choice situation with homogeneous tastes such a benchmark could obviously not be implemented. Nevertheless, we think it is important to study the psychological effects of the presence of a benchmark product in the most simple setting first, before moving on to more sophisticated and more realistic scenarios. Our third remark concerns the time constraint. In the real world, complex decisions often have no time constraint, but rather require the assessment of uncertain absolute and relative merits of multiple attributes of available options. In order to be able to objectively determine decision quality in our lab setting we presented the options with attributes that were relatively simple to assess. However, adding a time constraint made it difficult for subjects to process all relevant information and possibly forced them to concentrate on the most important attributes. This evaluation strategy is similar to strategies employed in real complex choices (see also Payne, Bettman, & Johnson, 1988).

The experiment ended with a questionnaire on demographics, buying behavior, self-reported personality traits (BFI list) and attitudes regarding financial products (AFM, 2014). See Appendix 3 for these survey questions. The survey was not incentivized but anonymity was guaranteed.

2.2 Experimental procedures

The experiments were conducted in the Experimental Laboratory for Sociology and Economics (ELSE) at Utrecht University. They were programmed and conducted with the experimental software ‘z-Tree’ developed by Fischbacher (2007). In seven sessions, a total of 158 subjects (average of 23 subjects per session) participated in the experiment.

The subjects were mainly undergraduate students from various fields at Utrecht University or Hogeschool Utrecht. Over 1000 potential subjects from the pool of the ELSE lab were approached by email to participate in the experiment, using the ORSEE recruitment system (Greiner, 2004). Upon arrival in the lab, subjects were randomly assigned a seat behind a computer. Subjects were randomly divided into 3 groups, where each of the three groups played a different treatment[15]. Treatments with and

---

[15] For groups 1,3 and 5 some an IT-related problem caused in a reduced number of tasks being implemented in periods 9-10 of the experiment (10 missing obs./group for period 9 and 20 missing obs. for period 10). We have controlled for these missing observations by testing our regression models both with and without

without the option to buy advice were administered in different sessions but on the same days in alternating order. This allows for comparison of the differences between the treatments and control for all other factors as these stay the same between treatments. Before the start of every experiment, general written instructions in English were given, which were kept identical across sessions (see the online appendix). Additional instructions were displayed on the screen. The first part of the experiment started when all subjects had fully read and understood the instructions. One full experimental session lasted on average 45 minutes and subjects earned an average of €10.90.

---

these periods and the results do not change significantly, We thus report the estimates for the large dataset including rounds 9-10.

3 Hypotheses and estimation strategy

3.1 Hypotheses

Traditional information-based interventions are based on the reasoning that the provision of information, for example information about the cost structure of all available products in the market, leads to objectively better decisions, such as the selection of the cheapest option. Information is assumed to lead to the explicit appraisal of costs and benefits related to different decisions and ultimately to changes in such behaviors. The predicted impact of information on behavior is consistent with the standard economic model that assumes boundless rationality: more information enables individuals to more accurately calculate the payoffs for each decision, taking into account that the effect of information may depend on cognitive ability and domain-specific expertise.

We therefore hypothesize that providing a benchmark product makes selection through attribute comparisons easier. We thus expect a positive effect on decision making:

> *Hypothesis 1:* In the treatments with absolute and relative benchmarks, the number of optimal decisions is higher, the number of suboptimal decisions is lower, and payoffs are higher than in the treatments without a benchmark.

The comparison to the benchmark is facilitated even more when all choice options are presented in *relative* values compared to the "average product" rather than in *absolute* values. We thus expect the positive effects on decision making to be stronger for the relative than for the absolute benchmark:

> *Hypothesis 2:* In the treatments with a relative benchmark, the number of optimal decisions is higher, the number of suboptimal decisions is lower, and payoffs are higher than in the treatments with an absolute benchmark.

With respect to the effect of costly information, the existing literature does not provide us with a clear direction of the expected effect. If advice is bought by only those subjects who benefit from it because they find the task too difficult, we expect the option to buy advice to have a positive effect on decision making:

> *Hypothesis 3:* In the treatments with the possibility to buy additional information (advice), the number of optimal decisions is higher, the number of suboptimal decisions is lower, and payoffs are higher than in the treatments without this option.

Also regarding the interaction effect of benchmark products and costly information, the existing literature does not provide us with a clear direction of the expected effect. However, if a benchmark simplifies decision making as hypothesized above and advice is costly, then one would expect advice to be less valuable in such simplified choice tasks.

> *Hypothesis 4:* In the treatments with absolute and relative benchmarks there is lower demand for costly information than in treatments without a benchmark.

3.3 Estimation strategy

We first test how payoffs are affected by the five respective treatment combinations (i) relative benchmark, (ii) absolute benchmark, (iii) advice, (iv) relative benchmark X advice, (v) absolute benchmark X advice. To this purpose we first estimate the following OLS model:

$$Y_{ij} = \beta_0 + \beta_1 BM_{ij} + \beta_2 A_{ij} + \beta_3 BM_{ij} * A_{ij} + \varepsilon_{ij} \qquad (2.1)$$

in which $Y_{ij,}$ represents the payoff individual *i* in decision-task *j*, BM_{ij} is a dummy for the benchmark treatment, A_{ij} is a dummy for the advice treatment, $BM_{ij}*A_{ij}$ is the interaction of the benchmark and advice treatment, ε are unobserved factors and $\beta_{0,}$ $\beta_{1,}$ β_2 are parameters to be estimated. Estimates of the parameter β_2 can be interpreted as the causal effect of benchmark treatment, which is randomly assigned to individuals. Hence, it is unlikely that unobserved factors will be correlated with BM. To account for the grouping of observations over subjects we cluster standard errors on subject level.

Next, we control for the possibility of non-random assignment of treatments to subjects by including a vector of subject-level control variables (C_{ij}): age, gender, self-reported preferences with respect to financial products and self-reported psychological traits form the survey and risk aversion as measured with the Holt-Laury lottery task.

$$Y_{ij} = \beta_0 + \beta_1 BM_{ij} + \beta_2 A_{ij} + \beta_3 BM_{ij} * A_{ij} + \beta_4 C_{ij} + \varepsilon_{ij} \qquad (2.2)$$

Subsequently, we investigate whether subjects require some practice to learn how to benefit from the benchmark. To this purpose, we add to the specification a dummy variable "post-round 6" ($R6_j$) with value equal to one if the decision-task j was after the sixth round of the experiment. We construct an interaction variable of this dummy and the benchmark treatment, "$BM_{ij}*R6_j$" to test whether the effect of the treatment differs between the earlier and later rounds of the experiments. Estimates of the parameter β_6 can be interpreted as the learning effect for the benchmark treatment.

$$Y_{ij} = \beta_0 + \beta_1 BM_{ij} + \beta_2 A_{ij} + \beta_3 BM_{ij}*A_{ij} + \beta_4 C_{ij} + \beta_5 R6_j + \beta_6 BM_{ij}*R6_j + \varepsilon_{ij} \qquad (2.3)$$

Next, we investigate whether the similarity of the various options moderates the effect of the benchmark treatment. To this purpose we construct an "attractiveness score" for each product, which is a function of the number of attributes with below average cost minus the number of attributes with above average cost. We then compute the between-product standard deviation of this score to construct a variable to proxy for product dissimilarity; this variable is called "PDS". We interact this variable with the benchmark treatment to construct the variable "PDS*BM".

$$Y_{ij} = \beta_0 + \beta_1 BM_{ij} + \beta_2 A_{ij} + \beta_3 BM_{ij}*A_{ij} + \beta_4 C_{ij} + \beta_5 R6_j + \beta_6 BM_{ij}*R6_j + \beta_7 PDS_{ij} + \beta_8 BM_{ij}*PDS_{ij} + \varepsilon_{ij} \qquad (2.4)$$

Having tested the effect of the treatment on payoffs, we then turn to exploring how the treatments affect the actual decision. To this purpose, we apply a probit model following specification (2.4) on the following dependent variables:

Best choice: Whether the subject selects the most optimal, i.e. least expensive, product

$$Y(opt)_{ij} = \beta_0 + \beta_1 BM_{ij} + \beta_2 A_{ij} + \beta_3 BM_{ij} * A_{ij} + \beta_4 C_{ij} + \beta_5 R6_j + \beta_6 BM_{ij} * R6_{ij} + \beta_7 PDS_{ij} + \beta_8 BM_{ij} * PDS_{ij} + \varepsilon_{ij} \qquad (2.5)$$

Worst choice: Whether the subject selects the least optimal, i.e. most expensive, product

$$Y(subopt)_{ij} = \beta_0 + \beta_1 BM_{ij} + \beta_2 A_{ij} + \beta_3 BM_{ij} * A_{ij} + \beta_4 C_{ij} + \beta_5 R6_j + \beta_6 BM_{ij} * R6_{ij} + \beta_7 PDS_{ij} + \beta_8 BM_{ij} * PDS_{ij} + \varepsilon_{ij} \qquad (2.6)$$

Choice: Whether the subjects selected any product at all within the allocated 30 seconds.

$$Y(choice)_{ij} = \beta_0 + \beta_1 BM_{ij} + \beta_2 A_{ij} + \beta_3 BM_{ij} * A_{ij} + \beta_4 C_{ij} + \beta_5 R6_j + \beta_6 BM_{ij} * R6_{ij} + \beta_7 PDS_{ij} + \beta_8 BM_{ij} * PDS_{ij} + \varepsilon_{ij} \qquad (2.7)$$

4 Results

4.1 Descriptive statistics

Table 3 describes the demographics of the sample across all sessions and treatments. The average age in our sample is 23, and 60% of the subjects are female. Subjects were from various academic backgrounds ranging from natural science to social sciences & humanities (not shown). Furthermore, there are no statistically significant differences in terms of age and sex, implying that randomization was successful.

Table 3: Demographics by treatment

Group	# obs.	Age (yr)	Female %
Control	27	23.03	0.60
Advice	25	22.40	0.60
Benchmark, absolute	27	23.17	0.52
Benchmark, absolute + Advice	25	23.04	0.56
Benchmark, relative	28	24.08	0.68
Benchmark, relative + Advice	26	22.92	0.65
Total	158	23.10	0.60

Table 4 provides a summary of choices and performance per treatment. The column “Decision time” indicates the average decision time in seconds, and the column “In-time” indicates the fraction of subjects who successfully made any decision within the allocated 30 seconds. The Column "Payoff" indicates the average payoff for subjects, and finally the Column "Net payoff" indicates the average payoffs corrected for the cost of buying advice. The Column "Best choice" indicates the percentage of subjects that chose the financial product with the lowest costs. The Column "Worst choice" indicates the percentage of subjects that chose the financial product with the highest costs. As can be seen for the groups in the non-advice treatment (Rows 1, 3 and 5) the payoff is equal to the net payoff, because no advice was available.

Table 4: Decisions and payoffs by treatment

Group	Obs.	Decision time	In-time	Payoff	Payoff net	Best choice	Worst choice
Control	240	22.58	0.82	1.13	1.13	0.53	0.32
Advice	250	22.93	0.92	3.52	3.07	0.73	0.06
Absolute benchmark (ABM)	238	22.69	0.84	1.99	1.99	0.53	0.21
ABM + Advice	250	22.58	0.94	3.14	2.45	0.69	0.09
Relative benchmark (RBM)	250	21.67	0.87	1.65	1.65	0.52	0.22
RBM + Advice	260	20.33	0.97	3.08	2.45	0.65	0.08
Total	1488	22.11	0.11	2.48	2.16	0.61	0.16

We see several results in Table 4. First, the advice treatment is associated with improved decisions, as indicated by the high fraction of "best choices" and lower fraction of "worst choices" compared to the non-advice treatment groups. Second, the improved decision-making in the advice treatment also results in higher payoffs, even if corrected for the cost of buying advice, as can be seen by comparing payoffs and net payoffs for the treatment group with the advice-only group (second row). Third, we find that – in comparison to the control group - both the absolute benchmark (ABM) and relative benchmark treatments are associated with higher payoffs, with a corresponding reduction in the fraction of worst choices. Fourth, we observe that both payoffs and the fraction of optimal choices is higher in the advice treatment than in the benchmark treatments. Finally, we observe that the combined advice and benchmark treatments (ABM and RBM) are associated with somewhat lower payoffs, fewer optimal decisions and more worst decisions than the advice treatment. In the next section we test the statistical significance of these differences.

4.2 Main estimation results

We now turn to the central question of this paper: whether benchmarks and costly information lead to improved decision-making in a complex choice environment. Hypothesis 1 states that the benchmark treatments lead to improved decision making and higher payoffs.

Relative benchmarks

We first consider the *relative* benchmark treatment. A Mann-Whitney U-test shows that this treatment is associated with higher payoffs than the control group, but this difference is only marginally significant (z=-1.63, Pr.>|z|= 0.1017, N=490). We find that this difference in payoffs is mainly driven by a reduction in the probability of subjects choosing the worst option (z=2.770, Pr.>|z|=0.0056, N=490) whereas there is no significant difference between the groups in terms of the probability of choosing the optimal product (z=-0.415, Pr.>|z|=0.6778, N=490).

We then turn to a series of regression analyses presented in Table 5. As per specification (1) we test the effect of the relative benchmark and its interaction with the advice treatments. We find that the relative benchmark has no significant effect on payoffs, whereas the advice treatment has a significant and positive effect. In addition we find that the benchmark and advice treatments do not interact, as indicated by the coefficient on the term "RBM*advice". These results are robust to controlling for demographics, risk-attitudes and self-reported preferences as per specification (2), indicated in the second column.

Subsequently, we investigate whether treatment effects increase as the experiment progresses. As per specification (3) we find that indeed subjects benefit more from the relative benchmark in later rounds of the experiments, as can be seen by the positive coefficient on the interaction term "RBM*post-period 6" in Column 4.

In the next specification we investigate whether the treatment effect depends on the similarity of the products. As per specification (4) in the fourth column of Table 5 we find that this interaction effect – indicated by the term "RBM*product dissimilarity" is positive and statistically significant, indicating that the benefit from relative benchmark treatment on payoffs increases as products become more dissimilar.

Table 5: Regression table of effect of relative benchmarks on payoffs

	(1) Payoff	(2) Payoff	(3) Payoff	(4) Payoff
Relative Benchmark (RBM)	0.522	0.705	0.346	-0.299
	(0.916)	(0.828)	(0.832)	(0.879)
Advice	2.386***	2.079**	2.060**	2.025**
	(0.845)	(0.826)	(0.821)	(0.824)
RBM*advice	-0.960	-0.955	-1.022	-1.126
	(1.045)	(1.159)	(1.152)	(1.143)
Post-Period 6 (PP6)			0.198	0.201
			(0.267)	(0.265)
RBM*PP6			1.053***	1.038***
			(0.356)	(0.357)
Product dissimilarity				0.195
				(0.136)
RBM*product dissimilarity				0.384**
				(0.181)
Controls for age, sex, self-reported preferences, risk attitude, decision time	No	Yes	Yes	Yes
Constant	1.129	-6.690	-6.797	-6.929
	(0.742)	(5.154)	(5.118)	(5.106)
Observations	895	895	895	895
R-squared	0.063	0.146	0.159	0.174

Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

We then investigate the mechanisms by which the relative benchmark influences payoffs. In particular, we are interested whether the observed effects on payoffs result from more optimal choices, fewer suboptimal choices, fewer instances of subjects not choosing at all within the allocated time, or a combination thereof. We use the fourth specification from the previous model, and using a probit model we regress this on the probability that subjects selected: (i) the optimal product, i.e. the product with lowest total costs, (ii) the worst product, i.e. the product the highest total costs and (iii) the probability that subjects did not select any product within the 30 seconds that were allocated.

Table 6: Regression table of effect of relative benchmarks on decision-making

	(1) Best choice	(2) Worst choice	(3) No decision
Relative Benchmark (RBM)	-0.084	0.021	-0.022
	(0.088)	(0.058)	(0.046)
Advice	0.185**	-0.188***	-0.071
	(0.083)	(0.066)	(0.043)
RBM*advice	-0.114	0.090	-0.033
	(0.115)	(0.092)	(0.051)
Post-Period 6 (PP6)	0.023	0.015	-0.066***
	(0.039)	(0.025)	(0.020)
RBM*PP6	0.048	-0.137***	0.010
	(0.051)	(0.036)	(0.030)
Product dissimilarity	0.025	-0.019	-0.030***
	(0.019)	(0.012)	(0.011)
RBM*product dissimilarity	0.059**	-0.021	0.010

	(0.027)	(0.017)	(0.015)
Controls for age, sex, self-reported preferences, risk attitude, decision time	Yes	Yes	Yes
Observations	895	895	990
Pseudo R2	0.1087	0.225	0.2317

Marginal effects of probit regression. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

The first two regression analyses shown in Table 6 indicate that the relative benchmark has no significant effect on the probability of subjects choosing an optimal or worst product, as can be seen by the insignificant effect of the first term "relative benchmarks" in the first two model specifications. However, we find that the relative benchmarks interacts with between-product variation in the first specification, indicating that the probability of subjects selecting the optimal product increased as the products in the choice set become more dissimilar in terms of the proxy for similarity presented in the previous section. Taken together these results suggest that the relative benchmark treatment mainly improved decision making by making dissimilarity between products more salient and thereby making it easier for subjects to select the optimal product.

The third column in Table 6 indicates that the relative benchmark treatment did not have a significant effect on the probability of not making a decision at all within the allocated 30 seconds. We do find that both learning effects and between-product dissimilarity reduce the probability of subjects not choosing in time.

In sum, these results indicate that the main mechanism through which relative reference points improve decision-making in complex choices is by making between-product dissimilarity more salient, thereby increasing the fraction of optimal choices.

Absolute benchmarks

We then consider the effect of the *absolute* benchmark treatment. A Mann-Whitney U-test shows that this treatment is not associated with higher payoffs (z=-1.515, Pr.>|z|= 0.1299, N=480), nor with an increased probability of choosing the optimal product (z=-0.276, Pr.>|z|=0.7829, N=480). However, the absolute benchmark is associated with a significantly lower probability of subjects selecting the worst product (z=2.525, Pr. > |z|=0.0116, N=480). We then test with the same regression models as presented in the previous section the robustness of these results.

As can be seen in the first row of Table 7 below, we find that the absolute benchmark treatment did not affect payoffs under increasing product dissimilarity. However, in contrast to the relative benchmark treatment, the absolute benchmarks also did not improve payoffs in later rounds of the experiment, as can be seen by the insignificant coefficient on the interaction term "ABM*post-period 6", nor in choice sets with more between-product variation, as can be seen by the insignificant coefficient on the interaction term "ABM*product similarity". In addition, we find that the absolute benchmark treatment did not have a significant effect on the probability of subjects choosing the optimal product, nor on the probability of choosing the worst product, nor on the probability of making any choice within 30 seconds (see Table 10 in the appendix).

Table 7: Regression table of effect of absolute benchmarks on payoffs

	(1) Payoff	(2) Payoff	(3) Payoff	(4) Payoff
Absolute Benchmark (ABM)	0.858	0.657	0.667	0.383
	(0.890)	(0.806)	(0.812)	(0.891)
Advice	2.386***	2.251***	2.228***	2.171***
	(0.845)	(0.791)	(0.790)	(0.802)

ABM*advice	-1.238	-0.614	-0.614	-0.630
	(1.063)	(1.116)	(1.118)	(1.125)
Post-Period 6 (PP6)			0.222	0.224
			(0.264)	(0.264)
ABM*PP6			-0.016	-0.023
			(0.386)	(0.384)
Product dissimilarity				0.195
				(0.132)
ABM*product dissimilarity				0.163
				(0.212)
Controls for age, sex, self-reported preferences, risk attitude, decision time	No	Yes	Yes	Yes
Constant	1.129	-5.915	-5.996	-6.426
	(0.742)	(4.871)	(4.872)	(4.916)
Observations	862	862	862	862
R-squared	0.059	0.146	0.147	0.153

Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

We can conclude on Hypothesis 1 that it is not supported in our data. That is, the presence of a benchmark in itself does not improve decision making significantly. We find some support, however, that a relative benchmark has a positive effect when we account for learning effects and when relative benchmarks can make differences between dissimilar products more salient.

Comparing relative and absolute benchmarks

Hypothesis 2 states that the relative benchmark treatment leads to higher payoffs and better decisions than the absolute benchmark treatment. To test this hypothesis, we thus compare the sample of the *relative* and the *absolute* benchmark treatments over all periods. A Mann-Whitney U-test shows no significant difference between the groups in terms of any of our variables of interest. However, when restricting the sample to the later decision rounds (post period 6) the relative benchmark treatment is associated with a lower probability of subjects choosing the worst option (z = -2.315, Pr. > |z| = 0.0206, N=364) than in the absolute benchmark treatment. We consider this a weak confirmation of our hypothesis. It seems indeed to be the case that presenting the options' attributes in relative values compared to a reference product has a greater (positive) effect on decision-making than presenting them in absolute values, but only after some learning has taken place. Our subjects clearly had to get used to the way the information was presented to them.

Advice

We then turn to Hypothesis 3 and analyze whether the *advice* treatment improved decision-making. As suggested by the descriptive statistics, a series of Mann-Whitney U-tests confirm that advice treatment results in more "best" choices (z = -2.156, Pr. > |z| = 0.0311), fewer "worst" choices (z = 4.526, Pr. > |z| = 0.0000) and higher gross payoffs (z = -2.898, Pr. > |z| = 0.0038), as well as higher payoffs net of advice costs (z = 2.156, Pr. > |z| = 0.0311). As can be observed in regression Tables 5, 6 and 7, the main effect of the advice treatment on payoffs is robust and positive in the presence of absolute and relative benchmark products. In Table 8 below we introduced several more controls and test the effect on payoffs, payoffs net of advice cost, as well as on the probability of choosing: (i) the best product, (ii) the worst product or (iii) no product.

Table 8: Regression table of effect of advice on payoffs and decision-making

	(1) Payoff	(2) Payoff net	(3) Best choice	(4) Worst choice	(5) No choice

Advice	2.695***	2.084**	0.319***	-0.214***	-0.118**
	(0.878)	(0.889)	(0.094)	(0.071)	(0.052)
Benchmark	0.675	0.637	0.008	-0.057	-0.041
	(0.744)	(0.749)	(0.072)	(0.047)	(0.034)
Advice*benchmark	-0.840	-0.988	-0.058	0.072	0.032
	(0.970)	(0.965)	(0.103)	(0.088)	(0.047)
Post-Period 6 (PP6)	0.817***	0.806***	0.068**	-0.056***	-0.028*
	(0.255)	(0.255)	(0.030)	(0.020)	(0.015)
Advice*PP6	-0.357	-0.329	-0.047	0.028	-0.067**
	(0.320)	(0.319)	(0.044)	(0.032)	(0.030)
Product dissimilarity	0.560***	0.551***	0.094***	-0.031***	-0.031***
	(0.136)	(0.135)	(0.020)	(0.011)	(0.008)
Advice*product dissimilarity	-0.281	-0.273	-0.061**	0.008	0.025
	(0.172)	(0.175)	(0.026)	(0.016)	(0.016)
Controls for age, sex, self-reported preferences, risk attitude, decision time	Yes	Yes	Yes	Yes	Yes
Constant	-3.018	-1.956			
	(3.651)	(3.564)			
Observations	1,331	1,331	1,331	1,331	1,468
(Pseudo) r-squared	0.115	0.098	0.0779	0.1357	0.1422

OLS results (Column 1-2), marginal effects of probit model (Columns 3-5). Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 8 confirms that the positive effects of the advice treatment on payoffs and decision-making are robust to inclusion of various controls and interactions. Advice improves both payoffs and payoffs net of cost, improves the probability of choosing the optimal product and reduces the probability of choosing the worst product. Furthermore, the advice treatment reduces the probability of not making a choice within the allocated timeframe. Contrary to the effect of the relative benchmark, the effect of the advice treatment on the probability of choosing the optimal option *increases* in product similarity. In other words, subjects benefit more from the option to buy advice in situations where the products are more similar. This is an intuitive result. It is harder to identify the optimal product in choice sets with great product similarity, so the advice – which is in effect simplifying the choice set to two options and thus reducing the cognitive effort needed for optimal decision making – has greater marginal effect in more complex situations. It also suggests that (relative) benchmarks and advice are in fact complements.

Benchmarks & demand for costly information

Finally, we turn to Hypothesis 4, i.e. the question whether the benchmark treatments affect demand for advice. To test this, we estimate a probit model in which we regress the dummy variable which indicates whether subjects purchased advice (1=purchased advice, 0=did not purchase advice) on the benchmark treatments.

The first column in Table 9 reports the effect of the absolute benchmark treatment. The positive coefficient on the first row shows that this treatment increased demand for costly advice. This effect did not change over the course of the experiment, as indicated by the insignificant coefficient on the "benchmark*post-period 6" interaction term.

The second column shows that the relative benchmark did not have a statistically significant effect on demand for advice. Finally, when pooling the results from both absolute and relative benchmarks we find again a statistically significant and positive effect on demand for advice, albeit only at the P<0.10 level.

Table 9: Regression table of effect of benchmark on demand for costly information

	(1) Buy advice	(2) Buy advice	(3) Buy advice
Benchmark (per type)	0.157**	0.033	0.134*
	(0.065)	(0.120)	(0.071)
Post-round 6 (PR6)	0.025	0.027	0.026
	(0.033)	(0.033)	(0.033)
BM*PR6	-0.035	-0.069	-0.052
	(0.044)	(0.046)	(0.040)
Controls for age, sex, self-reported preferences, risk attitude, decision time	Yes	Yes	Yes
Observations	490	500	750
Pseudo R2	0.3154	0.2985	0.2583

Marginal effects of probit regression. Robust standard errors in parentheses. Column 1 provides estimates for marginal effect of absolute benchmark (ABM), Column 2 for relative benchmark (RBM) and Column 3 for the combined set (BM) *** p<0.01, ** p<0.05, * p<0.1

In sum, these results indicate that only the absolute benchmark treatment influenced the demand for advice. We will discuss our findings in more detail in the next section.

5 Discussion

In this experiment we investigate how different kinds of choice architectures affect the quality of decision-making in complex choices. Subjects are presented a series of decision-tasks where they are asked to select the cheapest of four products that each consist of five cost-components. In addition, a time constraint of 30 seconds is imposed to increase the cognitive load. This setup serves to simulate complex choices that are ubiquitous in economic life, for example selecting a complex financial product such as a mortgage, which often has numerous cost-components and small print conditions.

We compare the effect of two types of interventions. In the benchmarks treatments subjects are informed about "market averages" of each attribute, where the values for the various cost-components are expressed either in absolute terms or relative to this "market average". In the costly information treatments subjects can receive information regarding the optimality of the options at a fixed cost.

Our results indicate that absolute benchmarks do not affect quality of decision-making. In contrast, relative benchmarks do improve decision-making as options in the choice set become increasingly dissimilar in terms of the number of optimal and suboptimal attributes. Our result suggests that providing a relative benchmark improves decision quality by making optimal and suboptimal cost-components more salient, enabling subjects to rank products in terms of their respective number of optimal and suboptimal cost-components. Such interventions would therefore be most useful when a large variety of products is on offer. This result builds on previous studies that show that information processing and decision-making can be steered by modifying attribute salience (Jarvenpaa, 1990; Mandel and Johnson, 2002; Sun et al., 2010; Lurie and Mason, 2007). For the health insurance market, Ericson and Starc (2016) show that the standardization of health insurance plans indeed allows consumers to more accurately differentiate between plans, but also that changes in the choice set are complementary to changes in the information interface.

Furthermore, we find that relative benchmarks not only affect decision quality, but also decision quantity, i.e. the probability that any of the four products is selected within the imposed time constraint. In line with results about decision quality, we find that the probability of a decision being made at all increases as products become more dissimilar in terms of the number of optimal and suboptimal attributes. This finding is line with previous studies that show that choice architecture can counter the problem of "choice inertia" without limiting the size of the choice set (Besedeš et al., 2015). In light of the policy debate on benchmarks and standard products our results allow for the tentative conclusion that their potential positive effect will depend on the specificities of the presentation and on product complexity and variety. Providing or imposing a (relative) benchmark is most likely to be effective in markets where a wide range of products with widely different attributes are being sold.

Our results furthermore indicate that the option to receive costly information improves the quality of decision making. Different from the effect of the relative benchmark, the effect of costly information on decision quality *increases* in product similarity. This finding suggests that financial guidance services (when abstracted from trust issues) may indeed have positive effects. In addition, we test how benchmarks affected the demand for advice. This research question is inspired by the fact that consumers that face complex financial choices, such as purchasing a mortgage, commonly have the option of purchasing advice from financial experts. Our results indicate that the provision of relative benchmarks does not affect the demand for costly information. Furthermore, we find that absolute benchmarks do increase demand for advice. One possible reason for this is that whereas relative benchmarks aid decision-makers under sufficient product dissimilarity, absolute benchmarks cause some confusion, which in turn could promote demand for costly advice.

The results regarding the positive effect of providing costly information raise the question of how and by whom such information should be provided. It must be noted here that our operationalization of the costly information treatments differs from real-life advice purchasing in two important aspects. In our treatments, subjects were given the option to pay in order to have two of the three suboptimal options being revealed as such. In contrast to real-life advice purchasing, subjects had to make their purchase decision before the actual choice set was revealed. Moreover, this information was not provided by another person, but directly by the computer. As such, agency problems and trust between advisor and advisee should not play a role in our experiment whereas such factors have been shown to significantly affect demand for advice in real markets (Inderst and Ottaviani, 2012; Bonaccio et al., 2006). Still, by offering trustworthy advice of a known quality (two options will be eliminated), we have established that such advice would improve decision maing across the board and if anything is complementary, not substitute to (relative) benchmarks.

In sum, our study shows that decision-making in complex choices can be improved through benchmarks if these are presented in relative terms. Furthermore, we show that demand for advice is not influenced by such benchmarks. This being the first study on benchmarks in the context of complex choices, further research is warranted to explore the role of contextual factors in this process, and to shed more light on the psychological mechanisms by which relative benchmarks aid decision-making. Moreover, to increase external validity, several subsequent steps could be taken. In a first step, our lab experiment could be repeated with a subject pool of representative financial services consumers. In case our findings are confirmed, a field experiment may be conducted to focus on complex financial decision making without time pressure, allowing for heterogeneous preferences and risk attitudes, and possibly also introducing the element of trust in financial advice.

References

Agnew, Julie R., and Lisa R. Szykman (2005). Asset allocation and information overload: The influence of information display, asset choice, and investor experience. *The Journal of Behavioral Finance* 6 (2), 57-70.

Barr, M. S., Mullainathan, S., & Shafir, E. (2008). Behaviorally Informed Financial Service Regulation. *Asset Building Program Policy Paper. New America Foundation.*

Besedeš, Tibor, Cary Deck, Sudipta Sarangi, and Mikhael Shor (2012a). Age Effects and Heuristics in Decision Making. *Review of Economics and Statistics*, 94(2), 580–595.

Besedeš, Tibor, Cary Deck, Sudipta Sarangi, and Mikhael Shor (2012b). Decision-making Strategies and Performance among Seniors. *Journal of Economic Behavior and Organization*, 81(2), 524–533.

Besedeš, Tibor, Cary Deck, Sudipta Sarangi, and Mikhael Shor. 2015. Reducing Choice Overload without Reducing Choices. *Review of Economics and Statistics* 97 (4), 793–802.

Bettman, James R., and Pradeep Kakkar (1977). Effects of information presentation format on consumer information acquisition strategies. *Journal of Consumer Research* 3(4), 233-240.

Bonaccio, S. and R. S. Dalal (2006). Advice Taking and Decision-Making: An Integrative Literature Review, and Implications for the Organizational Sciences. *Organizational Behavior and Human Decision Processes* 101, 127–151.

Calcagno, Riccardo, and Chiara Monticone (2015). Financial literacy and the demand for financial advice. *Journal of Banking & Finance* 50, 363-380.

Chater, Nick, Steffen Huck, and Roman Inderst (2010). Consumer Decision-Making in Retail Investment Services: A Behavioural Economics Perspective. Report to the European Commission/SANCO

Chernev, A., Böckenholt, U., Goodman, J. (2015). Choice overload: A conceptual review and meta-analysis. *Journal of Consumer Psychology* 25, 333–358.

Collins, J. Michael (2012). Financial advice: A substitute for financial literacy? *Financial Services Review* 21(4), 307.

Deelstra, Y., Nooteboom, S. G., Kohlmann, H. R., Van den Berg, J., & Innanen, S. (2003). Using knowledge for decision-making purposes in the context of large projects in The Netherlands. *Environmental Impact Assessment Review*, 23(5), 517-541.

Dijksterhuis, A., Bos, M.W., Nordgren, L.F., Van Baaren, R.B. (2006). On making the right choice: The deliberation-without-attention effect. *Science* 311, 1005–1007.

Ericson, K. M. M., & Starc, A. (2016). How product standardization affects choice: Evidence from the Massachusetts Health Insurance Exchange. *Journal of Health Economics*, 50, 71-85.

Fischbacher, U. (2007). z-Tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics* 10(2), 171-178.

Gigerenzer, G., & Todd, P. M. (1999). Fast and frugal heuristics: The adaptive toolbox. In Simple heuristics that make us smart (pp. 3-34). Oxford University Press.

Gino, F., & Moore, D. A. (2007). Effects of task difficulty on use of advice. *Journal of Behavioral Decision Making*, 20(1), 21-35.

Goldsmith, D. J., & Fitch, K. (1997). The normative context of advice as social support. *Human Communication Research*, 23(4), 454-476.

Goldsmith, D. J. (2000). Soliciting advice: The role of sequential placement in mitigating face threat. *Communications Monographs*, 67(1), 1-19.

Greifeneder, R., Scheibehenne, B., & Kleber, N. (2010). Less may be more when choosing is difficult: Choice complexity and too much choice. *Acta Psychologica*, 133(1), 45-50.

Hackethal, A., Haliassos, M., & Jappelli, T. (2012). Financial advisors: A case of babysitters?. *Journal of Banking & Finance*, 36(2), 509-524.

Heiss, F., Leive, A., McFadden, D., & Winter, J. (2013). Plan selection in Medicare Part D: Evidence from administrative data. *Journal of Health Economics*, 32(6), 1325-1344.

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

Hung, Angela, and Joanne Yoong (2010). Asking for help: Survey and experimental evidence on financial advice and behavior change. mimeo.

Inderst, R., & Ottaviani, M. (2012). Financial advice. *Journal of Economic Literature*, 50(2), 494-512.

Iyengar, S. S., & Lepper, M. R. (2000). When choice is demotivating: Can one desire too much of a good thing? *Journal of Personality and Social Psychology*,*79*(6), 995.

Iyengar, S. S., Huberman, G., & Jiang, W. (2004). How much choice is too much? Contributions to 401 (k) retirement plans. Pension design and structure: New lessons from behavioral finance, 83, 84-87..

Iyengar, S. S., & Kamenica, E. (2006). Choice overload and simplicity seeking. University of Chicago Graduate School of Business Working Paper, 87.

Jarvenpaa, Sirkka L. (1990). Graphic displays in decision making—the visual salience effect. *Journal of Behavioral Decision Making*, 3(4), 247-262.

Kirsch, I. S. (1993). Adult literacy in America: A first look at the results of the National Adult Literacy Survey. US Government Printing Office, Superintendent of Documents, Washington, DC 20402 (Stock No. 065-000-00588-3).

Lurie, Nicholas H., and Charlotte H. Mason (2007). Visual representation: Implications for decision making. *Journal of Marketing,* 71(1), 160-177.

Lusardi, A. and O. Mitchell (2006). Financial literacy and planning: Implications for retirement wellbeing. PRC Working Paper No. 1/2006.

Mandel, N., & Johnson, E. J. (2002). When web pages influence choice: Effects of visual primes on experts and novices. *Journal of Consumer Research*, 29(2), 235-245.

Payne, J. W., Bettman, J. R., & Johnson, E. J. (1988). Adaptive strategy selection in decision making. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 14(3), 534.

Reyna, V. F., & Brainerd, C. J. (2007). The importance of mathematics in health and human judgment: Numeracy, risk communication, and medical decision making. *Learning and Individual Differences*, 17(2), 147-159.

Robb, C. A., Babiarz, P., & Woodyard, A. (2012). The demand for financial professionals' advice: The role of financial knowledge, satisfaction, and confidence. *Financial Services Review*, 21(4), 291.

Soll, J. B., Keeney, R. L., & Larrick, R. P. (2013). Consumer misunderstanding of credit card use, payments, and debt: causes and solutions. *Journal of Public Policy & Marketing*, 32(1), 66-81.

Balz, J. P., Thaler, R. H., & Sunstein, C. R. (2014). Choice architecture. E. Shafir, The behavioral foundations of public policy, 428-439.

Yaniv, I. (2004). The benefit of additional opinions. *Current directions in psychological science*, 13(2), 75-78.

Yaniv, I. (2004). Receiving other people’s advice: Influence and benefit. *Organizational Behavior and Human Decision Processes*, 93(1), 1-13.

Yaniv, I., & Kleinberger, E. (2000). Advice taking in decision making: Egocentric discounting and reputation formation. *Organizational Behavior and Human Decision Processes*, 83(2), 260-281.

Zur, H. B., & Breznitz, S. J. (1981). The effect of time pressure on risky choice behavior. *Acta Psychologica*, 47(2), 89-104.

Appendix

Table 10: Regression table of effect of absolute benchmarks on product choice

	(1) best choice	(2) worst choice	(3) no decision
Absolute Benchmark (ABM)	-0.058	-0.034	-0.080
	(0.090)	(0.063)	(0.049)
Advice	0.221***	-0.194***	-0.092**
	(0.076)	(0.069)	(0.037)
ABM*advice	-0.006	0.044	0.024
	(0.111)	(0.103)	(0.051)
Post-Period 6 (PP6)	0.024	0.018	-0.071***
	(0.039)	(0.025)	(0.022)
ABM*PP6	-0.025	-0.029	0.044
	(0.057)	(0.037)	(0.030)
Product dissimilarity	0.026	-0.019*	-0.026**
	(0.019)	(0.011)	(0.011)
ABM*product dissimilarity	0.032	-0.010	0.015
	(0.031)	(0.016)	(0.021)
Controls for age, sex, self-reported preferences, risk attitude, decision time	Yes	Yes	Yes
Observations	862	862	958
Pseudo R2	0.1032	0.1956	0.2325

Marginal effects of probit regression. Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Online Appendix: Choice Complexity, Benchmarks and Costly Information

1 Instructions

V1 Treatment without advice

Welcome to this experiment!

PLEASE READ THE FOLLOWING INSTRUCTIONS AND THE INSTRUCTIONS ON YOUR SCREEN VERY CAREFULLY AS IT WILL AFFECT YOUR PERFORMANCE AND PROFITS IN THIS EXPERIMENT.

The experiment

The experiment consists of three parts. In the first part you have to repeatedly choose a product. During the second part of the experiment you are asked to choose repeatedly in which lottery you want to take part. The last part of the experiment is a questionnaire.

Part 1: Choosing the optimal financial product

In the first part of this experiment you are presented a table with four different products (product A, B, C and D) and your goal is to select the optimal product. The optimal product is the product with the lowest total costs.

Suppose the different products are financial products and that the maturity time of each product is one year. This means that you buy the product for a period of one year. The total costs of the products are not given, however you are presented with four different sub costs and a tax deduction of each product:

1. *Starting costs:* you have to pay these costs once when you buy the product.
2. *Monthly costs:* these are monthly costs which you have to pay every month for the duration of one year.
3. *Maturity costs:* costs that have to be paid at the end of your contract.
4. *Management fee (presented as a percentage of starting costs):* fee is paid once
5. *Tax deduction (presented as a percentage of monthly costs):* a tax saving once a year

These costs will be presented in a table similar to Table 1.

Table 1

	Product A	Product B	Product C	Product D
Start costs				
Monthly costs				
Maturity costs				
Management fee (%)				
Tax deduction (%)				

Knowing only the sub costs of all four products, it is your task to determine within 30 seconds which product is the optimal product. The time limit is presented in the upper right corner of your screen. If you do not choose a product within this time constraint, the computer will automatically choose the product with the highest total costs.

Payoff of Part 1

Your payoff depends on how well you make your decisions. You will be informed about how well you made your decisions only at the end of this part of the experiment. You are given an initial endowment of €8 and profits will be added to this when you choose optimal products and money will be deducted if you choose suboptimal products. At the end of the experiment the computer will randomly select one decision round and your decision in that specific round determines your payoff.

Decision	**Payoff**
Optimal	+ €5
Second best	+ €2.50
Second worst	- €2.50
Worst	- €5

Screen instructions

This hand-out sheet provides you with general information for the experiment. More specific information will be shown on your screen at the start of the experiment. Read these specific instructions very carefully as they influence your investment decisions.

Part 2: Lottery

In this part of the experiment 10 pairs of lotteries are presented to you and you have to choose in which lottery you want to take part in, lottery A or B. There is no time constraint in this part of the experiment.

Payoff of Part 2

The payoff you will receive in this part of the experiment is an extra payoff on top of what you already earned in the first part. The payoff depends on the lottery you participate in and on the outcome of the lottery. The outcome of the lottery is randomly decided by the computer. Also the payoff period is randomly determined. At the end of the experiment the computer will randomly select one of the pairs of lotteries and your decision for that specific pair determines your payoff. The possible payoffs are €3, €1.50, €1 and €0.

Part 3: Questionnaire

At the end of the experiment there will be a questionnaire for you to fill in. Please take your time and fill in this questionnaire truthfully. In the meantime we will prepare your payments from the previous parts of the experiment.

Good luck!

V2 Treatments with advice

Welcome to this experiment!

PLEASE READ THE FOLLOWING INSTRUCTIONS AND THE INSTRUCTIONS ON YOUR SCREEN VERY CAREFULLY AS IT WILL AFFECT YOUR PERFORMANCE AND PROFITS IN THIS EXPERIMENT.

The experiment

The experiment consists of three parts. In the first part you have to repeatedly choose a product. During the second part of the experiment you are asked to choose repeatedly in which lottery you want to take part. The last part of the experiment is a questionnaire.

Part 1: Choosing the optimal financial product

In the first part of this experiment you are presented a table with four different products (product A, B, C and D) and your goal is to select the optimal product. The optimal product is the product with the lowest total costs.

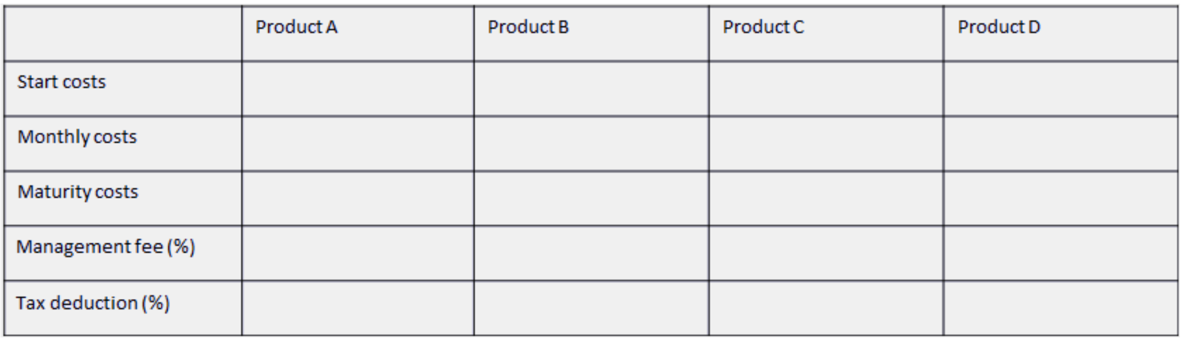
Suppose the different products are financial products and that the maturity time of each product is one year. This means that you buy the product for a period of one year. The total costs of the products are not given, however you are presented with four different sub costs and a tax deduction of each product:

1. *Starting costs:* you have to pay these costs once when you buy the product.
2. *Monthly costs:* these are monthly costs which you have to pay every month for the duration of one year.
3. *Maturity costs:* costs that have to be paid at the end of your contract.

4. *Management fee (presented as a percentage of starting costs):* fee is paid once
5. *Tax deduction (presented as a percentage of monthly costs):* a tax saving once a year

These costs will be presented in a table similar to Table 1.

Table 1



	Product A	Product B	Product C	Product D
Start costs				
Monthly costs				
Maturity costs				
Management fee (%)				
Tax deduction (%)				

Knowing only the sub costs of all four products, it is your task to determine within 30 seconds which product is the optimal product. The time limit is presented in the upper right corner of your screen. If you do not choose a product within this time constraint, the computer will automatically choose the product with the highest total costs.

Buying advice

Before the above table is shown you will be offered the opportunity to buy advice. The price of advice is €2.50 and this will be deducted from the payoff if you decide to buy advice. Buying advice will significantly increase your chances of choosing the optimal product, as the least optimal product is automatically eliminated and also another suboptimal product is indicated. This leaves you with only two options. Please note that advice is bought for each round separately. Buying advice in round 1 will only provide you with advice for round 1. The choice to buy advice in round 2 will again be presented prior to round 2.

Payoff of Part 1

Your payoff depends on how well you make your decisions. You will be informed about how well you made your decisions only at the end of this part of the experiment. You are given an initial endowment of €8 and profits will be added to this when you choose optimal products and money will be deducted if you choose suboptimal products. At the end of the experiment the computer will randomly select one decision round and your decision in that specific round determines your payoff.

Decision	**Payoff (without advice)**	**Decision**	**Payoff (with advice)**
Optimal	+ €5	Optimal	+ €2.50
Second best	+ €2.50	Second best	+ €0
Second worst	- €2.50	Second worst	- €5
Worst	- €5	Worst	- €7.50

Screen instructions

This hand-out sheet provides you with general information for the experiment. More specific information will be shown on your screen at the start of the experiment. Read these specific instructions very carefully as they influence your investment decisions.

Part 2: Lottery

In this part of the experiment 10 pairs of lotteries are presented to you and you have to choose in which lottery you want to take part in, lottery A or B. There is no time constraint in this part of the experiment.

Payoff of Part 2

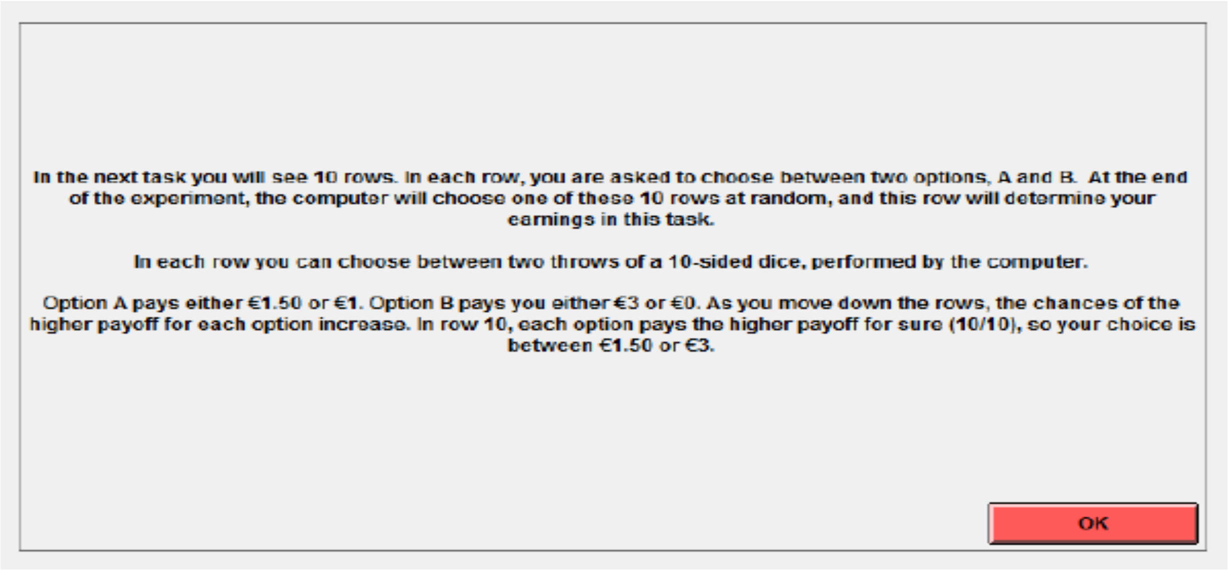
The payoff you will receive in this part of the experiment is an extra payoff on top of what you already earned in the first part. The payoff depends on the lottery you participate in and on the outcome of the lottery. The outcome of the lottery is randomly decided by the computer. Also the payoff period is randomly determined. At the end of the experiment the computer will randomly select one of the pairs of lotteries and your decision for that specific pair determines your payoff. The possible payoffs are €3, €1.50, €1 and €0.

Part 3: Questionnaire

At the end of the experiment there will be a questionnaire for you to fill in. Please take your time and fill in this questionnaire truthfully. In the meantime we will prepare your payments from the previous parts of the experiment.

Good luck!

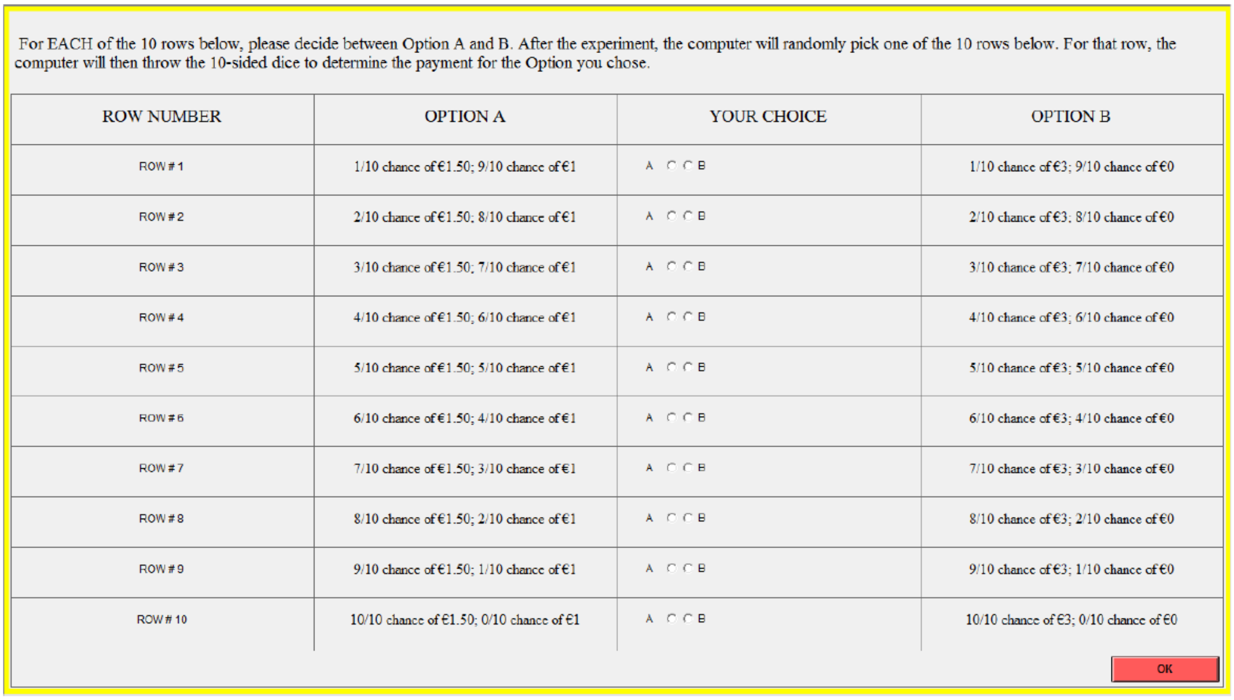
2 Holt & Laury task



In the next task you will see 10 rows. In each row, you are asked to choose between two options, A and B. At the end of the experiment, the computer will choose one of these 10 rows at random, and this row will determine your earnings in this task.

In each row you can choose between two throws of a 10-sided dice, performed by the computer.

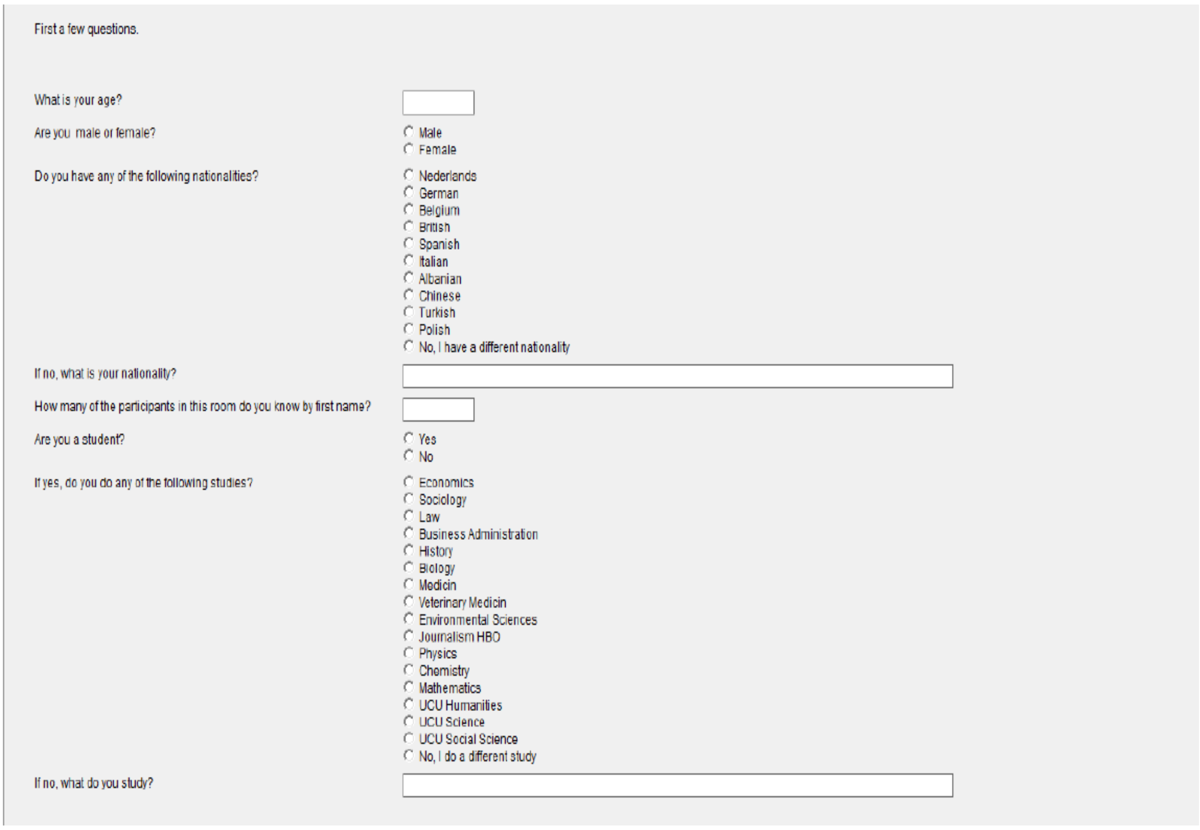
Option A pays either €1.50 or €1. Option B pays you either €3 or €0. As you move down the rows, the chances of the higher payoff for each option increase. In row 10, each option pays the higher payoff for sure (10/10), so your choice is between €1.50 or €3.



For EACH of the 10 rows below, please decide between Option A and B. After the experiment, the computer will randomly pick one of the 10 rows below. For that row, the computer will then throw the 10-sided dice to determine the payment for the Option you chose.

ROW NUMBER	OPTION A	YOUR CHOICE	OPTION B
ROW # 1	1/10 chance of €1.50; 9/10 chance of €1	A ○ ○ B	1/10 chance of €3; 9/10 chance of €0
ROW # 2	2/10 chance of €1.50; 8/10 chance of €1	A ○ ○ B	2/10 chance of €3; 8/10 chance of €0
ROW # 3	3/10 chance of €1.50; 7/10 chance of €1	A ○ ○ B	3/10 chance of €3; 7/10 chance of €0
ROW # 4	4/10 chance of €1.50; 6/10 chance of €1	A ○ ○ B	4/10 chance of €3; 6/10 chance of €0
ROW # 5	5/10 chance of €1.50; 5/10 chance of €1	A ○ ○ B	5/10 chance of €3; 5/10 chance of €0
ROW # 6	6/10 chance of €1.50; 4/10 chance of €1	A ○ ○ B	6/10 chance of €3; 4/10 chance of €0
ROW # 7	7/10 chance of €1.50; 3/10 chance of €1	A ○ ○ B	7/10 chance of €3; 3/10 chance of €0
ROW # 8	8/10 chance of €1.50; 2/10 chance of €1	A ○ ○ B	8/10 chance of €3; 2/10 chance of €0
ROW # 9	9/10 chance of €1.50; 1/10 chance of €1	A ○ ○ B	9/10 chance of €3; 1/10 chance of €0
ROW # 10	10/10 chance of €1.50; 0/10 chance of €1	A ○ ○ B	10/10 chance of €3; 0/10 chance of €0

3 Questionnaire



First a few questions.

What is your age?

Are you male or female?
○ Male
○ Female

Do you have any of the following nationalities?
○ Nederlands
○ German
○ Belgium
○ British
○ Spanish
○ Italian
○ Albanian
○ Chinese
○ Turkish
○ Polish
○ No, I have a different nationality

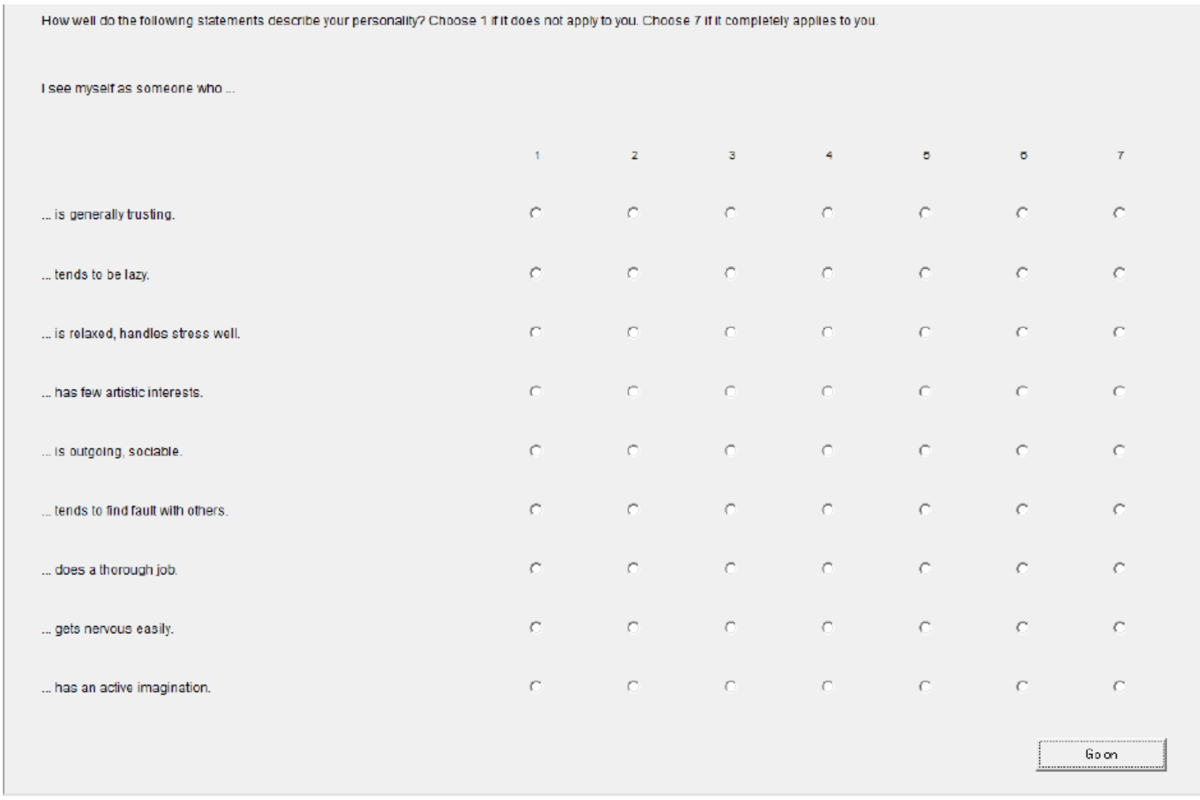
If no, what is your nationality?

How many of the participants in this room do you know by first name?

Are you a student?
○ Yes
○ No

If yes, do you do any of the following studies?
○ Economics
○ Sociology
○ Law
○ Business Administration
○ History
○ Biology
○ Medicin
○ Veterinary Medicin
○ Environmental Sciences
○ Journalism HBO
○ Physics
○ Chemistry
○ Mathematics
○ UCU Humanities
○ UCU Science
○ UCU Social Science
○ No, I do a different study

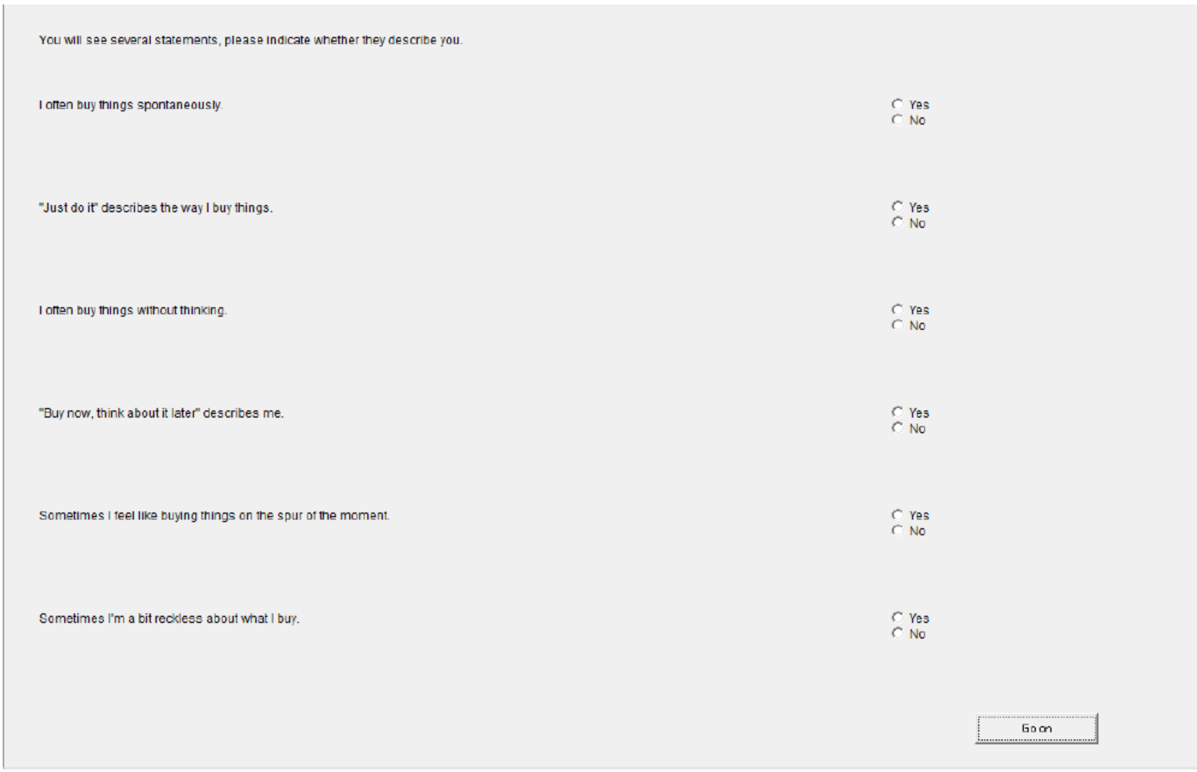
If no, what do you study?



How well do the following statements describe your personality? Choose 1 if it does not apply to you. Choose 7 if it completely applies to you.

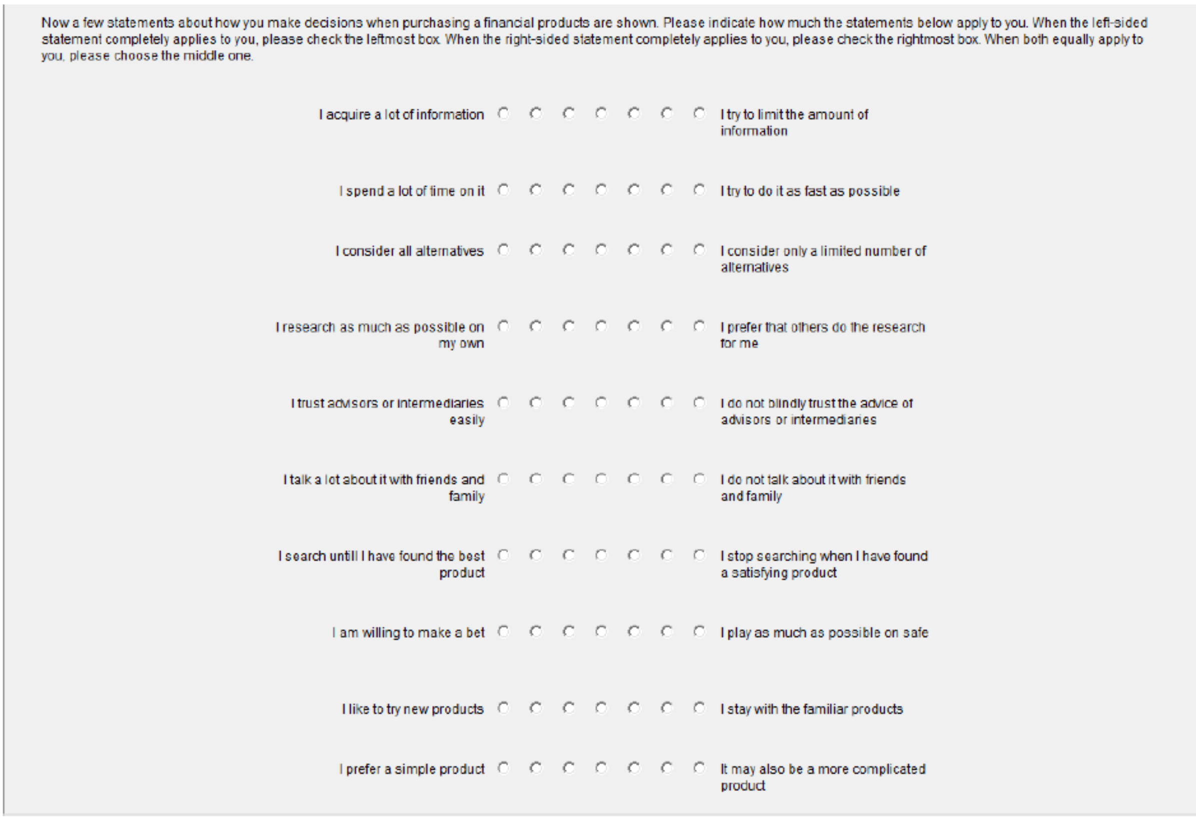
I see myself as someone who ...

	1	2	3	4	5	6	7
... is generally trusting.	○	○	○	○	○	○	○
... tends to be lazy.	○	○	○	○	○	○	○
... is relaxed, handles stress well.	○	○	○	○	○	○	○
... has few artistic interests.	○	○	○	○	○	○	○
... is outgoing, sociable.	○	○	○	○	○	○	○
... tends to find fault with others.	○	○	○	○	○	○	○
... does a thorough job.	○	○	○	○	○	○	○
... gets nervous easily.	○	○	○	○	○	○	○
... has an active imagination.	○	○	○	○	○	○	○



You will see several statements, please indicate whether they describe you.

I often buy things spontaneously.	○ Yes ○ No
"Just do it" describes the way I buy things.	○ Yes ○ No
I often buy things without thinking.	○ Yes ○ No
"Buy now, think about it later" describes me.	○ Yes ○ No
Sometimes I feel like buying things on the spur of the moment.	○ Yes ○ No
Sometimes I'm a bit reckless about what I buy.	○ Yes ○ No



Now a few statements about how you make decisions when purchasing a financial products are shown. Please indicate how much the statements below apply to you. When the left-sided statement completely applies to you, please check the leftmost box. When the right-sided statement completely applies to you, please check the rightmost box. When both equally apply to you, please choose the middle one.

I acquire a lot of information	○	○	○	○	○	○	○	I try to limit the amount of information
I spend a lot of time on it	○	○	○	○	○	○	○	I try to do it as fast as possible
I consider all alternatives	○	○	○	○	○	○	○	I consider only a limited number of alternatives
I research as much as possible on my own	○	○	○	○	○	○	○	I prefer that others do the research for me
I trust advisors or intermediaries easily	○	○	○	○	○	○	○	I do not blindly trust the advice of advisors or intermediaries
I talk a lot about it with friends and family	○	○	○	○	○	○	○	I do not talk about it with friends and family
I search until I have found the best product	○	○	○	○	○	○	○	I stop searching when I have found a satisfying product
I am willing to make a bet	○	○	○	○	○	○	○	I play as much as possible on safe
I like to try new products	○	○	○	○	○	○	○	I stay with the familiar products
I prefer a simple product	○	○	○	○	○	○	○	It may also be a more complicated product

Heading 1

Heading 2

Heading 3

Normal text

Caption