Misperceptions of Uncertainty and Their Applications to Prevention

ISBN 978 90 361 0602 3

Cover design: Crasborn Graphic Designers bno, Valkenburg a.d. Geul

This book is no. 759 of the Tinbergen Institute Research Series, established through cooperation between Rozenberg Publishers and the Tinbergen Institute. A list of books which already appeared in the series can be found in the back.

Misperceptions of Uncertainty and Their Applications to Prevention

Misvattingen van onzekerheid en hun toepassing op preventie

Thesis

to obtain the degree of Doctor from the
Erasmus University Rotterdam
by command of the rector magnificus

Prof.dr. R.C.M.E. Engels

and in accordance with the decision of the Doctorate Board.

The public defence date shall be held on
Thursday, February 6th, 2020 at 11.30 hrs

by

AYSIL EMIRMAHMUTOGLU

born in Kahramanmaras, Turkey

Erasmus University Rotterdam

Doctoral Committee:

Promotor:	Prof.dr. A. Baillon
Other members:	Prof.dr. P.P. Wakker
	Ass.Prof. L. Berger
	Dr. J. Delfgaauw
Co-promotor:	Dr. J.T.R. Stoop

Acknowledgements

The Netherlands is the land of my childhood dreams. When I was a kid whenever someone asked me where I would like to visit my answer was ready:"The Netherlands". It is because my favorite book as a little girl was "Hans Brinker, or, the silver skates: a story of life in Holland" by Mary Mapes Dodge. It blew my mind that a country below sea level could exist without being under water. The fact that people could ice-skate on the canals and go from one town to the other sounded just so much fun. What I read in that book was above my imagination as a little girl. Little did I know that I would live in that dream country one day. I first came to the Netherlands in 2010 for 5 months as an exchange student at Erasmus University. I would not have imagined that later I would come back and live there for another 6 years. Seems like the Netherlands is the place to be for me. It is funny though that in all those times I lived there I neither saw the dikes nor did I ice-skate on a canal. Why? Because I discovered that there are more than just the two. What made my time in the Netherlands special were the people that I met there.

There were no question marks for me about in which field of Economics to pursue my PhD. And I could not have a better supervisor than Prof. Aurélien Baillon in this journey. I knew immediately that I wanted to work with you when you presented your project to our cohort. Thank you, Aurélien, for teaching me everything with great enthusiasm, being always positive and supportive, helping me to find my way, encouraging me when I was pessimistic. I also consider myself privileged to have had the chance to be a part of Prof. Peter Wakker's group. I will never forget the day when I had my longest exam ever and how you grilled me. Thank you, Peter, for all you taught me as an economist and as a human being as well as for always bringing the group together at movie nights. Prof. Jan Stoop played a very critical role in the final push of my PhD when I started to question my path by telling me that opportunities are endless. Thank you, Jan, for our walks in the infinity loop, for your motivational talks and helping me gather my thoughts. I also had great colleagues that I consider "friends", Chen and Vitalie, who made my Rotterdam days nicer with their sweet energy.

I should thank my co-authors Han, Johannes, Richard, Bram, Evelien and Jeroen for their contribution and support in my research. Especially to you Richard, my first ever trip to the US could not have been more welcoming.

Completing the MPhil at Tinbergen Institute before even starting the PhD was only possible with sharing those long nights and hard work with great company. Those companions turned into amazing friends. Silvia, when it comes to work we were so similar in many ways. It meant a lot to me that I could always talk to you and you showed empathy. Gabriele, perfect example for "Work hard, play hard" with a big smile on your face. Thank you for making it possible to dance to reggaeton at almost all parties we had been together. And thank you both, for great hospitality in Andorra and Roma. Gavin always kept it cool. Thank you for always being there for us and your patience with our chaos. Andrej, thank you for always updating me about footy game results and the league chart :)

I lived in many different flats à la Ayşıl and I shared them with some of those great people from TI. Thank you, Coen, for the long dinners on the table with great conversations, David for always updating me with the political news from Turkey. And of course, Malin, you were the best third wheel on earth! Thank you for bringing so much fun to our tooth brushing sessions and being always caring to me.

Through TI I also met Veliyana and Ro. Sharing the most special moments in our lives with you two, first in Plovdiv, then in Istanbul and then in Amsterdam created a bond that is to be cherished for life. Thank you both for the great memories and making us feel gezellig at all times.

I have to thank Fatoş Abla for being my Turkish family in the Netherlands and for always making me feel at home.

My friends in Turkey, Mamiçe girls, thank you for always trying to make time for me in the limited time I stay in Istanbul. My family in Austria, thank you for always being very supportive and comforting.

I left home for studies already at the age of 14 moving from my hometown Kahramanmaraş to Istanbul. Over the years I only moved further away from home. Now that I am a mother as well I understand how big a sacrifice it is from my parents' side. Hardest of all was to be away from my sisters, something that I feel like I will never really get used to. But I hope this dissertation and PhD diploma compensate for it all a little bit. Thank you for all your support, for always encouraging me to aim higher and to bear with me while I was too stressed and busy.

Elif, you are the most beautiful thing that ever happened to me. You deserve a big thank as well because without you I would not have finished my last chapter and wrap this dissertation up that quickly. It was just too painful not having my mind fully focused on you when you were only tiny that I had to get it done.

Finally, the biggest thank is to you, Paul. I would not even have started the PhD in TI without your determination. We started this journey together in the Netherlands 6 years ago in a complete complexity privately. Over the years we managed to stand by each other's side, we got engaged, married, and had our little sunshine. You were always there for me, giving power in my most stressful moments, being proud in my successes. I cannot thank you enough…

Contents

Introduction

I care about understanding what leads people to do what they do and how we can help them to make better decisions. With that motivation in mind, in this dissertation I mainly focus on how individuals misperceive risks and uncertainty and how these misperceptions lead to suboptimal decision making. And I use different tools to do that. While Chapters 1 and 4 are lab experiments, Chapter 2 is a theoretical contribution and Chapter 3 is a field experiment. This is because I truly believe that none of the tools alone would be enough by itself in explaining people's behavior.

A significant part of my research is about how people make decisions about something that they do not even know whether it is ever going happen or not, in other words *ambiguity attitudes*. In 1962 Ellsberg provided a paradox with a thought experiment that showed the importance of ambiguity attitudes as follows:

Consider two urns one of which contains 50 red balls and 50 black balls in it, and the other contains 100 balls in red and black but we do not know in which proportion. We can call the first urn a "risky urn" since we know the distribution of it, and the second one an "ambiguous urn" since the distribution is unknown. We tell a person "You will receive 10 dollars if you pick a red ball. Do you prefer to pick the ball from the first urn or the second?". Imagine that the person chooses the first urn, the risky urn. Then we can argue that she might believe the probability of red balls in the ambiguous urn to be less than the probability of red balls in the risky urn which is 50%. Next, we tell the same person "Now you will receive 10 dollars if you pick a black ball. Which urn do you prefer?". Imagine that she chooses the risky urn again. Does it mean that she believes the probability of black balls in the ambiguous urn to be less than 50%? But then the probabilities of red and black balls in the ambiguous urn do not sum to 100%. How is it possible?

With this thought experiment Ellsberg proposed that people prefer to bet on things that they have more information about. Ever since, understanding ambiguity attitudes has been a very much debated topic in the literature. However, some components of ambiguity attitudes were still to be discovered, for example, attitudes for ambiguous rare events such as natural disasters. They are rare because they do not happen often, and they are ambiguous because we do not know exactly how often they happen. Still they are very important since the consequences of such events are mostly much bigger than things that we deal with in our everyday lives such as being caught in the rain without an umbrella. How do people react to rare events? Do they ignore them? Or do they attach so much importance to them? This is what I am searching for in Chapter 1.[1] I show by means of a lab experiment that people perceive rare events bigger than what they actually are when those events are presented to them separately rather than all together. We can think about many insurance packages that are sold separately rather than all-in-one insurance package that covers all risks. When people overweigh rare events this might lead them to overinsure.

While insurance is one way of coping with negative shocks people face, another solution is to prevent negative shocks in the first place. Sometimes there are risks that people can simply prevent. Simple lifestyle choices such as smoking, eating well, or exercising are the causes for the majority of preventable deaths. This means there is not enough prevention. People do not do enough. Isn't it puzzling that people buy health insurance but at the same time do not live a healthy life-style? One explanation could be that people do not know that they can prevent. However, everybody knows that smoking and obesity are objectively harmful and exercising is good for people. Moral hazard is also not enough to explain underprevention because insurance only covers financial risks whereas avoiding poor health has value in itself. Chapter 2 helps to solve this puzzle with an alternative and more plausible explanation.[2] I will show you that it is actually the same phenomenon that makes people both overinsure and prevent little, namely *probability weighting*.

What can we do to increase prevention and reduce probability weighting? What about just giving people more information about the risks they face? Or maybe there is a way to make them "feel" the risks better? Chapter 3 answers these questions with an application to cybersecurity at the organizational level.[3] Cybersecurity is one of the most challenging and hot topics of today's world because of rapidly changing technology every single day. One of the most common threats to cybersecurity for organisations is phishing attacks, attacks that aim at getting personal and/or financial information electronically. It is especially a big concern for governmental organisations such as ministries since they have so much confidential data. In Chapter 3, in collaboration with the Dutch Ministry of Economic Affairs I test whether communicating information in a more effective way or letting employees experience a simulated phishing attack help to reduce falling for phishing attacks. Both strategies turn out to be quite effective.

Simply providing information about cybersecurity risks is enough to increase awareness and reduce falling for phishing attacks. However, informing people about the importance of condom usage is arguably less effective when it comes to prevention of sexually transmitted diseases or unwanted pregnancies. This is because people's judgements of risk might differ in different contexts. Chapter 4 deals with this issue and answers the question of whether sexual context has an impact on ambiguity attitudes. There have been studies looking at sexual decision making, however, the

[1] co-authored with Aurélien Baillon

[2] co-authored with Aurélien Baillon, Han Bleichrodt, Johannes Jaspersen and Richard Peter

[3] co-authored with Aurélien Baillon, Jeroen de Bruin, Evelien van de Veer, Bram van Dijk

focus was on sexual decisions made in sexual moments. While decisions made in the moment are also important, especially the decisions that need to be made in advance are the ones that provide the grounds for prevention. For example, one cannot decide whether to use a condom in the moment without having bought that condom in the first place. Therefore, not only how people make decisions when they are sexually aroused is important but also whether being in a situation that has sexual associations alter decisions is. I show that sexual context has an impact on how people perceive likelihoods.

To sum up, my research shows that the majority of people are not as rational as they are assumed to be against a risk they face, however, their behavior is not random either. People have imperfect understanding of risks and likelihoods but those imperfections are systematic. The studies in this PhD dissertation help to understand those systematic misperceptions: what they are, what mechanisms generate them, how they affect behavior and what can be done to help people make better decisions.

Chapter 1

Zooming in on Ambiguity Attitudes*

Joint with

Aurélien Baillon

1.1. Introduction

Ambiguous rare events are pervasive in various fields of economics. Ambiguous rare events related to losses are events against which people may wish to insure. Policies for preventing or coping with environmental catastrophes also concern ambiguous rare events. Neglect of rare events can explain recent financial crises, as argued by Taleb (2007) in his book *The Black Swan*. An example of a rare event in the gain domain is to find a so-called 'unicorn', a start-up whose value exceeds one billion dollars. The occurrence of bubbles in the evaluation of high-tech start-ups, such the 'dot-com bubble' at the end of the 90s or the more recent Silicon Valley tech bubble, can be a sign that these rare events are overweighted by investors.

Kahneman and Tversky (1979) observed that rare events are typically either completely neglected or overweighted. For decision-making under risk, the common view in the literature is that low-probability events are overweighted (e.g., Tversky and Kahneman, 1992; Gonzalez and Wu, 1999). However, the picture is not so clear when we consider other decision paradigms. Recent research in psychology has shown that if unlikely events are not described but rather experienced by agents, such events tend to be partially neglected or underweighted (see, for instance, Barron and Erev, 2003; Hertwig et al., 2004; Hertwig and Erev, 2009). Regarding ambiguity, Ellsberg noted in his thesis (republished in 2011) that the common finding of ambiguity aversion might be due to the focus on moderate-likelihood events and gains and that the results might be reversed if unlikely events were to be considered. Several papers have confirmed this conjecture for events with likelihoods in the range of 10 to 30% (e.g., Ert and Trautmann, 2014; Baillon and Bleichrodt, 2015; Dimmock et al., 2015), but Ert and Trautmann (2014) also showed that experiencing unlikely ambiguous events made them less attractive.

---

*Published as " Baillon, A. and A. Emirmahmutoglu (2018). Zooming in on ambiguity attitudes. *International Economic Review, 59*(4), 2107-2131."

In this paper, we zoom in on *very* unlikely events[1] in an Ellsberg-like experiment. Despite the numerous studies on ambiguity attitudes conducted in recent decades (see Trautmann and Van De Kuilen, 2015, for a survey), rare events have been virtually ignored to date because of three major challenges. The first challenge is to provide sufficiently high incentives to consider such rare events.[2] In our experiment, subjects could either lose an initial endowment of €300 (loss frame) or win an equal amount (gain frame). The second challenge lies in identifying ambiguity attitudes generated on top of risk attitudes when moving from decision-making under risk to decision-making under uncertainty. We use *matching probabilities* to address this issue. The matching probability for an event E is the objective probability p that makes a decision-maker indifferent between receiving a nonzero outcome €300 with probability p and receiving €300 if event E occurs. Dimmock et al. (2015) formally showed that matching probabilities can directly capture ambiguity attitudes without requiring correction for utility or probability weighting. We generalize their result and develop an approach that is valid for all ambiguity models and all decision models under risk. The third challenge is related to controlling for people's unknown beliefs. A decision-maker may truly believe that an event is impossible, and we should not misinterpret such behavior as neglecting a rare event. Therefore, we do not use arbitrary benchmarks to assess overweighting or ignorance of very unlikely events. Instead, we compare matching probabilities only with themselves and study their internal consistency, as proposed by Baillon et al. (2018). We do this through the use of additivity measures. Intuitively, if an unlikely event is neither ignored nor overweighted, it should be assigned the same subjective value (matching probability) either in isolation or as part of a larger event. Hence, matching probabilities should be additive under ambiguity neutrality. If unlikely events are weighted more strongly in isolation (overweighting), then the matching probabilities will be said to be subadditive. Neglecting or underweighting would result in the opposite violation of additivity (superadditivity).[3]

Below, we begin by defining the theoretical framework and highlighting its advantages (section 2). We show that our approach is as general as possible and does not rely on any strong assumptions. After describing the experiment (section 3), we pursue our minimal-assumption approach for empirical analysis as well (section 4). We first use non-parametric tests to examine whether additivity is violated and, if so, whether it is violated differently in the gain frame and in the loss frame. Our analysis reveals that very unlikely events were not ignored but rather were overweighted overall, and more so in the loss frame. Second, to study heterogeneity of behavior,

[1]Throughout the paper, we use the terms "very unlikely event" and "rare event" interchangeably

[2]Einhorn and Hogarth (1986) simply avoided this problem by conducting a hypothetical survey.

[3]Thus far, the only study to address the first challenge (incentives) of which we are aware was conducted by Schade et al. (2012). However, they did not address the two other challenges and could not draw clear conclusions about ambiguity attitudes.

we use latent profile analysis with completely free parameters. This approach allows us to extract several behavioral profiles from the data without ex ante assumptions regarding what these profiles should be. Simultaneously, subjects are classified in accordance with these profiles. One of these profiles is close to ambiguity neutrality and represents approximately one third of the subjects. The other profiles consist of mild and extreme deviations from ambiguity neutrality, all in the sense of the overweighting of rare events. Finally, we discuss the interpretation and possible consequences of our results (section 5). Agents who assign greater weight to events in isolation than combined might be exploited in the form of money pumping, for instance, by splitting an insurance contract into subcontracts. The conclusion is presented in section 6.

1.2. Theoretical Framework

1.2.1. Notation and matching probabilities

The *state space* is finite and is denoted by S. It contains all possible states of nature. Only one state is realized, but it is unknown which one. Subsets of S are called *events*, and each event is denoted by E. The complementary event to E is denoted by E^C. The possible outcomes are monetary amounts from the set {-300, 0, 300}. *Bets* assign a nonzero outcome to an event and a zero outcome to the complement of that event. An *uncertain bet* is denoted by x_E0, with $x \in \{-300, 300\}$, and concerns an event E whose probability is uncertain. It yields €x if the true state belongs to E and yields nothing otherwise (if event E^C is realized). A *risky bet* is denoted by x_p0, with $x \in \{-300, 300\}$ and $p \in [0, 1]$. It pays an amount €x with an objective probability of p and pays nothing with an objective probability of $(1 - p)$.

Preferences over bets are denoted by $\succsim$. We assume that the decision-makers have preferences over all possible combinations of uncertain and risky bets, for all $E \subset S$, $x \in \{-300, 300\}$, and $p \in [0, 1]$. We also assume that the decision-makers' preferences exhibit monotonicity, such that they prefer more to less. Strict preference ($\succ$) and indifference ($\sim$) are defined as usual.

When a decision-maker is indifferent between two bets such that

$$x_E0 \qquad \sim \qquad x_p0, \tag{1}$$

we call p a *matching probability* of event E and denote it by $m^s(E) = p$, where the superscript $s \in \{+, -\}$ denotes the sign of x. Hence, the matching probability for event E and sign s is the objective probability p that makes a decision-maker indifferent between an uncertain bet on E and a risky bet that pays out the same amount x with probability p. We require minimal rationality of the decision-makers and assume that they satisfy $m^s(\varnothing) = 0$ and $m^s(S) = 1$ for all s. The function m^s is *increasing* if $E \subset F$ implies $m^s(E) \leq m^s(F)$. It is *additive* if, for all E and F

such that $E \cap F = \varnothing$, $m^s(E \cup F) = m^s(E) + m^s(F)$.

1.2.2. Using ambiguity neutrality as a benchmark

Following Ellsberg (1961), we say that an agent is *ambiguity averse (ambiguity seeking)* if, for all E, there exists a p such that $x_E0 \prec (\succ) x_p0$ and $x_{E^C}0 \prec (\succ) x_{1-p}0$. In a typical Ellsberg experiment, E corresponds to drawing a red ball from an opaque urn containing red and black balls in unknown proportion. The matching probability for event E can then be compared against a probability of 50% to characterize the decision-maker's ambiguity attitude, under the assumption that the decision-maker does not have any reason to believe that red is more likely than black. A more advanced approach consists of also measuring the matching probability of E^C (drawing a black ball) and then testing whether the sum of the matching probabilities is 1. Such an approach does not rely on any assumption regarding the decision-maker's beliefs. Below, we generalize this approach.

A decision-maker satisfying the *subjective expected utility* model (Savage, 1954) evaluates the uncertain bet x_E0 as $P(E)U(x)$, where $U(x)$ is the utility function and $P(E)$ is the additive subjective probability of E. *Probabilistic sophistication* relaxes the functional form assumption of subjective expected utility and holds when a probability measure $P(.)$ exists such that the decision-maker evaluates the uncertain bet x_E0 as $V(x_{P(E)}0)$, where $V(.)$ represents the decision-maker's preference over risky bets. Hence, under probabilistic sophistication, an individual is indifferent between x_E0 and $x_{P(E)}0$, and therefore, $m^s(E) = P(E)$. In other words, if probabilistic sophistication holds, the matching probabilities are sign-independent and additive. Following the ambiguity literature, we define *ambiguity neutrality* as probabilistic sophistication.[4] If V is the expected utility functional, then ambiguity neutrality is equivalent to subjective expected utility. A wide range of ambiguity models assumes expected utility under risk, most notably Schmeidler's (1989) Choquet expected utility, Gilboa and Schmeidler's (1989) maxmin expected utility, Klibanoff et al.'s (2005) smooth ambiguity model, Maccheroni et al.'s (2006) variational preferences, Siniscalchi's (2009) vector expected utility, and Cerreia-Vioglio et al.'s (2011) uncertainty-averse preferences. Observation 1 directly follows from the definition of ambiguity neutrality, even if V is not the expected utility functional.

Observation 1 *Assume any ambiguity model and any decision model under risk V. Ambiguity neutrality holds if and only if m^+ and m^- are additive and $m^+ = m^-$.*

Our approach will therefore consist of testing whether m^+ and m^- are additive and/or differ. This approach has four main advantages:

---

[4]Epstein (1999) proposed the adoption of this definition in a context in which no objective probabilities are available.

Isolation of ambiguity attitudes Matching probabilities isolate ambiguity attitudes from risk attitudes, as shown by Dimmock et al. (2015). They directly capture what uncertainty adds to risk (to V), which is a commonly accepted definition of ambiguity.

No restrictions regarding ambiguity models Our approach is not restricted to any specific ambiguity model. Deviations from probabilistic sophistication constitute evidence of ambiguity attitudes in all ambiguity models, and our approach is as general as possible.[5]

No restrictions regarding risk V can be any decision model under risk. It need not be expected utility nor prospect theory, as in Dimmock et al. (2015).

No assumptions regarding beliefs Additivity properties can be tested regardless of people's beliefs. Our approach does not rely on any assumptions regarding what people (should) believe.

Note that decision-makers may be neither ambiguity averse nor ambiguity seeking (i.e., may have matching probabilities of complementary events that sum to 1) while still deviating from ambiguity neutrality. This can occur, for instance, if they overweight unlikely events to the same extent that they underweight very likely events. This is why it is important to study the various ways in which decision-makers may deviate from neutrality. We do so by considering three indices of additivity, similar to those of Kilka and Weber (2001) and Baillon and Bleichrodt (2015). These indices represent three different preference conditions, namely, *binary complementarity, lower additivity* and *upper additivity*, and describe different patterns of ambiguity attitudes. Ambiguity neutrality predicts that all indices should be zero and, therefore, should also be the same for gains as for losses. The binary complementarity index captures ambiguity aversion in the sense of Ellsberg, that is, a general dislike for ambiguity regarding both events and their complements. The upper and lower additivity indices focus on deviations from ambiguity neutrality for very unlikely events and for very likely events, respectively.

1.2.3. Binary Complementarity

The first index that we use directly follows from the idea of Ellsberg's (1961) paradox as presented above and simply checks whether the matching probabilities of two complementary events sum to 1. Accordingly, our *binary complementarity index (BC index)* measures the distance from unity of the sum of the matching probabilities of an event and its complementary event as follows:

$$BC^s(E) = 1 - m^s(E) - m^s(E^C). \tag{2}$$

[5] Dimmock et al.'s (2015) was restricted to ambiguity models in which a subjective probability measure exists.

We say that binary complementarity holds if $BC^s(E) = 0$. $BC^s(E) > 0$ indicates ambiguity aversion for gains and ambiguity seeking for losses. Symmetrically, $BC^s(E) < 0$ indicates ambiguity seeking for gains and ambiguity aversion for losses.

The remaining two indices are based on preference conditions introduced by Tversky and Wakker (1995).

1.2.4. Lower Additivity

Consider two disjoint events $E, F \subset S$. *Lower additivity* holds if an event has the same influence when added to the empty event and when added to a non-empty event. In terms of preference conditions, under the assumption that we observe the following preferences (where equation (3) is trivial)

$$x_{\varnothing}0 \sim x_0 0 \tag{3}$$

$$\text{and} \quad x_E 0 \sim x_p 0, \tag{4}$$

lower additivity means that

$$\text{if} \quad x_F 0 \sim x_q 0 \tag{5}$$

$$\text{then} \quad x_{E\cup F} 0 \sim x_{p+q} 0 \tag{6}$$

whenever $E \cup F$ is bounded away from S.[6] Hence, lower additivity holds if a change from $\varnothing$ to E has the same impact on the matching probability as a change from F to $E \cup F$. In other words, the difference in matching probability between (3) and (4) should be equal to the difference between (5) and (6), i.e., $m^s(E) - m^s(\varnothing) = p = m^s(E \cup F) - m^s(F)$ with $m^s(\varnothing) = 0$. This equality gives rise to the *lower additivity index (LA index)*:

$$LA^s(E, F) = m^s(E) + m^s(F) - m^s(E \cup F). \tag{7}$$

Note that the LA index is commutative in its arguments E and F, i.e., $LA^s(E, F) = LA^s(F, E)$. Lower additivity holds if $LA^s(E, F) = 0$, lower *subadditivity* holds if $LA^s(E, F) > 0$, and lower *superadditivity* holds if $LA^s(E, F) < 0$.

1.2.5. Upper Additivity

The third and last index is the symmetric of the LA index, measuring the effect of removing an event from the universal event instead of adding an event to the empty event. *Upper additivity*

[6] See Wakker (2010), section 10.4.2, for a formal definition.

holds if the impact of removing an event from the universal event is the same as that of removing it from a non-universal event. If a decision-maker exhibits the following preferences (where equation (8) is trivial)

$$x_S 0 \quad \sim \quad x_1 0 \tag{8}$$

$$\text{and} \quad x_{S-E} 0 \quad \sim \quad x_{1-p} 0, \tag{9}$$

then upper additivity means that

$$\text{if} \quad x_{E \cup F} 0 \quad \sim \quad x_{p+q} 0 \tag{10}$$

$$\text{then} \quad x_F 0 \quad \sim \quad x_q 0 \tag{11}$$

whenever F is bounded away from $\varnothing$. Hence, upper additivity holds if a change from S to $S - E$ has the same impact on the matching probability as a change from $E \cup F$ to F, i.e., $m^s(S) - m^s(S - E) = p = m^s(E \cup F) - m^s(F)$ with $m^s(S) = 1$. We define the *upper additivity index (UA index)* as

$$UA^s(E, F) = [1 - m^s(S - E)] - [m^s(E \cup F) - m^s(F)], \tag{12}$$

where upper additivity is satisfied when $UA^s(E, F) = 0$. Upper *subadditivity* holds if $UA^s(E, F) > 0$, and upper *superadditivity* holds if $UA^s(E, F) < 0$. Note that in contrast to the LA index, the UA index is not commutative.

1.3. Method

In this section, we outline the experiment. Further details about the experiment are given in Appendix 1.A.

1.3.1. Subjects and procedure

Students from a Dutch university[7] participated in the experiment. The subjects were recruited from among students who had expressed a desire to be invited to participate in experiments. In total, $N = 99$ subjects consisting of 37 female and 62 male subjects with a median age of 23 participated in the experiment. The majority of the subjects were studying economics.[8] A maximum of two

[7] All students were taking courses at Erasmus University Rotterdam; however, due to joint programs, some were officially registered at the University of Amsterdam or the Free University of Amsterdam.

[8] The subjects were recruited by sending e-mail invitations to potential participants chosen randomly from a pool of students who had registered via a platform to participate in experiments conducted at Erasmus University. At this university, the two largest faculties are the business school and the school of economics. In our experiment, 84% of the

subjects participated in each session of the experiment. The sessions were held in a room with two cubicles (formed by panels placed on tables) to prevent communication between the subjects. The subjects took approximately 35 minutes to read the instructions, complete the experiment, and receive payment. Two between-subject treatments were applied: the "gain frame" and the "loss frame". The treatments were assigned randomly to the sessions.

1.3.2. Stimuli

The experiment was a variant of the original Ellsberg experiment. Ambiguity was created by means of a bag that contained multiple tickets marked with numbers ranging from 1 to 200. Neither the number of tickets in the bag nor the selection of numbers used to mark the tickets was known to the subjects. A ticket would be drawn from the bag, and the number on the ticket would determine the subjects' payment. Hence, the state space was $S = \{1, \cdots, 200\}$. We elicited the matching probabilities of various events from S.

To determine these events, the subjects were asked to assign different numbers to 6 symbols: $\triangle$, ☆, $\square$, $\bigcirc$, $\diamondsuit$, and ‡. Using these symbols, we created the events described in Table 1. In Set A, the specified events were very unlikely to occur, whereas they were very likely to occur in Set B. We elicited the matching probabilities for all events described in Table 1. In the gain frame (*gain treatment*), each of the specified events would lead to a €300 payoff. In the loss frame (*loss treatment*), these same events would each lead to the loss of a €300 endowment that the subjects initially received. The order of elicitation of the matching probabilities was randomized and different for each subject.[9]

We used choice lists to elicit the matching probabilities. Figure 1 shows an example of the choice list related to the event $\{☆, \square\}$ in Set A for the gain treatment. The ranges of probabilities in the choice lists for the risky bets (Option 2) were kept constant within each set of questions, but they varied between the sets. The probabilities ranged from 0% to 5.5% in Set A and from 94.5% to 100% in Set B. Therefore, a subject whose matching probabilities lay outside of these ranges constituted censored observations. The matching probability of each uncertain event was calculated by taking the midpoint between the highest (lowest) p such that $x_E0 \succ x_p0$ and the lowest (highest) p such that $x_E0 \prec x_p0$ for gains (losses).

subjects were studying "Economics", "Economics and Business Economics" or "Economics and Law", whereas 16% were studying "Management", "Law", "Psychology", "Sociology", or other disciplines. Economics students may be expected to behave more in line with traditional economic models (Marwell and Ames, 1981; Carter and Irons, 1991); in our case, this would suggest that they would be likely to be closer to ambiguity neutrality.

[9]In Appendix 1.A, we describe an additional set of questions, Set C, concerning events of moderate likelihood. Note that Set C was never a crucial component of the experiment because it did not concern rare events. As we show in the appendix, one of the four events in Set C was often misunderstood by the subjects; consequently, we decided to disregard Set C as a whole.

Table 1: Events and their descriptions

Set	Event	**Description:** the ticket drawn from the bag is marked with...
A	$\{\triangle\}$	...the number chosen by the subject for $\triangle$.
	$\{☆, \square\}$	...the number chosen by the subject for ☆ or $\square$.
	$\{\bigcirc, \diamondsuit, \ddagger\}$	...the number chosen by the subject for $\bigcirc$, $\diamondsuit$ or $\ddagger$.
	$\{\triangle, \bigcirc, \diamondsuit, \ddagger\}$	...the number chosen by the subject for $\triangle$, $\bigcirc$, $\diamondsuit$ or $\ddagger$.
	$\{☆, \square, \bigcirc, \diamondsuit, \ddagger\}$	...the number chosen by the subject for ☆, $\square$, $\bigcirc$, $\diamondsuit$ or $\ddagger$.
	$\{\triangle, ☆, \square, \bigcirc, \diamondsuit, \ddagger\}$	...the number chosen by the subject for $\triangle$, ☆, $\square$, $\bigcirc$, $\diamondsuit$ or $\ddagger$.
B	$\{\triangle\}^C$	...any number between 1 and 200 except for the number chosen by the subject for $\triangle$.
	$\{☆, \square\}^C$	...any number between 1 and 200 except for the numbers chosen by the subject for ☆ and $\square$.
	$\{\bigcirc, \diamondsuit, \ddagger\}^C$	...any number between 1 and 200 except for the numbers chosen by the subject for $\bigcirc$, $\diamondsuit$ and $\ddagger$.
	$\{\triangle, \bigcirc, \diamondsuit, \ddagger\}^C$	...any number between 1 and 200 except for the numbers chosen by the subject for $\triangle$, $\bigcirc$, $\diamondsuit$ and $\ddagger$.
	$\{☆, \square, \bigcirc, \diamondsuit, \ddagger\}^C$	...any number between 1 and 200 except for the numbers chosen by the subject for ☆, $\square$, $\bigcirc$, $\diamondsuit$ and $\ddagger$.
	$\{\triangle, ☆, \square, \bigcirc, \diamondsuit, \ddagger\}^C$	...any number between 1 and 200 except for the numbers chosen by the subject for $\triangle$, ☆, $\square$, $\bigcirc$, $\diamondsuit$ and $\ddagger$.

1.3.3. Incentives

Subjects received a €5 show-up fee. In addition, they had the chance to play one of their choices for real. The subjects in the gain treatment were each physically shown the €300 prize and were told that they could both win it, but the money (€600 in total) remained on the experimenter's desk. Given that we could not cause the subjects to lose €300 of their own money, we gave €300 to each subject in the loss treatment before they started reading the instructions. We placed the money on their desks and told them that it was theirs. All subjects signed a consent form, but the consent form for the loss treatment also asked the subjects to acknowledge that they were given €300. Nevertheless, it is possible that the subjects in the loss treatment considered only the final outcome (€0 or €300). Strictly speaking, the loss treatment was a loss-frame treatment. The fact that we displayed cash money was also the reason why we conducted each session with a maximum of two subjects. This way, we did not need to have more than €600 on hand at a time.

Because the amount to be gained or lost was high, it was important that the subjects did not believe that the experimenters could influence the outcome. Such mistrust could develop if the experimenters were to decide which specific number(s) had to be drawn from the bag to determine the outcomes of the uncertain bets. For example, a subject in the gain treatment might have believed that the specific numbers chosen by the experimenters would never be in the bag and that it would therefore be impossible to win. We prevented such mistrust in two ways. First, the bag was prepared before the subjects entered the room and thus before the subjects had selected their own

SET A

Question A2

Which one do you prefer?				
Option 1			**Option 2**	
			Win €300 with the following probability	**Code**
Win €300 if the ticket drawn from the bag is marked with number ☆ **or** □	☐	☐	0%	R2-1
	☐	☐	0.5%	R2-2
	☐	☐	0.6%	R2-3
	☐	☐	0.7%	R2-4
	☐	☐	0.8%	R2-5
	☐	☐	0.9%	R2-6
	☐	☐	1%	R2-7
	☐	☐	1.1%	R2-8
	☐	☐	1.2%	R2-9
	☐	☐	1.3%	R2-10
	☐	☐	1.4%	R2-11
	☐	☐	1.5%	R2-12
	☐	☐	2%	R2-13
	☐	☐	2.5%	R2-14
	☐	☐	3%	R2-15
	☐	☐	3.5%	R2-16
	☐	☐	4%	R2-17
	☐	☐	4.5%	R2-18
	☐	☐	5%	R2-19
	☐	☐	5.5%	R2-20

Figure 1: Sample question for the event $\{☆, □\}$ in the gain treatment

numbers to be drawn from the bag to determine the outcomes of the uncertain bets. Second, we asked the subjects to make a decision both for a given event (Set A) and for the complement of the same event (Set B), and we explicitly stated this in the instructions. Risky bets were implemented with dice.

At the beginning of the experiment, the subjects drew envelopes containing codes that would be opened at the end of the experiment and would determine which choice would be played for real. In the gain treatment, some envelopes would result in no choice being played at all; in the loss treatment, some envelopes would result in a guaranteed loss of the initial endowment. The exact list of envelopes is given in Appendix 1.A. The purpose of allowing the subjects to draw an envelope first and making them aware that only one choice would ultimately be played for real was to convince them to consider each choice as if it were the choice in the envelope (Johnson et al., 2014). The average payment was €14, but with a highly skewed distribution (three subjects left with €305).

1.3.4. Analysis

The analysis was conducted on 1114 observations from the 99 subjects (49 and 50 subjects each in the gain and loss treatments, respectively).[10]

Since we had censored observations, we chose to use non-parametric tests in our analysis. The Wilcoxon signed-rank test was used to test whether the indices were significantly different from zero. The differences in indices between the treatments were tested using the Mann-Whitney U test. These analyses were conducted separately for each event.

To study the heterogeneity of behavior, we performed a latent profile analysis (LPA).[11] By identifying shared characteristics in the subjects' indices, the LPA produced different endogenous groups (called profiles) and assigned a probability of being in each group to each subject. Hence, via the LPA, we could answer both the questions of what proportion of the subjects deviated from ambiguity neutrality and by how much. We implemented the LPA by means of the expectation maximization (EM) algorithm.[12] The EM algorithm has the advantages of simplicity, guaranteed convergence, and numerical stability. We conducted the LPA separately for the LA and UA indices. In our estimation procedure, we assumed only that the indices had multivariate normal distribution. For each profile, we estimated the mean of each index and the covariance matrix. We did not impose any restrictions on the parameter values. Therefore, we did not force the profiles to represent certain characteristics but rather let the data speak for themselves.

1.4. Results

In what follows, we multiply the matching probabilities by 100 and therefore report the results as percentages.

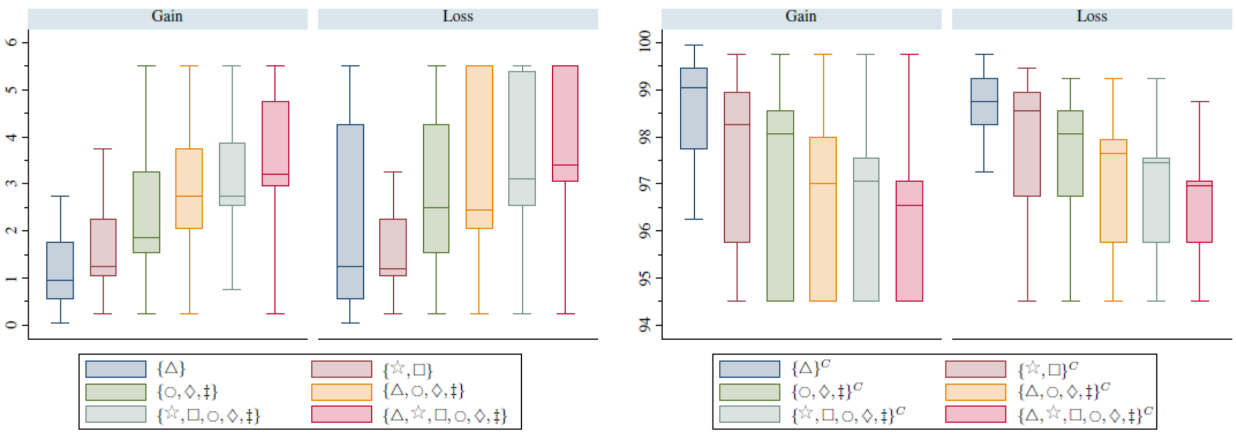
1.4.1. Aggregate Results

Figure 2 displays the matching probabilities for all events in Table 1. An ambiguity-neutral subject who believes all numbers between 1 and 200 have an equal chance of being drawn from the bag would have matching probabilities of $0.5\%, 1\%, ..., 3\%$ in Panel a and $99.5\%, 99\%, ..., 97\%$ in Panel b. Figure 2 shows that a vast majority of the subjects had matching probabilities higher than

[10]Of the 1188 total observations, 74 were excluded from the analysis. See Appendix 1.A for details on the exclusion criteria.

[11]LPA is also known as "finite mixture modeling"

[12]We applied the EM algorithm with 1000 iterations. The starting values of the profile means and covariances were obtained from three random subsamples in each iteration. The results of the iteration with the highest log-likelihood value were taken to be optimal.



(a) very unlikely events

(b) very likely events

Numbers are reported as percentages.
Every boxplot has lines at Q1, the median, and Q3. The adjacent lines show the most extreme values within 1.5 times the IQR of the nearer quartile.

Figure 2: Median matching probabilities

these thresholds in Panel a and lower in Panel b. These findings are consistent with the overweighting of rare events. However, this interpretation relies on the assumption that the subjects believed that all numbers were equally likely to be drawn from the bag. This assumption might be violated for subjects who chose numbers for the symbols △, ☆, □, ○, ◇, and ‡ that they expected to be more likely (or less likely) to be drawn. To avoid any assumptions regarding the subjects' beliefs, we studied the additivity indices defined in Section 2.

Binary Complementarity:

The median of the sum of the matching probabilities of an event and its complementary event was always very close to 100%, resulting in BC indices equal to 0 in most cases (see the upper part of Table 2). The BC indices differed from 0 in only four of the twelve cases (once in the gain treatment and three times in the loss treatment). In all these cases, the difference was in the direction consistent with ambiguity aversion. The significance tests for the differences between the two treatments show that the BC indices in the loss treatment were significantly or marginally lower than those in the gain treatment in four of the six indices. This pattern is also consistent with ambiguity aversion, which predicts higher BC indices in the gain frame than in the loss frame.[14]

[13]Since matching probabilities for all events were not available for all subjects (see Appendix 1.A - Matching probabilities), not all of the indices could be calculated for all subjects. This situation resulted in different sample sizes for each test.

[14]This is because ambiguity aversion predicts $BC^+ > 0$ and $0 > BC^-$.

Table 2: Median analysis

	Gain			Loss			Comparison	
	Median	z score	N	Median	z score	N	Conclusion	z score
$BC(\{\triangle\})$	0.00	1.08	42	0.00	-1.22	45	G>L $^{+}$	1.63
$BC(\{\star, \square\})$	0.00 $^{+}$	1.70	44	0.00	0.00	42		1.06
$BC(\{\bigcirc, \diamondsuit, \ddagger\})$	0.00	0.94	44	-0.10*	-2.42	46	G>L *	2.11
$BC(\{\triangle, \bigcirc, \diamondsuit, \ddagger\})$	0.00	0.98	44	0.00	-0.78	46		1.27
$BC(\{\star, \square, \bigcirc, \diamondsuit, \ddagger\})$	0.00	0.76	45	-0.10 *	-2.22	47	G>L *	1.98
$BC(\{\triangle, \star, \square, \bigcirc, \diamondsuit, \ddagger\})$	0.00	1.62	46	-0.10 *	-2.53	46	G>L **	2.73
$LA(\{\triangle\}, \{\bigcirc, \diamondsuit, \ddagger\})$	0.15 **	2.97	44	0.75 ***	4.30	44	G<L *	-1.76
$LA(\{\triangle\}, \{\star, \square, \bigcirc, \diamondsuit, \ddagger\})$	0.45 ***	3.55	44	0.95 ***	4.73	46	G<L *	-1.65
$LA(\{\star, \square\}, \{\bigcirc, \diamondsuit, \ddagger\})$	0.45 **	2.94	45	0.80 ***	3.90	42		-1.08
$UA(\{\triangle\}, \{\bigcirc, \diamondsuit, \ddagger\}^C)$	0.70 ***	3.60	40	0.70 ***	3.96	46		0.02
$UA(\{\triangle\}, \{\star, \square, \bigcirc, \diamondsuit, \ddagger\}^C)$	0.75 ***	3.97	42	1.05 ***	5.28	45		-0.79
$UA(\{\star, \square\}, \{\bigcirc, \diamondsuit, \ddagger\}^C)$	1.55 ***	4.16	41	0.65 ***	4.41	46		1.03

N represents the number of subjects in a particular test[13]

Numbers are reported as percentages

Stars indicate the significance levels of the Wilcoxon signed-rank test ("Gain" and "Loss" columns) and the Mann-Whitney U test for the BC (LA or UA) distribution for gains being higher (lower) than that for losses ("Comparison" column)

$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Lower additivity:

Lower additivity was defined as the difference in impact between adding a rare event to the empty event and adding the same rare event to a non-empty event. Table 2 shows that the median lower additivity indices were positive in all cases for both the gain frame and the loss frame. Hence, lower subadditivity held. Most subjects assigned higher weights to rare events when they were described in isolation than when they were part of larger events. There was also evidence that the overweighting of rare events was stronger in the loss frame than in the gain frame.

Upper additivity:

After analyzing lower additivity, we examined upper additivity by looking at the difference in impact between removing a rare event from the universal event and removing it from a non-universal event. The upper additivity indices were positive in all cases for both treatments (see the bottom of Table 2), indicating upper subadditivity and, thus, the overweighting of rare events that were removed from the universal event. Equivalently, events that were almost certain were underweighted. The between-treatment comparison revealed no significant difference.

1.4.2. Latent Profile Analysis

The aggregate analysis highlighted the occurrence of strong and consistent overweighting of rare events. To further understand this behavior, we established behavioral profiles via LPA. As a preparatory step, we generated kernel density plots for the LA and UA indices (Figure 3). Such plots can visualize the heterogeneity of behavior and whether this heterogeneity arises from different groups of subjects.

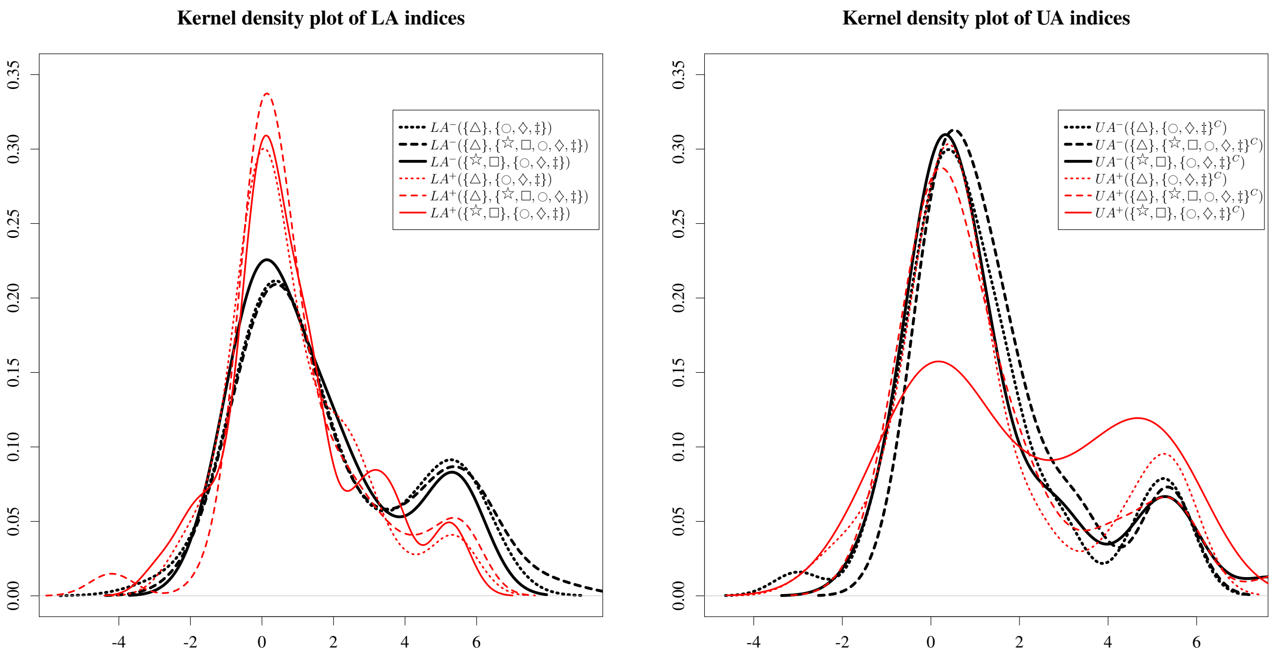


Figure 3: Kernel density plots

Examination of the kernel density plots revealed that the LA and UA indices were not homogeneous. Although there was a clear accumulation around a value of zero for all indices, there were also at least two other accumulations, a small one around two for several indices and a larger one around five for most indices. To account for these three accumulations, we performed LPA using the EM algorithm for three profiles with no restrictions on the class parameters (the means and covariances of the indices for each profile). It is possible that there were more than three profiles; however, we preferred this parsimonious specification because the consideration of each additional class would necessitate the estimation of ten more parameters. The LPA was performed independently for the LA and UA indices, over the set of subjects for whom we could calculate all indices in each case.[15]

[15] The numbers of subjects with complete observations for the LA and UA indices were 86 and 81, respectively.

[16] The Hessian matrix was calculated using the numDeriv package in R with the options d=.1 and r=4.

Table 3: Latent profile analysis

			$LA(\{\triangle\},\{\bigcirc,\diamondsuit,\ddagger\})$		$LA(\{\triangle\},\{\star,\square,\bigcirc,\diamondsuit,\ddagger\})$		$LA(\{\star,\square\},\{\bigcirc,\diamondsuit,\ddagger\})$	
Profile	p	S.E.	Mean	S.E.	Mean	S.E.	Mean	S.E.
1	51.64%***	0.10	0.24^{+}	0.14	0.23*	0.11	0.31*	0.15
2	39.06%***	0.09	1.56***	0.43	1.87***	0.49	1.46***	0.39
3	9.30%***	0.03	5.53***	0.00	5.44***	0.00	5.38***	0.03
			$UA(\{\triangle\},\{\bigcirc,\diamondsuit,\ddagger\}^C)$		$UA(\{\triangle\},\{\star,\square,\bigcirc,\diamondsuit,\ddagger\}^C)$		$UA(\{\star,\square\},\{\bigcirc,\diamondsuit,\ddagger\}^C)$	
Profile	p	S.E.	Mean	S.E.	Mean	S.E.	Mean	S.E.
1	42.89%***	0.08	0.42***	0.11	0.22^{+}	0.12	0.30**	0.10
2	47.24%***	0.04	1.16**	0.35	2.09***	0.30	2.23***	0.42
3	9.87%	0.08	5.27***	0.02	4.06***	0.01	5.24***	0.00

Standard errors were obtained from the Hessian matrix calculated at the estimated parameter values[16]
Numbers are reported as percentages
$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 3 presents the LPA results. They are consistent with our expectations from the kernel density plots. For Profile 1, all indices have means near zero with very small variances (the covariance matrices can be found in Tables 5 and 6 in Appendix 1.B). Although the means are (marginally) significantly different from 0, they are still very small, and Profile 1 is very close to ambiguity neutrality. For Profile 2, the indices have means near 2 with high variances, and for Profile 3, the means are near 5. The profiles for the LA and UA indices are very similar despite having been estimated independently. All profiles confirm the absence of upper or lower superadditivity. The subjects either overweighted rare events or were ambiguity neutral; there was no group of subjects who neglected rare events. The LPA assigned to each subject a probability of belonging to each group. Except in a few cases, each subject was assigned to one of the profiles with near certainty. Half of our sample was assigned to Profile 1 for LA and 43% for UA, indicating that many subjects were very close to ambiguity neutrality. Only approximately 10% of the subjects were assigned to Profile 3 (extreme overweighting).

Although the above analysis was conducted only for the subjects for whom complete observations were available regarding the indices, it was also possible to calculate the profile membership probabilities for subjects for whom at least one index value was available. Therefore, using the estimated profile means and covariances shown in Table 3, we calculated the probabilities for those subjects as well and included them in our analysis presented below. In this way, we could make use of all of the information available from the data.[17]

From the classification of the subjects into profiles, we could study whether subjects given one

[17]This approach resulted in the consideration of 92 and 91 subjects in the LA and UA analyses, respectively, 90 of whom were considered in both analyses.

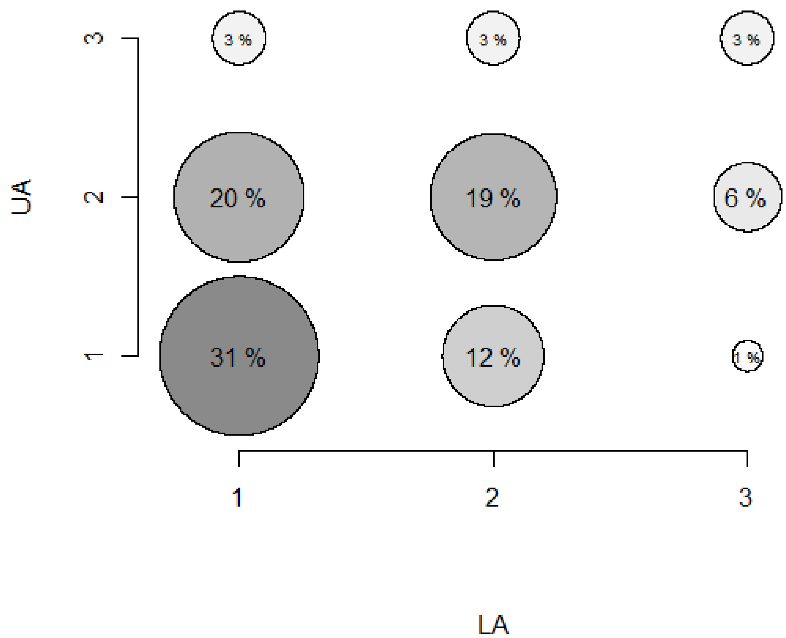


Figure 4: Cross-tabulation of LA and UA categorizations

treatment were more likely to be assigned to a specific profile than subjects given the other treatment. For the LA index, the subjects given the loss treatment had probabilities of 48%, 37% and 15% of being assigned to Profiles 1, 2, and 3, respectively, whereas the corresponding probabilities were 52%, 44% and 4% for the subjects given the gain treatment. Hence, although the proportion of Profile 1 subjects was the same for both treatments, the subjects given the loss treatment more often belonged to Profile 3 (extreme overweighting) than did the subjects given the gain treatment. For the UA index, the probabilities of belonging to each profile were 47%, 46% and 7% for the loss treatment and 35%, 51% and 14% for the gain treatment. Hence, the subjects tended to show upper subadditivity more often in the gain treatment than in the loss treatment.

Finally, each subject was categorized into one of the three profiles for the LA and UA indices independently. We were also interested in whether a subject who was categorized as showing lower additivity (subadditivity) would also be categorized as showing upper additivity (subadditivity). Figure 4 presents the proportions of the subjects assigned to the different combinations of profiles for LA and UA. The majority of the subjects were assigned to the same profile for LA as for UA.[18] Almost no one was categorized as Profile 1 (ambiguity neutrality) for LA and Profile 3 (extreme overweighting) for UA or vice versa. Overall, approximately one third of the subjects

[18]All subjects were assigned to the profile for which they had a $p > 50\%$. There were no subjects whose profile probabilities were below 50% for all profiles. Only 14 subjects exhibited a maximum p value below 70% for either the LA or UA profile.

were classified as Profile 1 for both LA and UA, meaning that they were consistently close to ambiguity neutrality. The fact that a substantial proportion of the subjects were ambiguity neutral is not surprising with regard to the findings of Ahn et al. (2014), who did not reject ambiguity neutrality for the majority of subjects.

1.5. Discussion

Our analysis revealed only weak ambiguity aversion as measured by the BC index. However, the LA and UA indices were significantly positive, leading to the rejection of the hypothesis of ambiguity neutrality and consistent with the overweighting of rare events. Three interpretations of such overweighting can be found in the literature. In the first, ambiguity attitudes are regarded as dependent on likelihood and outcomes (Hogarth and Einhorn, 1990). This leads to a *four-fold pattern* of ambiguity attitudes: ambiguity-seeking attitudes for very unlikely gains and very likely losses and ambiguity-averse attitudes for very likely gains and very unlikely losses (Trautmann and Van De Kuilen, 2015). The second interpretation explains the overweighting of unlikely events (and the equivalent underweighting of very likely events) as a consequence of likelihood insensitivity (Abdellaoui et al., 2011; Wakker, 2010). According to this interpretation, ambiguity decreases people's ability to discriminate between likelihood levels. In the extreme case, some people might assign the same weight to all events, hence overweighting rare events and underweighting very likely ones.

The third interpretation differentiates between ambiguity perception and ambiguity aversion, as in one of the best-known ambiguity models, the alpha-maxmin model (Ghirardato et al., 2004). In that model, ambiguity perception is represented by a set of priors. The decision-maker maximizes a linear combination of the best and worst expected utilities that can be obtained over this set of priors. The weight assigned to the worst case is denoted by alpha. Ambiguity aversion is then defined as an alpha value larger than 0.5. Baillon and Bleichrodt (2015) showed how the perception of ambiguity (a set of priors that is not a singleton) can lead to the overweighting of rare events for any alpha value other than 0 or 1. Hence, our results are compatible with the alpha-maxmin model.[19] Each profile identified in the LPA can then be interpreted as corresponding to a different degree of perceived ambiguity (almost none, mild, and extreme).

The experiment was conducted using choice lists because this approach has the advantage of making the incentive system easier to explain to the participants. As seen in Figure 1, the probability that would be chosen by an ambiguity-neutral person with uniform beliefs was more salient because of the smaller steps around it. This ensured both higher precision in this probability region and more conservative results. However, systematically switching from the ambiguous

[19] This compatibility requires a median alpha value that is very slightly above 0.5 to account for the BC results.

prospects to the risky prospects in the middle of the choice list (i.e., exhibiting a *middle bias*) tends to result in positive LA and UA indices and a zero BC index. Although there was no obvious evidence for a middle bias in the raw data, we ran an online experiment to check the robustness of our elicitation method. In this robustness check, we focused on our weakest results from the gain treatment ($LA(\{\triangle\}, \{\bigcirc, \diamondsuit, \ddagger\})$ and $UA(\{\triangle\}, \{\bigcirc, \diamondsuit, \ddagger\}^C)$, which had the lowest medians) and tested whether they could be replicated with a method immune to middle bias. We used a bisection method, in which subjects are presented with one choice after another and therefore cannot be influenced by any middle bias. The methodological details of the robustness check are described in Appendix 1.C. We obtained slightly lower medians for $LA(\{\triangle\}, \{\bigcirc, \diamondsuit, \ddagger\})$ and $UA(\{\triangle\}, \{\bigcirc, \diamondsuit, \ddagger\}^C)$ (0.05 and 0.55 instead of 0.15 and 0.75), with the former deviation being marginally significant[20] and the latter being significant at $p < 0.001$. We again obtained stronger deviations from 0 for the UA index than for the LA index.

Using random incentives in an ambiguity experiment (i.e., paying out for at most one choice among many choices with a given probability) is a common practice in the literature. Concretely, it means offering a compound lottery, with the first stage being the objective probability that a choice is played for real at all and the second stage being the chosen option, possibly an ambiguous bet. As noted by several authors (e.g., Oechssler and Roomets, 2014; Bade, 2015), if subjects mentally reverse the order of the compound lottery, perceiving first the ambiguous events and then the lottery, they may act as if they are ambiguity neutral even if they are not, leading to an overestimation of ambiguity neutrality. By contrast, Baillon et al. (2015) showed that perceiving the original order (objective probability followed by ambiguity) does not imply ambiguity neutrality. In our experiment, we made clear that the random incentives were implemented before the uncertainty was resolved, and even before the subjects made their choices, by using envelopes that the subjects drew at the beginning of the experiment to determine which choice would be played for real, as described in Section 3.3.

The results regarding the BC index show only limited support for ambiguity aversion, although the LA and UA results demonstrate that the subjects were not ambiguity neutral. This may seem at odds with the ambiguity literature, in which there is ample evidence for ambiguity aversion (see, for instance, the references in Table 3.4 of Trautmann and Van De Kuilen, 2015). Perhaps our implementation of the random incentive system may not have fully prevented some subjects from perceiving it as a way to hedge against the ambiguity. We may therefore have underestimated the level of ambiguity aversion. Furthermore, despite our efforts to make the outcomes salient (placing

[20]The significance was marginal according to a two-sided test; this deviation would be significant at the 5% level according to a one-sided test, which would be justified because we had a clear hypothesis regarding the direction of the effect after the main experiment.

the cash on the desk), the subjects also may have still considered the incentives to be low.[21] An alternative explanation, arising from the work done by Tversky in the 1990s (and formalized by Tversky and Wakker, 1995) is that, for more extreme likelihood levels, upper and lower subadditivity matter more than binary complementarity. In other words, for rare events, overweighting is stronger than ambiguity aversion.

In the experimental literature, it is common to provide an initial endowment to cover the losses with which subjects are faced during an experiment (see, among many others, Cohen et al., 1987; Eisenberger and Weber, 1995; Mason et al., 2005; Kermer et al., 2006; Harbaugh et al., 2010). However, we cannot know for sure whether the subjects perceived actual losses or whether they mentally integrated the loss with the initial endowment and thus considered only gains. Several papers have provided some evidence that subjects who are endowed with a monetary amount consider it theirs (e.g., Mason et al., 2005; Kermer et al., 2006). However, we acknowledge that there are better ways to investigate 'real' losses (not simply a loss frame). Abdellaoui and Kemel (2013) provided subjects with a monetary endowment but caused them to lose time, such that they could not easily integrate the loss with the endowment. Kocher et al. (2013) implemented "losses from posterior endowment", with subjects later winning back what they initially lost. Bosch-Domènech and Silvestre (2010) provided the initial endowment months before the experiment to encourage subjects to feel that it belonged to them. Unfortunately, we could not credibly implement any of these solutions for the large monetary amount used in this experiment.[22] As a positive consideration, Etchart-Vincent and l'Haridon (2011) found no differences in behavior in an experiment comparing hypothetical losses, real losses, and prior endowments. For this reason, we felt it suitable to choose the prior endowment approach, but we are still careful to refer to the implemented treatment as a loss frame to emphasize that it is a matter of framing/wording.

Ideally, we would expect that the LPA would extract from the data a profile of ambiguity-neutral subjects with mean indices not significantly different from zero. Indeed, the LPAs for both LA and UA yielded a profile very close to ambiguity neutrality, with very low indices and very low variance within this group. However, the mean indices were still significantly positive for this profile. Alternatively, we could have pre-defined the profiles and forced one profile to exhibit parameters equal to zero. This approach would result in an obvious but forced ambiguity-neutral profile. We believe that by allowing the data to speak for themselves, we were able to obtain results that are less perfect but more powerful.

Previous research regarding unlikely events has not extended much further than events with

[21] Surprisingly, we did observe mild ambiguity aversion in the robustness check conducted online and reported in Appendix 1.C, although the incentives in that experiment were less salient.

[22] It is difficult to force subjects to lose €300 worth of time, or to gain consent to participate if there is no guarantee that they will have to pay such an amount, or to be sure that they will show up after receiving such a large amount in advance.

likelihoods of approximately 5% or 10%. Such studies either have found that events were overweighted (Chipman, 1960; Kahn and Sarin, 1988; Curley and Yates, 1989; Casey and Scholz, 1991) or have not rejected ambiguity neutrality (Curley and Yates, 1985; Sarin and Weber, 1993). Two studies used stimuli more similar to ours but did not address the three challenges identified in the introduction to this paper (incentives, isolation of ambiguity attitudes, and control for beliefs). In a hypothetical experiment, Einhorn and Hogarth (1986) did not reject ambiguity neutrality for extremely unlikely events, but their study also did not control for beliefs. Schade et al. (2012) reported that subjects were more willing to pay for insurance in an ambiguous scenario than in a risky one, but this difference cannot be interpreted as a manifestation of ambiguity attitudes because they could not properly control for beliefs.

Crucially, our results rely on the events of interest being explicitly described. We cannot infer people's behavior when the relevant events are implicit, such that people may be unaware or at least not fully aware of them. Such unawareness is certainly pervasive in real life, but there are also many situations in which events are explicitly described, for instance, in insurance contracts.

Rare events exert greater influence in isolation. The subjects in our experiment assigned higher weights to events when they were isolated than when they were combined. Such behavior can be exploited, following arguments presented by Rabin and Thaler (2001) for loss aversion. These authors explained how myopic behavior (making decisions one after another without considering the overall consequences) and loss aversion can lead to money-pumping situations if people make small-scale insurance decisions one after another instead of considering the overall impact in terms of risk reduction. Overweighting of rare events will reinforce this pattern. Myopic agents will be prone to overinsure if they are presented with each possible (negative) event one after another, in isolation, instead of as a package.

1.6. Conclusion

Very unlikely events loom larger than they are. In an experiment, while controlling for risk attitudes and beliefs, we zoomed in on rare events. Using non-parametric tests, we found that the effect of a change from "no gain" to "some possibility of gain" was larger than the effect of a change from "some possibility of gain" to "greater possibility of gain". Similarly, the effect of a change from "certain gain" to "some possibility of gain" was larger than the effect of a change from "some possibility of gain" to "less possibility of gain". Both patterns were similar for losses. By means of latent profile analysis, we examined the heterogeneity in ambiguity attitudes. The results revealed that one third of our sample was consistently close to ambiguity neutrality, whereas the remaining two thirds showed mild or extreme overweighting. Such behavior is consistent with the mere perception of ambiguity in models such as the alpha-maxmin model. It can also lead to suboptimal situations such as overinsurance or overinvestment in long shots.

Appendix 1.A Additional information about the main experiment

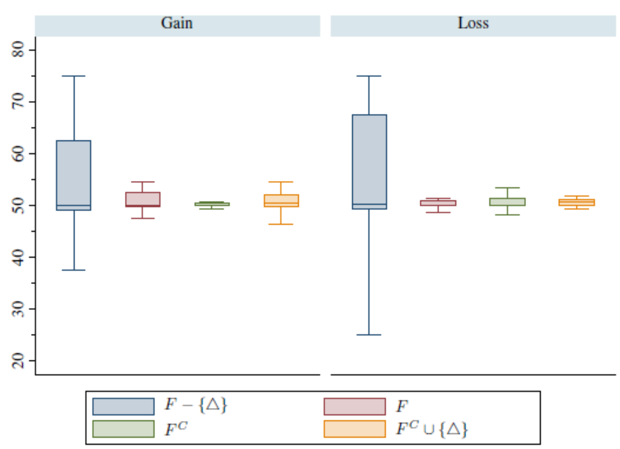
1.A.1 Additional questions

The full experiment consisted of three sets of questions. Sets A and B have been described in Table 1. Set C is described in Table 4. To implement Set C, we asked the subjects to choose values for $\triangle$ between 1 and 100 (while the values for all other symbols could be between 1 and 200).

Table 4: Events and their descriptions

Set	Event	**Description:** the ticket drawn from the bag is marked with...
C	$F - \{\triangle\}$	...any number between 1 and 100 except for the number chosen by the subject for $\triangle$.
	F	...any number between 1 and 100.
	F^C	...any number between 101 and 200.
	$F^C \cup \{\triangle\}$	...any number between 101 and 200 or the number chosen by the subject for $\triangle$.

Figure 5 shows the distributions of the matching probabilities for Set C. The event $F - \{\triangle\}$ seems not to have been understood by many participants; this led to a high variation in the answers, much higher than for any of the other matching probabilities. For this reason, we decided to exclude the entire set from the analysis.



Numbers are reported as percentages.
Every boxplot has lines at Q1, the median, and Q3. The adjacent lines show the most extreme values within 1.5 times the IQR of the nearer quartile.

Figure 5: Median matching probabilities for moderate-likelihood events

1.A.2 Stimuli: Choice list

In each question in Set A and Set B, the subjects were given a list of choices placing one fixed uncertain bet (Option 1) against 20 different risky bets (Option 2). The subjects were then asked to make a choice between the uncertain and risky bets for each of these 20 cases. The questions in Set C were prepared similarly; however, the number of risky bets in the choice list was 30 instead of 20. The reason for this difference was to ensure that the subjects would make the same number of choices in each set, namely, 120. An equal number of choices in each set was necessary for our incentive system, in which we used envelopes to determine the choice to be played for real (see Figure 6 for a sample question from Set C).

An ambiguity-neutral subject believing that all numbers were equally likely to be drawn from the bag would assign a probability of 0.5% to each number. For such a subject, the uncertain events in Sets A, C, and B would have probabilities ranging from 0.5% to 3%, from 49.5% to 50.5%, and from 97.5% to 99.5%, respectively. The lists of probabilities in the risky bets presented in our questions were constructed such that the probability differences between the two rows decreased as the rows approached the subjective probability of an ambiguity-neutral subject believing all numbers were equally likely (ambiguity-neutral probability). This arrangement might have assisted the subjects in interpreting the likelihoods of given events and might have led them toward ambiguity neutrality. If this was the case, our results are conservative.

1.A.3 Incentives

The chances of playing for real were determined by drawing from a box containing 50 envelopes. The distribution of the envelopes in this box was as follows:

- 47 envelopes containing a blank ticket
- 1 envelope containing a ticket for Set A
- 1 envelope containing a ticket for Set B
- 1 envelope containing a ticket for Set C

Here, a blank ticket meant that the subject would win nothing in the gain treatment and would lose €300 in the loss treatment. Any one of the other three tickets meant that a choice from the set specified on the ticket would be played for real.

A second box containing 120 envelopes was also prepared to determine which specific choice would be played if the envelope drawn from the first box contained a non-blank ticket. The tickets in the second box were marked with the codes that appeared in each row of the choice list in each question. Hence, the combination of the tickets drawn from the two boxes uniquely determined

SET C

Question C2

Which one do you prefer?				
Option 1			**Option 2**	
			Win €300 with the following probability	**Code**
	☐	☐	25%	R2-11
	☐	☐	30%	R2-12
	☐	☐	35%	R2-13
	☐	☐	40%	R2-14
	☐	☐	45%	R2-15
	☐	☐	46%	R2-16
	☐	☐	47%	R2-17
	☐	☐	48%	R2-18
	☐	☐	48.5%	R2-19
	☐	☐	49%	R2-20
	☐	☐	49.5%	R3-1
	☐	☐	49.6%	R3-2
	☐	☐	49.7%	R3-3
	☐	☐	49.8%	R3-4
Win €300 if the ticket drawn from the bag is marked with **any number between 1 and 100**	☐	☐	49.9%	R3-5
	☐	☐	50%	R3-6
	☐	☐	50.1%	R3-7
	☐	☐	50.2%	R3-8
	☐	☐	50.3%	R3-9
	☐	☐	50.4%	R3-10
	☐	☐	50.5%	R3-11
	☐	☐	51%	R3-12
	☐	☐	52%	R3-13
	☐	☐	53%	R3-14
	☐	☐	54%	R3-15
	☐	☐	55%	R3-16
	☐	☐	60%	R3-17
	☐	☐	65%	R3-18
	☐	☐	70%	R3-19
	☐	☐	75%	R3-20

Figure 6: Sample question from Set C

the choice to be played for real.

The ambiguous bags were created separately for each session and each subject. Therefore, communication between the subjects outside of the laboratory was irrelevant, and the outcomes for different subjects did not depend on each other.

The probabilities associated with the risky bets (Option 2) were implemented using three ten-sided dice that together could generate all possible numbers from 00.0 to 99.9 up to one decimal place. Hence, for example, a subject in the gain treatment who chose Option 2 for an X% probability of winning €300 would win if the dice showed a number strictly below X. Both the dice and the bags were shown to the subjects before they began answering the questions.

If a subject was to play a choice for real, it was done privately, and the payments were kept anonymous.

Out of the 99 subjects, no one decided to withdraw from the experiment once they had read the instructions. In total, 5 subjects drew a non-empty envelope from the first box, and 3 of them were paid €300 in accordance with the outcome of the play.

1.A.4 Matching probabilities

There were 22 subjects who gave at least one counterintuitive answer. We defined a counterintuitive answer for a subject in the gain treatment as a preference for a 0% probability of gain over some uncertain chance of winning or a refusal of a 100% probability of gain. Similarly, for a subject in the loss treatment, a counterintuitive answer was defined as the rejection of a 0% probability of loss or a preference for a 100% probability of loss over some uncertain chance of no loss. We treated such answers as erroneous and concluded that subjects who made such mistakes did not answer the questions carefully. However, it should be noted that because of the nature of the experiment, it was easy to select a column incorrectly by mistake, and the majority of these subjects made only one such mistake. Therefore, for the optimal usage of the available data, instead of completely eliminating those subjects from the analysis, we eliminated only the observations corresponding to such answers.

We did not impose a maximum one-time switching rule. Therefore, the subjects could switch between the two bets in a given choice list as many times as they wished. In total, 5 subjects switched multiple times. Calculating the matching probabilities for such observations would require strong assumptions, from which we refrained for the entire analysis. Hence, we treated these subjects in the same manner as those discussed above and did not include the observations for which there were multiple switches in our analysis.

For censored observations (when the subjects never switched between the bets and always chose the same option in a given choice list), the matching probabilities were set equal to the highest probability available in the choice list (5.5%) for the events in Set A (very unlikely events) or to the lowest probability available in the choice list (94.5%) for the events in Set B (very likely events). For the events in Set C, the applied correction depended on the treatment. If Option 1 was always chosen, then the matching probabilities for the gain (loss) treatment were set equal to the highest (lowest) available probability (75% (25%)), and the opposite probability assignments were made if Option 2 was always chosen.

Appendix 1.B LPA covariance matrices

Table 5: Estimated variance-covariance matrices for the LA index

	Profile 1		
	$LA(\{\triangle\},\{\bigcirc,\diamondsuit,\ddagger\})$	$LA(\{\triangle\},\{\star,\square,\bigcirc,\diamondsuit,\ddagger\})$	$LA(\{\star,\square\},\{\bigcirc,\diamondsuit,\ddagger\})$
$LA(\{\triangle\},\{\bigcirc,\diamondsuit,\ddagger\})$	0.61** (0.20)		
$LA(\{\triangle\},\{\star,\square,\bigcirc,\diamondsuit,\ddagger\})$	0.14 (0.09)	0.28* (0.12)	
$LA(\{\star,\square\},\{\bigcirc,\diamondsuit,\ddagger\})$	0.36* (0.15)	0.07 (0.08)	0.66*** (0.18)

	Profile 2		
	$LA(\{\triangle\},\{\bigcirc,\diamondsuit,\ddagger\})$	$LA(\{\triangle\},\{\star,\square,\bigcirc,\diamondsuit,\ddagger\})$	$LA(\{\star,\square\},\{\bigcirc,\diamondsuit,\ddagger\})$
$LA(\{\triangle\},\{\bigcirc,\diamondsuit,\ddagger\})$	4.44*** (1.19)		
$LA(\{\triangle\},\{\star,\square,\bigcirc,\diamondsuit,\ddagger\})$	2.25* (0.94)	4.69*** (1.23)	
$LA(\{\star,\square\},\{\bigcirc,\diamondsuit,\ddagger\})$	2.54** (0.92)	0.42 (0.87)	4.22*** (1.15)

	Profile 3		
	$LA(\{\triangle\},\{\bigcirc,\diamondsuit,\ddagger\})$	$LA(\{\triangle\},\{\star,\square,\bigcirc,\diamondsuit,\ddagger\})$	$LA(\{\star,\square\},\{\bigcirc,\diamondsuit,\ddagger\})$
$LA(\{\triangle\},\{\bigcirc,\diamondsuit,\ddagger\})$	0.04*** (0.00)		
$LA(\{\triangle\},\{\star,\square,\bigcirc,\diamondsuit,\ddagger\})$	-0.06*** (0.00)	0.08*** (0.00)	
$LA(\{\star,\square\},\{\bigcirc,\diamondsuit,\ddagger\})$	0.03*** (0.00)	-0.04*** (0.00)	0.07*** (0.00)

Numbers are reported as percentages

Standard errors are shown in parentheses

$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 6: Estimated variance-covariance matrices for the UA index

	Profile 1		
	$UA(\{\triangle\},\{\bigcirc,\diamondsuit,\ddagger\}^C)$	$UA(\{\triangle\},\{\star,\square,\bigcirc,\diamondsuit,\ddagger\}^C)$	$UA(\{\star,\square\},\{\bigcirc,\diamondsuit,\ddagger\}^C)$
$UA(\{\triangle\},\{\bigcirc,\diamondsuit,\ddagger\}^C)$	0.33***		
	(0.09)		
$UA(\{\triangle\},\{\star,\square,\bigcirc,\diamondsuit,\ddagger\}^C)$	0.24**	0.37**	
	(0.09)	(0.13)	
$UA(\{\star,\square\},\{\bigcirc,\diamondsuit,\ddagger\}^C)$	0.14*	0.04	0.27***
	(0.06)	(0.07)	(0.08)

	Profile 2		
	$UA(\{\triangle\},\{\bigcirc,\diamondsuit,\ddagger\}^C)$	$UA(\{\triangle\},\{\star,\square,\bigcirc,\diamondsuit,\ddagger\}^C)$	$UA(\{\star,\square\},\{\bigcirc,\diamondsuit,\ddagger\}^C)$
$UA(\{\triangle\},\{\bigcirc,\diamondsuit,\ddagger\}^C)$	3.85***		
	(1.06)		
$UA(\{\triangle\},\{\star,\square,\bigcirc,\diamondsuit,\ddagger\}^C)$	1.40*	3.10***	
	(0.69)	(0.79)	
$UA(\{\star,\square\},\{\bigcirc,\diamondsuit,\ddagger\}^C)$	1.72^{+}	-1.37	6.00***
	(0.95)	(0.84)	(1.43)

	Profile 3		
	$UA(\{\triangle\},\{\bigcirc,\diamondsuit,\ddagger\}^C)$	$UA(\{\triangle\},\{\star,\square,\bigcirc,\diamondsuit,\ddagger\}^C)$	$UA(\{\star,\square\},\{\bigcirc,\diamondsuit,\ddagger\}^C)$
$UA(\{\triangle\},\{\bigcirc,\diamondsuit,\ddagger\}^C)$	0.20***		
	(0.00)		
$UA(\{\triangle\},\{\star,\square,\bigcirc,\diamondsuit,\ddagger\}^C)$	0.93***	6.75***	
	(0.00)	(0.01)	
$UA(\{\star,\square\},\{\bigcirc,\diamondsuit,\ddagger\}^C)$	0.16***	1.26***	0.24***
	(0.00)	(0.00)	(0.00)

Numbers are reported as percentages

Standard errors are shown in parentheses

$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Appendix 1.C Robustness check of the elicitation method

We describe here the online experiment mentioned in the Discussion section. It was conducted to test the robustness of the results against middle bias.

1.C.1 Method

1.C.1.1 Subjects and procedure

The subjects were recruited from the same platform as for the experiment presented in the main text. The experiment was conducted online, with $N = 61$ subjects (27 female, median age 22). We used an online experiment because it was run during a period when most students had no obligation to be on campus. As in the main experiment, a large majority of the subjects (84%) were studying economics. Each subject took approximately 10-15 minutes to complete the online questionnaire.

1.C.1.2 Stimuli

We attempted to replicate our weakest results from the gain treatment, i.e., $LA(\{\triangle\}, \{\bigcirc, \diamondsuit, \ddagger\})$ and $UA(\{\triangle\}, \{\bigcirc, \diamondsuit, \ddagger\}^C)$. We elicited the matching probabilities of the events $\{\triangle\}$, $\{\bigcirc, \diamondsuit, \ddagger\}$ and $\{\triangle, \bigcirc, \diamondsuit, \ddagger\}$ and those of their complements. Instead of choice lists, we used a bisection method. We asked the subjects to indicate their preferences between an uncertain bet and a risky bet, starting with the probability that an ambiguity-neutral subject with uniform beliefs would choose (see Figure 7). If a subject preferred the uncertain bet (the risky bet), we increased (decreased) the probability associated with the risky bet in the next question. We continued to increase/decrease the probability associated with the risky bet (Option 2) in three subsequent questions after the first one. Once again, the probability increments were smaller around the ambiguity-neutral probability and increased in size as we moved away from it. If we were to represent the probabilities in a list, the probabilities that could potentially be queried against the event $\{\triangle, \bigcirc, \diamondsuit, \ddagger\}$ would be as shown in Figure 8 below. Hence, the procedure was very similar to the approach used in the main experiment, but it avoided any middle bias.

A disadvantage of the bisection method is that it takes more of the subjects' time. Because this second experiment was conducted online, it could not be too long; this is why we focused on replicating only the weakest findings of the main experiment.

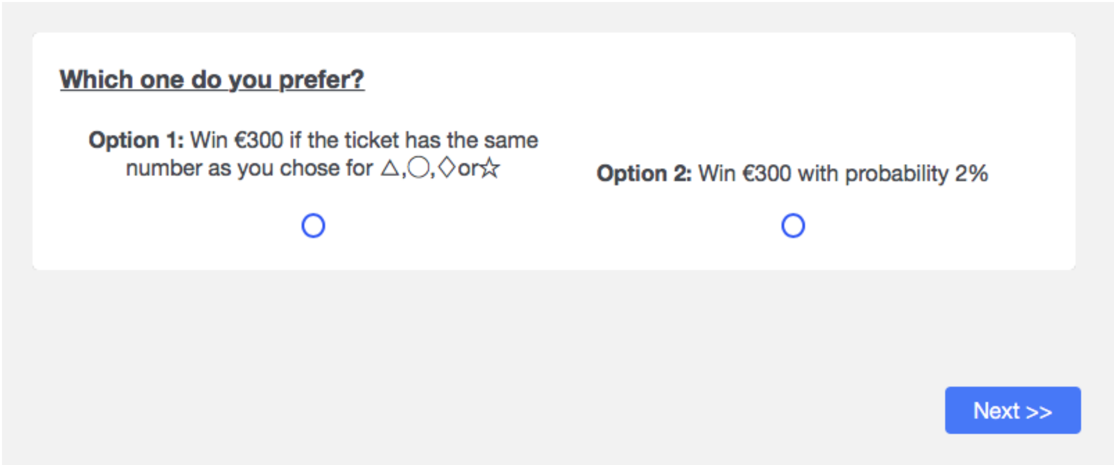


Figure 7: Screenshot of the first question for the event $\{\triangle, \bigcirc, \diamondsuit, \ddagger\}$

1.C.1.3 Incentives

The subjects were given the chance to play one of their choices for real. To prevent strategic behavior such as always choosing Option 1 to increase the probability in Option 2,[23] we did not randomly select one of their actual choices; instead, we randomly selected one question (e.g., as displayed in Figure 8), independently of whether this question was answered. Hence, the subjects' answers had no impact on which question was used to determine the payment. If the subject had answered the chosen question in the online experiment, the corresponding choice was directly implemented. If not, the choice was inferred from the answers given by the subject to the other questions using stochastic dominance[24] (see Johnson et al., 2014, for more details on this method).

Ultimately, 5% of the subjects (3 subjects) were randomly selected and invited to play for real. Which choice to play was again determined by envelopes that they drew upon arrival. In this case, there was only one box of envelopes containing 82 envelopes, inside which were written all possible event-probability pairs. The uncertain and risky bets were implemented with a bag and dice, as in the first experiment. Since the subjects filled out the questionnaire online, we could only tell them that they could win €300. We physically showed the €300 only to the subjects who were invited to play for real.

1.C.1.4 Matching probabilities

The matching probabilities were calculated in the same way as in the main experiment. We did not observe any subject to present counterintuitive answers, as in the first experiment. No subject

[23]This is a typical problem that arises when using random incentives with the bisection method.

[24]For example, if a subject preferred a 2% chance of winning €300 over a specific bet, it was inferred that they also preferred a 3% chance of winning €300 over the same bet.

Option 1	Option 2
Win €300 if the ticket has the same number as you chose for △, ○, ♢, ‡	Win €300 with probability 4%
	Win €300 with probability 3.5%
	Win €300 with probability 3%
	Win €300 with probability 2.5%
	Win €300 with probability 2.3%
	Win €300 with probability 2.2%
	Win €300 with probability 2.1%
	Win €300 with probability 2%
	Win €300 with probability 1.9%
	Win €300 with probability 1.8%
	Win €300 with probability 1.7%
	Win €300 with probability 1.5%
	Win €300 with probability 1%
	Win €300 with probability 0.5%
	Win €300 with probability 0%

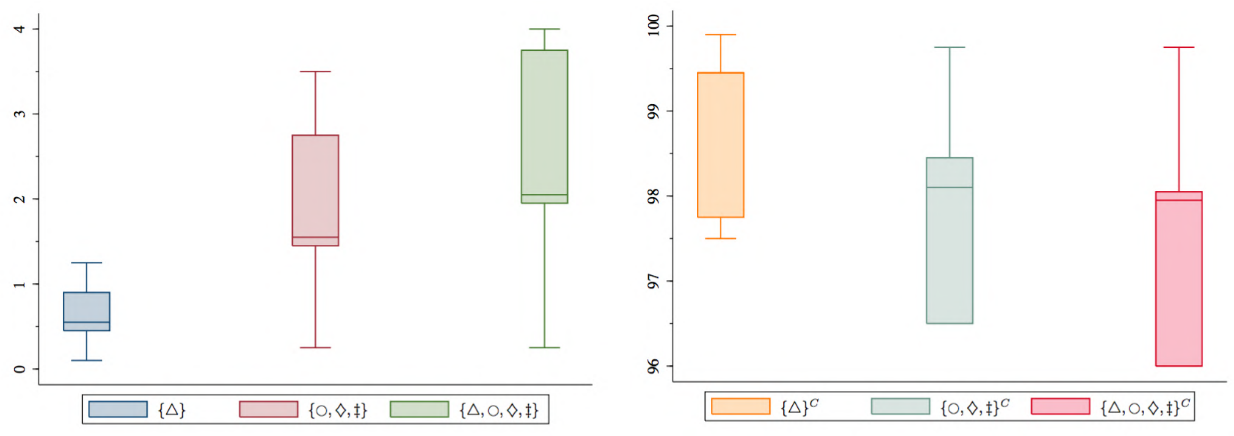
Figure 8: List representation of the potential probabilities for the event $\{\triangle, \bigcirc, \diamondsuit, \ddagger\}$

preferred a 0% probability of gain to some uncertain chance of winning or refused a 100% probability of gain. This suggests that we could eliminate mistakes that occurred as a consequence of using choice lists. However, the bisection method also introduces its own complications.

By the nature of the bisection design, the first question answered for each event determines the direction of the probabilities presented in Option 2 in the subsequent questions. For a subject who answers the first question mistakenly, it is not possible to change the direction back. Hence, to check whether any subjects were forced to answer questions in the direction opposite to their matching probabilities, we asked a 5th question for every event. This 5th question asked what they would be asked if they were to answer the first question for that event differently. We checked whether the answer to the 5th question was in the opposite direction from the answer to the first question. In this way, we could detect subjects who had potentially made a mistake on the very first question for a given event and for whom we had potentially miscalculated the matching probabilities. We excluded such observations from our analysis.[25]

1.C.2 Results

Figure 9 shows the distributions of the matching probabilities. Once again, we see that the majority of the subjects had matching probabilities equal to or greater than the probabilities an ambiguity-neutral subject with uniform beliefs would have in Panel a and lower matching probabilities in Panel b. These findings suggest the overweighting of rare events.



(a) very unlikely events (b) very likely events

Numbers are reported as percentages.
Every boxplot has lines at Q1, the median, and Q3. The adjacent lines show the most extreme values within 1.5 times the IQR of the nearer quartile.

Figure 9: Median matching probabilities for Experiment 2

The BC values differ significantly from 0 in all cases and always in the direction consistent with ambiguity aversion. Both the LA and UA indices are slightly decreased in magnitude compared with our previous analysis (see Table 2, where the median values are 0.15 for LA and 0.75 for

[25]Out of the 366 total observations, 20 were eliminated from the analysis in this way.

Table 7: Median analysis for Experiment 2

	Mean	Standard deviation	Median	z score	N
$BC(\{\triangle\})$	0.26	0.96	0.00 *	2.29	53
$BC(\{\bigcirc, \diamondsuit, \ddagger\})$	0.34	1.04	0.10 **	3.12	56
$BC(\{\triangle, \bigcirc, \diamondsuit, \ddagger\})$	0.13	0.99	0.00 *	2.10	55
$LA(\{\triangle\}, \{\bigcirc, \diamondsuit, \ddagger\})$	0.34	1.07	0.05 $^{+}$	1.90	52
$UA(\{\triangle\}, \{\bigcirc, \diamondsuit, \ddagger\}^C)$	0.87	1.07	0.55 ***	5.14	51

N represents the number of subjects in a particular test
Numbers are reported as percentages
Stars indicate the significance levels of the Wilcoxon signed-rank test
$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

UA); however, we still see evidence of both lower and upper subadditivity. Hence, we were able to replicate the weakest results of our main experiment even with a design that avoids middle bias.

Given the small number of index values extracted here, it is not possible to conduct an LPA using the data from our second experiment. However, we can use the results from the first experiment to gain a rough idea of what the results of such an analysis would be. Using the means and standard errors of the related indices $LA(\{\triangle\}, \{\bigcirc, \diamondsuit, \ddagger\})$ and $UA(\{\triangle\}, \{\bigcirc, \diamondsuit, \ddagger\}^C)$ for Profile 1 (see Table 3), we categorized the subjects into "Profile 1" and "others". We found that the proportions of subjects assigned to Profile 1 and others were respectively 79% and 21% for the LA index and 53% and 47% for the UA index.[26] Compared with our previous results, the proportion of subjects assigned to near ambiguity neutrality is higher for the LA index, whereas the result for the UA index is similar. This finding is not surprising since the only LA index considered here is the one that was the closest to lower additivity in the previous analysis. As before, we observe more upper subadditivity than lower subadditivity.

[26] A subject was assigned to Profile 1 with respect to a given index if their index value was strictly smaller than mean $+ 2 \times$ Standard Error. The numbers of subjects for whom we could calculate the LA and UA indices were 51 and 50, respectively.

Chapter 2

When risk perception gets in the way: Probability weighting and underprevention *

Joint with

Aurélien Baillon, Han Bleichrodt, Johannes Jaspersen, and Richard Peter

2.1 Introduction

Keeney (2008) pointed out that most deaths in the United States, and probably elsewhere, are attributable to people's personal decisions, such as smoking and eating too much or exercising too little. People undertake too little prevention (or self-protection[1]), which leads to increased mortality and impaired quality of life. Underprevention is a likely cause of the explosion in health care spending which takes up a sizable part of GDP in developed economies (e.g. 17.9% in the US in 2016[2]). Two possible reasons for underprevention are moral hazard, people expecting to get health care anyhow, and a misperception of the risks involved. The issue of moral hazard was discussed in depth by Arrow (1963) in his classical paper, which marked the start of health economics.

In that paper, Arrow first established a theory of "ideal insurance" and then discussed the impact of various market failures, like moral hazard. In a perfect setting, a risk-neutral insurer should adopt the insured's risk and the actions of the insured should not affect the probability of the risk occurring. Such a perfect setting is unrealistic for health risks, as pointed out by Arrow, especially due to information asymmetries. Arrow's analysis led to the widely-held belief that insurance and prevention are substitutes: more insurance leads to less self-protection. Ehrlich and Becker (1972) challenged this thesis by showing that insurance and prevention can be complements. To make their point, they introduced the first model of prevention.

*Published as "Baillon, A., H. Bleichrodt, A. Emirmahmutoglu, J.G. Jaspersen and R. Peter (forthcoming). When risk perception gets in the way: Probability weighting and underprevention. *Operations Research*."

[1]By prevention, we refer to activities that reduce the probability of a bad outcome. In the literature, it is sometimes also called self-protection or primary prevention.

[2]Centers for Medicare & Medicaid Services, see https://www.cms.gov/Research-Statistics-Data-and-Systems/Statistics-Trends-and-Reports/NationalHealthExpendData/NationalHealthAccountsHistorical.html

It is questionable whether moral hazard can explain underprevention. Health insurance only covers the financial aspects of poor health but avoiding poor health has value in itself (Cutler and Zeckhauser, 2000; Jaspersen and Richter, 2015).[3] This leaves misperceived risks as an alternative, more plausible explanation. Ehrlich and Becker and later prevention models assume perfectly rational, expected utility agents. This assumption is justified in economics by taking "rational behavior as a first approximation to actual behavior"(Arrow, 1951). Puzzling actual behavior, such as the underprevention highlighted by Keeney (2008), calls for going beyond the first, rational approximation. Arrow (1982) discussed how deviations from rationality (relaxing expected utility and Bayesianism) can explain anomalies such as the low uptake of flood insurance and financial-market puzzles. In the present paper, we study how people's imperfect perception of probabilities can lead to suboptimal prevention efforts.

Kahneman and Tversky (1979) demonstrated that people tend to be insufficiently sensitive to changes in probabilities, unless certainty or impossibility is involved, in which case they tend to overreact. Such behavior is described by an inverse S-shaped probability weighting function (Tversky and Kahneman, 1992). Prevention efforts differ from full insurance in that they reduce the risks of bad outcomes but do not entirely exclude them. Consequently, probability insensitivity blurs the benefits of prevention. The same probability weighting function that can lead to overinsurance can create underprevention. In only reducing the risks of bad outcomes, prevention is similar to probabilistic insurance, which as Kahneman and Tversky (1979) and Wakker et al. (1997) showed, is usually considered highly unattractive by agents.

We use a general setting, where the risk borne by the agent can affect his financial situation and/or his health and where the effects of effort need not be separable from the utility over wealth and health. We compare the prevention efforts of an agent with inverse S-shaped probability weighting with those of his rational twin, who has the same utility but does not weigh probabilities and behaves according to expected utility. The benchmark for optimal prevention is therefore personal and is determined by the agent's utility. Deviations from optimality are determined by his individual probability weighting function. Our paper characterizes the range of probabilities for which underprevention will occur. Combining the literature on probability weighting and the literature on prevention, our results indicate that in many cases, people's prevention efforts will be too low.

We conclude by considering the case where the probabilities of bad outcomes are not precisely

[3]Using a model of ex-ante moral hazard, Dionne (1982) showed that the effect of insurance coverage on prevention is less important in health insurance than in property insurance unless the marginal utility of money is significantly larger in the bad health state than in the good health state. Seog (2012) showed that when treatment has a preventive aspect, the effect of moral hazard will likely become even less important and insurance cover can actually increase prevention. This is also supported by the empirical evidence in Finkelstein et al. (2012) who showed that health insurance increases the amount of preventive care accessed by individuals. It is thus unlikely that health insurance is the key explanation for underprevention.

known. We show that in this case ambiguity amplifies the effects of probability weighting and leads to further underprevention. Our results imply that an important goal of health policy is to reduce probability weighting, for instance by diminishing the ambiguity surrounding prevention benefits. Helping people to correctly assess changes in probabilities can lead to more prevention and, consequently, to less future treatment.

2.2 A model of prevention and probability weighting

2.2.1 Optimal prevention

We consider an agent who faces a binary risk. Each possible outcome consists of two attributes, a financial and a health attribute. We write (x_g, h_g) for the outcome in the good state and (x_b, h_b) for the outcome in the bad state with x_g being preferred to x_b (that is $x_g \geq x_b$), h_g being preferred to h_b and at least one preference being strict.[4] The agent can engage in costly prevention efforts to reduce the probability of the bad outcome. We denote the level of prevention by $e \in [0, \overline{e}]$ and assume the probability of the bad outcome to be given by $p(e)$, where p is a decreasing function of e, that is $p'(e) < 0$. Consequently, the probability of the good outcome is $(1 - p(e))$ and is an increasing function of effort.

Preferences over wealth, health and effort are represented by a trivariate utility function $u(x, h; e)$ with $u_x > 0$, $u(x, h_g; e) \geq u(x, h_b; e)$ for any x and e, and $u_e < 0$, where subscripts denote partial derivatives. We take expected utility as the rational benchmark. The objective of the rational agent, indicated by superscript 'r', is then given by

$$\max_{e \in [0, \overline{e}]} U^r(e) = p(e)u(x_b, h_b; e) + (1 - p(e))u(x_g, h_g; e). \tag{1}$$

Due to continuity, U^r attains a maximum on $[0, \overline{e}]$, and we assume throughout the analysis that solutions are interior, in which case the associated first-order condition (FOC) holds:[5]

$$-p'(e^r)\left[u(x_g, h_g; e^r) - u(x_b, h_b; e^r)\right] + \left[p(e^r)u_e(x_b, h_b; e^r) + (1 - p(e^r))u_e(x_g, h_g; e^r)\right] = 0. \tag{2}$$

e^r denotes the rational agent's optimal level of prevention and p^r the associated probability of

[4] Our analysis relies on ordinal properties of the health variable only as in many applications health is not cardinal. The two attributes are positively correlated across states, which makes it clear which outcome is good and which one is bad. The techniques in this paper are also applicable when the two attributes correlate negatively. Distinguishing the good from the bad outcome then requires knowledge of the agent's preferences beyond monotonicity.

[5] If no effort were optimal, probability weighting could only lead to more effort, and if the maximum possible effort were optimal, probability weighting could only lead to less effort. In both cases, the question we ask in this paper has a trivial answer for lack of economic trade-off.

the bad outcome, that is $p^r = p(e^r)$. The first term in Eq. (2) represents the marginal benefit of prevention in the sense that effort increases the odds of the good state, which raises expected utility. The second term in Eq. (2) represents the marginal cost of prevention in the sense that effort is costly, which lowers expected utility. In equilibrium, the marginal benefit equals the marginal cost, and the agent has no incentive to deviate from his level of prevention, consistent with his choice being optimal.

2.2.2 Probability weighting

Prevention reduces risks by lowering the probability of the bad state but it typically does not eliminate them. Therefore agents need to compare two different risky situations to determine the value of prevention. Such comparisons are susceptible to behavioral biases which lead to violations of expected utility. Probability weighting can explain several deviations from expected utility. Before we explore the effects of probability weighting on optimal prevention, we recall the core properties of probability weighting functions and present some common classes of probability weighting functions. We start with the following standard definition.

Definition 1. *A* probability weighting function *is a strictly increasing function w defined on probabilities p in $[0, 1]$ with $w(0) = 0$ and $w(1) = 1$.*

Definition 1 states that probability weighting functions respect monotonicity. They allow to introduce nonlinearity in probability, a common empirical regularity of choice under risk (see Camerer, 1989; Wu and Gonzalez, 1996). We state two more properties.

Definition 2. *A probability weighting function is called:*

(i) Regressive *if it intersects the diagonal only once and from above.*

(ii) Cavex *if it is first concave and then convex.*

If a probability weighting function satisfies both (i) and (ii), we call it inverse S-shaped.

The term "regressive" to characterize Property (i) was used by Prelec (1998). This property is also commonly referred to as overweighting of small and underweighting of large probabilities, with a fixed point $p^* \in (0, 1)$ that separates small from large probabilities. This fixed point is about $1/3$, a property that Prelec (1998) called asymmetry. Property (ii) implies that changes in probability have less impact as one moves away from the endpoints of the unit interval. If it holds, we can find an inflection point $\widetilde{p} \in (0, 1)$ that separates the concave from the convex portio of the probability weighting function. Prelec (1998) called Property (ii) 'S shaped'. In the literature, inverse S-shaped is now more commonly used than S-shaped but authors do not always distinguish Property (i) and Property (ii). We use Wakker's (2010) term cavexity for Property (ii) to avoid confusion. Notice that the two properties are not equivalent because there are regressive probability weighting functions with multiple changes in curvature and cavex probability weighting functions that do not intersect the diagonal.

There are many classes of probability weighting functions that are inverse S-shaped for certain parameter values, and we list the most common ones in Table 1. The TK form was proposed by Tversky and Kahneman (1992). Monotonicity requires $\alpha \geq 0.28$, and an inverse S-shape holds for $\alpha < 1$. The case $\alpha = 1$ corresponds to expected utility because then the probability weighting function is the identity, $w(p) = p$. Goldstein and Einhorn (1987, GE) introduced a generalized two-parameter specification of this function.[6] Here, the parameter α primarily controls curvature whereas parameter β primarily controls elevation. Monotonicity holds for any $\alpha, \beta > 0$, and the probability weighting function is inverse S-shaped whenever $\alpha < 1$. Starting from an axiomatization, Prelec (1998) defined two probability weighting functions. The one-parameter function Prelec-1 is monotonic for $\alpha > 0$ and inverse S-shaped for $\alpha < 1$. For this class of probability weighting functions the fixed point coincides with the inflection point at $1/e$. In the corresponding two-parameter version Prelec-2, the parameter α again controls curvature and β elevation. It is monotonic for $\alpha, \beta > 0$ and inverse S-shaped for $\alpha < 1$.

Abbreviation	Reference	Equation
TK-92	Tversky and Kahneman (1992)	$w(p) = \frac{p^\alpha}{[p^\alpha + (1-p)^\alpha]^{\frac{1}{\alpha}}}$
GE	Goldstein and Einhorn (1987)	$w(p) = \frac{\beta p^\alpha}{\beta p^\alpha + (1-p)^\alpha}$
Prelec-1	Prelec (1998)	$w(p) = \exp(-(-\ln p)^\alpha)$
Prelec-2	Prelec (1998)	$w(p) = \exp(-\beta(-\ln p)^\alpha)$
Neo	Wakker (2010)	$w(p) = \begin{cases} 0, & \text{if } p = 0, \\ \frac{\alpha-\beta}{2} + (1-\alpha)p, & \text{if } p \in (0,1), \\ 1, & \text{if } p = 1, \end{cases}$

Table 1: Common parametric forms of $w(p)$ as put forward in the literature.

All these examples of probability weighting functions are continuous on $[0, 1]$ and are twice differentiable on $(0, 1)$, even in those cases when their shape is not inverse S. Because these conditions will be helpful in our analysis, we introduce the following definition.

Definition 3. *A probability weighting function is called* regular *if it is continuous on* $[0, 1]$ *and twice differentiable on* $(0, 1)$.

We use this definition to obtain the following Lemma, which introduces an important region in case of an inverse S-shaped probability weighting function.[7]

[6]See also Lattimore et al. (1992), Tversky and Fox (1995) and Gonzalez and Wu (1999).

[7]All mathematical proofs are in Appendix **??**.

Lemma 1. *For any regular inverse S-shaped probability weighting function w, we can find $p_1 \in (0, \min\{p^*, \widetilde{p}\})$ and $p_2 \in (\max\{p^*, \widetilde{p}\}, 1)$ such that $w'(p) > 1$ for $p \in (0, p_1) \cup (p_2, 1)$, $w'(p) = 1$ for $p \in \{p_1, p_2\}$ and $w'(p) < 1$ for $p \in (p_1, p_2)$.*

Inverse S-shaped probability weighting functions capture the idea of likelihood oversensitivity for probabilities near the corners of the unit interval and likelihood insensitivity for intermediate probabilities. We dub (p_1, p_2) the *likelihood insensitivity region* of probability weighting function w.[8] The Lemma shows that this region covers at least all probabilities between the fixed point and the inflection point. l'Haridon and Vieider (2018) elicited probability weighting functions in 30 countries all over the world. On a global average, the insensitivity region according to their parameter estimates ranges from 8% to 84%. Table 5 in Appendix 2.B displays the likelihood insensitivity regions for each country individually. Table 6 in Appendix 2.B gives the likelihood insensitivity regions obtained from some other studies.[9] For instance, Booij et al. (2010), reported weighting functions for a large, representative sample of the Dutch population. Their insensitivity region is $[9\%, 86\%]$ (using Prelec's two-parameter function).

A probability weighting function that pushes the idea of likelihood insensitivity to the limit is the neo-additive class shown in the last row of Table 1 (Neo, see Wakker, 2010). Its parameters are $\alpha \in (0, 1)$ and $\beta \in (-\alpha, \alpha)$. Parameter α captures the slope and was called the insensitivity index by Abdellaoui et al. (2011). It is 0 for an expected utility agent and converges to 1 for an individual who perceives all probabilities to be the same, e.g., 50-50. Parameter β is an index of pessimism, measuring the average elevation of the weighting function. It is also 0 under expected utility. An individual with extreme pessimism, acting as if the worst is always almost certain, has a β parameter converging to -1. Several papers have emphasized the importance of this probability weighting function (e.g., Cohen, 1992; Loomes et al., 2002; Chateauneuf et al., 2007; Teitelbaum, 2007), and Wakker (2010) argued that it provides an optimal trade-off between parsimony and fit. It is regressive with overweighting of probabilities smaller than $(\alpha - \beta)/(2\alpha)$ and underweighting of probabilities larger than $(\alpha - \beta)/(2\alpha)$. It is not regular, and we cannot apply Lemma 1. It is, however, immediate that $w'(p) = (1 - \alpha) < 1$ for any $p \in (0, 1)$ so that the insensitivity region is given by the entire open unit interval, $(0, 1)$. In our analysis of optimal prevention we will discuss

[8] By defining the insensitivity region based on the derivative of w, we deviate from Wakker (2010) who defined the likelihood insensitivity region based on the properties of lower and upper subadditivity (Tversky and Wakker, 1995). Wakker's definition is also applicable to non-differentiable functions w. Our definition of the insensitivity region captures the same intuition, but it is not equivalent to Wakker's definition. It is possible to construct counterexamples where the region based on $w'(p) < 1$ is not a likelihood insensitivity region in the sense of Wakker (2010). The definition of Wakker (2010) does not rely on derivatives and therefore, does not rely on infinitesimal changes. This feature is important for empirical work in which infinitesimal changes cannot be observed. However, Wakker's definition is not well-suited for the type of analysis conducted in this paper, which precisely makes use of derivatives. See footnote 10 for an example where our definition is better suited to the analysis of prevention.

[9] We selected studies that elicited weighting functions for losses (either independently or combined with gains).

neo-additive weighting functions separately.

2.3 Applying probability weighting to prevention decisions

2.3.1 Some general results

We will now return to the problem of choosing an optimal level of prevention. To facilitate comparability, we study an agent who is identical to the rational agent in terms of the binary risk exposure, the prevention technology and utility over wealth, health and effort. The only difference is that his perception of risk is affected by probability weighting, which we indicate by superscript 'w'. Given that the good state gives higher utility than the bad state and that prevention aims to mitigate the potential loss of utility, the objective function is now given by:

$$\max_{e\in[0,\overline{e}]} U^w(e) = w(p(e))u(x_b, h_b; e) + (1 - w(p(e)))u(x_g, h_g; e). \tag{3}$$

The first-order condition for an optimal solution is

$$\begin{aligned} U_e^w(e^w) \; = \; & - w'(p(e^w))p'(e^w)\left[u(x_g, h_g; e^w) - u(x_b, h_b; e^w)\right] \\ & + \left[w(p(e^w))u_e(x_b, h_b; e^w) + (1 - w(p(e^w)))u_e(x_g, h_g; e^w)\right] = 0, \end{aligned} \tag{4}$$

where e^w denotes the optimal level of prevention under probability weighting. When comparing Eq. (2) and (4), we see that probability weighting affects both the marginal benefit and the marginal cost of prevention. For probabilities in the likelihood insensitivity region, the marginal benefit of prevention is lower than for the rational agent because the agent underappreciates the reduction in the probability of the bad state. The reverse is true for probabilities in the likelihood oversensitivity region, for which the agent overvalues the reduction of the probability of the bad state. The effect of probability weighting on the marginal cost of prevention is unclear and depends on the effect of effort on utility. If the probability of the bad state is overweighted, then there is more weight on the marginal cost of effort in that state, and if the probability of the bad state is underweighted, then the marginal cost of effort in the good state is overweighted. However, it is not clear in which state the marginal effect of effort on utility is larger, so the overall effect remains inconclusive.

We can still compare the rational agent's level of prevention with that under probability weighting, depending on the magnitude of the rational agent's equilibrium probability. Technically, our proof relies on $U^w(e)$ being concave on $[0, \overline{e}]$, which we assume to hold. We summarize our result in the following proposition.

Proposition 1. *Assume a regular inverse S-shaped probability weighting function. We can find* $\underline{p} \in (0, \min\{p^*, \widetilde{p}\})$ *and* $\overline{p} \in (\max\{p^*, \widetilde{p}\}, 1)$ *such that probability weighting lowers prevention if* *and only if the rational agent's probability of the bad state lies between* $\underline{p}$ *and* $\overline{p}$*.*

We dub $(\underline{p}, \overline{p})$ the *underprevention region* because in this region probability weighting leads to less preventive effort compared to the rational benchmark. The consequence is too much risk exposure caused by the agent's distorted perception of the costs and benefits of prevention. Proposition 1 says that probability weighting results in underprevention for intermediate loss probabilities. Closer to the corners of the unit interval the reverse occurs, which is why we call $(0, \underline{p}) \cup (\overline{p}, 1)$ the *overprevention region*.

Notice the symmetry between Lemma 1 and Proposition 1. For intermediate loss probabilities, inverse S-shaped probability weighting functions exhibit likelihood insensitivity and lead to underprevention compared to the rational agent. Likewise, for probabilities towards the boundaries of the unit interval, inverse S-shaped probability weighting functions exhibit likelihood oversensitivity and lead to overprevention compared to the rational agent. This raises the question how the likelihood insensitivity region and the underprevention region are related. The following proposition addresses this question and shows that its answer depends on the marginal utility of effort in the good and in the bad state.

Proposition 2. *Under the assumptions in Proposition 1 it holds that:*

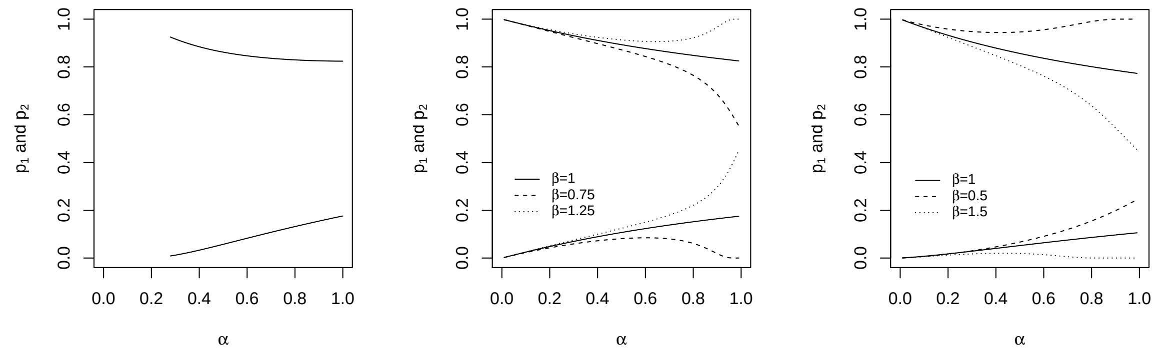
(i) $\underline{p} < p_1 < \overline{p} < p_2$ *if the marginal cost of effort is higher in the bad state than in the good state.*

(ii) $\underline{p} = p_1$ *and* $\overline{p} = p_2$ *if the marginal cost of effort is the same in both states.*

(iii) $p_1 < \underline{p} < p_2 < \overline{p}$ *if the marginal cost of effort is higher in the good state than the bad state.*

Proposition 2 informs us that the likelihood insensitivity region and the underprevention regions coincide if and only if it is equally costly at the margin to exert effort in the good state and the bad state.[10] Intuitively, if the marginal cost of prevention does not depend on the realization of the binary risk, it is also unaffected by the perception of this risk. In other words, probability weighting only alters the marginal benefit of prevention, and our discussion of Eq. (4) already identified likelihood in- and oversensitivity as the drivers for a lower or higher marginal benefit compared to the rational agent. Table 5 in Appendix 2.B gives examples of insensitivity regions for the Prelec-2 function. To further illustrate Lemma 1 and Proposition 2(ii), we determine the likelihood insensitivity region numerically for the different classes of regular inverse S-shaped probability weighting functions in Section 2.2.2. Figure 1 provides an overview. For the TK probability weighting function, an increase in the parameter α increases p_1 and decreases p_2 so that

[10]Here we see an advantage of defining the insensitivity region from derivatives. With Wakker's (2010) definition, we would not obtain such equivalence between insensitivity and underprevention when the marginal cost of effort is constant across states.



(a) TK weighting function (b) GE weighting function (c) Prelec-2 weighting function

Figure 1: Likelihood insensitivity regions for different probability weighting functions

the likelihood insensitivity region is contracted. Unlike the TK weighting function, the GE function and Prelec-2 function separate curvature and elevation. The parameter β controls elevation in both classes with higher values of β corresponding to more elevation in the GE and less elevation in the Prelec-2 probability weighting functions. Panels (b) and (c) show that more elevation increases both p_1 and p_2 so that the likelihood insensitivity region shifts to the right. At low elevation (low β for GE and high β for Prelec-2), lowering α first increases p_1 and then reduces it whereas it always increases p_2. Thus, an increase in curvature first moves the likelihood insensitivity region to the left and then expands it. Likewise, at high elevation (high β for GE and low β for Prelec-2), lowering α decreases p_1 whereas it first decreases p_2 and then increases it. Here, an increase in curvature first moves the likelihood insensitivity region to the left and then expands it.

If the marginal cost of effort is affected by the risk, the underprevention region will either be more to the left or more to the right than the likelihood insensitivity region and its width can also differ. To further explore this issue, we denote the difference between the marginal cost of effort in the bad state and the good state by d, that is $d = |u_e(x_b, h_b; e^r)| - |u_e(x_g, h_g; e^r)|$. Whether this difference is positive or negative depends on the underlying preferences, and Table 2 provides an overview.

	$u_{xe} < 0$	$u_{xe} = 0$	$u_{xe} > 0$
$u_e(x, h_g; e) < u_e(x, h_b; e)$	$-$	$-$	$+/-$
$u_e(x, h_g; e) = u_e(x, h_b; e)$	$-$	0	$+$
$u_e(x, h_g; e) > u_e(x, h_b; e)$	$+/-$	$+$	$+$

Table 2: Sign of d as a function of preferences

If bad financial and/or bad health outcomes make efforts more costly at the margin, we are in case (i) of Proposition 2 and the underprevention region is located to the left of the likelihood insensitivity region. We believe that this case is the most plausible as will be explained in Section 2.4.1 when we discuss how effort may affect the agent's utility derived from wealth and health. In this case, likelihood insensitivity is neither necessary nor sufficient for underprevention as Figure 2 illustrates. We can find probabilities between $\underline{p}$ and p_1 where underprevention occurs despite the agent's likelihood oversensitivity and we can also find probabilities between $\overline{p}$ and p_2 where overprevention occurs despite the agent's likelihood insensitivity.

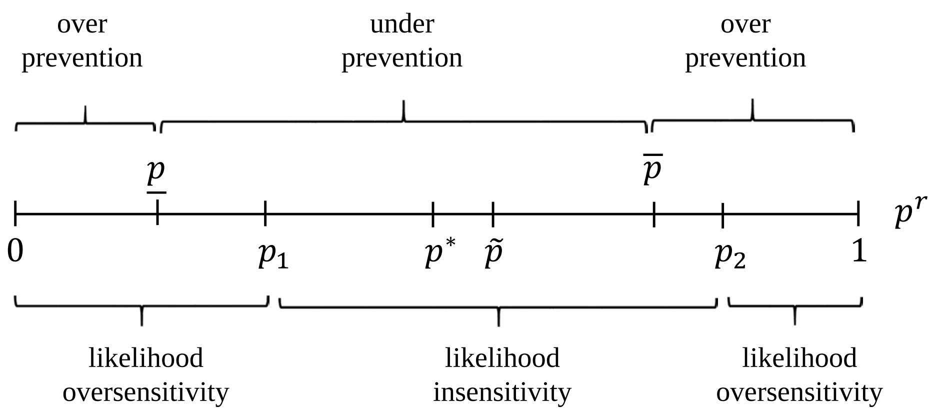


Figure 2: Likelihood insensitivity and underprevention region in Proposition 2(i)

2.3.2 Comparative statics

Proposition 1 characterizes how probability weighting affects the agent's prevention decision by identifying the underprevention region. Proposition 2 shows that the location of the underprevention region relative to the likelihood insensitivity region depends on how the marginal cost of effort varies with the state of nature. To gain further insight into the determinants of the underprevention region, we investigate how it is affected by the exogenous parameters of our model.

We focus on case (i) in Proposition 2, but similar techniques apply to case (iii). Let $\kappa = u_e(x_b, h_b; e^r)/u_e(x_g, h_g; e^r)$ be the ratio of the marginal cost of effort in the bad state to the marginal cost of effort in the good state. It measures how much more costly effort is at the margin when going from the good to the bad state of nature. We are in case (i) if $u_{xe} \geq 0$ and $u_e(x, h_g; e) \geq u_e(x, h_b; e)$ with at least one inequality strict (see Table 2). Then $\kappa > 1$. Several of our comparative statics predictions require comparing effort elasticities, which we define next.

Definition 4. *We introduce the following effort elasticities:*

(i) $\varepsilon_p(e) = -e\frac{p'(e)}{p(e)}$ *for the loss probability.*

(ii) $\varepsilon_{u_x}(x, h; e) = e\frac{u_{xe}(x,h;e)}{u_x(x,h;e)}$ *for the marginal utility of wealth.*

(iii) $\varepsilon_{\Delta h}(x, h_g, h_b; e) = e\frac{u_e(x,h_g;e)-u_e(x,h_b;e)}{u(x,h_g;e)-u(x,h_b;e)}$ *for the utility gain from better health.*

(iv) $\varepsilon_{u_e}(x, h; e) = e\frac{u_{ee}(x,h;e)}{u_e(x,h;e)}$ *for the marginal cost of effort.*

For the loss probability, the marginal utility of wealth and the marginal cost of effort, we use point elasticities because we can take derivatives with respect to the underlying attributes. The effort elasticity of the loss probability has a natural interpretation because it measures how a percentage increase in effort translates into a percentage decrease of the probability of the bad state.[11] The effort elasticities of the marginal utility of wealth and of the marginal cost of effort will be discussed in further detail in Section 2.4.1. For health, we define an arc elasticity to reflect that health may be an ordinal variable. Our results are summarized in the following proposition.

Proposition 3. *Assume a regular inverse S-shaped probability weighting function.*

(i) The underprevention region $(\underline{p}, \overline{p})$ *is negatively associated with* κ.

Furthermore, let $u_{xe} \geq 0$ *and* $u_e(x, h_g; e) \geq u_e(x, h_b; e)$ *with at least one inequality strict (case (i) of Proposition 2), and assume that the marginal cost of effort is more elastic in the bad state than in the good state (i.e.,* $\varepsilon_{u_e}(x_b, h_b; e) \geq \varepsilon_{u_e}(x_g, h_g; e)$*). Then:*

(ii) An increase in x_g *shifts the underprevention region to the left.*

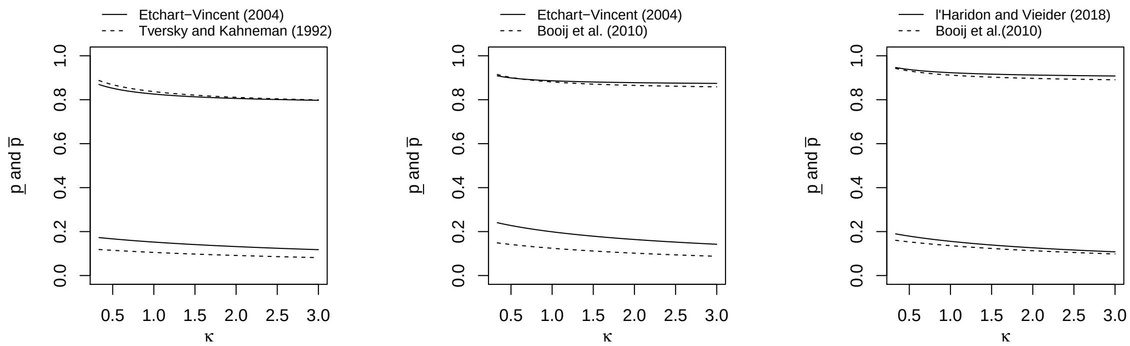
(iii) An increase in x_b *shifts the underprevention region to the right if* $\varepsilon_p(e) > \varepsilon_{u_x}(x, h; e)$.

(iv) An improvement in h_g *shifts the underprevention region to the left.*

(v) An improvement in h_b *shifts the underprevention region to the right if* $\varepsilon_p(e) > \varepsilon_{\Delta h}(x, h_g, h_b; e)$.

Proposition 3(i) says that the underprevention region moves to the left if exerting effort becomes more costly in the bad than the good state. It is therefore more likely to contain small probabilities, which is typical of many real-world prevention problems. A similar phenomenon occurs when comparing small to large risks. Assume a spread in consumption levels in the good and the bad state, that is an increase in x_g and a decrease in x_b. Proposition 3(ii) and (iii) say this also shifts the underprevention region to the left so that it contains more small probabilities. In other words, risks with more severe financial consequences are more likely to result in underprevention when the agent weighs probabilities. Similarly, a spread in health, an improvement in h_g coupled with a deterioration in h_b, shifts the underprevention region to the left as implied by results

[11] Hofmann and Peter (2015) and Courbage et al. (2017) showed the usefulness of this measure in the comparative statics analysis of several intertemporal prevention problems.



(a) TK weighting function (b) GE weighting function (c) Prelec-2 weighting function

Figure 3: Underprevention region as a function of κ for different probability weighting functions

(iv) and (v) of Proposition 3. Again, more severe health risks are more prone to underprevention when the agent weighs probabilities. Underprevention tends to be more concerning for severe than for mild risks.

We provide an illustration in Figure 3. Given that κ is a complex function of the rational agent's preferences over wealth and health, the endowments in both attributes and the available prevention technology, we opt for a stylized analysis by letting κ vary between $1/3$ and 3. We use the parameter estimates from Tversky and Kahneman (1992), Etchart-Vincent (2004), l'Haridon and Vieider (2018), and Booij et al. (2010) for the TK, the GE and the Prelec-2 probability weighting functions. Consistent with Proposition 3, an increase in κ lowers both $\underline{p}$ and $\overline{p}$ so that the underprevention region as a whole shifts to the left. The lines are mildly convex so that changes in κ have a larger effect when κ is low than when it is high. This also implies that changes in κ affect the width of the underprevention region, not only its location. The strongest effect of changes in κ occurs for Etchart-Vincent's parameter estimates of the GE probability weighting function. Overall though, changes in the parameters of the probability weighting function (see Figure 1) appear to have a stronger influence on the location of the insensitivity region, and hence the location of the underprevention region, than changes in κ. This highlights the importance of probability weighting in explaining underprevention.

2.3.3 Neo-additive probability weighting

As mentioned in Section 2.2.2, neo-additive probability weighting functions maximize the likelihood insensitivity region and are flatter than the identity function on the entire open unit interval. They could be classified as weakly cavex and do not satisfy regularity due to discontinuities at 0 and 1. As a result, Propositions 1 to 3 do not apply, and we study the effect of neo-additive probability weighting on optimal prevention separately in this section. The following proposition summarizes our results.

Proposition 4. *Let w be a neo-additive probability weighting function as defined in the last row of Table 1 with parameters $\alpha \in (0, 1)$ and $\beta \in (-\alpha, \alpha)$. Then:*

(i) Probability weighting leads to underprevention for all p^r.

(ii) An increase in α (likelihood insensitivity) leads to more underprevention.

(iii) A decrease in β (pessimism) leads to more underprevention if and only if the marginal cost of effort is higher in the bad state than the good state.

Neo-additive probability weighting affects both the marginal benefit and the marginal cost of prevention relative to the rational agent. The marginal benefit is reduced by $(1 - \alpha)$ due to likelihood insensitivity. The effect on the marginal cost is twofold: It is also reduced by a factor of $(1 - \alpha)$ but there is an additional effect arising from over- and underweighting of probabilities. As a result, the marginal cost of prevention under neo-additive probability weighting can be higher or lower than that of the rational agent. Part (ii) of Proposition 4 shows, however, that even when the marginal cost of prevention falls, it will not outweigh the drop in the marginal benefit. Consequently, an increase in α always leads to an increase in underprevention.

The parameter α measures likelihood insensitivity in neo-additive probability weighting functions, see Baillon et al. (2017). Part (ii) of Proposition 4 shows that more likelihood insensitivity leads to more underprevention. The parameter β can be interpreted as a measure of pessimism, with lower values corresponding to more pessimism and more overweighting of the probability of a bad outcome (Baillon et al., 2017). Pessimism does not affect the marginal benefit of prevention, but it reduces the marginal cost of prevention whenever the marginal cost of effort is higher in the bad state than the good state. This results in more underprevention.

Neo-additive probability weighting functions have specific characteristics that put our previous results into perspective. The underprevention region is maximal and equal to $(0, 1)$. It does *not* depend on the parameters of the neo-additive probability weighting function, unlike in the case of regular inverse S-shaped probability weighting functions. The amount of underprevention does depend on the parameters of the neo-additive probability weighting function, but the underprevention region always coincides with the likelihood insensitivity region, independent of how the marginal cost of effort varies with the state of nature. Comparative statics questions about its location therefore become vacuous. If we view neo-additive probability weighting functions as linear approximations of more complex inverse S-shaped probability weighting functions, we can interpret underprevention as the first-order effect of likelihood insensitivity. The existence of non-empty overprevention regions at the corners of the unit interval is a direct consequence of regularity, and specifically continuity, of all parametric inverse S-shaped probability weighting functions. This continuity "forces" the probability weighting function to exhibit likelihood oversensitivity close to the boundaries 0 and 1, which increases the marginal benefit of prevention relative to the rational agent. If likelihood insensitivity is the main characteristic of an agent's risk perception, as the neo-additive weighting function suggests, underprevention will imply that the agent faces too much risk.

2.4 Applications

2.4.1 Specific cost structures

Our model of optimal prevention in Eq. (1) is very general in that it allows for both financial and non-financial attributes of utility as well as a non-separable effect of effort on utility. Most extant research on optimal prevention imposes more specific assumptions. We discuss some of these more restrictive specifications and explain how we can obtain them as special cases of our analysis.

Disutility. If the cost of effort is measured in units of utility, we can write

$$u(x, h; e) = v(x, h) - \phi(e), \tag{5}$$

where v is utility over wealth and health and $\phi(e)$ is a measure of the disutility of effort with $\phi'(e) > 0$. Such a specification is also referred to as a separable or non-tangible cost of effort. A special case is that of intertemporal prevention where an agent incurs a monetary cost of effort in the first period to mitigate a future risk exposure that is not resolved until the second period (see Menegatti, 2009). It is immediate that the marginal cost of effort does not depend on the state of the world in such a situation, and we are in case (ii) of Proposition 2. The underprevention region coincides with the likelihood insensitivity region for any regular inverse S-shaped probability weighting function. Questions of comparative statics with respect to determinants of the rational agent's choice problem become vacuous in such a case.

Monetary cost, monetary risk. If the cost of effort is monetary with an increasing cost function $c(e)$ and the health attribute is constant (that is $h_g \sim h_b$), we obtain

$$u(x, h; e) = v(x - c(e)). \tag{6}$$

This case is also referred to as a non-separable or tangible cost of effort. Peter (2017) showed that intertemporal prevention problems with an upfront monetary cost lead to the same results as two-period prevention problems with endogenous saving. In this situation, the marginal cost of effort is larger in the bad state than in the good state for a risk-averse agent because

$$-u_e(x_b, h_b; e) = c'(e) \cdot v'(x_b - c(e)) > c'(e) \cdot v'(x_g - c(e)) = -u_e(x_g, h_g; e) \tag{7}$$

whenever v is concave. Thus, we are in case (i) of Proposition 2 and the underprevention region is to the left of the likelihood insensitivity region for any regular inverse S-shaped probability weighting function. The effort elasticities of the marginal utility of wealth and the marginal cost of effort simplify to

$$\varepsilon_{u_x}(x,h;e) = ec'(e)\left[-\frac{v''(x-c(e))}{v'(x-c(e))}\right] \text{ and } \varepsilon_{u_e}(x,h;e) = e\frac{c''(e)}{c'(e)} + ec'(e)\left[-\frac{v''(x-c(e))}{v'(x-c(e))}\right], \tag{8}$$

and are directly related to the Arrow-Pratt measure of absolute risk aversion. Requiring the marginal cost of effort to be more elastic in the bad than in the good state amounts to assuming that utility function v exhibits decreasing absolute risk aversion, a familiar condition in utility theory about which Arrow (1971, pp. 96) wrote that it "seems supported by everyday observation."

Monetary cost. If health does matter, we can write

$$u(x,h;e) = v(x-c(e),h), \tag{9}$$

where v is utility over wealth and health. The marginal cost of effort is larger in the bad state than in the good state if and only if

$$v'(x_b - c(e), h_b) - v'(x_g - c(e), h_b) > v'(x_g - c(e), h_g) - v'(x_g - c(e), h_b). \tag{10}$$

The left-hand side of Eq. (10) is the risk aversion effect encountered before. It implies that the marginal cost of effort is higher in the bad state than in the good state. The right-hand side compares the marginal utility of wealth in good and in bad health. The extant research suggests that the marginal utility of wealth is increasing in health for severe injuries (Viscusi and Evans, 1990; Sloan et al., 1998; Finkelstein et al., 2013) but decreasing in health for minor injuries (Evans and Viscusi, 1991). In the first situation the risk aversion effect and the effect of health changes on the marginal utility of consumption have opposite effects and no clear conclusion can be drawn. The second situation corresponds to case (i) of Proposition 2 and, consequently, the underprevention region is to the left of the likelihood insensitivity region for any regular inverse S-shaped probability weighting function.

Disutility and monetary cost. We can also combine a monetary and a non-monetary cost of effort so that

$$u(x,h;e) = v(x-c(e),h) - \phi(e). \tag{11}$$

This case is similar to the previous one. Whether the marginal cost of effort is higher in the bad or in the good state of nature depends on whether the risk aversion effect exceeds the effect of health state on the marginal utility of wealth. Given the above considerations, case (i) of Proposition 2 seems to be applicable to several relevant cases and we thus deem it as the most plausible for most applications.

2.4.2 Mortality

Our model of optimal prevention can be interpreted in terms of mortality reduction. If h_g denotes the "alive" state and h_b the "deceased" state, we can interpret $u(x, h_g; e)$ as consumption utility when alive and $u(x, h_b; e)$ as the utility of bequests. All our results continue to hold when allowing $u_x(x, h_b; e) \geq 0$, that is $u_x(x, h_b; e) = 0$ is not excluded (see Eeckhoudt and Hammitt, 2001). When assuming $u_x(x, h_g; e) > u_x(x, h_b; e)$, as is usual in the Value of a Statistical Life literature (see Jones-Lee, 1974), the marginal cost of effort will be larger in the alive state than in the deceased state whenever the bequest motive is sufficiently mild. Then, by Proposition 2 case (iii), the underprevention region lies to the right of the likelihood insensitivity region for any regular inverse S-shaped probability weighting function.

2.4.3 Other Reference Points

In Tversky and Kahneman's (1992) prospect theory, the probabilities of the most extreme outcomes are weighted first. Consider a binary risk yielding a large loss with probability p and a smaller loss otherwise. The utility of the large loss will be weighted by $w(p)$ and the utility of the smaller loss gets the remainder, $(1 - w(p))$. Equation (3) can be interpreted as prospect theory for losses, with $(x_g, h_g; 0)$ being the reference point. This is a rather optimistic reference point (ending up in the good state without any prevention). Then, $(x_b, h_b; e)$ is the most extreme loss, but $(x_g, h_g; e)$ is a loss as well. In this model, w is the weighting function for losses, and we based our empirical illustrations such as Figure 1 and Appendix 2.B on studies that elicited w for losses. A pessimistic reference point could be $(x_b, h_b; \bar{e})$, that is, being in the bad state in spite of having provided maximum effort. With this reference point, the model should be translated to the gain domain. In that case, the probability of the best outcome is weighted first. Hence, $w(p)$ in Equation (3) should be replaced by $1 - w^+(1 - p)$, with w^+ the weighting function for gains. Note that if $w^+(p)$ is inverse S-shaped, then $1 - w^+(1 - p)$ is also inverse S-shaped (as a function of p). Consequently, all our theoretical results still hold after proper recoding of w and p in terms of w^+ and $(1 - p)$, but the values obtained in the empirical illustrations may (slightly) differ.

One may wonder whether agents really consider all situations as losses (or all as gains) and what our results would look like in a mixed framing. Even though we cannot answer this question in general (for *any* reference point), we can study a form of stochastic reference points introduced by Kőszegi and Rabin (2007), a reference point that includes various possible outcomes. Kőszegi and Rabin (2007) put forward a model of a stochastic reference point in combination with so-called reference-dependent preferences with loss aversion. We will briefly explain how we can derive versions of our results under their model using their concept of a choice acclimated personal equilibrium. According to Kőszegi and Rabin (2007), this equilibrium is appropriate if the agent has enough time between the prevention decision and the resolution of uncertainty, that the decision itself becomes the (stochastic) reference point. Specifically, an agent providing effort e would use $(x_b, h_b; e)$ as reference point with probability $p(e)$ and $(x_g, h_g; e)$ with probability $(1 - p(e))$. The agent will therefore experience a gain with probability $p(e)(1 - p(e)$ (the reference point is $(x_b, h_b; e)$ but he ends up in the good state) and a loss with the same probability (the reference point is $(x_g, h_g; e)$ but he ends up in the bad state). We assume axiom A3' of Kőszegi and Rabin (2007) such that the agent decides based on the weighted sum of consumption utility and piece-wise linear gain-loss utility with loss aversion λ and weight η. In this setting, Barseghyan et al. (2013) showed that the agent's preferences can be represented as those of a probability weighting agent with probability weighting function $w(p) = p(1 + \eta(\lambda - 1)(1 - p))$. Since loss aversion implies $\lambda > 1$, this probability weighting function is concave on $[0, 1]$ and likelihood insensitive for $p > 0.5$. Even though the exact underprevention region will depend on the difference between the marginal cost of effort in the bad state and in the good state, we can see that reference dependent preferences will likely only lead to underprevention if the rational agent's probability of the bad state is relatively high. The reference-dependent loss aversion model of Kőszegi and Rabin (2007) is thus less suitable to explain underprevention in situations with moderate loss probabilities than our model based on inverse S-shaped probability weighting.

2.4.4 The role of ambiguity

So far we have considered risk, i.e., situations in which probabilities are known. In most real-world prevention decisions probabilities cannot be easily defined (Snow, 2011; Alary et al., 2013). Since Ellsberg (1961), the absence of probabilistic information is referred to as ambiguity. The theoretical literature on ambiguity tends to focus on ambiguity aversion but the experimental literature also highlighted that ambiguity reinforces insensitivity (e.g., Tversky and Fox, 1995; Abdellaoui et al., 2011). The intuition is that ambiguity has both a motivational effect, ambiguity aversion, and a cognitive effect, with people having more difficulty to discriminate between likelihood levels when likelihoods are imprecise.

Our results for risk indicate that more likelihood insensitivity leads to more underprevention. We used the data of Abdellaoui et al. (2011), who found more likelihood insensitivity for ambiguity than for risk, to illustrate that ambiguity reinforces underprevention. The size of the insensitivity region, obtained from their data, is 3.4 points larger for ambiguity than for risk and shifts it to the left.[12] Ambiguity about the impact of prevention efforts can therefore increase the underprevention

[12] Risk: 12.60%-84.87% for gains (15.13%-87.40% for losses using the dual function). Ambiguity: 6.41%-82.08% for gains (17.92%-93.59% for indirect losses).

region.

We can also derive results for another ambiguity model *alpha-maxmin*, which is proposed in Arrow and Hurwicz (1972) and axiomatized by Ghirardato et al. (2004). According to this model, people have a set of priors in mind (a set of probabilities that they view as possible) and take a linear combination of the best and the worst expected utility they may get. The idea that agents have a set of priors was first suggested by Wald (1950) and Arrow (1951, pp. 429). We assume that the set of priors is built by ε-*contamination* (as in, e.g., Epstein and Wang, 1994), around the "true" probability p. An agent with such preferences considers p but also all probabilities q around p given by $(1-\varepsilon)p \leq q \leq \varepsilon + p(1-\varepsilon)$. In other words, people assign a confidence weight of $(1-\varepsilon)$ to p but also a weight ε to extreme possibilities 0 and 1. The parameter ε gives the size of the set of priors. It is 0 for a singleton as under expected utility, and 1 if the agent considers the whole interval $[0,1]$ and perceives full ambiguity. The objective function becomes:

$$\begin{aligned}\max_{e\in[0,\bar{e}]} U^a(e) = \delta &\min_{q\leq\varepsilon+(1-\varepsilon)p(e)} \left(qu(x_b,h_b;e) + (1-q)u(x_g,h_g;e)\right) \\ &+ (1-\delta) \max_{q\geq(1-\varepsilon)p(e)} \left(qu(x_b,h_b;e) + (1-q)u(x_g,h_g;e)\right),\end{aligned} \tag{12}$$

with $\varepsilon, \delta \in (0,1)$. As shown by Chateauneuf et al. (2007), this model is equivalent to neo-additive weighting, with $\varepsilon = \alpha$ and $\delta = \frac{1}{2} - \frac{\beta}{2\alpha}$. In this model, δ is traditionally interpreted as capturing ambiguity aversion. The lower δ, the more weight on the worst case scenario. In the present setting, the worst case scenario means that prevention efforts have little effect on the risk faced. Proposition 4 then implies that perceiving greater ambiguity about the costs and benefits of prevention increases the amount of underprevention.

Corollary 1. *Assume the objective function as defined in* (12) *with parameters* $\varepsilon, \delta \in (0,1)$. *Then:*

(i) The presence of ambiguity leads to less prevention for all p^r.

(ii) An increase in ε *leads to more underprevention.*

(iii) A decrease in δ *leads to less underprevention if and only if the marginal cost of effort is higher in the bad state than in the good state.*

Corollary 1 illustrates how ambiguity about the effect of prevention can decrease the perceived benefits and reduce efforts. Statements (i) and (ii) are direct reformulations of the corresponding statements in Proposition 4. Statement (iii) shows that more ambiguity aversion can reduce underprevention if and only if efforts are more costly at the margin in the bad state than in the good state. It is interesting to note that more perceived ambiguity (higher ε) increases the amount of underprevention whereas more ambiguity aversion can reduce it.

2.5 Discussion

2.5.1 Related theoretical literature

One of the original motivations for prospect theory was people's dislike of probabilistic insurance (Kahneman and Tversky, 1979). Bleichrodt and Eeckhoudt (2006) derived that the effect of probability distortion on the willingness to pay to reduce health risks depends on the degree of likelihood insensitivity and on the degree of pessimism. Their model is a special case of Eq. (3) where the cost of effort is monetary and there are only two health states. Rheinberger et al. (2016) extended Bleichrodt and Eeckhoudt to three health states. Assuming linear utility for wealth Rheinberger et al. (2016) showed that probability weighting leads to underprevention for the prevailing incidence and mortality rates of severe diseases in the US. Their model also measures prevention through willingness to pay and does not include a separate effort term. It is the special case of Eq. (3) with three health states.

Etner and Jeleva (2014) defined the notion of fatalism, which is equivalent to $w(p) \leq p$ and $w'(p) \leq 1$ for all $p \in (0, 1)$. They showed that fatalism is a sufficient condition for underprevention. Their results do not apply to inverse S-shaped probability weighting, which are more common empirically. Their utility specification does not include health and only admits a monetary cost of effort. Their result of universal underprevention under fatalism also holds under neo-additive probability weighting in our model. This shows that $w'(p) < 1$ is the crucial property underlying fatalism, not $w(p) \leq p$, which again highlights the role of likelihood insensitivity.

Snow (2011) and Alary et al. (2013) both studied the impact of ambiguity on prevention under the two-stage smooth ambiguity model proposed by Klibanoff et al. (2005). Snow (2011) found that ambiguity raises the optimal level of prevention for an ambiguity-averse agent, and that this level increases with ambiguity aversion. He obtained a similar result with probability weighting but considered only convex weighting functions. Alary et al. (2013) found that the effect of ambiguity aversion on optimal prevention was indeterminate in a multi-state model of prevention.

2.5.2 Prevention and health risks

Keeney (2008) showed that smoking and obesity are the main causes of premature death caused by bad personal decisions. Even though their proportion is decreasing smokers still make up about a sixth of the U.S. population. In many developing countries the proportion of smokers is still increasing.[13] Keeney (2008) estimated that smoking causes roughly 450,000 deaths per year in the US. Table 3, based on the data from Van Baal et al. (2006), shows the probabilities of

[13]The share of smokers in Egypt, for example, increased from 36% to 50% between 2000 and 2015 (World Health Organization, 2018).

Dutch infants born in 2006 to live to various ages. Smoking reduces individual life expectancy significantly. Taking 70 as the expected retirement age for this cohort, the effort of not smoking reduces the probability of dying before retirement by 16 percentage points for men and by 10 percentage points for women. The reduction in mortality risk is even more pronounced at later ages.

Age	Healthy		Smoking		Obese	
	Men	Women	Men	Women	Men	Women
70	89.46%	92.27%	72.78%	82.27%	79.67%	84.90%
75	81.89%	87.29%	58.09%	72.08%	67.90%	76.69%
80	69.03%	78.29%	39.59%	57.15%	51.74%	64.14%

Table 3: Probability of living to retirement (and 5 and 10 years into it) by behavioral type. Table based on the analysis of Van Baal et al. (2006).

Smoking also leads to significant increases in the likelihood of bad health. The probability of being diagnosed with cancer in one's life is 56% for male and 62% for female smokers in the US compared to 36% and 33% for non-smokers.[14] While not all cancers lead to death, treatment is usually painful and leads to a substantial decrease in quality of life.

Unlike smoking, obesity is a growing health concern and tripled worldwide between 1975 and 2016 (World Health Organization, 2018). In the US more than one third (36.5%) of adults and 17% of youth were obese in 2014 (Ogden et al., 2015). Obesity increases the risk of severe diseases (e.g. heart conditions, stroke, diabetes and types of cancer), and therefore also increases mortality risk. Table 3 shows that obesity increases the probability of dying before retirement by roughly 10 percentage points. Similarly, the likelihood of coronary heart disease increases by almost 12 percentage points for obese individuals.[15] The benefits from prevention efforts such as a healthier diet and increased physical activity are thus large.

Individuals' choices depend on their circumstances and preferences. The decision to prevent premature death, as every other prevention decision, involves a (personal) trade-off between costs and potential benefits. However, given the sizable benefits of at least 10 expected life-years, it is puzzling why simple steps such as eating healthier or smoking less are not taken. Our analysis provides an answer. Because the above probabilities probably lie in the underprevention region, probability distortion is a plausible explanation why many people do not undertake such preventive actions.

Many strategies can prevent people from falling ill and underprevention holds for many of

[14] We calculated these probabilities using multiple data sources. See Appendix **??** for details.

[15] Own calculations based on the baseline probability provided in Mozaffarian et al. (2015), the relative risk given in Lhachimi et al. (2012) and the obesity prevalence reported by Ogden et al. (2015).

them. Up to one half of the patients fail to comply with their treatment schedule (Wright, 1993) risking even worse outcomes than their current health state. In all these examples the preventive effort is not monetary justifying our approach of modeling effort as a separate variable.

Our results help to improve the efficacy of prevention policies. For instance, they suggest that solely communicating the benefits of prevention may have little impact if the risk reduction occurs in the insensitivity region. This region is larger and insensitivity is higher if people are unfamiliar with the risks or if they perceive them as ambiguous (Tversky and Fox, 1995; Abdellaoui et al., 2011). Effective policies should try to reduce likelihood insensitivity by giving precise information about the risks involved. Additionally, governments may consider different ways of communicating the risks. For example, Gigerenzer and Hoffrage (1995) showed that people find frequencies easier to handle than probabilities and behave more like Bayesians for frequencies.

2.5.3 Other application domains

The model presented in this paper can shed light on other situations in which efforts can reduce future risks. An example is education. Today's generation has a 50% chance of not reaching their parents' real income (Chetty et al., 2017). A university degree, particularly in subjects like Computer Science and Natural Sciences increases expected future earnings and thus decreases the probability of falling short of parents' lifestyle. In 2015, the probability of ending up in the upper 50 percentile of the US income distribution was around 65% for college graduates against 36% for high school graduates.[16,17] It is thus surprising why not more young adults strive for a college education. Our results suggest that this may be because the probability of falling short of parents' lifestyle lies in the insensitivity region and probability weighting blurs the perceived benefits of higher education.

Climate change is arguably the most important risk society faces today. In a survey by ComRes on behalf of Global Challenges Foundation (2017) with around 8,000 participants from around the world, 84% of the participants considered climate change as a global catastrophic risk and 88% were willing to change their current living standards to prevent it. Still prevention efforts at the individual and societal level are limited at best (Intergovernmental Panel on Climate Change, 2013). One reason is the considerable ambiguity about the extent and effects of climate change and the effectiveness of the measures taken against it (Berger et al., 2016). Our results suggest that this ambiguity leads to more likelihood insensitivity and thus to more underprevention.

[16]Source: OECD, data obtained from `https://stats.oecd.org/Index.aspx?DataSetCode=EAG_EARNINGS#`.

[17]Other factors, like parents' income and education, also play a role in the prospects of young adults (Chetty et al., 2014), but the positive effects of college education on future income, particularly in a technical field, are widely accepted (Acemoglu and Autor, 2011).

2.5.4 Empirical evidence and predictions

In our model, the cost of effort can be monetary, non-monetary or a combination of both. Using the model for predictions might require a quantification of this cost. When it is monetary, such quantifications are often readily available on the market. The empirical literature has shown that monetary costs indeed play a role in individuals' prevention decisions. Cohen and Dupas (2010) used a field experiment manipulating the cost of mosquito nets as a protection against malaria. They showed that a decrease from 100% to 90% of the subsidy for mosquito nets decreases uptake by 60 percentage points. In a large randomized field experiment in the United States, Finkelstein et al. (2012) find that the provision of health insurance coverage markedly increased the utilization of preventive medical services, such as blood cholesterol measurements or mammographies. Regarding non-monetary effort, empirical quantifications are more difficult, but still possible. Mullahy (1999), for example, studied the time cost of flu shots and found that working people have a higher opportunity cost (larger impact of e on u) than non-working people, but that they also have more to gain from flu shots (larger utility difference between good state and bad state).

Our model features both monetary and health consequences. It is a current debate how probability weighting is impacted by the nature of the considered consequences. Most empirical evidence on probability weighting concerns monetary outcomes and some health outcomes but very few studies compare weighting functions across domains. Wakker and Deneffe (1996) provided indirect evidence of stronger deviations from expected utility for life-duration than for money. Rottenstreich and Hsee (2001) showed that affect-rich outcomes can amplify likelihood insensitivity. Abdellaoui and Kemel (2013) compared winning and losing money with saving and losing time. Time as an outcome yielded more insensitivity and more elevation than money. The insensitivity region for time was especially large. Kemel and Paraschiv (2018) found less elevated weighting functions when outcomes are human lives than when outcomes are monetary. Other factors influence probability weighting. Identifying them provides directions to design interventions targeting those who are more prone to underprevention. Age (Booij et al., 2010), gender (Fehr-Duda et al., 2010), affective states (Kliger and Levy, 2008) and numeracy (Petrova et al., 2014) have been linked to the extent to which probability weighting appears. Abdellaoui et al. (2011) show that the amount of likelihood insensitivity varies with the familiarity that agents have with the risk. Since probability insensitivity tends to favor underprevention, we can thus speculate that underprevention is a larger problem for risks with which agents are unfamiliar.

These results offer several avenues for designing more effective policy interventions to increase prevention efforts in society. Since numeracy and familiarity are negatively associated with the degree of likelihood insensitivity, one way forward could be to make agents more familiar with the risks they are facing. Such interventions would reduce ambiguity and increase subject specific numeracy (often measured by concepts such as financial or technological literacy) and would thus likely reduce underprevention, as well. Specific examples come to mind. In health applications, using certain graphical displays fosters understanding of the risk and increases decision quality (Fagerlin et al., 2005; Smerecnik et al., 2010). In the education example, high school students could be confronted with specific information about the employment opportunities of different careers. Combining our results with those by Rottenstreich and Hsee (2001) and Kliger and Levy (2008) further renders potential guidance on how public information campaigns about prevention should be designed. In such campaigns, information is often communicated in very affect-rich contexts. An example for this are the graphic displays on cigarette cartons mandatory in the European Union. The result that more affect-rich environments lead to more likelihood insensitivity and thus by our model to less prevention, might call the efficiency of such campaigns into question.

Our model focuses on probability weighting to understand underprevention but alternative explanations exist, such as time preferences. Future benefits from prevention should be discounted but an agent prone to present bias may exert even less effort by applying an additional penalty to any future utility. Such present bias can lead to procrastination, with agents planning prevention but not following through. Empirical studies are needed to disentangle the effect of time preferences from the effect of risk perception on prevention. Such studies could involve eliciting time and risk preferences and subsequently monitoring agents in their prevention behavior. Alternatively, interventions could be designed to target probability weighting or the present bias and they could be compared in an experiment. For instance, an intervention to make agents more familiar with some risks could be compared with an intervention proposing commitment devices to help people carry out their prevention plans.

2.6 Conclusion

The anomalies in insurance and financial markets led Arrow (1982) to the conclusion that "systematic deviations from individual rational behavior [...] are consonant with evidence from very different sources collected by psychologists." Following his example, we showed that low prevention efforts are consonant with the evidence collected on risk perception and probability weighting. Likelihood insensitivity and ambiguity blur the benefits of prevention, inducing agents to exert less effort than their rational counterparts would. Probability weighting can thus contribute to explain the puzzle put forward by Keeney (2008), that agents' own decisions are so often responsible for their own death. Our theoretical model enables us to identify when and how people behave suboptimally, making it possible to improve the efficacy of prevention policies.

Appendix 2.A Proofs for Section 3

2.A.1 Proof of Lemma 1

Define $f(p) = w(p) - p$, which is continuous on $[0, 1]$ and twice differentiable on $(0, 1)$ due to the regularity of w. We know that $f(0) = f(p^*) = f(1) = 0$ from Definition 1 and Property (i) in Definition 2. From Rolle's Theorem we obtain $p_1 \in (0, p^*)$ and $p_2 \in (p^*, 1)$ such that $f'(p_i) = w'(p_i) - 1 = 0$ for $p_i \in \{p_1, p_2\}$. So $f'(p)$ has at least two zeros on $(0, 1)$. From the inverse S shape of $w(p)$ we can infer that $f''(p) = w''(p)$ is negative for $p < \widetilde{p}$, zero for $p = \widetilde{p}$ and positive for $p > \widetilde{p}$. Hence, $f'(p)$ is strictly decreasing at first, has a minimum at $p = \widetilde{p}$ and is strictly increasing thereafter. Consequently, $f'(p)$ can have no more than two zeros, and one is to the left of $\widetilde{p}$ and the other one to the right of $\widetilde{p}$. It follows that $w'(p) > 1$ for $p \in (0, p_1) \cup (p_2, 1)$, $w'(p_i) = 1$ for $p_i \in \{p_1, p_2\}$ and $w'(p) < 1$ for $p \in (p_1, p_2)$.

2.A.2 Proof of Proposition 1

Inserting the rational agent's optimal level of prevention e^r into the first-order expression under probability weighting yields:

$$\begin{aligned}
U_e^w(e^r) &= -w'(p^r)p'(e^r)\left[u(x_g, h_g; e^r) - u(x_b, h_b; e^r)\right] && (13)\\
&\quad + \left[w(p^r)u_e(x_b, h_b; e^r) + (1 - w(p^r))u_e(x_g, h_g; e^r)\right] \\
&= -w'(p^r)\left[p^r u_e(x_b, h_b; e^r) + (1 - p^r)u_e(x_g, h_g; e^r)\right] && (14)\\
&\quad + \left[w(p^r)u_e(x_b, h_b; e^r) + (1 - w(p^r))u_e(x_g, h_g; e^r)\right] \\
&= \left[1 - w'(p^r)\right]\left[p^r u_e(x_b, h_b; e^r) + (1 - p^r)u_e(x_g, h_g; e^r)\right] && (15)\\
&\quad + \left[w(p^r) - p^r\right]\left[u_e(x_b, h_b; e^r) - u_e(x_g, h_g; e^r)\right].
\end{aligned}$$

The second equality holds due to the FOC for e^r, see Eq. (2), and the third one is a simple rearrangement.

We investigate the last expression as a funtion of p^r and denote it by $g(p^r)$ while keeping all other parameters fixed. To determine the behavior of g as p^r approaches the endpoints of the unit interval, we first notice that w' is not necessarily defined at 0 and 1. However, it follows from the inverse-S shape of w that w' either converges to a limit above 1 or diverges to $+\infty$ as p^r approaches 0 or 1.[18] We denote these limits by $w'(0)$ and $w'(1)$, taking into account that they may be infinite.

[18] Look at the sequence $\{w'(1/n)\}$ for $n \geq 1/p_1$ for example; each element exceeds 1 and the sequence is monotonically increasing due to concavity of w on $(0, \widetilde{p})$. If it is bounded, the monotone convergence theorem establishes the existence of a limit above 1, if it is unbounded, it diverges to $+\infty$. A similar argument applies to $p^r \to 1$.

Then, it holds that

$$\lim_{p^r \to 0} g(p^r) = [1 - w'(0)]\, u_e(x_g, h_g; e^r) > 0 \text{ and} \tag{16}$$

$$\lim_{p^r \to 1} g(p^r) = [1 - w'(1)]\, u_e(x_b, h_b; e^r) > 0, \tag{17}$$

where both limits are potentially infinite if $w'(0)$ and $w'(1)$ are. From the regularity of w it follows that g is a differentiable function on $(0,1)$. Its derivative with respect to p^r is given by

$$g'(p^r) = -w''(p^r)\left[p^r u_e(x_b, h_b; e^r) + (1 - p^r) u_e(x_g, h_g; e^r)\right], \tag{18}$$

which is negative for $p^r < \widetilde{p}$, zero for $p^r = \widetilde{p}$ and positive for $p^r > \widetilde{p}$. As a result, g as a function of p^r is strictly decreasing at first, has a minimum at $p^r = \widetilde{p}$ and is strictly increasing thereafter. Additionally, it holds that

$$g(p^*) = [1 - w'(p^*)]\left[p^r u_e(x_b, h_b; e^r) + (1 - p^r) u_e(x_g, h_g; e^r)\right] < 0, \tag{19}$$

because $w(p^*) = p^*$ per definition and because p^* lies inside the likelihood insensitivity region per Lemma 1. This implies that $g(\widetilde{p}) \leq g(p^*) < 0$ since $\widetilde{p}$ is the global minimizer of g on $(0,1)$. All these properties together with the continuity of g ensure the existence of exactly two zeros in $(0,1)$, denoted by $\underline{p}$ and $\overline{p}$, with one of them being smaller than p^* and $\widetilde{p}$, and the other one being larger than p^* and $\widetilde{p}$.

So if $p^r \in (\underline{p}, \overline{p})$, we obtain $g(p^r) = U_e^w(e^r) < 0$ so that probability weighting induces the agent to choose a lower level of prevention than the rational agent, $e^w < e^r$. If, however, $p^r \in (0, \underline{p}) \cup (\overline{p}, 1)$, we obtain $g(p^r) = U_e^w(e^r) > 0$, and probability weighting induces the agent to choose a higher level of prevention than the rational agent, $e^w > e^r$. Notice that the assumption of U^w being concave in e is critical in this last step.

2.A.3 Proof of Proposition 2

We evaluate $g(p^r)$ at p_1 and p_2, taking into account that $w'(p_i) = 1$ for $i = 1, 2$. It holds that

$$g(p_1) = [w(p_1) - p_1]\left[u_e(x_b, h_b; e^r) - u_e(x_g, h_g; e^r)\right], \tag{20}$$

where the first bracket is positive because $p_1 < p^*$ from Lemma 1. The second bracket is negative (zero, positive) if the marginal cost of effort is smaller (identical, larger) in the good state compared to the bad state.[19] Then, it holds that $g(p_1) < \ (=,\ >)\ g(\underline{p})$, and g is decreasing for probabilities

[19] Effort lowers utility, $u_e < 0$, so an appropriate measure of the marginal cost of effort is given by $|u_e| = -u_e$.

less than $\widetilde{p}$ so that $\underline{p} < \ (=,\ >)\ p_1$.

Similarly, we obtain that

$$g(p_2) = [w(p_2) - p_2]\,[u_e(x_b, h_b; e^r) - u_e(x_g, h_g; e^r)]\,, \tag{21}$$

where the first bracket is negative because $p_2 > p^*$ from Lemma 1. The second bracket is negative (zero, positive) if the marginal cost of effort is smaller (identical, larger) in the good state compared to the bad state. Then, it holds that $g(p_2) > \ (=,\ <)\ g(\overline{p})$, and g is increasing for probabilities above $\widetilde{p}$ so that $\overline{p} < \ (=,\ >)\ p_2$.

2.A.4 Proof of Proposition 3

The proof proceeds in several steps, and to increase readability, we will formulate some of these steps as separate Lemmas. The boundaries of the underprevention region, $\underline{p}$ and $\overline{p}$, are obtained as the solution of $g(p^r) = 0$. Rearranging this condition yields:

$$\frac{(1 - w(p^r)) - w'(p^r)(1 - p^r)}{w'(p^r)p^r - w(p^r)} = \frac{u_e(x_b, h_b; e^r)}{u_e(x_g, h_g; e^r)} = \kappa. \tag{22}$$

We denote the left-hand side of (22) by $h(p^r)$ and will first characterize its behavior on $[0, 1]$ when w is a regular inverse S-shaped probability weighting function.

Lemma 2. *h has the following properties:*

(i) A unique zero at some $q_1 \in (p_1, \min\{p^*, \widetilde{p}\})$.

(ii) A singularity at some $q_2 \in (\max\{p^*, \widetilde{p}\}, p_2)$ with $\lim_{p^r \uparrow q_2} h(p^r) = -\infty$ and $\lim_{p^r \downarrow q_2} h(p^r) = \infty$.

(iii) $\lim_{p^r \to 0} h(p^r) = \infty$ and $\lim_{p^r \to 1} h(p^r) = 0$.

(iv) It is positive for $p^r \in (0, q_1) \cup (q_2, 1)$ and negative for $p^r \in (q_1, q_2)$.

(v) It is decreasing for $p^r \in (0, \min\{p^*, \widetilde{p}\}) \cup (\max\{p^*, \widetilde{p}\}, q_2) \cup (q_2, 1)$ and increasing for $p^r \in (\min\{p^*, \widetilde{p}\}, \max\{p^*, \widetilde{p}\})$.

Proof. We denote by $h_n(p^r)$ the numerator of h and by $h_d(p^r)$ its denominator, that is $h(p^r) = h_n(p^r)/h_d(p^r)$. It holds that $\lim_{p^r \to 0} h_n(p^r) = 1 - w'(0) < 0$, potentially $-\infty$, $\lim_{p^r \to 1} h_n(p^r) = 0$ and $h_n'(p^r) = -w''(p^r)(1-p^r)$ so that h_n is strictly increasing for $p^r < \widetilde{p}$, has a maximum at $p^r = \widetilde{p}$, and is strictly decreasing for $p^r > \widetilde{p}$. This proves the existence of a unique zero $q_1 < \widetilde{p}$ of h_n and therefore of h. $h_n(p_1) = p_1 - w(p_1) < 0$ so that $q_1 > p_1$, and $h_n(p^*) = (1 - p^*)(1 - w'(p^*)) > 0$ so that $q_1 < p^*$. This proves (i).

For the denominator it holds that $\lim_{p^r \to 0} h_d(p^r) = 0$, $\lim_{p^r \to 1} h_d(p^r) = w'(1) - 1 > 0$, potentially ∞, and $h_d'(p^r) = w''(p^r)p^r$ so that h_d is strictly decreasing for $p^r < \widetilde{p}$ has a minimum at $p^r = \widetilde{p}$ and is strictly increasing for $p^r > \widetilde{p}$. This proves the existence of a unique zero $q_2 > \widetilde{p}$ of h_d, which is a singularity of h. $h_d(p^*) = p^*(w'(p^*) - 1) < 0$ so that $q_2 > p^*$ and $h_d(p_2) = p_2 - w(p_2) > 0$ so that $q_2 < p_2$. Furthermore, $h_n(q_2) = 1 - w'(q_2) > 0$ by definition of q_2, which shows that h switches from $-\infty$ to $+\infty$ at q_2. This proves (ii).

The limits of h for $p^r \to 0$ and $p^r \to 1$ follow from the limits of the numerator and the denominator at 0 and 1. Notice that $h_d(p^r)$ approaches 0 from below for $p^r \to 0$. This shows (iii).

The signs of h follow from the signs of h_n and h_d according to the following table, which demonstrates (iv).

p^r	$(0, q_1)$	(q_1, q_2)	$(q_2, 1)$
h_n	$-$	$+$	$+$
h_d	$-$	$-$	$+$
h_n/h_d	$+$	$-$	$+$

Table 4: Signs of h_n, h_d and h_n/h_d

The derivative of h with respect to p^r is given by

$$h'(p^r) = \frac{w''(p^r)\left[w(p^r) - p^r\right]}{\left[w'(p^r)p^r - w(p^r)\right]^2}, \tag{23}$$

and (v) follows immediately.

Figure 4 illustrates h as a function of p^r for a Prelec-2 probability weighting function with parameters $\alpha = 0.5$ and $\beta = 1.2$. The properties established in Lemma 2 imply that any positive value of κ renders two solutions to Eq. (22), $\underline{p}$ and $\overline{p}$, one smaller than p^* and $\widetilde{p}$ and one larger than p^* and $\widetilde{p}$ consistent with Proposition 1. Given that h is decreasing whenever it is positive, both $\underline{p}$ and $\overline{p}$ are negatively associated with κ, which shows claim (i) in Proposition 3.

To prove statements (ii) to (v), we first analyze how the rational agent's optimal level of prevention depends on x_g, x_b, h_g and h_b, and then proceed to show how changes in the exogenous parameters affect the ratio of the marginal cost of effort in the bad state over the marginal cost of effort in the good state, i.e., how κ is affected. Signing the effect on κ will then render the effect on the underprevention region via statement (i) of Proposition 3.

Lemma 3. *Assume $u_{xe} \geq 0$ and $u_e(x, h_g; e) \geq u_e(x, h_b; e)$. The rational agent's optimal level of prevention is:*

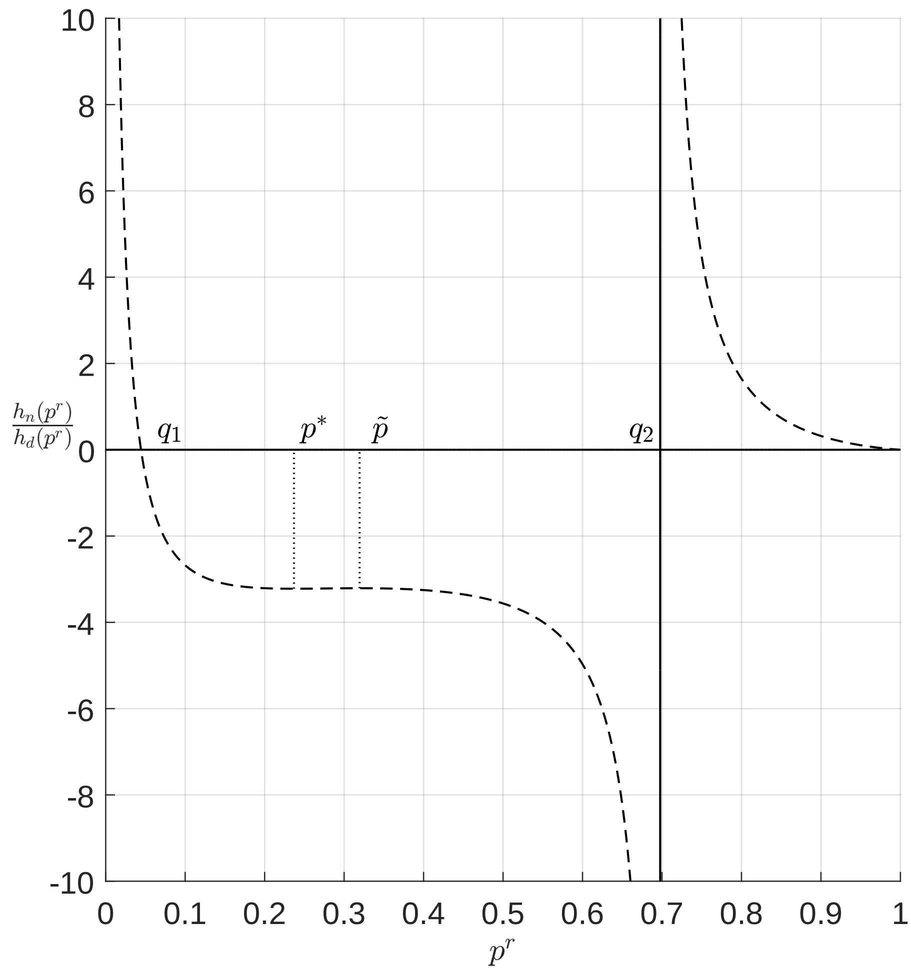


Figure 4: Shape of h as a function of p^r

a) Increasing in x_g.

b) Decreasing in x_b if the loss probability is more elastic than the marginal utility of wealth.

c) Increasing in h_g.

d) Decreasing in h_b if the loss probability is more elastic than the utility gain from better health.

Proof. We first show a) and b). Our assumption of e^r being an interior solution yields $U^r_{ee}(e^r) < 0$ for maximality. Per the implicit function rule, exogenous changes affect the optimal level of prevention according to

$$\frac{\mathrm{d}e^r}{\mathrm{d}k} = -\frac{U^r_{ke}(e^r)}{U^r_{ee}(e^r)},\ k \in \{x_g, x_b\}, \tag{24}$$

so that $\operatorname{sgn}(\frac{\mathrm{d}e^r}{\mathrm{d}k}) = \operatorname{sgn}(U^r_{ke}(e^r))$. The relevant cross-derivatives are given by

$$U^r_{x_g e}(e^r) = -p'(e^r)u_x(x_g, h_g; e^r) + (1 - p(e^r))u_{xe}(x_g, h_g; e^r) > 0 \tag{25}$$

and

$$\begin{aligned} U^r_{x_b e}(e^r) &= p'(e^r)u_x(x_b, h_b; e^r) + p(e^r)u_{xe}(x_b, h_b; e^r) && (26)\\ &= \frac{1}{e^r}p(e^r)u_x(x_b, h_b; e^r) \cdot [\varepsilon_{u_x}(x_g, h_g; e^r) - \varepsilon_p(e^r)] < 0, && (27)\end{aligned}$$

where the signs follow from our assumptions. To demonstrate c) and d), we assume a change in health from h_g to $\hat{h}_g$ with $\hat{h}_g \succ h_g$ and a change in health from h_b to $\hat{h}_b$ with $h_g \succ \hat{h}_b \succ h_b$. If e^r denotes the rational agent's optimal level of prevention when health outcomes are h_g and h_b, we obtain that

$$\begin{aligned} & -p'(e^r)\left[u(x_g, \hat{h}_g; e^r) - u(x_b, h_b; e^r)\right] + \left[p(e^r)u_e(x_b, h_b; e^r) + (1 - p(e^r))u_e(x_g, \hat{h}_g; e^r)\right] && (28)\\ = \; & -p'(e^r)\left[u(x_g, \hat{h}_g; e^r) - u(x_g, h_g; e^r)\right] + (1 - p(e^r))\left[u_e(x_g, \hat{h}_g; e^r) - u_e(x_g, h_g; e^r)\right] > 0, && (29)\end{aligned}$$

where the equality holds by the FOC for e^r, see Eq. (2), and the sign follows from our assumptions. As a result, when the good health outcome improves from h_g to $\hat{h}_g$, the effort level e^r is no longer enough and the rational agent would increase it in response. Similarly, it holds that

$$\begin{aligned} & -p'(e^r)\left[u(x_g, h_g; e^r) - u(x_b, \hat{h}_b; e^r)\right] + \left[p(e^r)u_e(x_b, \hat{h}_b; e^r) + (1 - p(e^r))u_e(x_g, h_g; e^r)\right] && (30)\\ = \; & -p'(e^r)\left[u(x_b, h_b; e^r) - u(x_b, \hat{h}_b; e^r)\right] + p^r\left[u(x_b, \hat{h}_b; e^r) - u(x_b, h_b; e^r)\right] && (31)\\ = \; & \frac{1}{e^r}p^r\left[u(x_b, \hat{h}_b; e^r) - u(x_b, h_b; e^r)\right] \cdot \left[\varepsilon_{\Delta h}(x_b, \hat{h}_b, h_b; e^r) - \varepsilon_p(e^r)\right] < 0, && (32)\end{aligned}$$

where the equality holds by the FOC for e^r, see Eq. (2), and the sign follows from our assumptions. Hence, when the bad health outcome improves from h_b to $\hat{h}_b$, while maintaining its interpretation of being worse health than h_g, the effort level e^r is too high and the rational agent decreases it in response.

To complete the proof of statements (ii) to (v) of Proposition 3, we need to determine the effect of changes in x_g, x_b, h_g and h_b on κ in Eq. (22). It holds that

$$\frac{\mathrm{d}\kappa}{\mathrm{d}x_g} = \frac{u_e(x_g,h_g;e^r)u_{ee}(x_b,h_b;e^r)\frac{\mathrm{d}e^r}{\mathrm{d}x_g} - u_e(x_b,h_b;e^r)\left[u_{ee}(x_g,h_g;e^r)\frac{\mathrm{d}e^r}{\mathrm{d}x_g} + u_{xe}(x_g,h_g;e^r)\right]}{u_e(x_g,h_g;e^r)^2} \tag{33}$$

$$= \frac{u_e(x_b,h_b;e^r)}{e^r \cdot u_e(x_g,h_g;e^r)}\left[\varepsilon_{u_e}(x_b,h_b;e^r) - \varepsilon_{u_e}(x_g,h_g;e^r)\right]\frac{\mathrm{d}e^r}{\mathrm{d}x_g} - \frac{u_e(x_b,h_b;e^r)u_{xe}(x_g,h_g;e^r)}{u_e(x_g,h_g;e^r)^2} \geq 0, \tag{34}$$

using the assumptions stated in Proposition 3 and the result in Lemma 3a). So an increase in x_g increases κ, in which case the underprevention region shifts to the left per Proposition 3(i). Likewise, we obtain that

$$\frac{\mathrm{d}\kappa}{\mathrm{d}x_b} = \frac{u_e(x_b,h_b;e^r)}{e^r \cdot u_e(x_g,h_g;e^r)}\left[\varepsilon_{u_e}(x_b,h_b;e^r) - \varepsilon_{u_e}(x_g,h_g;e^r)\right]\frac{\mathrm{d}e^r}{\mathrm{d}x_b} + \frac{u_{xe}(x_g,h_g;e^r)}{u_e(x_g,h_g;e^r)} \leq 0, \tag{35}$$

from the assumptions in Proposition 3 and Lemma 3b). An increase in x_b lowers κ so that the underprevention region shifts to the right due to Proposition 3(i). If health in the good state improves from h_g to $\hat{h}_g$, this results in an increase of the rational agent's optimal effort level from e^r to $\hat{e}^r$, and as a result

$$\frac{u_e(x_b,h_b;e^r)}{u_e(x_g,h_g;e^r)} \leq \frac{u_e(x_b,h_b;\hat{e}^r)}{u_e(x_g,h_g;\hat{e}^r)} \leq \frac{u_e(x_b,h_b;\hat{e}^r)}{u_e(x_g,\hat{h}_g;\hat{e}^r)}. \tag{36}$$

The first inequality holds because the marginal cost of effort is assumed to be more elastic in the bad than the good state, and the second one is a consequence of our assumption that an improvement in health reduces the marginal cost of effort. So κ increases when health improves from h_g to $\hat{h}_g$, which shifts the underprevention region to the left per Proposition 3(i). Finally, when health in the bad state improves from h_b to $\hat{h}_b$ with associated effort levels e^r and $\hat{e}^r$, arguments that are by now familiar establish that

$$\frac{u_e(x_b,h_b;e^r)}{u_e(x_g,h_g;e^r)} \geq \frac{u_e(x_b,h_b;\hat{e}^r)}{u_e(x_g,h_g;\hat{e}^r)} \geq \frac{u_e(x_b,h_b;\hat{e}^r)}{u_e(x_g,\hat{h}_g;\hat{e}^r)} \tag{37}$$

under the assumptions made. So κ decreases, which shifts the underprevention region to the right per Proposition 3(i).

2.A.5 Proof of Proposition 4

We evaluate the agent's first-order expression when his probability weighting function is neo-additive at the rational agent's optimal level of prevention e^r. This yields

$$\begin{aligned}
U_e^w(e^r) &= -(1-\alpha)p'(e^r)\left[u(x_g,h_g;e^r)-u(x_b,h_b;e^r)\right] &(38)\\
&\quad+\left[\left(\frac{\alpha-\beta}{2}+(1-\alpha)p^r\right)u_e(x_b,h_b;e^r)+\left(1-\frac{\alpha-\beta}{2}-(1-\alpha)p^r\right)u_e(x_g,h_g;e^r)\right]\\
&= -(1-\alpha)\left[p^r u_e(x_b,h_b;e^r)+(1-p^r)u_e(x_g,h_g;e^r)\right] &(39)\\
&\quad+\left[\left(\frac{\alpha-\beta}{2}+(1-\alpha)p^r\right)u_e(x_b,h_b;e^r)+\left(1-\frac{\alpha-\beta}{2}-(1-\alpha)p^r\right)u_e(x_g,h_g;e^r)\right]\\
&= \frac{\alpha-\beta}{2}u_e(x_b,h_b;e^r)+\frac{\alpha+\beta}{2}u_e(x_g,h_g;e^r), &(40)
\end{aligned}$$

where the first equality follows from the definition of the neo-additive probability weighting function, the second by substituting from the FOC for e^r in Eq. (2), and the third one by combinining and simplifying terms. The parameters in the neo-additive probability weighting function are such that $(\alpha-\beta)>0$ and $(\alpha+\beta)>0$ so that $U_e^w(e^r)<0$, which shows that the agent with neo-additive probability weighting chooses a lower level of prevention than the rational agent (claim (i)).

For (ii) and (iii) we differentiate the agent's first-order condition with respect to α and β, respectively, and determine the sign. We obtain

$$\begin{aligned}
\frac{\partial U_e^w(e^w)}{\partial\alpha} &= p'(e^w)\left[u(x_g,h_g;e^w)-u(x_b,h_b;e^w)\right] &(41)\\
&\quad+\left[\left(\frac{1}{2}-p^w\right)u_e(x_b,h_b;e^w)-\left(\frac{1}{2}-p^w\right)u_e(x_g,h_g;e^w)\right]\\
&= \frac{1}{1-\alpha}\left[\left(\frac{\alpha-\beta}{2}+(1-\alpha)p^w\right)u_e(x_b,h_b;e^w)+\left(1-\frac{\alpha-\beta}{2}-(1-\alpha)p^w\right)u_e(x_g,h_g;e^w)\right]\\
&\quad+\left(\frac{1}{2}-p^w\right)\left[u_e(x_b,h_b;e^w)-u_e(x_g,h_g;e^w)\right] &(42)\\
&= \frac{1}{2(1-\alpha)}\left[(1-\beta)u_e(x_b,h_b;e^w)+(1+\beta)u_e(x_g,h_g;e^w)\right], &(43)
\end{aligned}$$

where the second equality follows from substituting in the FOC for e^w, see Eq. (4), and the third one by rearranging and combining terms. The parameters in the neo-additive weighting function are such that $(1-\alpha)>0$, $(1-\beta)>0$ and $(1+\beta)>0$ so that $\partial U_e^w(e^w)/\partial\alpha<0$. An increase in parameter α reduces e^w. Finally, we determine

$$\frac{\partial U_e^w(e^w)}{\partial\beta}=-\frac{1}{2}\left[u_e(x_b,h_b;e^w)-u_e(x_g,h_g;e^w)\right], \tag{44}$$

which is negative (zero, positive) if and only if $-u_e(x_b,h_b;e^w)$ exceeds (is equal to, falls below) $-u_e(x_g,h_g;e^w)$.

2.A.6 Proof of Corollary 1

Statements (i) and (ii) are direct reformulations of (i) and (ii) from Proposition 4. For (iii), we first notice $\frac{\alpha-\beta}{2} = \delta\varepsilon$ and $\frac{\alpha+\beta}{2} = (1-\delta)\varepsilon$. Hence,

$$U_e^w(e^w) = \delta\varepsilon u_e(x_b, h_b; e^w) + (1-\delta)\varepsilon u_e(x_g, h_g; e^w). \tag{45}$$

We then differentiate the agent's first-order condition with respect to δ and obtain

$$\frac{\partial U_e^w(e^w)}{\partial \delta} = \varepsilon \left[u_e(x_b, h_b; e^w) - u_e(x_g, h_g; e^w)\right], \tag{46}$$

which is positive (zero, negative) if and only if $-u_e(x_b, h_b; e^w)$ exceeds (is equal to, falls below) $-u_e(x_g, h_g; e^w)$.

Appendix 2.B Evidence of likelihood insensitivity regions

Tables 5 and 6 display the insensitivity regions calculated from the results of prior experimental studies. Table 5 is based on the parameter estimates of a Prelec-2 function by l'Haridon and Vieider (2018) 29 individual countries.

Table 6 uses the results from several different studies that estimate parameter values for commonly used probability weighting functions. We compute the insensitivity regions corresponding to the chosen probability weighting function at the reported point estimates.

Table 5: Insensitivity Regions for several countries

Country	Insensitivity region	Unit interval
Australia	12.43% - 88.34%	
Belgium	11.39% - 86.10%	
Brazil	13.02% - 85.96%	
Cambodia	3.70% - 85.11%	
Chile	11.02% - 87.07%	
China	10.95% - 85.44%	
Colombia	8.98% - 80.85%	
Costarica	8.34% - 81.20%	
Czech	15.40% - 88.44%	
Ethiopia	5.67% - 85.02%	
France	12.23% - 84.40%	
Germany	11.98% - 84.51%	
India	4.77% - 82.19%	
Japan	14.41% - 86.23%	
Kyrgyzstan	6.26% - 83.49%	
Malaysia	8.95% - 85.99%	
Nicaragua	5.31% - 85.55%	
Nigeria	2.67% - 91.44%	
Peru	6.95% - 84.94%	
Poland	9.35% - 81.16%	
Russia	7.12% - 78.45%	
Saudi	9.52% - 85.28%	
South Africa	8.00% - 81.03%	
Spain	12.52% - 86.02%	
Thailand	7.95% - 75.85%	
Tunisia	6.09% - 85.60%	
UK	6.05% - 82.05%	
USA	12.62% - 84.13%	
Vietnam	9.00% - 81.82%	
Global	**7.69% - 84.43%**	

Insensitivity regions are calculated using the two-parameter Prelec function as estimated in l'Haridon and Vieider (2018). Guatemala is removed from the list since the weighting function was not inverse S-shaped. The last column represents the insensitivity regions as bars in the unit interval.

Table 6: Estimates of the insensitivity region based on empirical results

Weighting function	Study	Domain	Insensitivity region	Unit interval
TK	Tversky and Kahneman (1992)	Loss	10.52% - 83.69%	
$w(p) = \frac{p^\alpha}{[p^\alpha+(1-p)^\alpha]^{\frac{1}{\alpha}}}$	Abdellaoui (2000)	Loss	10.77% - 83.61%	
	Etchart-Vincent (2004)	Loss	15.23% - 82.58%	
	Andersen et al. (2006)	Combined	12.22% - 83.17%	
	Berns et al. (2007)	Loss	10.40% - 83.74%	
	Etchart-Vincent (2009)	Loss	12.22% - 83.17%	
	Harrison and Rutström (2009)	Combined	15.67% - 82.53%	
	Kemel and Paraschiv (2013)	Loss	10.28% - 83.78%	
GE	Abdellaoui (2000)	Loss	10.19% - 84.25%	
$w(p) = \frac{\beta p^\alpha}{\beta p^\alpha+(1-p)^\alpha}$	Etchart-Vincent (2004)	Loss	19.90% - 88.63%	
	Abdellaoui et al. (2005)	Loss	26.40% - 95.61%	
	Fehr-Duda et al. (2006)	Loss	11.84% - 89.47%	
	Etchart-Vincent (2009)	Loss	14.65% - 89.21%	
	Booij et al. (2010)	Loss	12.47% - 88.08%	
	Fehr-Duda et al. (2010)	Loss	8.80% - 90.62%	
	Abdellaoui et al. (2011)	Loss	8.45% - 80.51%	
	Abdellaoui and Kemel (2013)	Loss	11.26% - 86.07%	
Prelec-1	Donkers et al. (2001)	Combined	4.45% - 87.07%	
$w(p) = \exp(-(-\ln p)^\alpha)$	Tu et al. (2005)	Loss	8.29% - 80.59%	
	Tanaka et al. (2010)	Loss	7.96% - 81.08%	
	Charupat et al. (2013)	Loss	7.67% - 81.55%	
	Kemel and Paraschiv (2013)	Loss	9.61% - 78.64%	
	Wibbenmeyer et al. (2013)	Loss	1.74% - 93.24%	
Prelec-2	Booij et al. (2010)	Loss	8.86% - 86.39%	
$w(p) = \exp(-\beta(-\ln p)^\alpha)$	Abdellaoui and Kemel (2013)	Loss	6.50% - 83.42%	
	Charupat et al. (2013)	Loss	6.56% - 81.21%	

"Loss" and "Combined" indicate that the study elicited parameter values for losses or for both losses and gains assuming they are equal, respectively.
The last column represents the insensitivity regions as bars in the unit interval.

Appendix 2.C Calculation of conditional cancer incidence probabilities in Section 5.2

We aim to calculate the lifetime probability of being diagnosed with any type of cancer conditional on the individual's smoking status. We denote the smoking status as S if the individual is a smoker and as $\neg S$ for a non-smoker[20]. We use the taxonomy from Howlader et al. (2017) to distinguish between 23 different types of cancer C_i with unconditional incidence rate $P(C_i), i \in \{1, ..., 23\}$. $P(C_\Sigma)$ denotes the probability of being diagnosed with any type of cancer. Smoking affects 12 types of cancer by a relative risk ratio of π_i (see Surgeon General, 2014) such that

$$P(C_i|S) = \pi_i P(C_i|\neg S). \tag{47}$$

Using the law of total probability, this becomes

$$P(C_i|S) = \pi_i \frac{P(C_i) - P(C_i|S)P(S)}{P(\neg S)}. \tag{48}$$

Solving for $P(C_i|S)$ then yields

$$P(C_i|S) = \frac{\pi_i P(C_i)}{P(\neg S) + \pi_i P(S)}, \tag{49}$$

to which we know all inputs. $P(C_i|\neg S)$ can then be calculated through the law of total probability.

Data from Howlader et al. (2017) informs us that cancer incidents are neither stochastically independent nor mutually exclusive so that neither $P(\neg C_\Sigma) = \prod_{i=1}^{23} P(\neg C_i)$ nor $P(C_\Sigma) = \sum_{i=1}^{23} P(C_i)$ holds. To aggregate cancer-specific incidence rates, we make two simplifying assumptions. First, we assume all types of cancer to be pairwise correlated with the same correlation coefficient ρ. Second, we assume that an individual cannot be diagnosed with more than two types of cancer. The latter assumption is certainly incorrect, but since the incidence rates are small, the associated bias is small enough to make it acceptable. Based on these assumptions, we find that

$$P(C_\Sigma|S) = \sum_{i=1}^{23} P(C_i|S) - \sum_{1 \le i < j \le 23} P(C_i \cap C_j|S) \tag{50}$$

where

$$P(C_i \cap C_j|S) = \rho\sqrt{P(C_i|S)(1 - P(C_i|S))P(C_j|S)(1 - P(C_j|S))} + P(C_i|S)P(C_j|S) \tag{51}$$

from the definition of the correlation coefficient. We calibrate the missing input factor ρ such as to solve Eq. (50) for the unconditional probabilities provided by Howlader et al. (2017).

The calculations are carried out for women and men separately. All inputs are given in Table

[20]The Centers for Medicare & Medicaid Services use a more granular codification, see `https://www.cms.gov/Regulations-and-Guidance/Legislation/EHRIncentivePrograms/downloads/9_Record_Smoking_Status.pdf`. We choose a binary distinction for smoking status for reasons of data availability.

7. Smoking prevalence, lifetime cancer incidence rates, and relative risk ratios pertain to the year 2014. Calculations on the Behavioral Risk Factor Surveillance System Survey data of the Center for Disease Control and Prevention (2015) were carried out using weighted sampling to make the sample representative.

Table 7: Input Parameters and Sources

Input	Description	Male	Female	Reference
$P(S)$	Probability of smoking	19.26%	15.25%	Center for Disease Control and Prevention (2015)
$P(C_\Sigma)$	Incidence Rate (IR) all cancer	39.66%	37.65%	Howlader et al. (2017)
$P(C_1)$	IR Bladder (includes in situ)	3.76%	1.12%	Howlader et al. (2017)
$P(C_2)$	IR Brain and nervous system	0.70%	0.54%	Howlader et al. (2017)
$P(C_3)$	IR Breast	0.12%	12.41%	Howlader et al. (2017)
$P(C_4)$	IR Cervix		0.62%	Howlader et al. (2017)
$P(C_5)$	IR Colon and rectum	4.49%	4.15%	Howlader et al. (2017)
$P(C_6)$	IR Esophagus	0.76%	0.22%	Howlader et al. (2017)
$P(C_7)$	IR Hodgkin disease	0.23%	0.19%	Howlader et al. (2017)
$P(C_8)$	IR Kidney and renal pelvis	2.09%	1.20%	Howlader et al. (2017)
$P(C_9)$	IR Larynx (voice box)	0.55%	0.12%	Howlader et al. (2017)
$P(C_{10})$	IR Leukemia	1.79%	1.26%	Howlader et al. (2017)
$P(C_{11})$	IR Liver and bile duct	1.39%	0.60%	Howlader et al. (2017)
$P(C_{12})$	IR Lung and bronchus	6.85%	5.95%	Howlader et al. (2017)
$P(C_{13})$	IR Melanoma of the skin	2.77%	1.72%	Howlader et al. (2017)
$P(C_{14})$	IR Multiple myeloma	0.89%	0.65%	Howlader et al. (2017)
$P(C_{15})$	IR Non-Hodgkin lymphoma	2.38%	1.87%	Howlader et al. (2017)
$P(C_{16})$	IR Oral cavity and pharynx	1.61%	0.68%	Howlader et al. (2017)
$P(C_{17})$	IR Ovary		1.27%	Howlader et al. (2017)
$P(C_{18})$	IR Pancreas	1.58%	1.54%	Howlader et al. (2017)
$P(C_{19})$	IR Prostate	11.55%		Howlader et al. (2017)
$P(C_{20})$	IR Stomach	1.05%	0.65%	Howlader et al. (2017)
$P(C_{21})$	IR Testicles	0.40%		Howlader et al. (2017)
$P(C_{22})$	IR Thyroid	0.63%	1.79%	Howlader et al. (2017)
$P(C_{23})$	IR Uterine corpus		2.85%	Howlader et al. (2017)
π_1	Relative risk of C_1	2.8	2.73	Gandini et al. (2008), indicated there as lower urinary tract
π_4	Relative risk of C_4		1.83	Gandini et al. (2008)
π_5	Relative risk of C_5	1	1	Evidence in the literature is mixed. Assumed 1.
π_6	Relative risk of C_6	2.52	2.28	Gandini et al. (2008)
π_8	Relative risk of C_8	1.59	1.35	Gandini et al. (2008)
π_9	Relative risk of C_9	6.98	6.98	Gandini et al. (2008), no analysis by gender reported

Continued on next page

Input Parameters and Sources (*continued*)

Input	Description	Male	Female	Reference
π_{11}	Relative risk of C_{11}	1.85	1.49	Gandini et al. (2008)
π_{12}	Relative risk of C_{12}	9.87	7.58	Gandini et al. (2008)
π_{16}	Relative risk of C_{16}	5.1	5.1	Gandini et al. (2008), no analysis by gender reported
π_{18}	Relative risk of C_{18}	1.63	1.73	Gandini et al. (2008)
π_{20}	Relative risk of C_{20}	1.64	1.64	Gandini et al. (2008), no analysis by gender reported
π_{23}	Relative risk of C_{23}		2	Sasco et al. (2004)

Chapter 3

Informing, simulating experience, or both: A field experiment on phishing risks*

Joint with

Aurélien Baillon, Jeroen de Bruin, Evelien van de Veer, and Bram van Dijk

3.1 Introduction

Phishing attacks, the attempt to deceptively acquire personal and/or financial information (usernames, passwords etc.) by electronic communication, pose a significant threat for organizations. Social networks are increasingly used in phishing attacks but phishing by emails remains the main risk in an organizational setting. This is due to the relative simplicity of designing and sending phishing emails and its potential to reach many individuals at the same time. The text of a phishing email mostly addresses the recipient with urgency cues, words that invoke feelings of vulnerability or threat, in order to try to force the recipient to act immediately and impulsively. These urgency cues are most deceitful, because they turn attention away from other cues that may potentially help the receiver to recognize a phishing email (Vishwanath et al., 2011). Attackers can also trick users into downloading malicious malware, after they click on a link embedded in the email (Ramanathan and Wechsler, 2013). In recent years, phishing emails have evolved from poorly-designed and untargeted texts into highly personalized and sophisticated messages, which has made recipients more likely to believe that the content is expected and legitimate (Blythe et al., 2011; Berghel, 2006).

In order to cope with increased information security threats and ensure information security, organizations actively take technical security measures (Gupta et al., 2018). Although these protective mechanisms contribute to improved information security (Ransbotham and Mitra, 2009), it is rarely enough to entirely rely on them (Bada et al., 2015). Organizations that deploy both technical and non-technical protective means are likely to be more successful in protecting against information security risks (Pahnila et al., 2007; Vroom and von Solms, 2004; Bulgurcu et al., 2010).

*Published as " Baillon A., J. de Bruin, A. Emirmahmutoglu, E. van de Veer and B. van Dijk (2019). Informing, simulating experience, or both: A field experiment on phishing risks. *PLoS ONE*, 14(12): e0224216. https://doi.org/10.1371/journal.pone.0224216"

Organizations create policies and procedures to ensure information security (Whitman and Mattord, 2003). Although constructing policies and procedures is an essential outset, it is not enough to make sure employees comply with them. Vulnerabilities of the human factor in information security are usually ascribed to non-intentional behavior. Some may simply lack knowledge, skills and abilities to protect themselves against threats, and to comply with existing policies (Albrechtsen, 2007).

Is it enough to increase knowledge by providing more information to employees or should companies look for alternative approaches, letting employees experience (a simulated version of) the threat? Cai and Song (2017) showed that insurance take-up increased after people played an insurance game. In many cases, people who are victims of phishing fraud never realize that they have been a victim of a phishing attempt or realize it too late, when the extremely negative consequences occur. Letting employees experience (simulated) phishing fraud can complement general information provision. It can make them more receptive to the information.

In a large field experiment, we studied the effect of information provision and experience in reducing the phishing risks. Although existing studies have examined the impact of training and simulation on susceptibility to phishing fraud (Kumaraguru et al., 2007; Sheng et al., 2007, 2010; Downs et al., 2006; Wright and Marett, 2010; Bowen et al., 2011; Burns et al., 2013), many studies involved role-play activities or lab experiments. In lab experiments, the possibility of phishing email tends to be salient and day-to-day distractions, which increase phishing susceptibility in the real world, are absent. Field experiments avoid these problems but are more difficult to organize. A large-scale field experiment was conducted on students and staff of a university by Mohebzada et al. (2012). The authors found that warnings about phishing risk were not sufficient to prevent users from responding to phishing emails. Other field experiments have been conducted to study the risk factors (Silic and Back, 2016; Halevi et al., 2015; Williams et al., 2018) and the effectiveness of phishing exercises in organizations (Siadati et al., 2017) when phishing emails vary in terms of persuasiveness.

Our experiment was conducted at the Dutch Ministry of Economic Affairs, with more than 10,000 subjects who were unaware that they participated in such an experiment. We were thereby able to avoid the many biases which arise in a laboratory experimental setting, and managed to observe actual behavior in a setup that closely mirrors an actual phishing attack. Our experiment consisted of a control and three treatments: information, experience, and both. Information provision aimed at increasing (procedural) knowledge whereas simulated experience could make employees more alert about the threat and more receptive to information. Many authors have argued that combining information and simulated experience leads to stronger effects (Bowen et al., 2011; Kumaraguru et al., 2007; Burns et al., 2013; Sheng et al., 2007). Our experimental design allowed us to test for the existence of such synergies. Besides testing the overall effectiveness of our interventions, we were able to study whether gender, age, and employment contract affect an individual's susceptibility to phishing fraud. This information is of great value to policymakers both within and outside of corporations, as they enable targeted interventions (Kleitman et al., 2018).

In the experiment, we sent a phishing email to measure the susceptibility of employees to click on a dubious link and then give away their password. About one third of the subjects clicked on the link in our control treatment and 22% gave their password. Informing subjects about the risks of phishing reduced the proportion of subjects clicking on the link by 7 points and the proportion of password given away by 6 points. A first experience with a phishing email reduced the proportion of subjects clicking on the link by 9 points and of providing their password by 8 points. Combining both the information campaign and the experience intervention did not substantially improve the results with respect to experience alone. Overall, in an organization of the size of the Dutch Ministry of Economic Affairs, letting all employees experience one phishing email can avoid that 800 passwords are given away.

3.2 Conceptual framework

Several models have been proposed to explain phishing susceptibility. They include individual factors (personality, perception, knowledge, motivation) and phishing email characteristics (Vishwanath et al., 2018; Musuva et al., 2019). The latter cannot be influenced by the (victim) organization and we therefore focus on the former.

Due to limited attentional resources, people often rely on automatic or heuristic processing when reading e-mails, increasing the likelihood that people click on links or download malignant software (Vishwanath et al., 2018). The level of heuristic processing are influenced by cyber-risk beliefs, which refer to the perception that people have about online threats, and which are influenced by the degree of experience, efficacy, and knowledge that people have on the subject. According to the suspicion cognition automaticity model (SCAM) (Vishwanath et al., 2018), suspecting a specific email to be a phishing email is directly and indirectly influenced by general cyber-risk beliefs, which include risk perception. Higher risk perception obviously makes people more suspicious about emails but also makes them rely less on heuristic processing and more on systematic processing when reading emails (Vishwanath et al., 2018). A deeper, more systematic processing of the information contained in the emails, including the cues signaling phishing threats, makes people better able to detect phishing attacks.

Drawing from the field of education science, a simulated experience is argued to be an effective substitute for learning from an actual experience, especially when simulations provide concrete and emotionally charged experiences (Zigmont et al., 2011). Similar to an experience of an actual phishing email, a simulated phishing email may in this way increase risk perception and subsequently the level of systematic processing and the degree of suspicion applied to future emails.

Knowledge, on the other hand, refers to the information the receiver has about the domain of the threat, the ways to detect it and the required actions (Musuva et al., 2019; Kleitman et al., 2018). Information provision may increase knowledge about the prevalence of phishing attacks and the risks these pose. The information can be about how to detect phishing emails and how to deal with them, which can raise efficacy. Increased knowledge and efficacy together may increase cyber-risk beliefs and subsequently lower the reliance on heuristic processing.

Many papers recommend to incorporate information and experience of phishing emails into one training material (Bowen et al., 2011; Kumaraguru et al., 2007; Burns et al., 2013; Sheng et al., 2007). Crucial is to embed the training in the employees' natural environment (Xiong et al., 2019). Arachchilage and Love (2014) found that combining conceptual knowledge and procedural knowledge is key in enhancing phishing detection and avoidance. A one-time simulated experience is more likely to affect conceptual knowledge ("know that") than the procedural knowledge ("know how").[1] Information provision can target both. Yet, giving information after employees have received a phishing email can increase the perceived relevance of the information. Information campaigns about procedures to avoid phishing have indeed been found to have greater impact after users have fallen for an attack (Alsharnouby et al., 2015). From the literature we therefore expect both simulated experience and information provision to be effective, but we also expect their combination to be most effective.

3.3 Method

To test the effectiveness of combining (one-time) experience and information provision on phishing risks, we conducted a large-scale field experiment. The main characteristics of the experiment are outlined below, with further details in Appendix 3.A.

3.3.1 Subjects and design

The subjects of this experiment were 10,929 employees of the Dutch Ministry of Economic Affairs, out of the 12,567 official employees of the Ministry. Reasons for exclusions were mostly technical (e.g., missing information), or related to the rank at the Ministry (Minister, Secretary General...). Details are reported in Appendix 3.A. Most subjects were males (60,6%), with an average age of 47 years (Table 2). Subjects learned that they were part of an experiment only after the experiment was conducted.

We implemented two interventions: information provision (*Info*) and experience (*Exp*). We used a 2x2 design and subjects were divided into four groups of roughly equal size: *Control* (2723 employees), *Info* (2740), *Exp* (2724) and *ExpInfo* (2742).

[1] Repeated simulated experience could also improve procedural knowledge if sufficient feedback is provided.

The experiment was organized around five specific dates $T \in \{1, ..., 5\}$ (see Table 1 for the exact dates). All four groups received an email that resembled a real phishing email at $T = 5$. In the following, we simply refer to this and all other such emails that we sent a simply phishing email. The first treatment group (*Info*) received an information email in $T = 2$, $T = 3$ and $T = 4$ with information about what phishing is and how it works ($T = 2$), how one can recognize phishing emails ($T = 3$), and what one should do when receiving a phishing email ($T = 4$). The second treatment group (*Exp*) was sent a phishing email at time $T = 1$ and a short debriefing email explaining that the sent email was fake at the end of the same day. No information was given that the emails were part of an experiment. The third treatment group (*ExpInfo*), received both interventions, thus the phishing email and debriefing email at $T = 1$ and the information emails at $T = 2$, $T = 3$ and $T = 4$. One day after the phishing mail at $T = 5$, all four groups received a general debriefing. Table 1 below summarizes the experimental timeline and gives the exact dates of the experiment.

The choice of the timeline followed several constraints. First, to be in line with our conceptual framework, experience had to precede information. We would then expect subjects from the *ExpInfo* treatment to be more alert about phishing and more receptive to information about it than those from the *Info* treatment, who were not exposed to the first phishing experience. Second, to avoid having one very long email that might discourage readers, the information provided was split between three emails, which were spread over three weeks. Third, we waited one additional week after the end of the *Info* and *ExpInfo* interventions before sending the final phishing email. Sending the final phishing email later was not possible because many employees would be on vacation (Christmas break). Sending it earlier was undesirable. An email on the same day or a day later than the *Info* and *ExpInfo* interventions might be too easily detectable, inflating the measured effectiveness of these treatments. A similar one-week delay was used for instance by Xiong et al. Xiong et al. (2019) to study the effect of training on phishing detection.

A privacy impact assessment was drawn up to identify potential issues concerning privacy and informed consent. Based on this, the following measures were taken: (1) the analysis was done on anonymized data, the reporting of the results is only on the basis of aggregated data; (2) prior to the experiment the general norm of Information Security System Policy compliance was posted on the intranet; (3) the Employees Council of the Ministry was informed; (4) passwords and other information given by employees were not recorded; and (5) after the experiment, employees were debriefed through an extensive email and were provided with contact details of the researchers. The secretary general and head of internal organization/chief information officer of the ministry gave their (written) approval of the research.

Table 1: **Experimental timeline**

	$T = 1$ 05/11/2015	$T = 2$ 19/11/2015	$T = 3$ 26/11/2015	$T = 4$ 03/12/2015	$T = 5$ 15/12/2015
Control					Phishing mail + debriefing
Info		Infographic 1	Infographic 2	Infographic 3	Phishing mail + debriefing
Exp	Phishing mail + short debriefing				Phishing mail + debriefing
ExpInfo	Phishing mail + short debriefing	Infographic 1	Infographic 2	Infographic 3	Phishing mail + debriefing

3.3.2 Group formation

We randomized the subjects at the level of the lowest known organizational unit, henceforth referred to as “basic unit”. In total we had 184 unique basic units, with an average of 61 subjects per basic unit. We have not opted for randomization at the individual level to avoid contamination of the results by intervention spillover effects. With randomization at the individual level, it would have been possible that two subjects working together were divided into different treatment groups. This could have led some subjects to be affected by more than one treatment.

Table 2 shows the distribution of subjects across the four groups. We ran Kruskall-Wallis tests to check if the subjects were equally divided between groups in terms of age, age-groups, gender, employment contract (internal/external), and (the five largest) organizational division (in this paper we refer to them as A, B, C, D, E). Test results showed no statistically significant differences between groups in all variables except, as could be expected, for organizational division (Appendix 3.B Table 5). Not all organizational divisions had the same number of basic units (with division A even having only one basic unit) and the size of the basic units varied substantially. We will control for divisional differences in our analysis but it is worth noting that our randomization was successful on all other aspects.

Unfortunately (and out of our control), after the first phishing email at $T = 1$, an online notification was posted for the employees in organizational division C, stating that the phishing email that some employees received was a fake one. Hence the subjects in the *Control* and *Info* groups received this notification as well. This may have affected the results as it created intervention spillover effects within that division. We would then expect treatments effects to be smaller for that division. In what follows, we will always report the analysis with and without this division. Including the division can be expected to give more conservative estimates. Table 3 shows the characteristics of the reduced sample. The distributions of different groups show significant differences on this sample (see Appendix 3.B Table 5 for Kruskall-Wallis test results).

Table 2: **Descriptive statistics**

Group	N. of subjects	Male	Age						Internal Employee	Organisational division				
			Mean	16-25	26-35	36-45	46-55	>55		A	B	C	D	E
Control	2723	60.52%	47.45	2.90%	10.54%	26.07%	35.51%	24.97%	80.21%	–	14.18%	33.79%	19.32%	32.72%
Info	2740	61.06%	47.35	2.04%	15.26%	25.69%	28.39%	28.61%	79.34%	9.56%	12.23%	19.60%	13.61%	45.00%
Exp	2724	59.99%	47.05	2.86%	13.07%	28.45%	29.22%	26.40%	80.76%	–	10.17%	27.09%	25.33%	37.41%
ExpInfo	2742	60.76%	47.31	2.12%	12.65%	26.81%	33.33%	25.09%	80.49%	–	12.47%	26.99%	28.05%	32.49%
Whole sample	**10929**	**60.58%**	**47.29**	**2.48%**	**12.88%**	**26.75%**	**31.61%**	**26.27%**	**80.20%**	**2.4%**	**12.26%**	**26.86%**	**21.58%**	**36.91%**

Table 3: **Descriptive statistics after the exclusion of division C**

Group	N. of subjects	Male	Age						Internal Employee	Organisational division			
			Mean	16-25	26-35	36-45	46-55	>55		A	B	D	E
Control	1803	60.68%	47.45	0.94%	11.15%	29.62%	34.33%	23.96%	70.27%	–	21.41%	29.17%	49.42%
Info	2203	56.83%	45.69	2.36%	17.43%	28.92%	28.92%	22.38%	74.58%	11.89%	15.21%	16.93%	55.97%
Exp	1986	58.91%	46.14	2.82%	14.20%	31.77%	28.30%	22.91%	74.42%	–	13.95%	34.74%	51.31%
ExpInfo	2002	60.24%	46.72	2.35%	13.34%	29.27%	31.72%	23.33%	73.53%	–	17.08%	38.41%	44.51%
Reduced sample	**7994**	**59.07%**	**46.46**	**2.15%**	**14.19%**	**29.87%**	**30.69%**	**23.10%**	**73.30%**	**3.28%**	**16.76%**	**29.50%**	**50.46%**

3.3.3 Procedure

Pre-intervention period (all groups)

We first ensured minimum knowledge about the information security policy of the ministry by posting a service notice on the intranet of all five organizational divisions of the ministry prior to the experiment, and visible to all subjects. This message explained the dangers of giving away personal details. Furthermore, the message stated that the Ministry or a division of the Ministry would never ask employees for their password, username etc.

T=1: First phishing email: simulating experience and feedback (Groups *Exp* and *ExpInfo*

Subjects from the *Exp* and *ExpInfo* treatments received an imitation of a real phishing email. The subject line was: "Economic Affairs – Mobile Password Recovery System". This email was sent by the operational management, and subjects were asked to link their account to their mobile phone number in order to recover their password easily if it was lost, or to change it.

The email contained several characteristics enabling receivers to assess the email as being fake/fraudulent, presenting more or less the same level of difficulty as phishing emails that were actually sent at that period. These characteristics were: (1) a misspell in the sender email, (2) inappropriate use of capital letters in the subject line, (3) a change in the logo and logo color, (4) an unusual form of salutation for the Ministry, (5) addressing the receiver in the formal form instead of the informal form, which is normally used, (6) a hyperlink in the email that refers to a vague website with an extension that would normally not be used within the Ministry (.net) and (7) two different but resembling fonts in the main text and the disclaimer (Appendix 3.C Figure 3).

We chose an email subject and sender, which we believed to be equally relevant to most subjects. The link in the email redirected the subjects to a "fake" website (`www.mobilepasswordrecoverysystem.net`). This website had a very basic design and contained a few elements of the governmental visual design style, with some modifications. In order to link their accounts and phone numbers, subjects were asked to fill in three personal details; (1) username, (2) password, and (3) phone number. After filling in the details, subjects were redirected to a second screen, thanking them for the registration and stating that the registration would be completed within five working days. It was not necessary to fill in all the three personal details. Even if a subject filled in only one field and clicked on "send", s/he was directed to the second screen that thanked for his/her registration (Appendix 3.C Figure 4 and 5).

At the end of the day, all subjects from treatments *Exp* and *ExpInfo* received a short debriefing explaining that the email was an "imitation" email designed to increase awareness for phishing fraud. No information was given that the email was part of an experiment or that there would be follow-up actions (Appendix 3.C Figure 6). The debriefing email focused on important it was that all employees contribute to a safer digital environment. By receiving that email, subject could also learn whether they had made a mistake or not. There was no information about how to recognize phishing, nor about how to react. Subjects from the *ExpInfo* treatment would receive such information in the following weeks, as described next.

T=2,3,4: Information Provision – Infographics (Groups *Info* and *ExpInfo*)

Subjects in treatments *Info* and *ExpInfo* received emails explaining ways to avoid falling for phishing attacks. This information provision occurred in three consecutive weeks, using colorful infographics to maximize the impact of the treatments. The first email explained what phishing is and how it works, the second how receivers could recognize them, and the third what actions receivers should undertake when they receive a phishing email (Appendix 3.C Figure 7). The infographics were designed in a way that (a) makes the subjects understand the risks, (b) keeps the message simple and short, (c) provides clear actionable items that subjects can easily adopt, and (d) uses story-based graphics as suggested by Kumaraguru et al., 2007 and Sheng et al., 2007 Kumaraguru et al. (2007); Sheng et al. (2007).

T=5: Second Phishing email (all groups)

All subjects of the four groups received a (second) phishing email forty days after the subjects in treatments *Exp* and *ExpInfo* had received a phishing email and twelve days after the subjects from the *Info* and *ExpInfo* treatments had received the last infographics. The second phishing email resembled the first one in terms of looks, length, and recognizable characteristics of phishing mails. Subjects from the *Info* and *ExpInfo*) treatments who would apply the recommendations they received in the infographics should recognize it as a phishing email.

This email was sent by the IT department of the Ministry (with a misspelling in the sender address: helpdesk@dlctu.nl instead of helpdesk@dictu.nl). Subjects were told that they had reached their maximum storage limit of Outlook and the limit had to be raised via a hyperlink to `www.verhooogjeopslaglimiet.net`, which can be translated as `www.increaseyourstoragelimit.net` (Appendix 3.C Figure 8). The email asked for an immediate action of the subjects. If they clicked on the link in the email, they were directed to the website. The website they would reach was again basic, with some visuals of Outlook Exchange, subjects were told that by filling in e-mail, username and password, limits could be raised up to 8 GB. If the subject indeed filled in the details, a pop-up screen was shown, stating that the registration was being processed and that it would be completed within five workdays (Appendix 3.C Figure 9 and 10).

Post-intervention period – Debriefing (all groups)

All subjects received a general debriefing the day after receiving the (second) phishing email. In this elaborate debriefing the subjects were told that the phishing email(s) and information mails were part of an experiment. They were given information about; (1) the cause and purpose of the research, (2) the design of the research, (3) which precautions had been taken in order to respect the privacy of employees and to protect (personal) details, and (4) where subjects could submit other questions and/or remarks.

Furthermore, they were informed that the experiment was part of the campaign iBewust-zijn (Information awareness). With this campaign the Ministry aimed to encourage and support its employees as much as possible in developing knowledge and awareness regarding information security. Also, it reassured the employees that the phishing mail was fake, such that no consequences were attached if subjects indeed had filled in personal details.

3.4 Analysis

Data was collected on whether a subject had clicked on the link and had filled in one or more personal details, and the time of completion. For privacy concerns, the content of what subjects had filled in was not registered. The analysis was conducted on 10,929 observations from the whole sample and on 7,994 observations when division C was excluded.

We measured falling for phishing with three dummy variables: *Visit*, *Fill* and $Fill|Visit$. *Visit* indicates whether the subject clicked on the link and visited the website. Irrespective of whether the subjects filled in personal details, clicking on a link embedded in a phishing email by itself can be very dangerous since such links may infect computers with malware. *Fill* takes value 1 if the subject filled in their password. Although the subjects could also fill in their username or mobile phone number/email address, we chose password as the variable of interest since we regard it as the most confidential data among all.[2] However, the results are robust to other variables as well since 99.15% of the subjects who filled in any field did so for all three fields asked. Finally, *Fill*|*Visit* is an indicator variable for whether subjects filled in the password given that they had clicked on the link (hence excluding subjects who did not visit the website). The average of *Fill*|*Visit* can be interpreted as the probability to fill in the password field conditional on visiting the phishing website. It informs us whether subjects recognized the phishing fraud only after they visited the website.

We performed two types of analysis that were chosen to account for cluster randomized trials (the clusters being basic units described in section 2.2). First, we tested the effectiveness of interventions by weighted t-tests on cluster averages of *Visit*, *Fill* and *Fill*|*Visit*. For each pair of treatments, the weighted t-test compared the percentages of subjects falling for phishing email of all clusters (basic units) of one treatment with those of the other treatment, but weighting the percentages by the cluster size. Next, we performed logistic regressions on all three variables of interests with standard errors clustered at basic unit level. We also controlled for treatment group, gender, employment contract, organizational division and age.

3.5 Results

3.5.1 Weighted t-tests

Fig. 1 and 2 display, for the whole and reduced sample respectively, the proportion of subjects falling for phishing email in each treatment group for each measure, and reports the significance level of the weighted t-tests. The detailed results of weighted-t tests are given in Appendix 3.B Table 7. We describe here the results for the whole sample. Treatment effects are larger when excluding division C. While a third of the subjects in the *Control* group failed to recognize that the received mail was a phishing mail and clicked on the link, the proportion of people visiting the phishing website dropped by 7 to 9 percentage points in the intervention groups.

Among those who visited the link, 68% also filled in their password in the *Control* group. This proportion was very similar in *Info*, being only reduced by 4 points, which was not significant. By contrast, *Fill*|*Visit* was 58% in *Exp*, and 54% in *ExpInfo*, both significantly lower than in the *Control* group. The difference between *Info* and *ExpInfo* was also significant at a 1% level. The differences between interventions, on the other hand, were not significant with the exception of *Info* and *ExpInfo* in Fig. 1C.

The unconditional frequency of subjects filling in their password, as measured by *Fill*, was

[2]There were no subjects who filled in only the password, supporting the idea that people are more reluctant to give this information away.

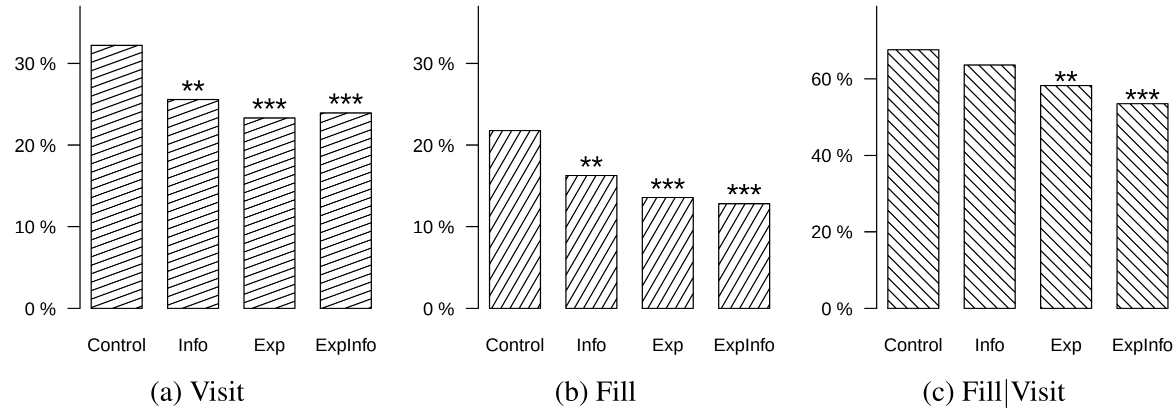


Figure 1: Percentages of subjects falling for phishing email (whole sample)
Stars indicating significance levels for difference of each treatment group compared to the control group with * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

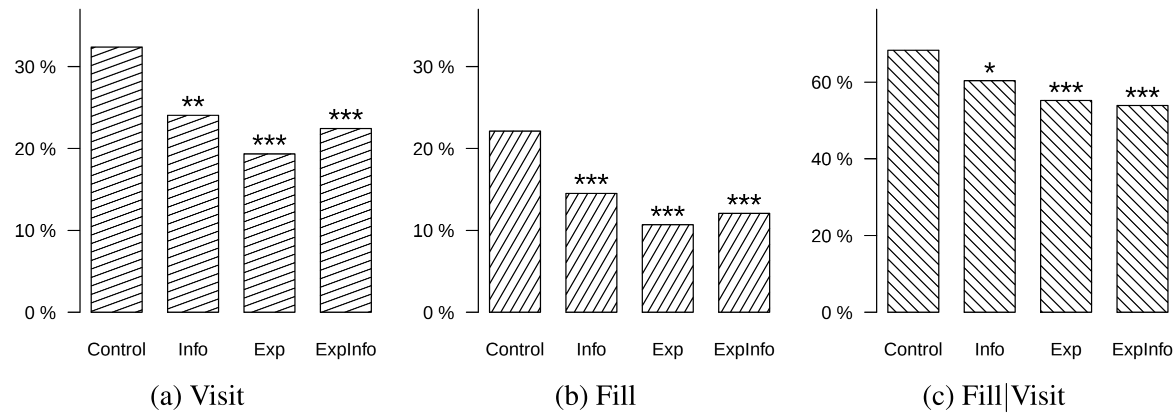


Figure 2: Percentages of subjects falling for phishing email (excluding division C)
Stars indicating significance levels for difference of each treatment group compared to the control group with * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

22% in *Control*, and dropped to 16%, 14%, and 13% in *Info*, *Exp*, and *ExpInfo* respectively. The effects were stronger after exclusion of Division C. A decrease of 8 percentage points in *Fill* means about 200 fewer passwords given away in each of the intervention treatments than in *Control*. Implementing these simple interventions for an organization of the size of the ministry can avoid that 800 passwords are given to hackers.

3.5.2 Regression analysis

The results above gave us a first glimpse on the effect of interventions on tackling phishing fraud. As explained above, randomization was only done at the level of basic units, and the size of those varied across organizational divisions. Moreover, as described in section 3.3.2 we found some differences between groups in some descriptive statistics when division C was excluded. To account for these differences, we additionally ran logistic regressions with clustered standard errors, controlling for descriptive variables and divisions. This second type of analysis also informed us on the characteristics of the employees who were more prone to click on a dubious link or give away their password. Table 4 presents the average marginal effects of each independent variable on the predicted probability of falling for the phishing email estimated on observation values. We discuss below the effects for the whole sample. The effect sizes were larger after excluding division C and we only describe them when they lead to different conclusions. As a robustness check we ran panel logistic regression with random effects using our basic units as panel variable (identifier). Our results were robust to this specification (see Appendix 3.B Table 8).

Table 4: **Logistic regression analysis - Average marginal effects**

	dy/dx - Whole sample			dy/dx - excluding division C		
	Visit	Fill	Fill|Visit	Visit	Fill	Fill|Visit
Treatment						
$Control$ reference						
Info	-0.062^{*}	-0.056^{**}	-0.050	-0.083^{***}	-0.082^{***}	-0.103^{***}
	(0.026)	(0.021)	(0.027)	(0.020)	(0.016)	(0.027)
Exp	-0.082^{***}	-0.078^{***}	-0.102^{***}	-0.124^{***}	-0.112^{***}	-0.135^{***}
	(0.023)	(0.018)	(0.031)	(0.019)	(0.012)	(0.037)
ExpInfo	-0.079^{***}	-0.087^{***}	-0.148^{***}	-0.097^{***}	-0.099^{***}	-0.154^{***}
	(0.019)	(0.017)	(0.030)	(0.016)	(0.015)	(0.034)
Gender						
male	0.042^{***}	0.017	-0.041	0.029^{*}	0.005	-0.054^{*}
	(0.010)	(0.011)	(0.024)	(0.012)	(0.013)	(0.027)
Employee contract						
Internal Employee	0.025	-0.000	-0.064^{*}	0.028	0.006	-0.048
	(0.019)	(0.012)	(0.027)	(0.018)	(0.012)	(0.028)
Age group						
16-25 reference						
26-35	0.035	0.027	0.073	0.029	0.001	-0.095
	(0.027)	(0.019)	(0.100)	(0.029)	(0.025)	(0.108)
36-45	0.093^{***}	0.067^{***}	0.114	0.072^{*}	0.024	-0.097
	(0.027)	(0.018)	(0.090)	(0.030)	(0.023)	(0.085)
46-55	0.147^{***}	0.118^{***}	0.192^{*}	0.131^{***}	0.075^{**}	-0.025
	(0.027)	(0.017)	(0.089)	(0.031)	(0.024)	(0.083)
>55	0.147^{***}	0.138^{***}	0.264^{**}	0.122^{***}	0.091^{**}	0.063
	(0.031)	(0.022)	(0.094)	(0.033)	(0.028)	(0.097)
Division dummies	Yes	Yes	Yes	Yes	Yes	Yes
Observations	10929	10929	2869	7994	7994	1947

Division dummies are added with division B as reference category.
Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

There is strong evidence that all treatments have a significant negative effect. The probability of visiting the phishing website decreased by around 6 points in the *Info* treatment and 8 points with the other two interventions. The marginal effects of the three interventions were about the same when we studied the probability to fill in the password (see column *Fill*) instead of the probability to click on the link (column *Visit*). The probability to fill in a password conditional on visiting the website was reduced by 10 points in the *Exp* group and by 15 points in the *ExpInfo* group. The effect of *Info* on $Fill|Visit$ was not significant on the whole sample, but it was when excluding division C. We tested the differences between the treatment effects and found no significant difference, with three exceptions. The effect of *Info* on $Fill|Visit$ was smaller than that of *ExpInfo* in the whole sample and the effects of *Info* on *Fill* and on *Visit* were smaller than those of *Exp* after excluding division C (see Appendix 3.B Table 9).

We found some evidence that men were more likely to click on the phishing link but less likely to fill in their password afterwards once they were on the website. Even if men were not more likely than women to give away their password overall (no gender effect on *Fill*), their propensity to click on phishing links could pose a threat when such links trigger malware download. The younger age group (16-25) was the least likely to visit the phishing website. Employees between 36 and 45 were 9% more likely to click on the link than the 16-25 age group and those above 46 were almost 15% more likely to click on the link than the youngest group. The effect of age on the probability to fill in the password conditional on visiting the website were not robust to excluding division C and we therefore refrain from commenting them.

3.6 Discussion

In our field experiment, we observed a non-negligible proportion of subjects falling into a phishing attack. An information campaign substantially reduced that risk, and letting subjects experience a phishing email tended to be at least as effective. Personal experience may have led people to see threats as more probable, and to view themselves as potential future victims, suggesting that cyber-risk beliefs were the most serious barrier to phishing detection. Moreover, people may be more likely to read phishing and/or debriefing emails than information emails. We were not able to investigate this because we could not check whether subjects opened/read the emails we sent. We cannot exclude that the experience intervention was effective because it made employees believe that their employer could know whether they easily give away their password.

It has been argued that information becomes more relevant if it is given after individuals experience the phishing email. Many suggested that combining the two types of intervention yields the best results by reinforcing the effect of each other Bowen et al. (2011); Kumaraguru et al. (2007); Burns et al. (2013); Sheng et al. (2007). Surprisingly, in our experiment, the effectiveness of combining both experience and information did not differ from the effectiveness of experience alone. A possible explanation was that the experience intervention increased employees' perceived risk enough for them to acquire knowledge on their own or activated previously existing knowledge. Another possible explanation for the absence of synergy is the existence of a ceiling effect. It may be difficult to reduce the proportion of password given away much below 10% with simple interventions as ours. Our results question whether it is worth piling up interventions against phishing fraud. Each intervention requires sending several emails to users. In organizations in which people complain about getting too many emails and cognitive overload, one may prefer to focus on the most effective intervention.

Info and *Exp* differed in terms of content (infographics versus a phishing email) but also in terms of dates in which this content was sent. It allowed us to combine both interventions in *Exp* but it decreased the comparability between *Info* and *Exp*. We should therefore be careful when comparing the effects of these two treatments. If anything, we would have expected the *Info* treatment, with more emails and those being sent closer in time to the final phishing email, to be more effective than the *Exp* treatment. This is not what was found. However, it is sound to compare the effect of infographics on its own with the effect of infographics when combined with experience. We expected the latter to be larger, but it was not the case.

The *Exp* treatment involved a first phishing email followed by a debriefing email. We dubbed the corresponding intervention ‘experience’ but it involved both experience and feedback. Our feedback was rather limited (see Appendix 3.C Figure 6) and independent of whether people had visited the phishing website and whether they had given their password. Further research could vary the degree of feedback and how personalized it is.

We have interpreted filling in the password field as giving away their password but we cannot know whether they provided their true password or not. What they provided was not saved, for privacy and safety reasons, and we did not have access to their true password anyhow. It could be that some people filled in fake information. We should therefore be cautious with this part of the results. Even if the provided information was not correct, having visited the link already posed a threat. Visiting fraudulent links makes people susceptible to malware attacks and this behavior should be eliminated in the first place.

The treatment effects we observed were slightly lower for the analysis on the whole sample than after excluding division C. In division C, a message was posted online after treatments *Exp* and *ExpInfo* received the first phishing email, thereby affecting the *Control* and *Info* groups as well. If anything, the results on the whole sample give us the lower bound. The effectiveness of experience could also be studied, but in a less controlled way, by comparing the proportion of subjects falling for phishing in the first email and the second email in the two treatments in which two phishing emails were sent. However, this would only be possible if the tests were identical. About 15% of subjects in *Exp* and *ExpInfo* groups clicked on the link in the first email while more than 20% did so in the second email. The difference can come from the second email being more difficult to detect than the first one but also from sending the debriefing email earlier after the first phishing email than after the second. By contrast, comparing the *Exp* and *ExpInfo* treatments with the control as we did in the result section does not suffer from such possible confounds.

Making people more aware of phishing threats and asking them to report suspicious emails may backfire in high number of false positives, i.e., employees misjudging genuine emails and reporting them to the IT department. Kleitman et al. Kleitman et al. (2018) studied which characteristics influence phishing susceptibility but also false positives. We do not have evidence about false positives in our experiment. The operations department of the ministry, in charge of information security, did not report that it was a problem at the time of the study. The official policy was that people should report any doubtful email, the cost of a false positive being judged much lower than that of successful phishing. However, this reasoning was based on the experience of the operations department and their cost-benefit analysis at the time of the experiment. In other instances, anti-phishing campaigns may create a burden on IT departments and generate other organizational costs if the rate of false positives upsurges.

3.7 Conclusion

In a field experiment, we studied the effect of experience, information, and their combination on employees' reaction to phishing emails. Our information treatment was designed to have a maximal impact, with infographics and clear messages. We could expect the infographics to be especially effective after a first (simulated) phishing experience. Each intervention in isolation had a large effect on the probability to click on a dubious link and to give away personal details. Combining both interventions did not substantially increase the effect of experience alone, even though subjects in the experience treatment were only exposed to one experience. Our results question the opportunity of piling up (costly) interventions.

Appendix 3.A Additional information about the method

3.A.1 Subjects

The total number of employees in the Dutch Ministry of Economic Affairs was 12,567. We excluded those employees with errors/flaws in the dataset, i.e. 40 employees due to their age (age < 16 & age > 70), 17 due to missing organizational unit information and 408 due to missing e-mail addresses. Furthermore, we excluded the departments which were covered by a different IT assistance. Finally, some employees were excluded due to their rank in the organization (Minister, State Secretary and Secretary General etc.).

3.A.2 Procedure

Pre-intervention

We made arrangements with several parties such as the IT helpdesk and the information security coordinators in case subjects contacted them during the experiment. We wanted to assure that none of these parties would inform the subjects that they were part of an experiment, because this could confound our results. Therefore, we provided them with standardized protocols for email and phone. This allowed these parties to answer the possible questions of subjects, without disturbance of the experiment.

Moreover, at the IT helpdesk a dedicated group of employees was formed, who were in first line of contact with the subjects. Subjects were redirected to this dedicated group through a choice option at the service line or by automatically forwarding emails to them. When subjects phoned or emailed them, they were told standardized answers, which were carefully constructed and suited for each specific question participants could ask. This way, we could; (1) inform/help subjects as much as possible, without informing them of the experiment, (2) control the outgoing messages.

T = 1

Participants who recognized the mail as being fraudulent and send it to the IT helpdesk (of the ministry of Economic Affairs), received an answer stating that the mail was indeed a phishing email, that the threat had receded, and that no further actions on behalf of the user were needed. If participants mailed the IT helpdesk, without adding the mail, they were still asked to send the phishing email. This was to establish with certainty that the notification indeed concerned our 'imitation' phishing email and not a 'real-non-experiment-related' phishing email. Employees who phoned the IT helpdesk to report the 'imitation' phishing email, were first asked the subject and sender of the email. Furthermore, if it indeed concerned our email, they were asked to send the email as an attachment in an email to the IT helpdesk. Employees who did not see the email as suspicious but asked substantive questions regarding the 'mobile password recovery system' were given the answer that they would receive an answer within three workdays (which is standard protocol for all IT related questions).

T = 2, 3, 4

The first information email starts with a short introduction of the director of operational management. The introduction states that the received mail is part of an information provision campaign, consisting of three information emails. Furthermore the subjects are given information about how phishing fraud works and the newer generation of phishing fraud, spear phishing attacks.

The second information mail starts with a short recap of the first information mail. Furthermore it provides the reader with six points of recognition by which he/she could determine whether emails are fraudulent or not, such as; (1) the mail address of the sender, (2) the salutation, (3) style of writing (grammatical or spelling errors), (4) the hyperlink in the mail, (5) look the sender up on the internet, (6) check the mail address in the signature. Also it lists the types of information that the ministry never will ask their employees. Finally, it states the three largest consequences (for organizations) of phishing fraud.

The third information mail starts again with a short recap of the first two information mails. Furthermore it provides information about, how you should act in the case that (you think that) you have received a phishing mail and how and where you could report a phishing mail. It concludes with some useful links to other (phishing awareness) campaigns of the Dutch Ministry of Economic Affairs (www.veiliginternetten.nl and iBewustzijn) where more information could be found, including some examples of phishing mails.

Moreover the three infographics had some synergetic power, as each of the three infographics starts with a short recap of the previous mail(s). This ensures the participants that each individual mail is understandable in itself, without necessity of reading them all. Furthermore, the infographics; (1) were readable on the mobile phone, laptop and tablet, (2) did not need to be downloaded first, and (3) were in line with the visual identity style of the Ministry. The story based graphics and annotations, were highly related to the core business activities of the employees and all images/pictures were copyright free, bought or self-made.

Appendix 3.B Additional statistics

Table 5: Kruskal-Wallis test results for equality groups formed

	Whole sample		Excluding div. C	
Attribute	χ^2	p	χ^2	p
Age	2.479	0.479	28.645	0.000
Age-Group	3.697	0.296	26.179	0.000
Gender	0.508	0.917	5.544	0.136
Employee Contract	0.932	0.818	6.822	0.078
Organisational division	21.256	0.000	15.274	0.002

Table 6: Percentages of subjects falling for phishing email

	Whole Sample			Excluding div. C		
	Visit	*Fill*	*Fill|Visit*	*Visit*	*Fill*	*Fill|Visit*
Control	32.21%	21.78%	67.62%	32.39%	22.13%	68.32%
Info	25.58%	16.28%	63.62%	24.06%	14.53%	60.38%
Exp	23.31%	13.58%	58.27%	19.34%	10.67%	55.21%
ExpInfo	23.92%	12.80%	53.51%	22.43%	12.09%	53.90%

Table 7: Weighted t-test results

(a) Whole sample

Visit

	Info t	*Info* p	*Exp* t	*Exp* p	*ExpInfo* t	*ExpInfo* p
Control	2.220	0.032	3.723	0.000	3.830	0.000
Info			0.747	0.459	0.579	0.566
Exp					-0.273	0.786

Fill

	Info t	*Info* p	*Exp* t	*Exp* p	*ExpInfo* t	*ExpInfo* p
Control	2.361	0.023	4.328	0.000	5.116	0.000
Info			1.163	0.251	1.577	0.123
Exp					0.450	0.654

Fill|Visit

	Info t	*Info* p	*Exp* t	*Exp* p	*ExpInfo* t	*ExpInfo* p
Control	1.103	0.276	2.542	0.014	4.176	0.000
Info			1.253	0.216	2.707	0.009
Exp					1.486	0.142

(b) Excluding div. C

Visit

	Info t	*Info* p	*Exp* t	*Exp* p	*ExpInfo* t	*ExpInfo* p
Control	2.642	0.012	5.159	0.000	3.971	0.000
Info			1.619	0.115	0.563	0.577
Exp					-1.404	0.167

Fill

	Info t	*Info* p	*Exp* t	*Exp* p	*ExpInfo* t	*ExpInfo* p
Control	3.227	0.003	5.781	0.000	4.794	0.000
Info			1.909	0.065	1.145	0.259
Exp					-0.829	0.411

Fill|Visit

	Info t	*Info* p	*Exp* t	*Exp* p	*ExpInfo* t	*ExpInfo* p
Control	1.870	0.070	2.767	0.009	3.618	0.001
Info			0.942	0.352	1.604	0.116
Exp					0.581	0.564

Table 8: XTLogistic regression analysis - Average marginal effects

	dy/dx - All observations			dy/dx - excluding NVMA		
	Visit	Fill	Fill|Visit	Visit	Fill	Fill|Visit
Treatment						
$Control$ omitted						
Info	-0.079^{*}	-0.069^{*}	-0.056	-0.096^{*}	-0.096^{**}	-0.120^{**}
	(0.037)	(0.029)	(0.041)	(0.043)	(0.033)	(0.043)
Exp	-0.088^{**}	-0.068^{**}	-0.082^{*}	-0.137^{***}	-0.113^{***}	-0.125^{**}
	(0.033)	(0.026)	(0.038)	(0.040)	(0.031)	(0.041)
ExpInfo	-0.087^{**}	-0.085^{***}	-0.139^{***}	-0.110^{**}	-0.106^{***}	-0.150^{***}
	(0.031)	(0.025)	(0.036)	(0.039)	(0.030)	(0.038)
Gender						
male	0.040^{***}	0.015^{*}	-0.040^{*}	0.032^{**}	0.010	-0.049^{*}
	(0.009)	(0.008)	(0.020)	(0.010)	(0.009)	(0.024)
Employee contract						
Internal Employee	0.005	-0.014	-0.065^{*}	0.009	-0.006	-0.046
	(0.014)	(0.011)	(0.030)	(0.013)	(0.011)	(0.030)
Age group						
16-25 omitted						
26-35	0.022	0.022	0.091	0.031	0.005	-0.089
	(0.024)	(0.016)	(0.085)	(0.030)	(0.024)	(0.110)
36-45	0.082^{***}	0.062^{***}	0.135	0.077^{**}	0.030	-0.089
	(0.024)	(0.016)	(0.081)	(0.030)	(0.023)	(0.107)
46-55	0.133^{***}	0.110^{***}	0.215^{**}	0.134^{***}	0.080^{***}	-0.017
	(0.024)	(0.017)	(0.081)	(0.030)	(0.024)	(0.106)
>55	0.137^{***}	0.132^{***}	0.285^{***}	0.129^{***}	0.098^{***}	0.070
	(0.025)	(0.018)	(0.081)	(0.031)	(0.025)	(0.107)
Division fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	10929	10929	2869	7994	7994	1947

Division fixed effects are added while base division is B.

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 9: χ^2 test for differences of groups in logistic regression

(a) Whole sample

Visit

	Exp		*ExpInfo*	
	χ^2	p	χ^2	p
Info	0.51	0.475	0.47	0.493
Exp			0.02	0.888

Fill

	Exp		*ExpInfo*	
	χ^2	p	χ^2	p
Info	1.08	0.299	2.49	0.115
Exp			0.33	0.565

Fill$|$Visit

	Exp		*ExpInfo*	
	χ^2	p	χ^2	p
Info	1.92	0.166	7.41	0.007
Exp			1.43	0.232

(b) Excluding div. C

Visit

	Exp		*ExpInfo*	
	χ^2	p	χ^2	p
Info	3.68	0.055	0.54	0.461
Exp			2.07	0.150

Fill

	Exp		*ExpInfo*	
	χ^2	p	χ^2	p
Info	4.87	0.027	1.26	0.262
Exp			0.86	0.354

Fill$|$Visit

	Exp		*ExpInfo*	
	χ^2	p	χ^2	p
Info	0.52	0.472	1.58	0.209
Exp			0.16	0.693

Appendix 3.C Additional experimental material

Figure 3: **T=1**, First phishing email (Translated from Dutch)

Sender: BusinessOperations@Mlnez.nl
Subject: Activate your personal Mobile Password Recovery System

Dear Economic Affairs colleague,

After a successful pilot of the department Business Operations, we recently started implementing the EZ - Mobile Password Recovery System (EZ-MPRS) for all EA employees. Via this system you can, at all times, retrieve and change your password for your user account. With this, we hope to serve you even better and faster.

Our data shows that you are not yet using the EZ - Mobile Password Recovery System. That is why we ask you, to link your account to your mobile number.

Activate here your EZ – Mobile Password Recovery System (MPRS)

For more information, see link below:

https://rijksweb.nl/ezmprs

Best regards,

Director Business Operations

Dit bericht kan informatie bevatten die niet voor u is bestemd. Indien u niet de geadresseerde bent of dit bericht abusievelijk aan u is gezonden, wordt u verzocht dat aan de afzender te melden en het bericht te verwijderen.
De Staat aanvaardt geen aansprakelijkheid voor schade, van welke aard ook, die verband houdt met risico's verbonden aan het elektronisch verzenden van berichten.

This message may contain information that is not intended for you. If you are not the addressee or if this message was sent to you by mistake, you are requested to inform the sender and delete the message.
The State accepts no liability for damage of any kind resulting from the risks inherent in the electronic transmission of messages.

Figure 4: First phishing email linked website (Translated from Dutch)

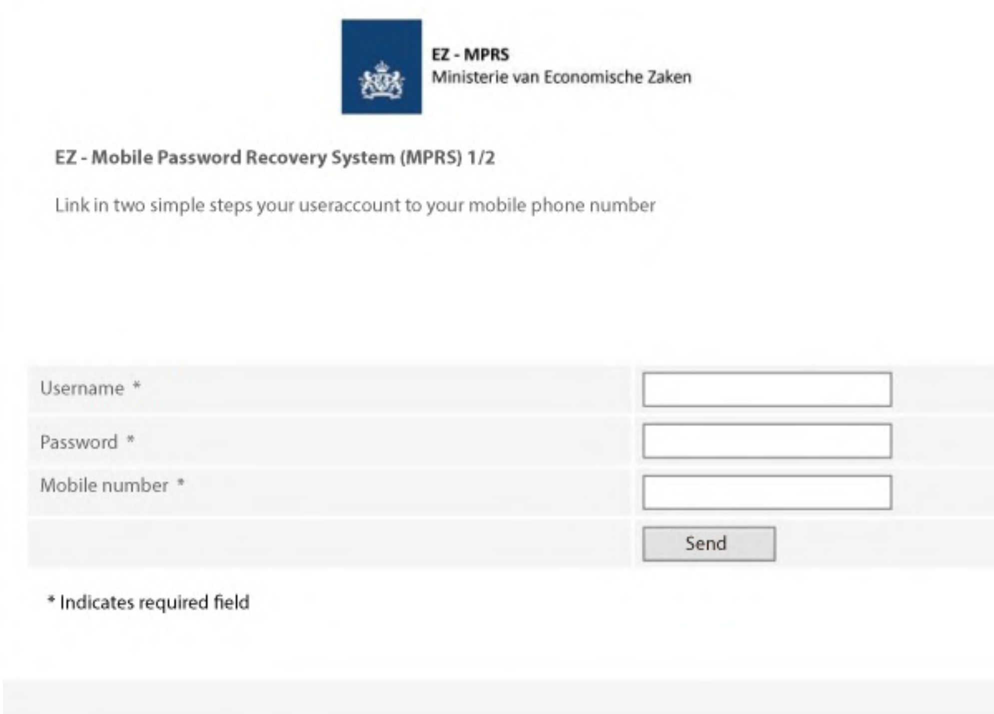


Figure 5: First phishing email Pop-up screen (Translated from Dutch)

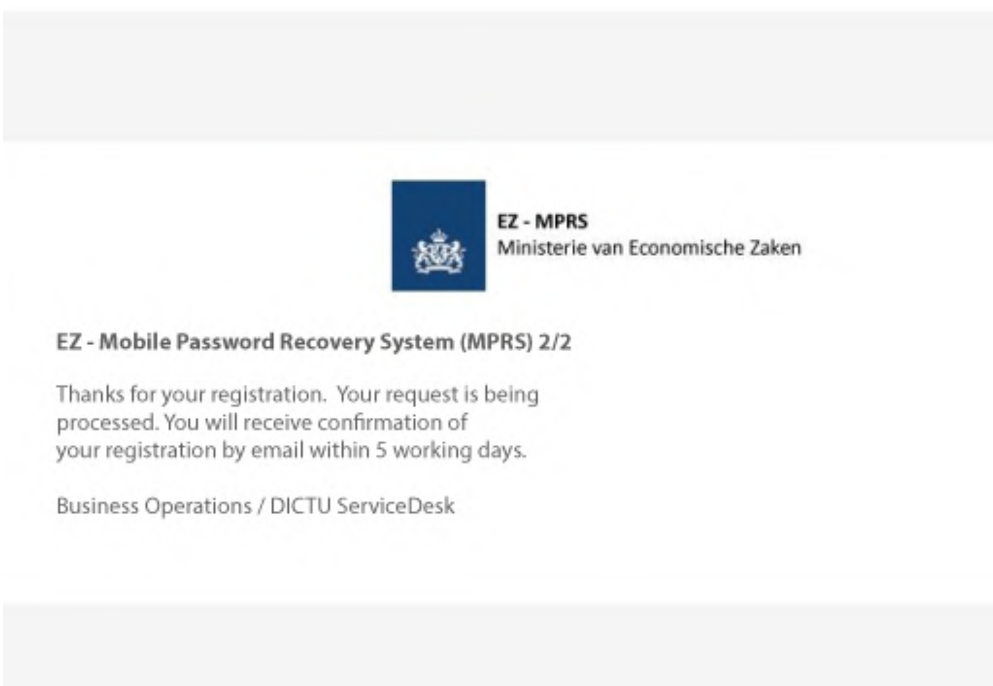


Figure 6: Short debriefing after first phishing mail (Translated from Dutch)

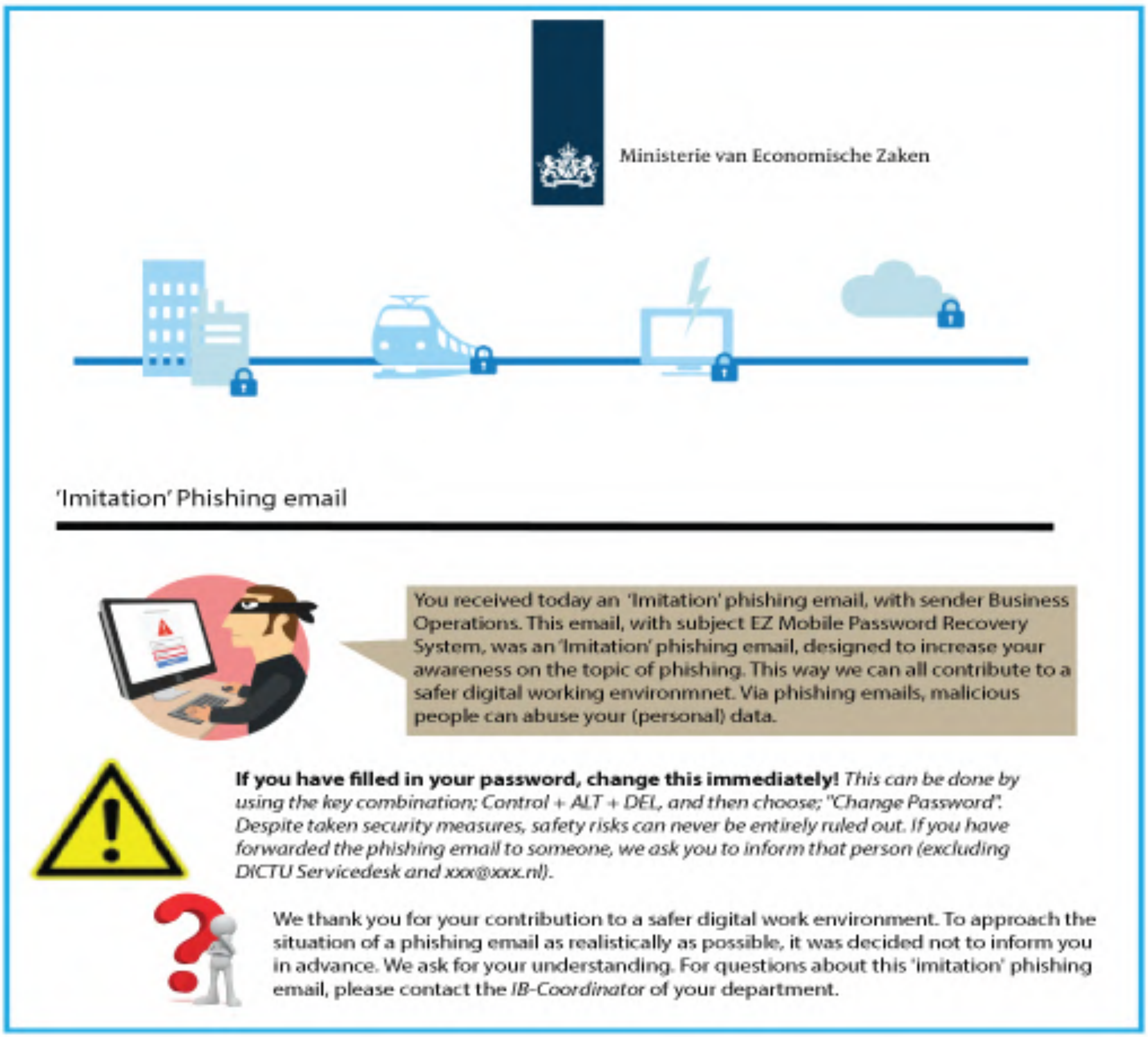
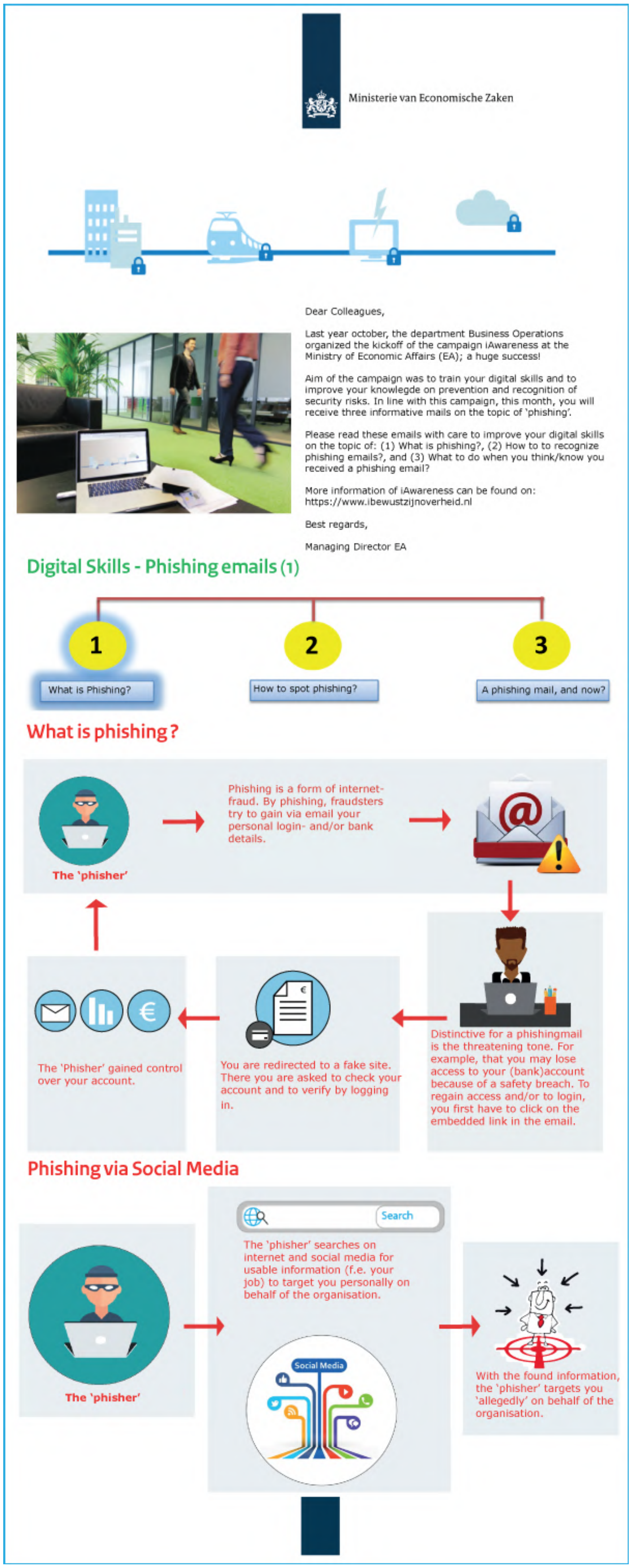
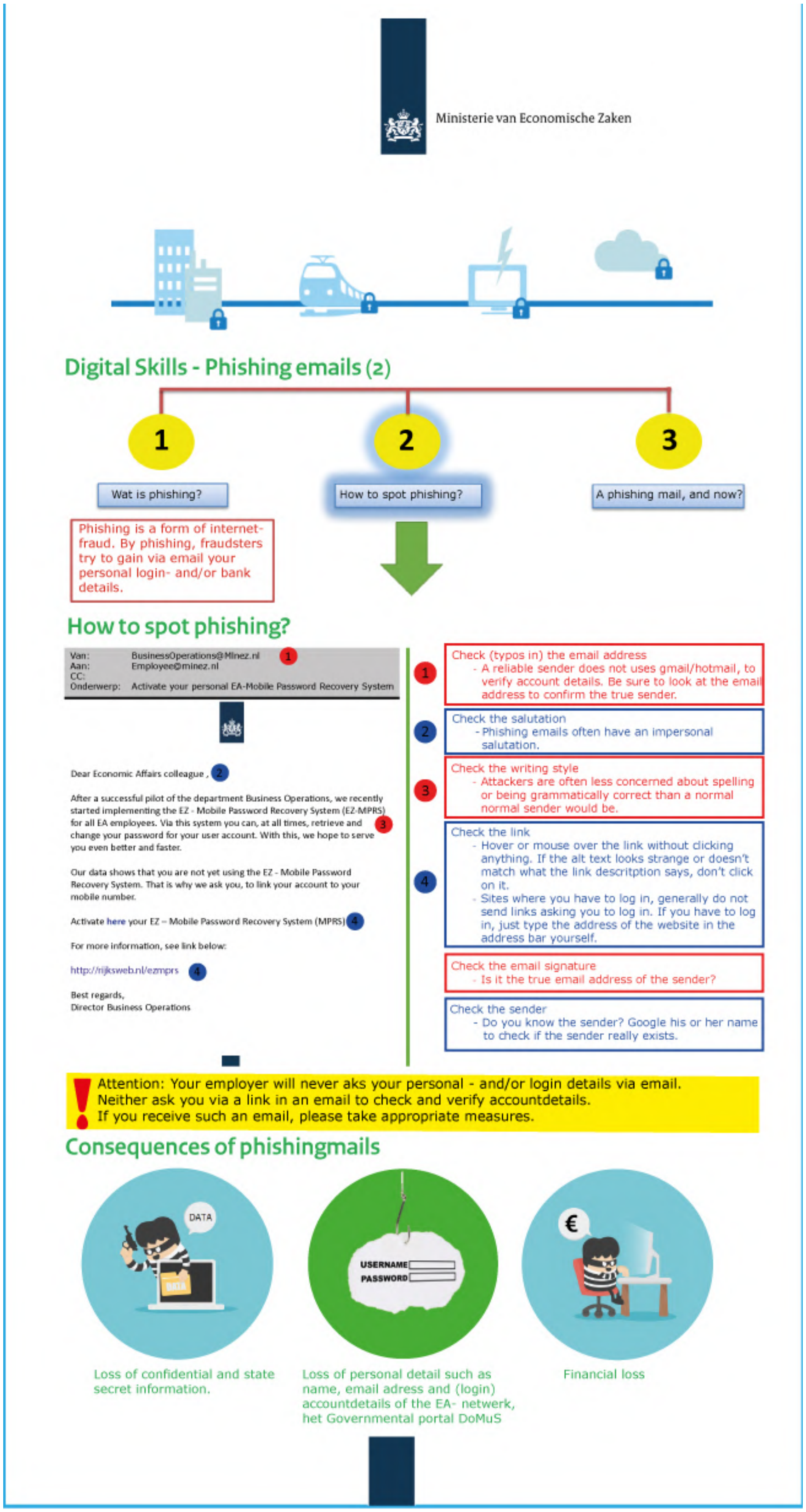


Figure 7: **T = 2, 3, 4**, Infographics
(Translated from Dutch and contact details are replaced for privacy concerns)




Ministerie van Economische Zaken
Digital Skills - Phishing emails (2)
1
2
3
Wat is phishing?
How to spot phishing?
A phishing mail, and now?
Phishing is a form of internet-fraud. By phishing, fraudsters try to gain via email your personal login- and/or bank details.
How to spot phishing?
Van: BusinessOperations@Mlnez.nl
Aan: Employee@minez.nl
CC:
Onderwerp: Activate your personal EA-Mobile Password Recovery System
Dear Economic Affairs colleague ,
After a successful pilot of the department Business Operations, we recently started implementing the EZ - Mobile Password Recovery System (EZ-MPRS) for all EA employees. Via this system you can, at all times, retrieve and change your password for your user account. With this, we hope to serve you even better and faster.
Our data shows that you are not yet using the EZ - Mobile Password Recovery System. That is why we ask you, to link your account to your mobile number.
Activate here your EZ – Mobile Password Recovery System (MPRS)
For more information, see link below:
http://rijksweb.nl/ezmprs
Best regards,
Director Business Operations
Check (typos in) the email address
- A reliable sender does not uses gmail/hotmail, to verify account details. Be sure to look at the email address to confirm the true sender.
Check the salutation
- Phishing emails often have an impersonal salutation.
Check the writing style
- Attackers are often less concerned about spelling or being grammatically correct than a normal normal sender would be.
Check the link
- Hover or mouse over the link without clicking anything. If the alt text looks strange or doesn't match what the link descritption says, don't click on it.
- Sites where you have to log in, generally do not send links asking you to log in. If you have to log in, just type the address of the website in the address bar yourself.
Check the email signature
- Is it the true email address of the sender?
Check the sender
- Do you know the sender? Google his or her name to check if the sender really exists.
Attention: Your employer will never aks your personal - and/or login details via email.
Neither ask you via a link in an email to check and verify accountdetails.
If you receive such an email, please take appropriate measures.
Consequences of phishingmails
DATA
USERNAME
PASSWORD
€
Loss of confidential and state secret information.
Loss of personal detail such as name, email adress and (login) accountdetails of the EA- netwerk, het Governmental portal DoMuS
Financial loss

Ministerie van Economische Zaken

Digital Skills - Phishing emails (3)

1 What is Phishing?

Phishing is a form of internet-fraud. By phishing, fraudsters try to gain via email your personal login- and/or bank details.

2 How to spot phishing?

Characteristics to spot phishing: (1) email address, (2) the salutation, (3) the writing style, (4) the link (5) the sender, and (6) the signature's email address

3 phishing mail, and now?

How to respond to a (potential) phishing email?

- Do NOT click on the embedded link.
- Never share login details or confidential Information.
- Send the email as attachment to **xxxx@xxx.nl**.
- **RVO:** CC to: xxxx2@xxx.nl
- **NVWA:** CC to xxxx3@xxx.nl

Questions regarding suspicious mails? - Contact details

- **DICTU Servicedesk**

Phone number:	xxx xxx	**voor:**	KD & SODM
	xxx xxx	**voor:**	RVO & NVWA
			AT & DICTU
Email address:	**xxxx4@xxx.nl**		

- **The IB - Coordinator(s) of your department**
- **Portal (Informationsecurity)**
  - Intranet of your department

More Information?

- https://www.ibewustzijnoverheid.nl
- https://veiliginternetten.nl

Figure 8: **T=5**, Second phishing email (Translated from Dutch)

Sender: helpdesk@dlctu.nl
Subject: Increase your maximum outlook exchange storage limit.

Dear Economic Affairs colleague,

Your mailbox has exceeded the maximum storage limit set by DICTU. You may not send or receive e-mail until you have upgraded the maximum limit. To increase your limit, click on the link below:

Increase your storage limit here

If you do not do this, you run the risk that your mail account will be locked. Thank you for your cooperation.

Best regards,

For more information, see the link below:

https://rijksweb.nl/exchangelimiet

The DICTU Helpdesk

Dit bericht kan informatie bevatten die niet voor u is bestemd. Indien u niet de geadresseerde bent of dit bericht abusievelijk aan u is gezonden, wordt u verzocht dat aan de afzender te melden en het bericht te verwijderen. De Staat aanvaardt geen aansprakelijkheid voor schade, van welke aard ook, die verband houdt met risicos verbonden aan het elektronisch verzenden van berichten.

This message may contain information that is not intended for you. If you are not the addressee or if this message was sent to you by mistake, you are requested to inform the sender and delete the message. The State accepts no liability for damage of any kind resulting from the risks inherent in the electronic transmission of messages.

Figure 9: Second phishing email linked website (Translated from Dutch)

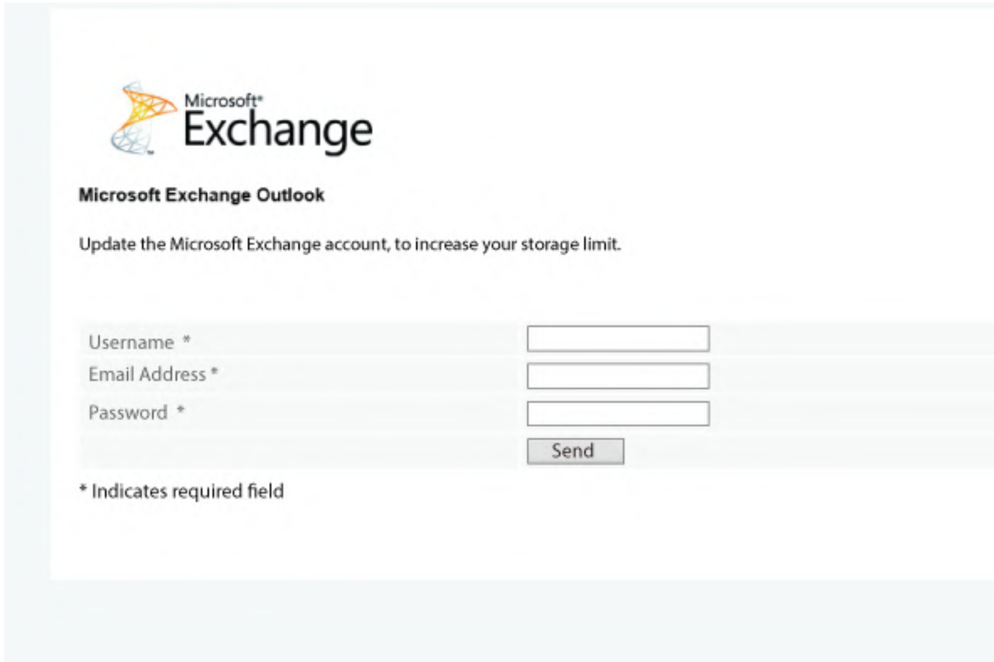
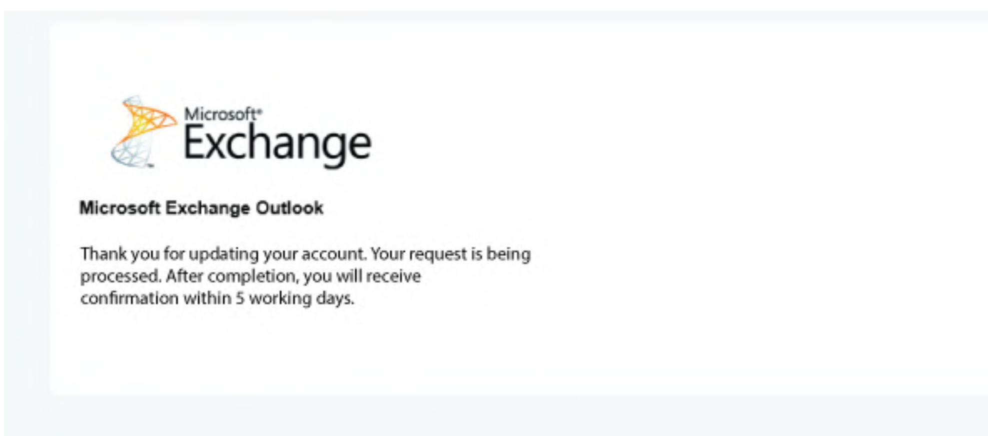


Figure 10: Second phishing email Pop-up screen (Translated from Dutch)



Conclusion

My PhD dissertation focuses on how people misperceive risk and uncertainty, and how this cognitive bias affects individuals' preventive actions.

In Chapter 1, we zoom in on *very* unlikely events in a lab experiment and investigate whether they are overweighted or neglected. In recent decades there have been many studies on ambiguity attitudes, however, the focus was on moderate likelihood events and gains. Rare events have been almost ignored so far because of three main challenges: providing sufficiently high incentives, disentangling ambiguity attitude from risk attitude and controlling for beliefs. In this study we address and control for all three challenges. We measure ambiguity attitudes through additivity indices. With these indices we capture whether a subject assigns the same subjective value to an event when it is in isolation or combined with other events. If that is the case than the subject is ambiguity neutral and the event is neither overweighted nor neglected. Any deviation from additivity gives us non-neutrality in ambiguity. We first use non-parametric tests to examine whether ambiguity neutrality is violated. The analysis reveals that very unlikely events are not ignored but rather are overweighted overall, being weighted more strongly in isolation than when part of larger events. We then use latent profile analysis to study heterogeneity of behavior. With this approach we classify the subjects into several behavioral profiles which are determined freely by the data without our intervention. One third behaves close to ambiguity neutrality consistently. The others exhibit overweighting of rare events. These results are important since overweighting of rare events might lead to overinsurance in potential losses and overinvestment in long shots.

Chapter 2, a theoretical contribution, studies how people's imperfect perception of probabilities can lead to suboptimal prevention efforts. We compare the optimal prevention (self-protection) level of an agent who does non-linear probability weighting to the optimal prevention level of another agent who has the same utility function but maximizes expected utility. In other words, we check how much probability weighting of an agent causes him to deviate from his rational self in his optimal prevention decision. We identify the probability range in which misperceptions regarding risk and the benefits of preventive actions generate underprevention. We show that a consequence of likelihood insensitivity is not only that agents do not perceive risks they face perfectly but also the benefits of prevention. This leads agents to put less effort than their rational selves would. Our results may provide a novel explanation for the widespread presence of objectively harmful activities such as smoking and obesity. Policies that aim at promoting better life styles, therefore, should not only focus on informing people about the risks they face but help them better assess those risks and how those risks change with prevention.

Chapter 3 analyses an intervention aiming at increasing prevention at the organizational level. Specifically, in a large field experiment we test the effect of information provision and simulated experience on reducing the risks of falling into a phishing attack, and thereby increasing cyber-security. Our experiment was conducted at the Dutch Ministry of Economic Affairs, with more than 10,000 subjects. The subjects were not aware that they participated in such an experiment as a result of which we could observe actual behavior in a real life setting. In the experiment, we sent a phishing email to measure the susceptibility of employees to click on a dubious link and then give away their password. Prior to this phishing email, subjects were assigned to a specific treatment group or instead the control group. The information treatment included infographics and very clear messages. In the experience treatment, employees experienced (simulated) phishing fraud and were subsequently informed about it, prior to receiving the actual phishing email. Both approaches substantially reduced the proportion of employees giving away their password in the actual phishing email. Compared to the control group, being informed about the risks of phishing reduced the percentage of subjects clicking the link and the percentage of subjects giving away their passwords by 7 and 6 percentage points, respectively. The reduction in these proportions were 9 and 8 percentage points for subjects who had a first experience with a phishing email and got feedback. Combining both interventions did not substantially increase the effect of simulated experience alone. This shows that doing more is not necessarily better and sometimes just simple and low cost interventions are effective.

Chapter 4 answers the question of whether sexual context has an impact on ambiguity attitudes in a lab experiment. Taking Ellsberg urn experiments with monetary outcome as the baseline, sexual context appears in the source of uncertainty and/or outcome. Sources of uncertainty vary in within-subject treatments. In one of those treatments instead of betting on the winning color of balls picked from an urn, the subjects bet on the winning color of condoms picked from an urn. Therefore, the source remains very similar to the one in baseline treatment but has a sexual connotation. In the other within-subject treatment sexual context appears in a more natural source of uncertainty. This time the subjects bet on the proportion of people having sex weekly. Sexual context in outcome is used in a between-subject treatment where subjects could win condoms instead of money. The results show that while ambiguity aversion is independent of sexual context ambiguity generated insensitivity is. Subjects become strongly ambiguity insensitive when sexual context appears in a natural source, and behave close to neutrality when the outcome is sexual and source is artificial. This study shows that people are affected from sexual connotations in the process of decision making even if decisions are not sex related.

The results in this PhD thesis show that misperceptions of risk and uncertainty in the form of weighting of likelihoods can explain the coexistence of both overinsurance and underprevention behavior. Simulated experience and efficient communication of information are proved to be effective tools for increasing prevention by potentially increasing risk perception in cybersecurity context. However, the results also show that the strength of weighting is context dependent. Further experimental research where subjects deal with situations that resemble real-life insurance and prevention decisions in different contexts would help in designing intervention methods tailored to those contexts.

Nederlandse samenvatting (Summary in Dutch)

Dit proefschrift bestudeert fouten die mensen maken bij het inschatten van risico en onzekerheid, en hoe deze "cognitieve bias" de preventieve acties van individuen beïnvloedt. Hoofdstuk 1 laat in een laboratoriumexperiment zien dat hoe zeldzame gebeurtenissen gepresenteerd worden van invloed is op hoe mensen die gebeurtenissen waarnemen. Ik laat zien dat mensen zeldzame gebeurtenissen groter ervaren dan ze in werkelijkheid zijn wanneer die gebeurtenissen afzonderlijk aan elkaar worden gepresenteerd in plaats van gezamelijk. Hoofdstuk 2 laat theoretisch zien dat dit verklaard kan worden aan de hand van kansweging, hetzelfde fenomeen dat er ook voor zorgt dat mensen tegelijkertijd oververzekerd zijn en weinig preventieve handelingen uitvoeren. Hoofdstuk 3 analyseert een interventie in cybersecurity gericht op het vergroten van preventie op organisatieniveau in een veldexperiment. In het experiment test ik of enerzijds het effectiever communiceren van informatie naar werknemers, of anderzijds een gesimuleerde phishing-aanval, helpt om phishing-aanvallen te voorkomen. Hoofdstuk 4 behandelt de kwestie dat de risicoinschatting van mensen kan verschillen met de context. Een laboratoriumexperiment toont aan dat een seksuele context invloed heeft op de ambiguïteit van attitudes.

BIBLIOGRAPHY

Abdellaoui, M. (2000). Parameter-free elicitation of utility and probability weighting functions. *Management Science 46*(11), 1497–1512.

Abdellaoui, M., A. Baillon, L. Placido, and P. P. Wakker (2011). The rich domain of uncertainty: Source functions and their experimental implementation. *American Economic Review 101*(2), 695–723.

Abdellaoui, M. and E. Kemel (2013). Eliciting prospect theory when consequences are measured in time units:"time is not money". *Management Science 60*(7), 1844–1859.

Abdellaoui, M., O. L'Haridon, and C. Paraschiv (2011). Experienced vs. described uncertainty: Do we need two prospect theory specifications? *Management Science 57*(10), 1879–1895.

Abdellaoui, M., F. Vossmann, and M. Weber (2005). Choice-based elicitation and decomposition of decision weights for gains and losses under uncertainty. *Management Science 51*(9), 1384–1399.

Acemoglu, D. and D. Autor (2011). Skills, tasks and technologies: Implications for employment and earnings. In *Handbook of Labor Economics*, Volume 4, pp. 1043–1171. Elsevier.

Ahn, D., S. Choi, D. Gale, and S. Kariv (2014). Estimating ambiguity aversion in a portfolio choice experiment. *Quantitative Economics 5*(2), 195–223.

Alary, D., C. Gollier, and N. Treich (2013). The effect of ambiguity aversion on insurance and self-protection. *The Economic Journal 123*(573), 1188–1202.

Albrechtsen, E. (2007). A qualitative study of users' view on information security. *Computers & security 26*(4), 276–289.

Alsharnouby, M., F. Alaca, and S. Chiasson (2015). Why phishing still works: User strategies for combating phishing attacks. *International Journal of Human-Computer Studies 82*, 69–82.

Anantanasuwong, K., R. Kouwenberg, O. S. Mitchell, and K. Peijnenberg (2019). Ambiguity attitudes about investments: Evidence from the field. Technical report, National Bureau of Economic Research.

Andersen, S., G. W. Harrison, and E. E. Rutström (2006). Choice behavior, asset integration and natural reference points. Technical report, Working Paper 06.

Arachchilage, N. A. G. and S. Love (2014). Security awareness of computer users: A phishing threat avoidance perspective. *Computers in Human Behavior 38*, 304–312.

Ariely, D. and G. Loewenstein (2006). The heat of the moment: The effect of sexual arousal on sexual decision making. *Journal of Behavioral Decision Making 19*(2), 87–98.

Arrow, K. J. (1951). Alternative approaches to the theory of choice in risk-taking situations. *Econometrica 19*(4), 404–437.

Arrow, K. J. (1963). Uncertainty and the welfare economics of medical care. *The American Economic Review 53*(5), 941–973.

Arrow, K. J. (1971). The theory of risk aversion. *Essays in the theory of risk-bearing*, 90–120.

Arrow, K. J. (1982). Risk perception in psychology and economics. *Economic Inquiry 20*(1), 1–9.

Arrow, K. J. and L. Hurwicz (1972). An optimality criterion for decision-making under ignorance. *Uncertainty and Expectations in Economics*, 1–11.

Bada, M., A. Sasse, and J. R. C. Nurse (2015). Cyber security awareness campaigns: Why do they fail to change behaviour? In *International Conference on Cyber Security for Sustainable Society*, pp. 118–131.

Bade, S. (2015). Randomization devices and the elicitation of ambiguity-averse preferences. *Journal of Economic Theory 159*, 221–235.

Baillon, A. and H. Bleichrodt (2015). Testing ambiguity models through the measurement of probabilities for gains and losses. *American Economic Journal: Microeconomics 7*(2), 77–100.

Baillon, A., H. Bleichrodt, U. Keskin, O. l'Haridon, and C. Li (2017). The effect of learning on ambiguity attitudes. *Management Science 64*(5), 2181–2198.

Baillon, A., Y. Halevy, C. Li, et al. (2015). Experimental elicitation of ambiguity attitude using the random incentive system. *University of British Columbia working paper*.

Baillon, A., Z. Huang, A. Selim, and P. P. Wakker (2018). Measuring ambiguity attitudes for all (natural) events. *Econometrica 86*(5), 1839–1858.

Barron, G. and I. Erev (2003). Small feedback-based decisions and their limited correspondence to description-based decisions. *Journal of Behavioral Decision Making 16*(3), 215–233.

Barseghyan, L., F. Molinari, T. O'Donoghue, and J. C. Teitelbaum (2013). The nature of risk preferences: Evidence from insurance choices. *The American Economic Review 103*(6), 2499–2529.

Berger, L., J. Emmerling, and M. Tavoni (2016). Managing catastrophic climate risks under model uncertainty aversion. *Management Science 63*(3), 749–765.

Berghel, H. (2006). Phishing mongers and posers. *Communications of the ACM 49*(4), 21–25.

Berns, G. S., C. M. Capra, S. Moore, and C. Noussair (2007). A shocking experiment: New evidence on probability weighting and common ratio violations. *Judgment and Decision Making 2*(4), 234–242.

Blanton, H. and M. Gerrard (1997). Effect of sexual motivation on men's risk perception for sexually transmitted disease: There must be 50 ways to justify a lover. *Health Psychology 16*(4), 374.

Bleichrodt, H. and L. Eeckhoudt (2006). Willingness to pay for reductions in health risks when probabilities are distorted. *Health Economics 15*(2), 211–214.

Blythe, M., H. Petrie, and J. A. Clark (2011). F for fake: Four studies on how we fall for phish. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 3469–3478. ACM.

Booij, A. S., B. M. Van Praag, and G. Van De Kuilen (2010). A parametric analysis of prospect theory's functionals for the general population. *Theory and Decision 68*(1-2), 115–148.

Bosch-Domènech, A. and J. Silvestre (2010). Averting risk in the face of large losses: Bernoulli vs. tversky and kahneman. *Economics Letters 107*(2), 180–182.

Bowen, B. M., R. Devarajan, and S. Stolfo (2011). Measuring the human factor of cyber security. In *Technologies for Homeland Security (HST), 2011 IEEE International Conference on*, pp. 230–235. IEEE.

Bulgurcu, B., H. Cavusoglu, and I. Benbasat (2010). Information security policy compliance: An empirical study of rationality-based beliefs and information security awareness. *MIS quarterly 34*(3), 523–548.

Burns, M. B., A. Durcikova, and J. L. Jenkins (2013). What kind of interventions can help users from falling for phishing attempts: A research proposal for examining stage-appropriate interventions. In *System Sciences (HICSS), 2013 46th Hawaii International Conference on*, pp. 4023–4032. IEEE.

Cai, J. and C. Song (2017). Do disaster experience and knowledge affect insurance take-up decisions? *Journal of Development Economics 124*, 83–94.

Camerer, C. (1989). An experimental test of several generalized utility theories. *Journal of Risk and Uncertainty 2*(1), 61–104.

Carter, J. R. and M. D. Irons (1991, June). Are economists different, and if so, why? *Journal of Economic Perspectives 5*(2), 171–177.

Casey, J. T. and J. T. Scholz (1991). Boundary effects of vague risk information on taxpayer decisions. *Organizational Behavior and Human Decision Processes 50*(2), 360–394.

Center for Disease Control and Prevention (2015). 2014 behavioral risk factor surveillance system survey data. Electronic Resource. Accessed: 08/05/2018.

Cerreia-Vioglio, S., F. Maccheroni, M. Marinacci, and L. Montrucchio (2011). Uncertainty averse preferences. *Journal of Economic Theory 146*(4), 1275–1330.

Charupat, N., R. Deaves, T. Derouin, M. Klotzle, and P. Miu (2013). Emotional balance and probability weighting. *Theory and Decision 75*(1), 17–41.

Chateauneuf, A., J. Eichberger, and S. Grant (2007). Choice under uncertainty with the best and worst in mind: Neo-additive capacities. *Journal of Economic Theory 137*(1), 538–567.

Chetty, R., D. Grusky, M. Hell, N. Hendren, R. Manduca, and J. Narang (2017). The fading American dream: Trends in absolute income mobility since 1940. *Science 356*(6336), 398–406.

Chetty, R., N. Hendren, P. Kline, E. Saez, and N. Turner (2014). Is the United States still a land of opportunity? Recent trends in intergenerational mobility. *American Economic Review 104*(5), 141–47.

Chew, S. H., R. P. Ebstein, and S. Zhong (2012). Ambiguity aversion and familiarity bias: Evidence from behavioral and gene association studies. *Journal of Risk and Uncertainty 44*(1), 1–18.

Chipman, J. S. (1960). Stochastic choice and subjective probability. In D. Willner (Ed.), *Decisions, Values and Groups vol.1*. New York: Pergamon Press.

Cohen, J. and P. Dupas (2010). Free distribution or cost-sharing? evidence from a randomized malaria prevention experiment. *The Quarterly Journal of Economics 125*(1), 1–45.

Cohen, M. (1992). Security level, potential level, expected utility: A three-criteria decision model under risk. *Theory and Decision 33*(2), 101–134.

Cohen, M., J.-Y. Jaffray, and T. Said (1987). Experimental comparison of individual behavior under risk and under uncertainty for gains and for losses. *Organizational behavior and human decision processes 39*(1), 1–22.

Courbage, C., H. Loubergé, and R. Peter (2017). Optimal prevention for multiple risks. *Journal of Risk and Insurance 84*(3), 899–922.

Curley, S. P. and J. F. Yates (1985). The center and range of the probability interval as factors affecting ambiguity preferences. *Organizational behavior and human decision processes 36*(2), 273–287.

Curley, S. P. and J. F. Yates (1989). An empirical evaluation of descriptive models of ambiguity reactions in choice situations. *Journal of Mathematical psychology 33*(4), 397–427.

Cutler, D. M. and R. J. Zeckhauser (2000). The anatomy of health insurance. In A. J. Culyer and J. P. Newhouse (Eds.), *Handbook of Health Economics*, Volume 1, pp. 563–643. Amsterdam: Elsevier.

Dimmock, S. G., R. Kouwenberg, O. S. Mitchell, and K. Peijnenburg (2016). Ambiguity aversion and household portfolio choice puzzles: Empirical evidence. *Journal of Financial Economics 119*(3), 559–577.

Dimmock, S. G., R. Kouwenberg, and P. P. Wakker (2015). Ambiguity attitudes in a large representative sample. *Management Science 62*(5), 1363–1380.

Dionne, G. (1982). Moral hazard and state-dependent utility function. *Journal of Risk and Insurance 49*(3), 405–422.

Ditto, P. H., D. A. Pizarro, E. B. Epstein, J. A. Jacobson, and T. K. MacDonald (2006). Visceral influences on risk-taking behavior. *Journal of Behavioral Decision Making 19*(2), 99–113.

Dodge, M. M. (1867). *Hans Brinker, or, the silver skates: a story of life in Holland*. James O'Kane.

Donkers, B., B. Melenberg, and A. Van Soest (2001). Estimating risk attitudes using lotteries: A large sample approach. *Journal of Risk and Uncertainty 22*(2), 165–195.

Downs, J. S., M. B. Holbrook, and L. F. Cranor (2006). Decision strategies and susceptibility to phishing. In *Proceedings of the second symposium on Usable privacy and security*, pp. 79–90. ACM.

Eeckhoudt, L. and J. Hammitt (2001). Background risks and the value of a statistical life. *Journal of Risk and Uncertainty 23*(3), 261–279.

Ehrlich, I. and G. S. Becker (1972). Market insurance, self-insurance, and self-protection. *Journal of Political Economy 80*(4), 623–648.

Einhorn, H. J. and R. M. Hogarth (1986). Decision making under ambiguity. *Journal of Business 59*, 225–250.

Eisenberger, R. and M. Weber (1995). Willingness-to-pay and willingness-to-accept for risky and ambiguous lotteries. *Journal of Risk and Uncertainty 10*(3), 223–233.

Ellsberg, D. (1961). Risk, ambiguity, and the Savage axioms. *The Quarterly Journal of Economics 75*(4), 643–669.

Ellsberg, D. (2011). Introduction to the symposium issue: notes on the origins of the ellsberg urns. *Economic Theory 48*(2), 221–227.

Epstein, L. G. (1999). A definition of uncertainty aversion. *Review of Economic Studies*, 579–608.

Epstein, L. G. and T. Wang (1994). Intertemporal asset pricing under Knightian uncertainty. *Econometrica 62*(2), 283–322.

Ert, E. and S. T. Trautmann (2014, Aug). Sampling experience reverses preferences for ambiguity. *Journal of Risk and Uncertainty 49*(1), 31–42.

Etchart-Vincent, N. (2004). Is probability weighting sensitive to the magnitude of consequences? An experimental investigation on losses. *Journal of Risk and Uncertainty 28*(3), 217–235.

Etchart-Vincent, N. (2009). Probability weighting and the 'level' and 'spacing' of outcomes: An experimental study over losses. *Journal of Risk and Uncertainty 39*(1), 45–63.

Etchart-Vincent, N. and O. l'Haridon (2011). Monetary incentives in the loss domain and behavior toward risk: An experimental comparison of three reward schemes including real losses. *Journal of risk and uncertainty 42*(1), 61–83.

Etner, J. and M. Jeleva (2014). Underestimation of probabilities modifications: Characterization and economic implications. *Economic Theory 56*(2), 291–307.

Evans, W. and K. Viscusi (1991). Estimation of state-dependent utility functions using survey data. *The Review of Economics and Statistics 73*(1), 94–104.

Fagerlin, A., C. Wang, and P. A. Ubel (2005). Reducing the influence of anecdotal reasoning on people's health care decisions: Is a picture worth a thousand statistics? *Medical decision making 25*(4), 398–405.

Fehr-Duda, H., A. Bruhin, T. Epper, and R. Schubert (2010). Rationality on the rise: Why relative risk aversion increases with stake size. *Journal of Risk and Uncertainty 40*(2), 147–180.

Fehr-Duda, H., M. Gennaro, and R. Schubert (2006). Gender, financial risk, and probability weights. *Theory and Decision 60*(2-3), 283–313.

Finkelstein, A., E. Luttmer, and M. Notowidigdo (2013). What good is wealth without health? The effect of health on the marginal utility of consumption. *Journal of the European Economic Association 11*(suppl_1), 221–258.

Finkelstein, A., S. Taubman, B. Wright, M. Bernstein, J. Gruber, J. P. Newhouse, H. Allen, K. Baicker, and Oregon Health Study Group (2012). The Oregon health insurance experiment: Evidence from the first year. *The Quarterly Journal of Economics 127*(3), 1057–1106.

Gandini, S., E. Botteri, S. Iodice, M. Boniol, A. B. Lowenfels, P. Maisonneuve, and P. Boyle (2008). Tobacco smoking and cancer: A meta-analysis. *International Journal of Cancer 122*(1), 155–164.

Ghirardato, P., F. Maccheroni, and M. Marinacci (2004). Differentiating ambiguity and ambiguity attitude. *Journal of Economic Theory 118*(2), 133–173.

Gigerenzer, G. and U. Hoffrage (1995). How to improve Bayesian reasoning without instruction: frequency formats. *Psychological Review 102*(4), 684–704.

Gilboa, I. and D. Schmeidler (1989). Maxmin expected utility with non-unique prior. *Journal of mathematical economics 18*(2), 141–153.

Global Challenges Foundation (2017). Attitudes to global risks and governance. Available at: `https://api.globalchallenges.org/static/files/ComRes.pdf`.

Goldstein, W. M. and H. Einhorn (1987). Expression theory and the preference reversal phenomena. *Psychological Review 94*(2), 236–254.

Gonzalez, R. and G. Wu (1999). On the shape of the probability weighting function. *Cognitive psychology 38*(1), 129–166.

Gupta, B. B., N. A. Arachchilage, and K. E. Psannis (2018). Defending against phishing attacks: taxonomy of methods, current issues and future directions. *Telecommunication Systems 67*(2), 247–267.

Halevi, T., N. Memon, and O. Nov (2015). Spear-phishing in the wild: A real-world study of personality, phishing self-efficacy and vulnerability to spear-phishing attacks. `http://dx.doi.org/10.2139/ssrn.2544742`.

Harbaugh, W. T., K. Krause, and L. Vesterlund (2010). The fourfold pattern of risk attitudes in choice and pricing tasks. *The Economic Journal 120*(545), 595–611.

Harrison, G. W. and E. E. Rutström (2009). Expected utility theory and prospect theory: One wedding and a decent funeral. *Experimental Economics 12*(2), 133–158.

Heath, C. and A. Tversky (1991). Preference and belief: Ambiguity and competence in choice under uncertainty. *Journal of risk and uncertainty 4*(1), 5–28.

Hertwig, R., G. Barron, E. U. Weber, and I. Erev (2004). Decisions from experience and the effect of rare events in risky choice. *Psychological Science 15*(8), 534–539.

Hertwig, R. and I. Erev (2009). The description–experience gap in risky choice. *Trends in cognitive sciences 13*(12), 517–523.

Hofmann, A. and R. Peter (2015). Multivariate prevention decisions: Safe today or sorry tomorrow? *Economics Letters 128*, 51–53.

Hogarth, R. M. and H. J. Einhorn (1990). Venture theory: A model of decision weights. *Management Science 36*(7), 780–803.

Howlader, N., A. Noone, M. Krapcho, D. Miller, K. Bishop, C. Kosary, M. Yu, J. Ruhl, Z. Tatalovich, A. Mariotto, D. Lewis, H. Chen, E. Feuer, and K. e. Cronin (2017). SEER Cancer Statistics Review, 1975-2014. Technical report, National Cancer Institute. Bethesda.

Intergovernmental Panel on Climate Change (2013). *Climate Change 2013*. Cambridge University Press.

Jaspersen, J. G. and A. Richter (2015). The wealth effects of premium subsidies on moral hazard in insurance markets. *European Economic Review 77*, 139–153.

Johnson, C., A. Baillon, H. Bleichrodt, Z. Li, D. Van Dolder, and P. P. Wakker (2014). Prince: An improved method for measuring incentivized preferences. *mimeo Erasmus University Rotetrdam*.

Jones-Lee, M. (1974). The value of changes in the probability of death or injury. *Journal of Political Economy 82*(4), 835–849.

Kahn, B. E. and R. K. Sarin (1988). Modeling ambiguity in decisions under uncertainty. *Journal of Consumer Research*, 265–272.

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

Keeney, R. L. (2008). Personal decisions are the leading cause of death. *Operations Research 56*(6), 1335–1347.

Kemel, E. and C. Paraschiv (2013). Prospect theory for joint time and money consequences in risk and ambiguity. *Transportation Research Part B: Methodological 56*, 81–95.

Kemel, E. and C. Paraschiv (2018). Deciding about human lives: an experimental measure of risk attitudes under prospect theory. *Social Choice and Welfare 51*(1), 163–192.

Kermer, D. A., E. Driver-Linn, T. D. Wilson, and D. T. Gilbert (2006). Loss aversion is an affective forecasting error. *Psychological science 17*(8), 649–653.

Kilka, M. and M. Weber (2001). What determines the shape of the probability weighting function under uncertainty? *Management science 47*(12), 1712–1726.

Kleitman, S., M. K. Law, and J. Kay (2018). It's the deceiver and the receiver: Individual differences in phishing susceptibility and false positives with item profiling. *PloS one 13*(10), e0205089.

Klibanoff, P., M. Marinacci, and S. Mukerji (2005). A smooth model of decision making under ambiguity. *Econometrica 73*(6), 1849–1892.

Kliger, D. and O. Levy (2008). Mood impacts on probability weighting functions:"large-gamble" evidence. *The Journal of Socio-Economics 37*(4), 1397–1411.

Kocher, M. G., A. M. Lahno, and S. T. Trautmann (2018). Ambiguity aversion is not universal. *European Economic Review 101*, 268–283.

Kocher, M. G., J. Pahlke, and S. T. Trautmann (2013). Tempus fugit: time pressure in risky decisions. *Management Science 59*(10), 2380–2391.

Kőszegi, B. and M. Rabin (2007). Reference-dependent risk attitudes. *The American Economic Review 97*(4), 1047–1073.

Kumaraguru, P., Y. Rhee, A. Acquisti, L. F. Cranor, J. Hong, and E. Nunge (2007). Protecting people from phishing: The design and evaluation of an embedded training email system. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pp. 905–914. ACM.

Lattimore, P., J. Baker, and A. Witte (1992). The influence of probability on risky choice: A parametric examination. *Journal of Economic Behavior & Organization 17*(3), 377–400.

Lhachimi, S. K., W. J. Nusselder, H. A. Smit, P. Van Baal, P. Baili, K. Bennett, E. Fernández, M. C. Kulik, T. Lobstein, J. Pomerleau, et al. (2012). DYNAMO-HIA – A dynamic modeling tool for generic health impact assessments. *PLoS One 7*(5), e33317.

l'Haridon, O. and F. M. Vieider (2018). All over the map: A worldwide comparison of risk preferences. *Quantitative Economics (forthcoming)*.

Li, C. (2017). Are the poor worse at dealing with ambiguity? *Journal of Risk and Uncertainty 54*(3), 239–268.

Li, Z., J. Müller, P. P. Wakker, and T. V. Wang (2017). The rich domain of ambiguity explored. *Management Science 64*(7), 3227–3240.

Loewenstein, G. (2000). Emotions in economic theory and economic behavior. *American economic review 90*(2), 426–432.

Loewenstein, G., D. Nagin, and R. Paternoster (1997). The effect of sexual arousal on expectations of sexual forcefulness. *Journal of Research in Crime and Delinquency 34*(4), 443–473.

Loomes, G., P. Moffatt, and R. Sugden (2002). A microeconometric test of alternative stochastic theories of risky choice. *Journal of Risk and Uncertainty 24*(2), 103–130.

Maccheroni, F., M. Marinacci, and A. Rustichini (2006). Ambiguity aversion, robustness, and the variational representation of preferences. *Econometrica 74*(6), 1447–1498.

Marwell, G. and R. E. Ames (1981). Economists free ride, does anyone else?: Experiments on the provision of public goods, iv. *Journal of public economics 15*(3), 295–310.

Mason, C. F., J. F. Shogren, C. Settle, and J. A. List (2005). Investigating risky choices over losses using experimental data. *Journal of Risk and Uncertainty 31*(2), 187–215.

Menegatti, M. (2009). Optimal prevention and prudence in a two-period model. *Mathematical Social Sciences 58*(3), 393–397.

Mohebzada, J., A. El Zarka, A. H. BHojani, and A. Darwish (2012). Phishing in a university community: Two large scale phishing experiments. In *Innovations in Information Technology (IIT), 2012 International Conference on*, pp. 249–254. IEEE.

Mozaffarian, D., E. J. Benjamin, A. S. Go, D. K. Arnett, M. J. Blaha, M. Cushman, S. De Ferranti, J.-P. Després, H. J. Fullerton, V. J. Howard, et al. (2015). Heart disease and stroke statistics - 2015 update: A report from the American Heart Association. *Circulation 131*(4), e29–e322.

Mullahy, J. (1999, 2). It'll only hurt a second? microeconomic determinants of who gets flu shots. *Health Economics 8*(1), 9–24.

Musuva, P. M., K. W. Getao, and C. K. Chepken (2019). A new approach to modelling the effects of cognitive processing and threat detection on phishing susceptibility. *Computers in Human Behavior 94*, 154–175.

Oechssler, J. and A. Roomets (2014). Unintended hedging in ambiguity experiments. *Economics Letters 122*(2), 243–246.

Ogden, C. L., M. D. Carroll, C. D. Fryar, and K. M. Flegal (2015). *Prevalence of obesity among adults and youth: United States, 2011-2014*. US Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Health Statistics.

Pahnila, S., M. Siponen, and A. Mahmood (2007). Employees' behavior towards is security policy compliance. In *System sciences, 2007. HICSS 2007. 40Th annual hawaii international conference on*, pp. 156b–156b. IEEE.

Peter, R. (2017). Optimal self-protection in two periods: On the role of endogenous saving. *Journal of Economic Behavior & Organization 137*, 19–36.

Petrova, D. G., J. van der Pligt, and R. Garcia-Retamero (2014). Feeling the numbers: On the interplay between risk, affect, and numeracy. *Journal of Behavioral Decision Making 27*(3), 191–199.

Prelec, D. (1998). The probability weighting function. *Econometrica 66*(3), 497–527.

Rabin, M. and R. H. Thaler (2001). Anomalies: risk aversion. *The Journal of Economic Perspectives 15*(1), 219–232.

Ramanathan, V. and H. Wechsler (2013). Phishing detection and impersonated entity discovery using conditional random field and latent Dirichlet allocation. *Computers & Security 34*, 123–139.

Ransbotham, S. and S. Mitra (2009). Choice and chance: A conceptual model of paths to information security compromise. *Information Systems Research 20*(1), 121–139.

Rheinberger, C., D. Herrera-Araujo, and J. Hammitt (2016). The value of disease prevention vs treatment. *Journal of Health Economics 50*, 247–255.

Rottenstreich, Y. and C. K. Hsee (2001). Money, kisses, and electric shocks: On the affective psychology of risk. *Psychological science 12*(3), 185–190.

Sarin, R. K. and M. Weber (1993). Effects of ambiguity in market experiments. *Management science 39*(5), 602–615.

Sasco, A., M. Secretan, and K. Straif (2004). Tobacco smoking and cancer: A brief review of recent epidemiological evidence. *Lung Cancer 45*(2), S3–S9.

Savage, L. J. (1954). *The foundations of statistics.* (1972 ed.). John Wiley & Sons.

Schade, C., H. Kunreuther, and P. Koellinger (2012). Protecting against low-probability disasters: The role of worry. *Journal of Behavioral Decision Making 25*(5), 534–543.

Schmeidler, D. (1989). Subjective probability and expected utility without additivity. *Econometrica: Journal of the Econometric Society*, 571–587.

Seog, S. H. (2012). Moral hazard and health insurance when treatment is preventive. *Journal of Risk and Insurance 79*(4), 1017–1038.

Sheng, S., M. Holbrook, P. Kumaraguru, L. F. Cranor, and J. Downs (2010). Who falls for phish?: A demographic analysis of phishing susceptibility and effectiveness of interventions. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 373–382. ACM.

Sheng, S., B. Magnien, P. Kumaraguru, A. Acquisti, L. F. Cranor, J. Hong, and E. Nunge (2007). Anti-phishing phil: The design and evaluation of a game that teaches people not to fall for phish. In *Proceedings of the 3rd symposium on Usable privacy and security*, pp. 88–99. ACM.

Siadati, H., S. Palka, A. Siegel, and D. McCoy (2017). Measuring the effectiveness of embedded phishing exercises. In *10th USENIX Workshop on Cyber Security Experimentation and Test (CSET 17)*. USENIX Association.

Silic, M. and A. Back (2016). The dark side of social networking sites: Understanding phishing risks. *Computers in Human Behavior 60*, 35–43.

Siniscalchi, M. (2009). Vector expected utility and attitudes toward variation. *Econometrica 77*(3), 801–855.

Sloan, F., K. Viscusi, H. Chesson, C. Conover, and K. Whetten-Goldstein (1998). Alternative approaches to valuing intangible health losses: The evidence for multiple sclerosis. *Journal of Health Economics 17*(4), 475–497.

Smerecnik, C. M. R., I. Mesters, L. T. E. Kessels, R. A. C. Ruiter, N. K. De Vries, and H. De Vries (2010). Understanding the positive effects of graphical risk information on comprehension: Measuring attention directed to written, tabular, and graphical risk information. *Risk Analysis 30*(9), 1387–1398.

Snow, A. (2011). Ambiguity aversion and the propensities for self-insurance and self-protection. *Journal of Risk and Uncertainty 42*(1), 27–43.

Surgeon General (2014). The health consequences of smoking – 50 years of progress. Technical report, U.S. Department of Health and Human Services.

Taleb, N. N. (2007). *The black swan: The impact of the highly improbable*. Random House.

Tanaka, T., C. F. Camerer, and Q. Nguyen (2010). Risk and time preferences: Linking experimental and household survey data from Vietnam. *American Economic Review 100*(1), 557–71.

Teitelbaum, J. (2007). A unilateral accident model under ambiguity. *The Journal of Legal Studies 36*(2), 431–477.

Trautmann, S. T. and G. Van De Kuilen (2015). Ambiguity attitudes. *The Wiley Blackwell handbook of judgment and decision making 1*, 89–116.

Tu, Q. et al. (2005). Empirical analysis of time preferences and risk aversion. Technical report, Tilburg University, School of Economics and Management.

Tversky, A. and C. Fox (1995). Weighing risk and uncertainty. *Psychological Review 102*(2), 269–283.

Tversky, A. and D. Kahneman (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and uncertainty 5*(4), 297–323.

Tversky, A. and P. Wakker (1995). Risk attitudes and decision weights. *Econometrica 63*(6), 1255–1280.

Van Baal, P. H., R. T. Hoogenveen, A. G. de Wit, and H. C. Boshuizen (2006). Estimating health-adjusted life expectancy conditional on risk factors: results for smoking and obesity. *Population Health Metrics 4*(1), 14.

Viscusi, K. and W. Evans (1990). Utility functions that depend on health status: Estimates and economic implications. *The American Economic Review 80*(3), 353–374.

Vishwanath, A., B. Harrison, and Y. J. Ng (2018). Suspicion, cognition, and automaticity model of phishing susceptibility. *Communication Research 45*(8), 1146–1166.

Vishwanath, A., T. Herath, R. Chen, J. Wang, and H. R. Rao (2011). Why do people get phished? Testing individual differences in phishing vulnerability within an integrated, information processing model. *Decision Support Systems 51*(3), 576–586.

Vroom, C. and R. von Solms (2004). Towards information security behavioural compliance. *Computers & Security 23*(3), 191–198.

Wakker, P. and D. Deneffe (1996). Eliciting von neumann-morgenstern utilities when probabilities are distorted or unknown. *Management science 42*(8), 1131–1150.

Wakker, P., R. Thaler, and A. Tversky (1997). Probabilistic insurance. *Journal of Risk and Uncertainty 15*(1), 7–28.

Wakker, P. P. (2010). *Prospect theory: For risk and ambiguity*. Cambridge University Press.

Wald, A. (1950). *Statistical decision functions.* Wiley.

Whitman, M. and H. Mattord (2003). Principles of information security: Thompson course technology. *Kennesaw State University*.

Wibbenmeyer, M. J., M. S. Hand, D. E. Calkin, T. J. Venn, and M. P. Thompson (2013). Risk preferences in strategic wildfire decision making: A choice experiment with US wildfire managers. *Risk Analysis 33*(6), 1021–1037.

Williams, E. J., J. Hinds, and A. N. Joinson (2018). Exploring susceptibility to phishing in the workplace. *International Journal of Human-Computer Studies 120*, 1–13.

World Health Organization (2018). Global health observatory data repository. Electronic Resource. Accessed: 08/05/2018.

Wright, E. (1993). Non-compliance - or how many aunts has Matilda? *The Lancet 342*(8876), 909–913.

Wright, R. T. and K. Marett (2010). The influence of experiential and dispositional factors in phishing: An empirical investigation of the deceived. *Journal of Management Information Systems 27*(1), 273–303.

Wu, G. and R. Gonzalez (1996). Curvature of the probability weighting function. *Management Science 42*(12), 1676–1690.

Wylie, K. (2009). A global survey of sexual behaviours. *Journal of Family and Reproductive Health*, 39–49.

Xiong, A., R. W. Proctor, W. Yang, and N. Li (2019). Embedding training within warnings improves skills of identifying phishing webpages. *Human factors 61*(4), 577–595.

Zigmont, J. J., L. J. Kappus, and S. N. Sudikoff (2011). Theoretical foundations of learning through simulation. In *Seminars in perinatology*, Volume 35, pp. 47–51. Elsevier.

Tinbergen Institute Research Series

The Tinbergen Institute is the Institute for Economic Research, which was founded in 1987 by the Faculties of Economics and Econometrics of the Erasmus University Rotterdam, University of Amsterdam and VU University Amsterdam. The Institute is named after the late Professor Jan Tinbergen, Dutch Nobel Prize laureate in economics in 1969. The Tinbergen Institute is located in Amsterdam and Rotterdam. The following books recently appeared in the Tinbergen Institute Research Series:

709 M.O. HOYER, *Social Preferences and Emotions in Repeated Interactions*
710 N. GHEBRIHIWET, *Multinational Firms, Technology Transfer, and FDI Policy*
711 H.FANG, *Multivariate Density Forecast Evaluation and Nonparametric Granger Causality Testing*
712 Y. KANTOR, *Urban Form and the Labor Market*
713 R.M. TEULINGS, *Untangling Gravity*
714 K.J.VAN WILGENBURG, *Beliefs, Preferences and Health Insurance Behavior*
715 L. SWART, *Less Now or More Later? Essays on the Measurement of Time Preferences in Economic Experiments*
716 D. NIBBERING, *The Gains from Dimensionality*
717 V. HOORNWEG, *A Tradeoff in Econometrics*
718 S. KUCINSKAS, *Essays in Financial Economics*
719 O. FURTUNA, *Fiscal Austerity and Risk Sharing in Advanced Economies*
720 E. JAKUCIONYTE, *The Macroeconomic Consequences of Carry Trade Gone Wrong and Borrower Protection*
721 M. LI, *Essays on Time Series Models with Unobserved Components and Their Applications*
722 N. CIURILĂ, *Risk Sharing Properties and Labor Supply Disincentives of Pay-As-You-Go Pension Systems*
723 N.M. BOSCH, *Empirical Studies on Tax Incentives and Labour Market Behaviour*
724 S.D. JAGAU, *Listen to the Sirens: Understanding Psychological Mechanisms with Theory and Experimental Tests*
725 S. ALBRECHT, *Empirical Studies in Labour and Migration Economics*
726 Y.ZHU, *On the Effects of CEO Compensation*
727 S. XIA, *Essays on Markets for CEOs and Financial Analysts*
728 I. SAKALAUSKAITE, *Essays on Malpractice in Finance*
729 M.M. GARDBERG, *Financial Integration and Global Imbalances.*
730 U. THŰMMEL, *Of Machines and Men: Optimal Redistributive Policies under Technological Change*
731 B.J.L. KEIJSERS, *Essays in Applied Time Series Analysis*
732 G. CIMINELLI, *Essays on Macroeconomic Policies after the Crisis*
733 Z.M. LI, *Econometric Analysis of High-frequency Market Microstructure*

734 C.M. OOSTERVEEN, *Education Design Matters*
735 S.C. BARENDSE, *In and Outside the Tails: Making and Evaluating Forecasts*
736 S. SÓVÁGÓ, *Where to Go Next? Essays on the Economics of School Choice*
737 M. HENNEQUIN, *Expectations and Bubbles in Asset Market Experiments*
738 M.W. ADLER, *The Economics of Roads: Congestion, Public Transit and Accident Management*
739 R.J. DÖTTLING, *Essays in Financial Economics*
740 E.S. ZWIERS, *About Family and Fate: Childhood Circumstances and Human Capital Formation*
741 Y.M. KUTLUAY, *The Value of (Avoiding) Malaria*
742 A. BOROWSKA, *Methods for Accurate and Efficient Bayesian Analysis of Time Series*
743 B. HU, *The Amazon Business Model, the Platform Economy and Executive Compensation: Three Essays in Search Theory*
744 R.C. SPERNA WEILAND, *Essays on Macro-Financial Risks*
745 P.M. GOLEC, *Essays in Financial Economics*
746 M.N. SOUVERIJN, *Incentives at work*
747 M.H. COVENEY, *Modern Imperatives: Essays on Education and Health Policy*
748 P. VAN BRUGGEN, *On Measuring Preferences*
749 M.H.C. NIENTKER, *On the Stability of Stochastic Dynamic Systems and their use in Econometrics*
750 S. GARCIA MANDICÓ, *Social Insurance, Labor Supply and Intra-Household Spillovers*
751 Y. SUN, *Consumer Search and Quality*
752 I. KERKEMEZOS, *On the Dynamics of (Anti) Competitive Behaviour in the Airline Industry*
753 G.W. GOY, *Modern Challenges to Monetary Policy*
754 A.C. VAN VLODROP, *Essays on Modeling Time-Varying Parameters*
755 J. SUN, *Tell Me How To Vote, Understanding the Role of Media in Modern Elections*
756 J.H. THIEL, *Competition, Dynamic Pricing and Advice in Frictional Markets: Theory and Evidence from the Dutch Market for Mortgages*
757 A. NEGRIU, *On the Economics of Institutions and Technology: a Computational Approach*
758 F. GRESNIGT, *Identifying and Predicting Financial Earth Quakes using Hawkes Processes*