Improving flood preparedness using insights from economic experiments

Mol, Jantsje M.

2021

document version
Publisher's PDF, also known as Version of record

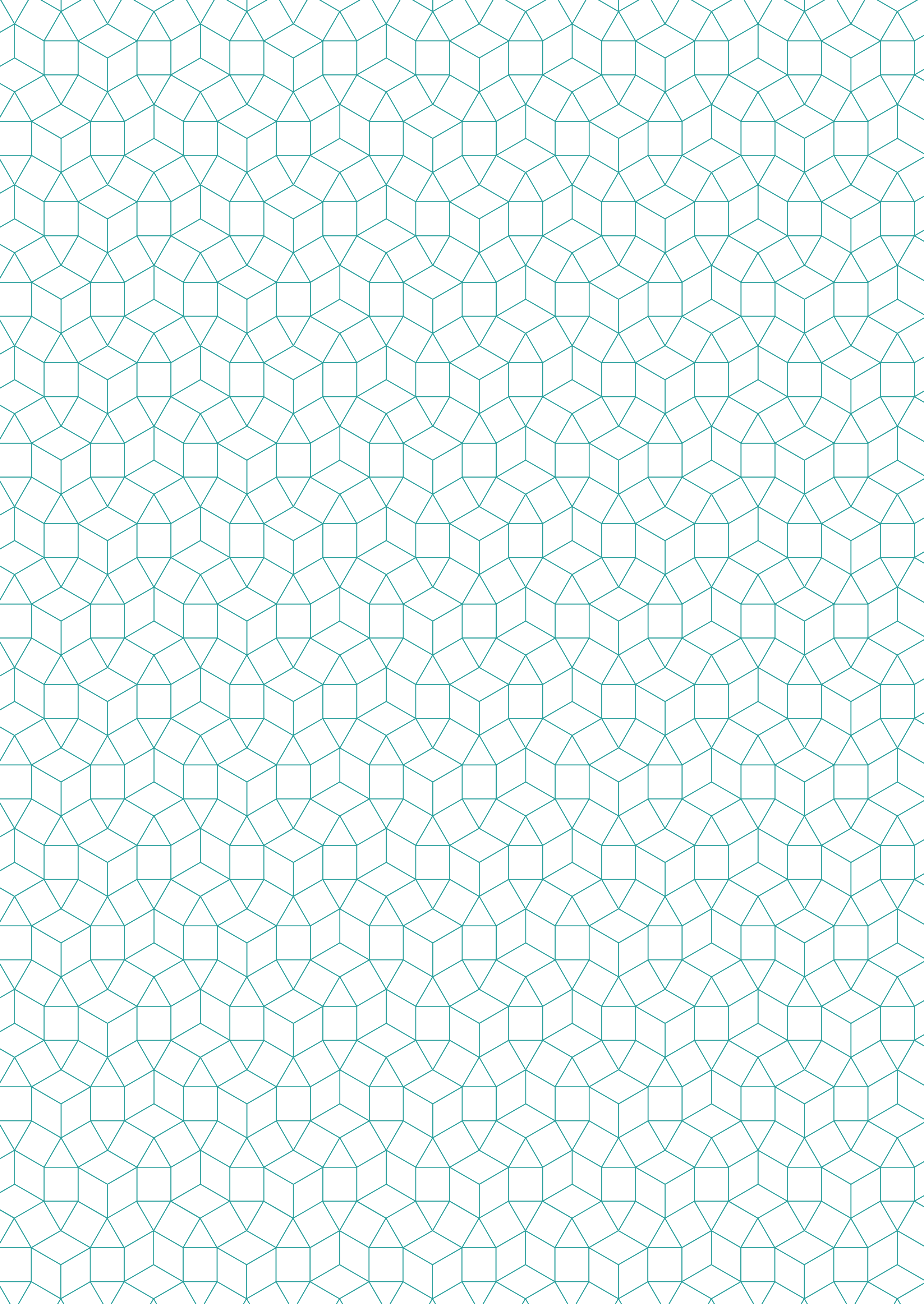
Jantsje M. Mol

Improving flood preparedness using insights from economic experiments

Jantsje M. Mol

IMPROVING FLOOD PREPAREDNESS

USING INSIGHTS FROM ECONOMIC EXPERIMENTS



Improving flood preparedness using insights from economic experiments

Jantsje M. Mol

Reading committee: prof.dr. Jeroen Aerts
prof.dr. Martijn van den Assem
prof.dr. Eline van der Heijden
dr. Özgür Gürerk
dr. Claudia Schwirplies

Publisher: Oma Maatje Mol-Krijger
Cover design: Leonie Staman (www.leoniestaman.nl)
Printed by: Gildeprint (www.gildeprint.nl)

Improving flood preparedness using insights from economic experiments
PhD thesis, Vrije Universiteit Amsterdam, The Netherlands
ISBN: 978-94-6419-245-2

Jantsje M. Mol, Amsterdam, September 2021
www.jantsje.nl

This research was conducted under the auspices of the Graduate School for Socio-Economic and Natural Sciences of the Environment (SENSE). The research for this thesis was carried out at the Institute for Environmental Studies (IVM), Vrije Universiteit Amsterdam.

VRIJE UNIVERSITEIT

Improving flood preparedness using insights from economic experiments

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad Doctor aan

de Vrije Universiteit Amsterdam,

op gezag van de rector magnificus

prof.dr. V. Subramaniam,

in het openbaar te verdedigen

ten overstaan van de promotiecommissie

van de Faculteit der Bètawetenschappen

op vrijdag 10 september 2021 om 13.45 uur

in de aula van de universiteit,

De Boelelaan 1105

door

Jantsje Maria Mol

geboren te Bolsward

promotor: prof.dr. W.J.W. Botzen

copromotor: dr. J.E. Blasch

The world cannot be understood without numbers. But the world cannot be understood with numbers alone.

HANS ROSLING

Contents

Samenvatting

Overstromingen horen tot de meest gevaarlijke natuurverschijnselen in de wereld. Jaarlijks worden er miljoenen mensen door geraakt, wat leidt tot vele doden, gewonden, evacuaties, emigratie en extreem veel schade. De verwachting is dat er in de toekomst door klimaatverandering extremere overstromingen zullen plaatsvinden, wat in combinatie met bevolkingsgroei en socio-economische ontwikkelingen in risicogebieden kan leiden tot rampzalige gevolgen. Deze ontwikkeling vraagt om onderzoek naar de mogelijkheden om risico's te verkleinen. De vraag is hoe verzekeringen moeten worden opgezet om de kosten te spreiden en wat er aan investeringen nodig is om schade te beperken. Daarbij kan gedacht worden aan grote infrastructurele projecten zoals de Deltawerken of het stimuleren van investeringen in schadebeperkende maatregelen door huiseigenaren.

Individuele huiseigenaren kunnen een aantal maatregelen nemen om wateroverlast ten gevolge van een overstroming te reduceren of zelfs te voorkomen. De eerste optie is het plaatsen van zandzakken of vloedschotten om te voorkomen dat water in het huis komt, oftewel het *droog overstromingsbestendig maken* van een huis. Een tweede is om aanpassingen aan het interieur en het gebouw te doen, zodat er zo min mogelijk schade ontstaat wanneer er water binnenstroomt. Denk aan het verplaatsen van de meterkast of de wasmachine naar een hogere verdieping en het vervangen van tapijt op de begane grond door een waterbestendige (tegel)vloer. Deze methode wordt ook wel *nat overstromingsbestendig maken* genoemd. Tenslotte is er de mogelijkheid complete huizen boven het waterniveau van overstromingen te bouwen, bijvoorbeeld op palen, wat voornamelijk effectief kan zijn bij nieuwbouwprojecten. Onderzoek heeft aangetoond dat schadebeperkende maatregelen door huiseigenaren kosteneffectief zijn op de lange termijn, bijvoorbeeld voor periodes met dezelfde looptijd als hypotheken.

Ondanks de steeds groter wordende dreiging van zeespiegelstijging, zijn er tot dusver nog maar weinig bewoners in overstromingsrisicogebieden die investeren in schadebeperkende maatregelen. Er is sprake van een relatief laag investeringsniveau, wat verklaard zou kunnen worden door *moreel*

risico, slechte risico-inschattingen en cognitieve beperkingen. Dit proefschrift probeert allereerst vast te stellen welke factoren dit lage investeringsniveau het beste kunnen verklaren. Verder richt het zich op het vinden en testen van prikkels die investeringen zouden kunnen stimuleren, zoals verzekeringsprikkels, risicocommunicatie in *virtual reality* en sociale *norm-nudges*. Het doel van dit proefschrift is om inzicht te verkrijgen in de effectiviteit van verschillende methoden die het aantal en de omvang van investeringen in schadebeperkende maatregelen kunnen vergroten. Dit wordt bereikt door het uitvoeren van experimenten, waaronder (1) experimenten met studenten in een computerlab; (2) online experimenten in vragenlijsten onder huiseigenaren in overstromingsrisicogebieden; en (3) een economisch experiment waarbij gebruik gemaakt wordt van een *virtual reality*-bril.

Hoofdstuk 2 gaat na of en in hoeverre financiële prikkels het nemen van schadebeperkende maatregelen onder een verplichte overstromingsrisico-verzekering kunnen bevorderen. Moreel risico (*moral hazard*) is een begrip uit de verzekeringstheorie. Het beschrijft een situatie waarin mensen met een verzekering zich minder voorzichtig gedragen dan onverzekerden, simpelweg omdat zij verwachten dat de verzekering schade zal vergoeden. De traditionele financiële prikkel uit de verzekeringswereld om moreel risico te verlagen, is het verhogen van het eigen risico, wat resulteert in een verlaagde dekking. Het onderliggende idee is dat een polishouder met een hoog eigen risico gemotiveerder is om zichzelf tegen risico's te beschermen, aangezien deze persoon een groot gedeelte van de schade uit eigen zak moet betalen.

Het labexperiment in dit hoofdstuk test verschillende niveaus van het eigen risico, namelijk 5%, 15% of 20%. De experimentele resultaten komen overeen met de verwachtingen op basis van theorie: een verhoogd eigen risico leidt tot een lichte toename in investeringen in schadebeperkende maatregelen. Een andere financiële prikkel is om de verzekerde een premiekorting aan te bieden, afhankelijk van de verwachte schadevermindering. Zo'n premiekorting wordt al wel geregeld door zorgverzekeraars toegepast, maar wordt zelden gebruikt in de context van natuurrampverzekeringen. De resultaten van het experiment tonen aan dat een premiekorting schadereducerende investeringen inderdaad kan doen stijgen, ongeveer evenveel als wanneer de kans op schade toeneemt van 1% tot 5%. Een bekend probleem voor huiseigenaren die willen investeren is dat zij de lumpsum-betaling niet in één keer kunnen of willen voldoen. Dit probleem kan worden opgelost door een lening onder gunstige rentevoorwaarden, waardoor de investeringskosten over een langere periode te spreiden zijn. Deze gunstige lening-optie zorgde in het labexperiment echter niet voor meer investeringen. Dat kan betekenen dat een dergelijke lening niet erg behulpzaam is bij het verhogen van investeringen in schadebeperking. Of, zoals onze studentensteekproef suggereert, kan de ineffectiviteit van de lening worden verklaard door een algemene afkeer van maken van schulden of door een gebrek aan externe validiteit: in het spel werden de investeringskosten gespreid over meerdere minuten, in plaats van over jaren, zoals in de echte wereld.

Hoofdstuk 3 beschrijft een ander experiment met betrekking tot financiële prikkels voor investeringen in schadebeperkende maatregelen. Ditmaal waren de proefpersonen echter geen studenten in een lab maar huiseigenaren in risicogebieden. Centraal stond het type verzekering: privaat of juist publiek. Een deel van de proefpersonen kreeg de mogelijkheid een verzekering te selecteren, terwijl het andere deel in de uitgangspositie met een verplichte verzekering geconfronteerd werd. Uit de resultaten blijkt dat mensen die vrijwillig bereid waren een overstromingsrisicoverzekering af te sluiten ongeveer 1,000 ECU (experimentele valuta) meer in schadebeperking investeerden dan personen die verplicht zo'n verzekering hadden moeten afsluiten. Dit hoofdstuk bevestigt verder dat een premiekorting investeringen kan vergroten, en dat deze financiële prikkel even effectief is onder beide soorten verzekeringen (vrijwillig/privaat versus verplicht/publiek).

In Hoofdstuk 4 zijn verschillende verkeerde inschattingen van overstromingsrisico's door Nederlandse huiseigenaren onderzocht. Het onder- of juist overschatten blijkt te maken te hebben met de risico's zelf (hoe extremer ze zijn, hoe groter de kans dat mensen ze verkeerd inschatten), maar ook met een aantal cognitieve denkfouten en iemands persoonlijkheid. Dit hoofdstuk toont aan dat 53% van de huishoudens de kans op een overstroming overschat, maar ook dat 54% tegelijkertijd juist onderschat hoe hoog het water bij een overstroming maximaal kan stijgen. De meeste respondenten weten de maximale schade goed in te schatten. De resultaten in dit hoofdstuk tonen verder aan dat er drie groepen mensen zijn die betere inschattingen maken van het risico, in die zin dat ze er minder ver naast zitten. Dat zijn mensen die een overstroming hebben meegemaakt, oudere mensen en mensen die veel vertrouwen hebben in dijkonderhoud, en mensen die veel vertrouwen hebben in dijkonderhoud, betere inschattingen maken van het risico (zij zitten er minder ver naast).

Uit eerder onderzoek is gebleken dat mensen die een overstroming hebben meegemaakt, over het algemeen meer investeren in bescherming en preventie dan mensen uit een soortgelijke omgeving die zoiets nog nooit hebben ervaren. Uiteraard is het geen goede oplossing om het risicobewustzijn van het grote publiek te vergroten door werkelijk dijken te laten doorbreken, maar nieuwe technologieën kunnen hier uitkomst bieden. Hoofdstuk 5 biedt een overzicht van de mogelijkheden van *high-immersive virtual reality*-technologie voor experimenteel economisch onderzoek, waarbij mensen met behulp van een *virtual reality* (VR) bril volledig worden ondergedompeld in een andere wereld. Deze methode geeft onderzoekers de mogelijkheid om op een veilige manier risicoperceptie te verbeteren, negatieve emoties op te roepen en vertrouwen in de effectiviteit van bepaalde maatregelen te vergroten.

Het doel van Hoofdstuk 6 is om na te gaan of en hoe een gesimuleerde overstroming in een virtual reality-omgeving ertoe kan leiden dat mensen zich beter op een overstroming voorbereiden. De resultaten van het experiment laten zien dat de risicoperceptie en de bereidheid tot het doen van investeringen in het overstromingsspel hoger waren bij de groep die de nagebootste overstroming in VR had meegemaakt dan bij de controlegroep die dat niet had gedaan. Deze effecten waren stabiel tot wel vier weken na de VR-ervaring. Er was echter geen verschil tussen beide groepen ten aanzien van de informatie die ze opvroegen of het aantal schadebeperkende maatregelen die ze thuis hadden genomen, binnen de vier weken na het experiment.

Een andere belangrijke factor in de omgang met overstromingsrisico die is blootgelegd door vragenlijstonderzoek betreft het gedrag van 'anderen'. Hoofdstuk 7 verkent de mogelijkheden van een zogenoemde *sociale norm-nudge*: een bericht met informatie over het gedrag van anderen dat zou kunnen stimuleren tot navolging. In dit hoofdstuk worden twee *norm-nudge* berichten getest in een online experiment met grote representatieve groepen huiseigenaren in Nederland en Spanje. Hier was geen sprake van enig effect van de interventie: blijkbaar konden de gekozen specifieke sociale norm-nudges mensen niet motiveren tot een betere voorbereiding op overstromingsrisico's. Wanneer zo'n *nudge* in een bepaald domein niet effectief blijkt te werken kan dit leiden tot het advies richting de overheid om steviger maatregelen toe te passen, zoals het voorzien in financiële prikkels of regulering in toezicht.

Dit proefschrift geeft verschillende aanbevelingen voor beleidsmakers. Allereerst is uit de ondernomen experimenten duidelijk geworden dat in scenario's met extreem kleine kansen mensen geen enkel moreel risico ervaren. Hieruit volgt dat het in zulke scenario's niet nodig is om een (hoog) eigen risico als optie aan te bieden. Dit proefschrift rechtvaardigt eerder het verplicht stellen van een overstromingsrisicoverzekering voor hypotheken in risicogebieden, juist omdat men geen moreel risico blijkt te ervaren en op basis van het gegeven dat respondenten niet erg geneigd waren gebruik te maken van een vrijwillige verzekering. Verder blijkt uit het onderzoek dat er een substantiële groep polishouders bestaat die bereid is om te investeren in schadebeperkende maatregelen, zelfs als zij al verzekerd zijn tegen overstromingsschade. Helder is ook dat het investeren in schadebeperkende maatregelen gestimuleerd kan worden door het aanbieden van premiekortingen. Deze resultaten ondersteunen de politieke hervormingen in de Europese Unie en in de Verenigde Staten die zijn gericht op het verbinden van overstromingsdekking aan risicoreductie.

Dit proefschrift geeft ook enkele aanbevelingen met betrekking tot informatiecampagnes die als doel hebben inwoners beter voor te bereiden op een mogelijke overstroming. De verkregen data tonen aan dat het beter is om de nadruk te leggen op de kosteneffectiviteit van maatregelen en risicogerelateerde emoties dan op het vergroten van het algemene besef dat een overstroming kan plaatsvinden. Deze campagnes kunnen het beste gericht worden op huiseigenaren in overstromingsgevoelige gebieden, aangezien zij op dit moment oververtegenwoordigd blijken in de groep die het risico onderschat. Een mogelijk effectieve manier om risicoperceptie te vergroten is om simulaties van een overstroming te laten zien in *virtual reality*. Vervolgonderzoek moet uitwijzen of goedkopere technologie, zoals eenvoudige lenzenhouders waarin smartphones geplaatst kunnen worden, tot hetzelfde resultaat kunnen leiden als het doen beleven van overstromingen met geavanceerde dure VR-brillen.

Een laatste aanbeveling aan beleidsmakers die uit dit proefschrift voortkomt is dat het belangrijk is om energie te steken in het aanspreken van persoonlijke normen. Uit verschillende hoofdstukken in dit proefschrift blijkt dat mensen met een sterk moreel besef (in vaktermen: *persoonlijke normen*) eerder geneigd zijn om schadebeperkende investeringen te doen. Toekomstig onderzoek moet uitwijzen hoe persoonlijke normen op dit terrein tot stand komen en hoe mensen effectief daarop aangesproken kunnen worden.

Summary

Flooding is one of the most dangerous natural hazards worldwide, causing widespread economic damage in coastal areas, thousands of deaths and injuries, and displacing millions of people every year. The impacts of flooding are projected to increase in the future as a result of combined socioeconomic development in exposed locations and climate change. This trend calls for research into flood risk reduction strategies, including disaster risk insurance (spreading risks over a large group of individuals) and damage reduction, by providing structural flood protection or stimulating individuals to invest in damage reduction measures.

Individual homeowners can take a number of measures to reduce potential flood damage to their homes. First, they can use sandbags or flood shields to prevent the water from entering their home, which is called *dry flood proofing*. Second, they can make amendments to their interior to minimize damage once the water enters their home, such as moving electrical appliances to a higher floor or replacing an expensive carpet by a tile floor. This method is called *wet flood proofing*. Finally, entire homes can be elevated, a method that may be most effective for new structures. Previous research has shown that damage reduction measures taken by homeowners can substantially limit the expected damages from flooding, which makes these measures cost-effective over time (e.g. the time-span of a mortgage).

Nevertheless, few people in flood-prone areas invest in these measures. This thesis examines several factors that could explain the lack of voluntary investment in individual damage reduction measures, including moral hazard, risk misperceptions and bounded rationality. Furthermore, this thesis investigates various incentives to stimulate investments in damage-reducing measures, such as insurance incentives, risk communication in virtual reality and social norm-nudges. The main goal of this thesis is to obtain insights into the effectiveness of different ways to stimulate investments in flood damage-reducing mitigation measures. This is achieved by experimental economics

methods: lab experiments with students, online experiments in surveys with homeowners in flood-prone areas, as well as an economic experiment using a *virtual reality* experience.

A central economic theory problem to be tackled in this context is moral hazard. Moral hazard occurs when losses are shared and incentives to limit risk are absent. For example, people who have travel insurance may take less care of their belongings, because they know they will be reimbursed by their insurer in case of loss, theft or damage. The problem of moral hazard has been documented in many contexts, but evidence on moral hazard in disaster insurance markets is scarce. This thesis addresses the moral hazard problem with experimental economics methods.

Chapter 2 develops an economic lab experiment to evaluate investments in damage-reducing measures in response to different treatments. Moral hazard is found in the scenarios where the probability of loss is high (15%), but not when the probability of loss is low (3%). This chapter further examines the impact of different financial incentives of flood risk mitigation measures under mandatory flood insurance. The traditional financial incentive applied by the insurance industry is increasing the deductible level, which will decrease coverage. The idea is that under reduced coverage, a policyholder will be more motivated to take care of the risk, seeing that a larger part of the damage has to be paid by him- or herself. The lab experiment in this chapter tests different deductible levels: varying between 5%, 15% and 20%. The results are in line with theoretical predictions: increasing the deductible leads to slightly higher investments in damage-reducing measures.

Another financial incentive to stimulate investments is to give policyholders a discount on their premium, based on the expected value of damage-reduction. Such a premium discount is already common practice in health insurance, but had not yet been tested in the context of disaster risk insurance. The experimental results show that a premium discount can indeed increase investments in individual damage-reducing measures to the same extent as an increase in probability from 1% to 5%. A potential problem for homeowners who want to invest in damage-reducing measures is that they simply cannot pay the high initial investment costs. One solution that has been proposed is a low interest loan that spreads investment costs over multiple periods. However, the experimental loan treatment did not encourage subjects to invest more in damage-reduction. It could be the case that such a loan is not very helpful in increasing investments in damage-reduction. Alternatively, the ineffectiveness of the loan treatment can be explained by the fact that our student sample disliked the idea of being in debt, or by a lack of external validity (the investment costs were spread over several minutes in the game, rather than over years, as in the real world).

Chapter 3 again examines financial incentives for investing in damage-reducing measures, but with homeowners in floodplains, rather than students in the lab. In this experiment, the type of insurance scheme was varied: some participants could select or reject insurance (a voluntary insurance scheme), while others were confronted with a mandatory insurance scheme. Respondents who were willing to pay for voluntary flood insurance coverage invested approximately 1,000 ECU (experimental currency units) more than those under mandatory insurance coverage. The results further confirm that a premium discount can increase investments in damage-reducing measures, and that this financial incentive is equally effective in the voluntary and the mandatory insurance markets.

In **Chapter 4**, possible flood risk misperceptions of Dutch floodplain inhabitants are studied. Whether respondents under- or overestimated flood risk was related to objective risk assessments, heuristics and personal characteristics. This chapter reveals that 53% of households overestimate the flood probability and 54% underestimate the maximum water level in case of a flood. Most respondents correctly estimate the maximum damage. The chapter further shows that experience of a flood, age and trust in dike maintenance seem to decrease flood risk misperceptions.

From past research, we know that people with direct flood experience generally invest more in protection than those who live in similar regions who have no such experience. Of course, breaking dikes is no real solution to tackle the problem of low flood preparedness, but novel virtual reality technology may offer some opportunities. **Chapter 5** investigates the possibilities of immersive virtual environments for experimental economics. This approach may allow researchers to safely boost risk perception, negative emotions and coping appraisal through an experience in a high-immersive virtual reality environment. The aim of **Chapter 6** is to examine whether a simulated flood can stimulate people to prepare for flooding. This research finds that risk perceptions and investments in the flood risk investment game were significantly larger for participants who experienced the virtual flood than for those in the control group. These effects are persistent up to 4 weeks after the VR experience. However, the results show no change between the groups in information search or number of measures installed at home.

Another important determinant in flood preparedness known from survey research is the behavior of others. **Chapter 7** explores whether the tendency to follow others can be stimulated through a so-called social norm-nudge: a message that informs individuals about the actions of others, which may stimulate people to copy this behavior. This chapter tested two norm-nudge messages in an online experiment with large representative samples of homeowners in two European countries. The results did not show any evidence of a treatment effect, which suggests that these social norm-nudges do not affect flood preparedness of respondents. If a nudge in the environmental domain proves ineffective, this may justify the use of stronger measures, such as financial incentives or regulations and bans.

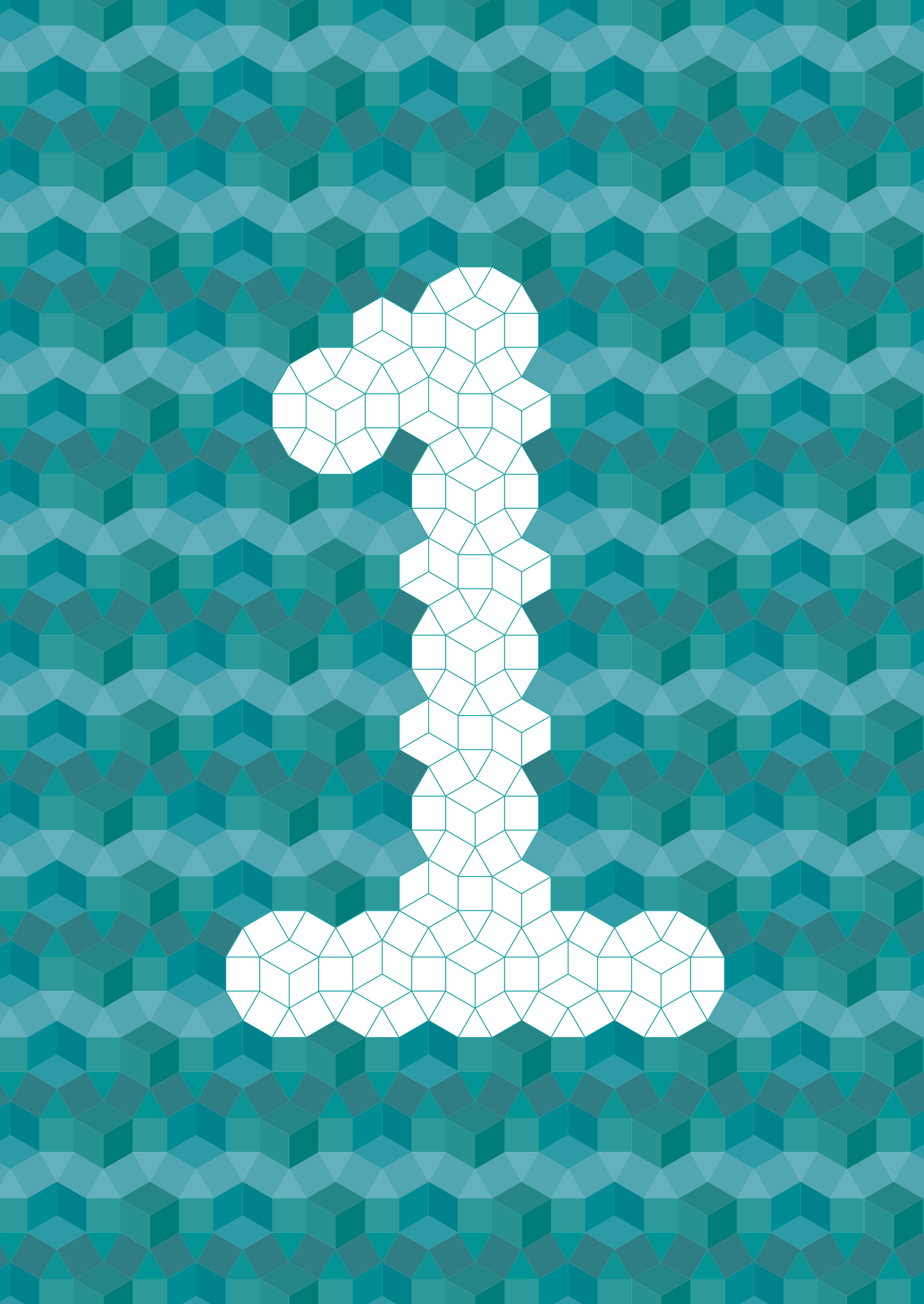
This thesis makes several key policy recommendations. First, the finding that there is no moral hazard throughout multiple experiments in disaster risk insurance context suggests that high deductibles are not necessary to limit such an effect. This thesis may justify the strengthening of purchase requirements for flood insurance, based on the lack of moral hazard effects and low voluntary take-up rates. Moreover, the experimental results showed that a substantial group of policyholders is willing to invest in damage-reducing measures, even if they are already insured. Flood preparedness can be stimulated further by rewarding policyholders who make such investment with discounts on their insurance premium. These results support the ongoing debates and reforms aimed at linking flood insurance coverage with risk reduction in the European Union and the United States.

This thesis also has some recommendations for future informational campaigns aimed at improving flood preparedness. The results showed that it is better to focus on explaining cost-effectiveness of protective measures and risk-related emotions than on increasing awareness about flood risk in general. These campaigns should specifically target homeowners in low-lying areas as they are currently over-represented in the share of under-estimators of flood risk. One potential way to increase risk perception is to use *virtual reality* simulations of floods. Future research could examine whether lower tech approaches, such as VR set-ups that rely on a smartphone, can be equally effective as expensive head-mounted displays.

Finally, policy makers should pay particular attention to activating personal norms, which were found to be associated with flood risk preparedness in several chapters throughout this thesis. Further work could examine the interactions between the antecedents of personal norms and message design to explore how personal norms can be effectively activated.

List of Figures

List of Tables


1

Introduction

Background: Increasing risk from floods

Flooding is one of the most significant natural disasters worldwide, responsible for 662 billion US dollars of recorded damages and affecting 2.3 billion people worldwide between 1995 and 2015 (UNISDR, 2015). The impacts of flooding on society have increased in the past decades and are expected to increase further in the future, due to a combination of climate change and increased development in disaster prone areas (IPCC, 2012; Munich RE, 2018). This trend leads to a growing interest in damage reduction strategies, which can be used to manage the financial risk for individuals and institutions (disaster risk insurance) or reduce the risk altogether (disaster risk reduction).

Insurance arrangements can be useful tools for limiting the costs of natural disasters by spreading risk intertemporally and geographically over a large group of policyholders and for providing financial compensation after a disaster to facilitate recovery. Several aspects of disaster risk insurance have been studied, such as demand Botzen and van den Bergh, 2012; Robinson and Botzen, 2018) and affordability (Hudson et al., 2016). Despite growing interest in insurance as a tool in disaster risk management, the design of such insurance arrangements is heavily debated among governments, which tend to focus on affordability and coverage, and the insurance industry, which tends to focus on risk-based pricing and risk reduction (Hudson et al., 2016).

With regard to disaster risk reduction, most research so far has focused on the role of governments in providing structural flood protection, such as dikes (Kreibich et al., 2015). The two main classes of risk reduction are defined as self-insurance (reducing the damage in case of a loss) and self-protection (reducing the probability of a loss occurring) (Ehrlich and Becker, 1972). Stimulating individuals to invest in self-insurance is a promising approach to decrease expected damages from floods (Den et al, 2017). Self-protection, on the other hand, is less applicable to flood risk mitigation, as it is almost impossible for individuals to lower the probability of a flood. Risk reduction through self-protection is more effective for other natural hazards, such as wildfires: by burning small patches of forest, the probability of a large wildfire decreases. The economically optimal damage reduction strategy for flooding is probably a combination of different elements, including disaster risk insurance, public disaster risk reduction and individual investments in damage reduction.

One relevant geographical area for studying flood damage reduction strategies is the Netherlands, a European country with long history of protection against flooding. Approximately half of the country is located behind dikes, including the metropolitan area where the main business districts and the government are situated. These low-lying areas (dike-rings) are protected from flooding by large dike infrastructures, leading to one of the highest flood safety standards across the globe (Scussolini et al., 2016). For example, some dike-rings at the coast have safety standards of 1:10000, which means that the dikes are designed to withstand an extreme flood event that may occur once in 10,000 years. Dike-rings including the main rivers have lower safety standards, ranging from 1:250 to 1:4000. Safety standards of dike-rings are set by law and were recently updated in view of a nationwide flood risk assessment (Vergouwe, 2015). The current safety standards conform to maximal acceptable failure probabilities per dike segment (Jonkman et al., 2018). Regional water authorities are in charge of inspections, maintenance and emergency measures to maintain these safety standards (Lendering et al., 2016). Despite the fact that the Dutch flood defenses have one of the highest safety standards in the world, defenses cannot guarantee 100% protection. In other words, a residual flood risk remains. The consequences of flooding in this area could be catastrophic, with potential damages up to 100 billion Euros (Aerts et al., 2008). Nevertheless, Dutch floodplain inhabitants might not be aware of the possibility of flooding, as the most recent severe river floods occurred in 1993 and 1995 and the most recent coastal flood dates back to 1953.

Strengthening flood defenses to the point at which flooding is totally harmless for society is economically inefficient, because the costs would exceed the benefits (Eijgenraam et al., 2014). Raising dike levels would decrease the probability of a flood, but increase potential flood damage amounts in the situation of a dike breach, because higher dikes allow for higher maximum water levels (Botzen et al., 2013). Moreover, higher dikes may give a false sense of safety, which could facilitate development in flood plains and hence increase potential damage. This principle has been called ‘the levee effect’ (Tobin, 1995). Most flood risk managers agree that residual flood risk should not be managed by structural measures only, and therefore the focus has shifted to alternative measures, such as insurance arrangements and individual damage reduction strategies.

In the European Union, compensation for flood losses varies across member states, where some offer public insurance which is often mandatory and others involve private market insurance which is often voluntary (Schwarze et al., 2011; Paudel et al., 2012). In various countries it has been debated whether these arrangements should be reformed to provide policyholders with stronger incentives to limit the risk. The interplay of insurance and self-protection has been extensively studied, both theoretically (Ehrlich and Becker, 1972) and empirically (Jaspersen, 2016). However, less is known about the interaction of insurance with individual self-insurance activities, which is one of the main topics of this thesis.

Explaining investments in damage-reducing measures

Individual homeowners can take a number of measures to reduce potential flood damage to their property. These measures fall into three broad categories: dry flood proofing (shielding a house to prevent water from entering), wet flood proofing (minimizing damage once water has invaded a house), and the elevation of structures. Recent evidence shows that damage reduction measures taken by private homeowners are cost-effective over time and can substantially limit the expected damages from flooding (Kreibich et al., 2015). Despite the availability of cost-effective damage reduction measures, few people in flood-prone areas invest in or implement them (Kreibich et al., 2015; Botzen et al., 2019a). The lack of voluntary investment in mitigation measures could be explained by several factors, including moral hazard, risk misperceptions and bounded rationality.

Moral hazard

A natural starting point for examining investments in self-insurance is economic theory. One important model of individual decision making under risk is expected utility theory (EUT), which assumes that individuals assess the likelihood and consequences of several choice alternatives, and subsequently choose the alternative that gives the highest expected utility (von Neumann and Morgenstern, 1947). When the objective likelihood is uncertain or unavailable, individuals may still maximize expected utility by using their own subjective estimates of probabilities and losses (Savage, 1954), which in our applications are the perceived flood probability and damage. From an expected utility theory perspective, self-insurance investments should increase when the probability of a loss increases and when the insurance coverage decreases through a higher deductible. The deductible is the amount of damage that must be paid by the policyholder before the insurer will cover any expenses, which provides a financial incentive to reduce risk for the policyholder. In other words, the deductible reduces a policyholder's level of insurance coverage, and reduced insurance coverage provides an incentive to increase self-insurance investments. On the other hand, investments in self-insurance will be lower in the presence of insurance, due to moral hazard and adverse selection (Winter, 2013).

Moral hazard is a well-known problem in economic theory, especially in the insurance domain. It occurs in the absence of incentives to take care and limit risk, when losses are shared. This behavior has been defined as ex ante moral hazard, as opposed to ex post moral hazard, which refers to the situation where the size of the loss is overstated to get higher compensation for the loss (Di Mauro, 2002). The problem of moral hazard has been documented in many contexts of asymmetric information, including insurance contracts, court settlements, tax evasion and work effort (Rowell and Connelly, 2012). However, so far there has not been much evidence on the existence of moral hazard in natural disaster insurance markets (Botzen, 2013). A related theoretical concept is adverse selection, where people who feel that they are vulnerable, purchase insurance, while those who feel secure do not. In this case, the pool of the insured consists mainly of highly exposed individuals, removing the spreading of risk between high and low risk types that typically characterizes insurance, which can result in unaffordable premiums and deficient coverage in case of a disaster. Conversely, advantageous selection has also been documented: risk averse individuals both purchase insurance and take risk mitigation measures, while risk seeking individuals omit both (de Meza and Webb, 2001). Despite the large theoretical literature on moral hazard, little is known about whether moral hazard or advantageous selection dominates in flood insurance markets, and how moral hazard effects differ between different types of (voluntary or mandatory) flood insurance arrangements.

Risk misperceptions

One explanation for the lack of investments in self-insurance, is that flood risk perceptions of homeowners differ considerably from objective estimates, skewing their assessment of the damage that can be avoided by risk reduction measures (Siegrist and Gutscher, 2008; Bubeck et al., 2013). Kunreuther and Pauly (2004) postulated based on the expected utility framework that individuals facing disaster risks expect a low return from searching for information about their risk, and hence are unlikely to be fully informed about the risk they face. As a result, perceptions of disaster risks are likely to be biased, but would still be related to the objective risk faced by individuals (Kunreuther and Pauly, 2004). This means that risk perceptions would at least partially relate to objective risk and, hence, the latter may relate to the degree to which people under- or overestimate their risk. Flood risk perceptions are important, as they may affect support for public investments in flood protection infrastructure (Ripberger et al., 2018). This leads to a growing interest in risk perception research, which is important for the design of effective risk communication campaigns that stimulate people to better prepare for increasing natural disaster risks (Botzen, 2013; Kellens et al., 2013). Previous studies have examined flood risk perception in relation to knowledge of the causes of flood events (Botzen et al., 2009a), distance to a perceived flood zone (O'Neill et al., 2016) and climate change information (de Boer et al., 2016). However, these studies did not examine how perceptions of the flood probability and damage differ from objectives estimates, and which factors explain misperceptions of flood risk in the Netherlands.

Decision heuristics

Various aspects of human behavior, especially in the domain of risky decision-making, are inconsistent with expected utility theory. As humans are not fully rational agents (bounded rationality, Simon, 1959) their behavior is better modeled by behavioral economic theories that do not assume full rationality (Kahneman, 2003). For example, some people downgrade or underweight the probability of risky prospects. Underweighting of risk can be accommodated by Prospect Theory (Kahneman and Tversky, 1979), which is a frequently used model for decisions under risk that has been used to explain behavior related to natural disasters (Page et al., 2014; Koetse and Brouwer, 2016). Under Prospect Theory, risk attitudes are defined by a combination of utility curvature, loss aversion and probability weighting.

Bounded rationality can cause several problems when individuals deal with natural disasters, such as floods. For example, individuals may be uninformed about risks they face from such events due to excessively high search costs of gathering information (Kunreuther and Pauly, 2004). Even when objective probabilities are available, for example when the government provides detailed flood maps, people might not process them rationally. To circumvent complex mental calculations, individuals sometimes fall back on certain heuristics or rules of thumb (Slovic et al., 2004), including threshold probabilities, the availability heuristic, intuitive feelings they have about risks and excessive risk aversion. In case of a substantial difference between rational predictions and behavioral findings, investigating behavioral motivations of (not) taking flood mitigation measures is relevant information for insurers as well as policy makers.

One example of deviations from rationality are threshold probabilities, which are related to difficulties understanding low-probability high-impact (LPHI) risks and underestimation of these risks in the absence of personal experience (Kunreuther and Pauly, 2004). Consequently, individuals might only respond to the risk when a certain threshold level of concern is reached (McClelland et al., 1993) or they might generally under-weigh the probability in their insurance decision. This decreases risk awareness, which represents an individual's subjective evaluation of an objective risk. Even if individuals are completely informed about the flood risk in their neighborhood, they may still neglect this risk until the flood probability exceeds a certain threshold (McClelland et al., 1993). Robinson and Botzen (2018) show that individuals who worry more about flooding report decreased threshold levels of concern. If an individual regards the flood probability as falling below their threshold level of concern, then this individual may not consider spending money on flood risk reduction. As a result, disasters with a probability smaller than the respective threshold level of concern will be neglected.

Behavioral explanations for low investments in self-insurance include systematic biases in judgment. One systematic decision bias related to the automatic and intuitive ways individuals process LPHI risks is myopia, or "the tendency to focus on overly short future time horizons when appraising immediate costs and the potential benefits of protective investments" (Kunreuther and Pauly, 2018, p.4). This means that these present-biased individuals appreciate value they have right now more than they expect to enjoy value in the future. As a result, the immediate upfront costs of protective investments loom larger than the predicted reduction of losses in the future.

Another important systematic bias which is examined in this thesis is herding, i.e. "the tendency to base choices on the observed actions of others" (Kunreuther and Pauly, 2018, p.4). The tendency to model behavior on what others do is a very common bias, especially under conditions of uncertainty. However, if others are not better informed than the individual him- or herself, which is often the case in the domain of natural hazards, herding will lead to suboptimal decisions. This has been illustrated in a survey of homeowners under flood and earthquake threats, where discussions with friends and neighbors were found to be more important factors for flood insurance demand than perceived risk (Kunreuther et al., 1978). Similarly, a survey of households in Australia found that perceived social norms had a greater influence on flood insurance purchases than homeowners' perceptions of flood risk (Lo, 2013). Note that previous research on social norms in the context of disaster risk is entirely survey-based and therefore correlational. This thesis is the first to examine the role of herding and social norms on damage-reducing investments in an experimental way, which allows for causal interpretations.

Other behavioral explanations

Finally, psychological theories may explain how individuals prepare for risks, such as protection motivation theory (PMT) (Rogers, 1975). PMT was originally developed to analyze preventive behavior in the health domain and has been applied effectively to other domains in the past decade, including flood risk preparedness (Bubeck et al., 2013; Grothmann and Reusswig, 2006). PMT captures two main cognitive processes that people experience when facing a threat: coping appraisal and threat appraisal. Threat appraisal describes the subjective evaluation of a certain risk by an individual, or how threatened one feels by the risk. Coping appraisal, on the other hand, refers to the cognitive process of the evaluation of possible responses to this threat, including their own ability to deal with the threat. Coping appraisal includes the perceived efficacy of mitigation measures (response efficacy), perceived ability to implement these mitigation measures (self-efficacy) and the perceived cost of mitigation measures (response cost) (Floyd et al., 2000). These coping values seem to be among the most important determinants of disaster preparedness (Bubeck et al., 2012; Botzen et al., 2019a) but it remains unclear how they can be enhanced. When subjects experience a risk in virtual reality, as well as the appropriate response, this might help them to realize that they are able to implement these measures (self-efficacy) and that the measures are indeed effective (response efficacy).

Stimulating investments in damage-reducing measures

To overcome these difficulties in the promotion of flood prevention measures, different incentives may be provided to floodplain inhabitants.

Insurance features and related incentives

The traditional way to offset moral hazard is through the use of a deductible. Other possibilities include premium discounts for households who invest in costly risk reduction measures and a loan to spread risk reduction investment costs (Poussin et al., 2014). The latter may reduce effects of individual time discounting and myopia which imply that large upfront mitigation costs weigh heavily compared with long term benefits of reduced risks, by dividing these upfront cost into smaller amounts to be paid in the future. Economic theory predicts that individuals invest less in self-insurance under insurance coverage, unless they are incentivized to make such investments through premium discounts (Ehrlich and Becker, 1972). However, individuals may respond differently to insurance features, such as a premium discount, when the insurance offered is mandatory (public insurance), rather than voluntary (market insurance), which is nearly impossible to study with non-experimental data. This thesis advances the literature by systematically studying moral hazard in relation to a variety of probability levels and deductibles, which has not been done yet in previous studies.

Risk communication in virtual reality

A large body of literature has revealed that individuals who have experienced a flood event, invest significantly more in preventive measures than those who live in analogous areas but lack direct flood experience (see e.g. Grothmann and Reusswig, 2006; Osberghaus, 2017). This relationship seems to be driven by strong negative emotions, while effectiveness and cost considerations also play a role (Siegrist and Gutscher, 2008). One drawback of lab experiments is that damage solely consists of a monetary component, which does not resemble the full experience of real-world flood damage, which may include a strong emotional component. A novel approach is to use virtual reality technology to examine whether a simulated flood can stimulate people to prepare for flooding. In a high-immersive virtual reality (VR) environment, users can interact with a computer simulated three-dimensional environment by using special equipment, such as a head mounted display with stereoscopic view (Innocenti, 2017). This approach allows the researchers to boost risk perception, coping appraisal, negative emotions and damage-reducing behavior through a flooding experience in a high-immersive VR environment. This approach has been successfully applied to find the psychological determinants of fire risk prevention (Jansen et al., 2020). Whereas early desktop VR games have been applied to the domain of flood risk (Zaalberg and Midden, 2013), a careful experiment using high-immersive VR to test the effects on risk perception, coping appraisal and behavior is missing from the disaster risk reduction literature.

Social norm-nudges

Previous research indicates that flood preparedness behavior is driven by the risk-reduction behaviors of others (Poussin et al., 2014; Grothmann and Reusswig, 2006). Along these lines, information about flood preparedness of others can increase flood preparedness. For example, Bubeck et al. (2013) showed a positive relationship between mitigation behavior and having neighbors and friends who have implemented flood mitigation measures. Hence, the herding bias can be used as a 'nudge', increasing rather than decreasing flood preparedness. Generally, nudges are a set of behavioral interventions that use cognitive boundaries, biases and habits in the presentation of choice alternatives, with the ultimate aim of improving welfare of those being 'nudged' (Thaler and Sunstein, 2008). One popular nudge based on social norms is called a norm-nudge (Bicchieri and Dimant, 2019), which encourages certain behavior by informing individuals about the actions of others, for example by showing energy conservation behavior of neighbors (Allcott, 2011) or tax compliance rates of fellow citizens (Hallsworth et al., 2017). Norm-nudges may stimulate people to copy this behavior, because humans are inclined to model behavior on what others do, or what they believe others do. Compared to traditional interventions such as taxes or regulations, norm-nudges are considered cheap, easy to implement and less prone to political resistance (Benartzi et al., 2017). Nevertheless, norm-nudges do not work in all circumstances and their effectiveness depends on the design of the norm-nudge (Hummel and Maedche, 2019). Moreover, there is a risk that a norm-nudge will elicit no effects (see e.g. Mackay et al., 2020; Chabé-Ferret et al., 2019) or even backfire, if not properly tailored to the population and context of interest (Hauser et al., 2018). So far, the effectiveness has not yet been tested in the context of flood preparedness, and little is known about how social norms nudges influence preventive behavior across countries characterized by different flood risk management regimes and different cultural backgrounds.

Research questions

Based on the aforementioned scientific research gaps, the main goal of this thesis is to obtain insights into the effectiveness of different ways to stimulate individuals to invest in flood damage-reducing mitigation measures. This objective can be met by answering the following six research questions:

1. To what extent are investments in damage-reducing measures determined by loss probabilities, deductibles and a moral hazard effect? (Chapter 2)
2. Are financial incentives from insurance effective in increasing investments in damage-reducing measures and does effectiveness vary with insurance scheme (public or private)? (Chapter 2 & 3)
3. Do households generally under- or overestimate flood risk and what factors explain these misperceptions? (Chapter 4)
4. What are the possibilities and challenges for experimental economics in high-immersive virtual environments? (Chapter 5)
5. Is it possible to increase flood preparedness with the experience of a flood in high-immersive virtual environments? (Chapter 6)
6. Could social norm-nudges help people in better preparing for flood risk and do they interact with individual characteristics and intercultural differences? (Chapter 7)

Method: Economic experiments

To examine preparedness behavior empirically, scientists have used field survey data, insurance market data and experimental methods. While the high external validity of field survey data is very valuable, the disadvantage of this type of research is that it is hard to find causal relationships, as different insurance plans are not allocated randomly to homeowners (endogeneity[1] bias). Insurance market data for natural hazard insurance markets is often not available and if it is, crucial data on preparedness behavior may not be documented. Moreover, it is very challenging if not impossible to disentangle moral hazard from adverse selection effects in insurance market data, and to examine behavioral mechanisms that drive demand for protection (Hudson et al., 2017). To address the potential confounds in field survey data and the lack of available market data, this PhD thesis applies experimental economics methods. The purpose of controlled lab experiments is to "examine why particular behavior outcomes occur in some situations and not in others" (Ostrom, 2010, p.647). Lab experiments have been beneficial to identify causal mechanisms in social science in general (e.g. Falk and Heckman, 2009) and in the insurance context in particular (Laury et al., 2009).

[1] Endogeneity occurs when the distribution of a predictor variable is correlated with the error term, for example due to a selection effect or an omitted variable.

Typology of economic experiments

Economic experiments are not limited to the laboratory, but can also be conducted in the field. Harrison and List (2004) created a taxonomy of four categories of experiments, ranging from very abstract to completely natural settings. Conventional laboratory experiments use abstract framing and a student subject pool to test predictions of (game) theory. Artefactual field experiments or lab-in-the-field experiments are more natural because they use a relevant population as subjects. Framed field experiments also use this relevant population and in addition take place in a natural environment, such as a school or a hospital. Finally, natural field experiments study the relevant population in a relevant setting, where participants do not know that they are in an experiment. A recent addition to the toolkit of experimental economists are immersive virtual reality experiments, which allow rich visualizations of the natural decision making environment, while controlling for various confounds. One important perception confound are previous experiences that participants may have in mind during economic games.

This thesis applies lab, lab-in-the-field and virtual reality experiments to examine the most promising ways to stimulate investments in damage-reducing measures. Framed field and natural field experiments would lack a certain level of experimental control necessary to differentiate between moral hazard and adverse selection. However, future research may test the validity of the results from this thesis in a field experiment.

Principles in experimental economics

Two important principles in experimental economics that are also followed in these thesis are the use of monetary incentives and the norm of no deception (Bardsley et al., 2010). To create an economic decision making situation under controlled preferences, participants are offered real monetary incentives. To gain full control over a participant's preferences in the experiment, three assumptions should be satisfied. First, participants prefer more money over less money (monotonicity). Second, earnings should be task related (salience). And finally, earnings should be sufficient to overcome other unobserved costs related to participation in the experiment (dominance). One problem with experiments related to natural hazards is that they normally include large losses. Such a large loss can be operationalized by implementing large payoffs, where the losses could be deducted from. However, the implementation of large payoffs is typically restricted by the budget of the experiment (Etchart-Vincent, 2004). This budget restriction can be fulfilled by implementing the random problem selection mechanism, which randomly selects one task and one subject to be paid at the end of the experiment (Kunreuther and Michel-Kerjan, 2015; Chaudhry et al., 2018).

A second principle in experimental economics is the proposition that experimental subjects should never be deceived, for example about the rules of the game or the payoff. The fact that no deception is used facilitates trust between experimental subjects and the experimenter, which is a necessary component of experimental control. For example, if a participant suspects deception, she or he might not believe that the payoff is task-dependent, which may cause this participant to play the game without full attention. This is the reason that many experimental labs have posted signs to explain about their no-deception policy.

Methods applied in the current thesis

The first experiment of this thesis was a lab experiment with students as subjects ($N = 357$) to examine moral hazard and self-insurance under different probability levels and deductibles. To this end, I developed a new individual investment game (the flood game) which allowed for variations in probability level, deductible and insurance. The experimental software used was oTree, which allowed for elaborated visuals and an easy extension of the game to an online platform (Chen et al., 2016). Different between-subjects treatments were aimed at identifying the effects of a premium discount and a mitigation loan. One important advantage of a lab experiment with a student sample is that it is no problem to have a task of a certain complexity. However, this comes at the cost of generalizability, because students are not the population of interest when it comes to flood preparedness decisions. For example, students are inexperienced with the purchase of homeowners insurance and their individual characteristics (such as risk attitudes and time preferences) may differ from the population. Nevertheless, the behavior of students can still allow for comparisons across the treatments. Given the multi-round structure of the data, I applied panel data analysis methods to analyze the data of the first experiment. Subjects were paid in cash at the end of the experiment, based on their results in the game.

To improve external validity, I developed the flood game further into a simple one-shot version, suitable for a representative sample of Dutch homeowners who could access the game through an online portal. This set-up allowed for a large sample size ($N = 2111$) of relevant decision-makers, namely homeowners living in areas under flood risk. The larger sample size allowed for an analysis of investments in self-insurance under (voluntary) market insurance, as it was expected that a small fraction of participants are willing to pay the premium for insurance against low probability flood risk. The multi-round design of the flood game in the first experiment was rather complex and repetitive for participants. Therefore, I anticipated that the consumer panel participants in the second study might be irritated or get bored when being asked to make their choice repeatedly, which could lead to lower completion rates and erratic choices. To examine self-selection into insurance, I developed a module in the flood game to assess willingness to pay for flood insurance. Based on their decisions in this module, participants could self-select into a treatment with or without insurance. Participants in the insurance treatments, both mandatory and self-selected, paid the same subsidized premium. All respondents were paid a fixed participation fee, while one participant was randomly selected for a large payment. This payment corresponded to the participant's bank balance at the end of the main scenario at a conversion rate of 100 ECU = €1, which could be up to €650.

The extensive post-experimental survey in the online homeowner experiment resulted in a rich data set on different aspects of flood risk perception (i.e. anticipated damage, expected water levels and return periods). I combined the data on risk perceptions with objective risk data derived from GIS (geographic information system) methods to examine the misperceptions of homeowners in the Dutch river delta. I estimated various regression models of objective risk variables, heuristics and personal characteristics on subjective flood probability and categorical expected damage from flooding.

For the virtual reality experiment, a scenario of a flooded home was developed in C Sharp via Unity 3D by the development team of the Network Institute of VU Amsterdam. In the VR experience, participants were first asked to protect their home by stacking sand bags to the doors and windows. Subsequently, they could experience a flood from inside their own protected home, as well as in the unprotected and flooded home of the neighbors. After the VR experience, participants were asked to complete the flood game on a desktop computer in the lab. An important novelty of this study is that all participants received a follow-up survey including another round of the flood game, a few weeks after the experiment, to test for the persistence of effects.

The lab-in-the-field experiment presented in the final chapter investigated the effect of social norm-nudges in the context of flood preparedness. To this end, two treatment groups were confronted with an empirical norm-nudge message with information about decisions of previous respondents. A third treatment group faced a focusing norm treatment, by eliciting beliefs about others' investment choices before participating in the investment game (Krupka and Weber, 2009). This experiment was also carried out online, with a large sample size in two different European countries ($N = 1200$ in the Netherlands and $N = 605$ in Spain). This set-up allowed for an assessment of differences in current flood risk management between those countries - with the Netherlands more focused on public flood protection through dikes and Spain on individual protection measures - influence risk attitudes and personal norms for protecting one's home.

Outline of the thesis

The contents of this thesis are divided over eight chapters. The purpose of Chapter 2 is to analyze the impact of different financial incentives of flood risk mitigation measures under mandatory flood insurance, answering research Question 1. In addition, the higher degree of experimental control allows for a clarification of the existence of moral hazard in this context (Question 2). Chapter 3 extends the first lab experiment to an online environment with floodplain inhabitants as participants, where a larger sample size allows for an analysis of individual risk and time preferences in the mitigation investment decision. This chapter further examines advantageous selection by investigating both mandatory and voluntary flood insurance purchases, answering Question 2. Next, Chapter 4 examines possible flood risk misperceptions of floodplain residents in the Netherlands, offers insights into factors that are related with the under- or overestimation of perceived flood risk. This chapter takes the analysis of flood risk misperceptions one step further by relating the type of misperception (over- versus under-estimation) to objective risk assessments, heuristics, and personal characteristics. Chapter 5 investigates the possibilities of immersive virtual environments for experimental economics, to answer Question 4 and to inform the experiment in the next chapter. Consequently, Chapter 6 applies the novel high-immersive virtual reality technology to examine the effect of experiencing a virtual flood on risk perception and damage-reducing investment behavior in an economic game (Question 5). Chapter 7 tests two empirical norm-nudge frames in an online experiment with large representative samples of homeowners in two European countries, to evaluate the possible interactions between norm-nudge effectiveness, individual characteristics and intercultural differences. This chapter answers Question 6. Finally, Chapter 8 concludes. Figure 1.1 provides a schematic overview of the thesis and the relationships between the chapters.

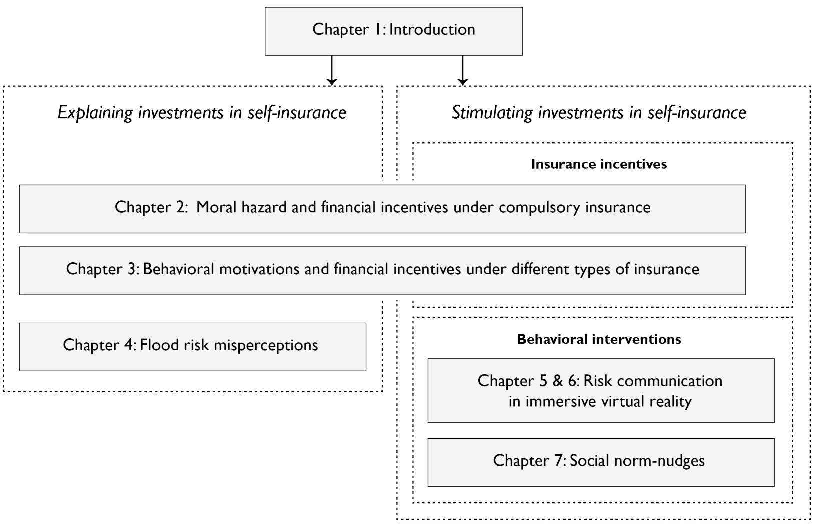


Figure 1.1: Schematic overview of the chapters of the thesis.


2

Risk reduction in compulsory disaster insurance: Experimental evidence on moral hazard and financial incentives

In a world in which economic losses due to natural disasters are set to increase, it is essential to study risk reduction strategies, including individual homeowner investments in damage-reducing (mitigation) measures. In this lab experiment (N = 357), we investigated the effects of different financial incentives, probability levels, and deductibles on self-insurance investments in a natural disaster insurance market with compulsory coverage. In particular, we examined how these investments are jointly influenced by financial incentives, such as insurance, premium discounts, and mitigation loans. We also studied the influence of behavioral characteristics, including individual time and risk preferences. We found that investments increase when the expected value of the damage increases (i.e., higher deductibles, higher probabilities). Moral hazard is found in the high-probability (15%) scenarios, but not in the low-probability (3%) scenarios. This suggests that moral hazard is less of an issue in an insurance market where probabilities are low. Our results demonstrate that a premium discount can increase investment in damage-reduction, as can a policyholder's risk aversion, perceived efficacy of protective measures, and worry about flooding.

This chapter is published as:
Mol, J. M., Botzen, W. J. W., & Blasch, J. E. (2020). Risk reduction in compulsory disaster insurance: Experimental evidence on moral hazard and financial incentives. *Journal of Behavioral and Experimental Economics,* 84(February), 101500.

2.1 Introduction

Economic losses due to low-probability/high-impact natural disaster events, such as floods, have increased in the past 25 years and this trend is likely to continue (IPCC, 2012; Munich RE, 2018). Insurance arrangements can be useful tools for limiting the costs of natural disasters by spreading risk intertemporally and geographically over a large group of policyholders[1] and for providing financial compensation after a disaster to facilitate recovery. Despite growing interest in insurance as a tool in disaster risk management, the design of such insurance arrangements is heavily debated among governments, which tend to focus on affordability and coverage, and the insurance industry, which tends to focus on risk-based pricing and risk reduction (Hudson et al., 2016).

Different options exist for policyholders to reduce risk, including self-insurance (reducing the damage in case of a loss) and self-protection (reducing the probability of a loss occurring). The interplay of insurance, self-insurance, and self-protection has been extensively studied, starting with an influential theoretical paper by Ehrlich and Becker (1972). Their model shows that market insurance and self-insurance are substitutes, whereas self-protection can be complementary to market insurance. Over the years, many experiments tested the normative predictions of insurance demand (see, e.g., Jaspersen, 2016, for a comprehensive review). While most of these papers investigate empirical regularities related to insurance demand, few focus on the interaction with risk reduction activities. This chapter relates to the empirical literature on self-insurance and self-protection, with a focus on the relevant dimensions of heterogeneity of self-insurance under compulsory insurance coverage for low-probability/high-impact risk. From an expected utility theory perspective, self-insurance investments should increase when the probability of a loss increases and when the insurance coverage decreases through a higher deductible. Investments in self-insurance should decrease in the presence of insurance due to moral hazard (Winter, 2013). Insurance arrangements could be further combined with explicit financial incentives to stimulate policyholders to install damage-reduction measures, such as premium discounts that reflect reduced risk. Our study aims to answer the following research questions: To what extent are investments in self-insurance under compulsory insurance coverage for low-probability/high-impact risk determined by loss probabilities, deductibles, and a moral hazard effect? Are financial incentives from insurance effective in increasing such investments?

Loss probabilities

Several previous studies have examined the value of self-insurance and self-protection under different probability levels, using an experimental methodology. In his seminal paper, Shogren (1990) studied individual

[1] For example the Caribbean Catastrophe Risk Insurance Facility: www.ccrif.org

responses to risk by self-insurance and self-protection, with experimental auctions under different probabilities (1%, 10%, 20%, and 40%). The study found higher investments in both risk reduction methods under increasing probabilities. Di Mauro and Anna (1996) examined the valuation of self-insurance and self-protection while varying the probability levels (3%, 20%, 50%, and 80%). They found higher bids on self-insurance and self-protection for increasing probabilities. Shafran (2011) examined preferences for self-protection against low and high probabilities of loss (1%, 2%, 20%, and 40%). In line with normative predictions from prospect theory, the study found that subjects were more likely to protect against risks with high probability than those with low probability and the same expected loss. Note that they examined self-protection rather than self-insurance, which is a key difference between this and our own study. More recently, Ozdemir (2017) compared the valuation of self-insurance and self-protection under risky and ambiguous prospects with different probabilities of loss (3%, 50%, and 80%) and found that the willingness to pay for self-insurance increases with probability, but only weakly.

Moral hazard

A potential difficulty in the promotion of damage-reduction measures is information asymmetry between the insurer and the policyholder regarding implemented measures. This asymmetry can lead to moral hazard, whereby insured individuals take fewer preventive measures, as these do not lower their premiums as long as the insurer cannot observe them (Arrow, 1963; Stiglitz, 1974; Arnott and Stiglitz, 1988). Many studies have empirically investigated moral hazard in insurance markets (see Cohen and Siegelman, 2010; Rowell and Connelly, 2012, for an overview), finding that it varies across markets, depending on the type of insurance product, amongst other factors. In this regard, studying the effect of insurance coverage on self-insurance in isolation from other factors enables getting insights into the moral hazard effect under different probabilities. Some researchers have examined moral hazard using an experimental approach (see Table A1). The contexts vary, including the principal-agent paradigm (work effort), field experiments on default in micro finance, and studies related to insurance. The closest to our experiment are Berger and Hershey (1994) and Di Mauro (2002), as they examine insurance contexts. These experiments show that moral hazard is less likely to occur under deterministic losses and low probability of compensation (amongst other circumstances).

Deductibles

To overcome the moral hazard problem, insurance companies have traditionally adopted deductibles to decrease the coverage of their clients (Winter, 2013). The deductible is the amount of damage that must be paid by the policyholder before the insurer will cover any expenses, which provides a financial incentive to reduce risk for the policyholder. In other words, the deductible reduces a policyholder's level of insurance coverage. Some studies used an experimental methodology to investigate insurance behavior under different levels of deductibles or insurance coverage. However, to the best of our knowledge, there is no previous research that examines the effect of different deductible levels on investment in risk reduction. Papon (2008) conducted an experiment on insurance demand with different levels of deductibles (full coverage, 10%, 30%, 50%, and no insurance) under low-probability risks and found that participants prefer extreme cases of coverage: No insurance or full insurance. Krieger and Felder (2013) conducted an experiment in the health insurance domain, where participants could select different levels of deductibles (full coverage, 20%, 30%, 40%, and 50%) under different types of information provision. The results indicate the presence of a status-quo bias in health insurance policies: Respondents chose their insurance policies based on the default offer. In a related laboratory experiment, Corcos et al. (2017) examined the demand for insurance coverage by presenting subjects with 20 equally-spaced deductible options, reaching from no insurance to full coverage. The results confirmed the bimodal pattern in flood insurance demand, with clear preferences for both extreme cases.

Financial incentives

In addition to deductibles, other financial incentives can be provided to stimulate damage-reduction investment by homeowners, such as premium discounts that reflect reduced damage due to policyholders investments in self-insurance (Kleindorfer et al., 2012; Poussin et al., 2014). Policymakers are increasingly using financial incentives to facilitate behavioral change in different domains of society, including health and financial decisions. However, recent research has shown that these incentives must be carefully designed to be effective (Patel et al., 2016; Hooker et al., 2018). Financial incentives have been used for decades in the insurance industry, but studies evaluating the effectiveness of these are relatively recent (Stevenson et al., 2018). This chapter contributes to the literature by evaluating the effectiveness of a premium discount and a mitigation loan on self-insurance in the context of disaster risk insurance. A premium discount serves as a financial reward for reducing potential damage, which is already common practice in health insurance (Tambor et al., 2016). Alternatively, low-interest mitigation loans may be provided by the government or other financial institutions to encourage investment in damage-reduction measures that have high upfront costs, such as flood proofing a house (Michel-Kerjan and Kunreuther, 2011). Loans spread the investment costs over time. This can encourage individuals with high discount rates (i.e., those who place more emphasis on immediate risk mitigation costs than on future risk mitigation benefits) to invest in damage reduction measures. We are not aware of any previous experimental work that directly tests the influence of these insurance incentives (premium discount and mitigation loan) on self-insurance investment.

This chapter advances the experimental literature on self-insurance by systematically studying the effects of different probability levels, deductibles, and other financial incentives on self-insurance investments. Moreover, to our knowledge, moral hazard has not been studied experimentally in relation to a variety of probability levels and deductibles. The current study aims to fill this gap by operationalizing investment in damage-reduction in a controlled lab experiment under different financial incentive treatments, starting from a baseline treatment without insurance and mitigation incentives. The results are likely to be useful for insurance companies and policymakers who aim to increase both insurance coverage and policyholder damage-reduction activities. Note that the dominant natural risk reduction strategy for individuals is self-insurance: One cannot prevent a flood or earthquake, but simple measures such as floodproofing may significantly decrease damage. Both theory and experiments have shown that policyholders respond differently to self-insurance than to self-protection (Ehrlich and Becker, 1972; Shogren, 1990). While most empirical papers concern self-protection, we cannot simply generalize these results to self-insurance. Rather, the drivers of self-insurance should be systematically examined; and this is an important contribution of the current chapter.

The remainder of this chapter is organized as follows: Section 2.2 describes the experimental design; Section 2.3 derives hypotheses for each of the treatments, based on simulations of a theoretical model; Section 2.4 presents results; Section 2.5 discusses policy implications; and Section 2.6 concludes.

2.2 Experimental design

We examined investment levels in damage-reduction under different financial incentives for mitigation of disaster risk. Participants were presented with six independent scenarios of an investment game under flood risk for multiple rounds. The experiment was framed in the context of insurance, thus all treatments (except "No Insurance") included a deductible.

The experiment consisted of several individual decision-making tasks, computerized in oTree (Chen et al., 2016). Earnings were in Experimental Currency Units (ECU) and converted back to euros at the end of the game. In the first stage, the initial endowment was earned and invested in a virtual house. As in Laury et al. (2009), participants were given a real effort task to earn this endowment, to overcome the "house money effect" (Thaler and Johnson, 1990). Participants were thus shown the prospect of losing rather than winning money (see Harrison and Rutström, 2008). One result of an earnings task in which initial earnings are determined by effort could be variability among subjects, with high performing subjects earning more than low performing subjects, leading to an unwanted stake effect (Dannenberg et al., 2012). Therefore, a new real effort task was developed in oTree,[2] in which participants were asked to collect ECU by clicking on a grid of 100 boxes which either contained money or did not. The money was randomly distributed by the software to 60 of the 100 boxes. When 30 boxes with money had been collected, the boxes were deactivated, such that all subjects finished with the same budget. To enhance a game-like situation, a timer was placed on the *Collect money* page, although there was no consequence of collecting quickly or slowly. (Screenshots of the new real effort task can be found on page 2 of the Online Supplementary Material.) After earning their starting capital, participants were asked to buy a virtual house (worth 240,000 ECU) with which to play the investment game. The remainder of the starting capital (75,000 ECU) was stored as "savings" and could be used to pay for investments, premiums, and damages. We explained to subjects that the house was prone to flood risk.

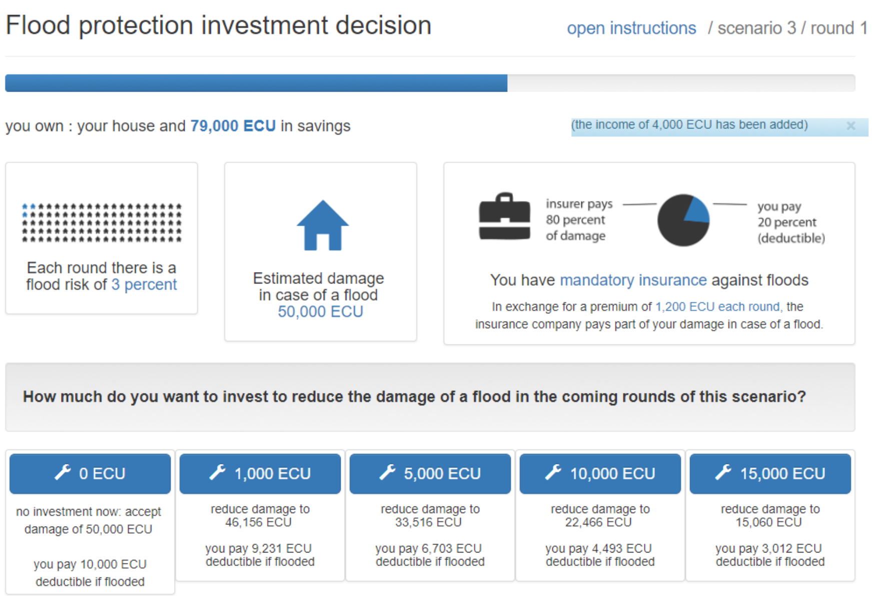


Figure 2.1: Investment decision screen in "Baseline Insurance" treatment.

[2] The task was based on the JavaScript code of the Bomb Risk Elicitation Task (Holzmeister and Pfurtscheller, 2016) with help of Mathijs Luger, a programmer of Vrije Universiteit Amsterdam.

2.2.1 Investment game

A scenario began with the introduction of the parameters: Flood probability, maximum damage, and deductible level. This lasted for 12 rounds. The sequence of pages in each round was *Invest*, *Pay premium*, then *Flood risk result*. The *Invest* page offered five discrete investment levels with accompanying benefits, as shown in Figure 2.1. Investments were effective for damage-reduction in all rounds of a scenario, beginning with the investment round. On the *Pay premium* page, subjects paid a fair premium (participants were price-takers). After each payment, the savings balance was adjusted accordingly. The *Flood risk result* page showed 100 houses, with the house of the participant indicated by a dotted square. The software selected the flooded house(s) at random, according to flood probability. The flooded house(s) was indicated in blue (see Figure 2.2). If a participant's house were flooded, the deductible (or damage, in the No Insurance treatment) was paid from the savings balance. After the *Flood risk result*, an income of 4000 ECU was added to the savings balance in each round. In each subsequent round, participants could either invest more or stay with the current investment (reducing the investment was not possible). Participants in the "Loan" treatment were offered a 1% interest loan to spread the investment costs over 10 rounds. When those participants chose a positive investment level, a *Pay loan cost* page was added between *Invest* and *Pay premium*. In the No Insurance treatment, the *Pay premium* page was skipped. The full experimental instructions can be found in the Online Supplementary Material.

The delivery of the instructions was followed by five rounds in a test scenario to ensure participants were familiar with the game. The instructions were available as a pop-up screen throughout the experiment. The test scenario was followed by comprehension questions. These questions were conditional on treatment and are listed in Appendix 2E. The answers could be retrieved from the (pop-up) instructions. The software kept track of the number of times a participant (re)opened the instructions, as well as the number of failed attempts to answer the comprehension questions. These were used as experimental control variables in the regression analysis. After answering the comprehension questions correctly, subjects began with the first scenario of the investment game.

2.2.2 Scenarios

Subjects played 6 different scenarios of 12 rounds each. Each of these scenarios contained a different combination of flood probabilities and deductibles. The order of the scenarios was randomly shuffled by the software and was saved to control for order effects. An overview of the scenarios is given in Table

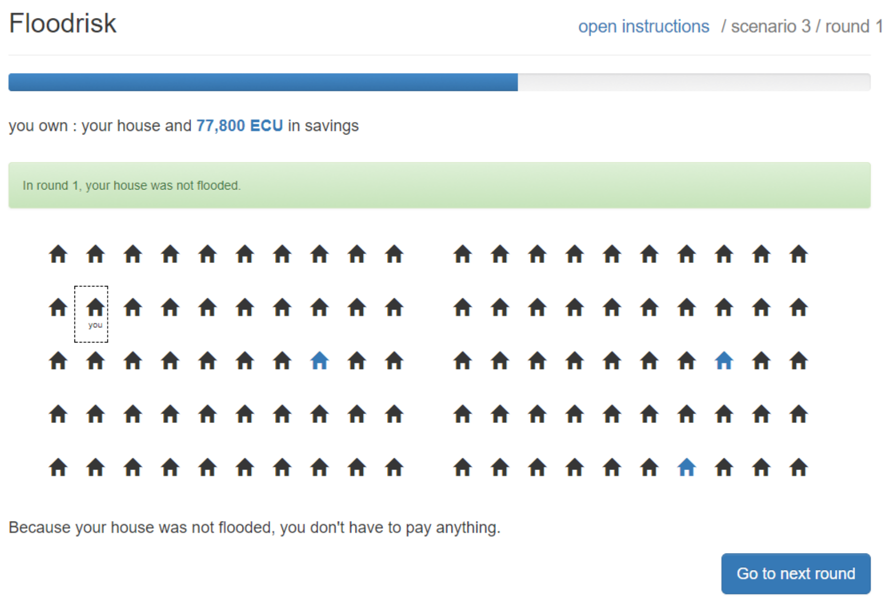


Figure 2.2: Flood risk screen under low probability. *Note:* Three houses are blue, indicating flooded, and participant is not flooded).

2.1. Participants were paid the final savings balance[3] of one randomly chosen scenario, at a conversion rate of 20,000 ECU = €1 (between €0 and €7 on top of the participation fee), and the independence of the scenarios was made salient by a pop-up screen at the start of each scenario. This screen also indicated the change since the previous scenario in flood probability, deductible, and premium. When a new scenario began, the savings balance was restored to the starting value of 75,000 ECU.

In addition to these payments, one participant was randomly selected from the full sample when all sessions had ended. This participant was rewarded with a large payment: His/her results in one random scenario or the additional time preferences task were paid at a conversion rate of 200 ECU = €1. The fact that each subject had a chance to earn up to €700 based on the results in the investment game was stated on all payment pages, thus highlighting the high stakes of the experiment. Figure 2.3 gives a schematic overview of the experiment.

[3] Savings balance = starting value (75,000 ECU) + income - premiums - deductibles - damages - investments.

Table 2.1: Overview of scenarios by treatment, deductible and probability

Treatment	Deductible	Probability					
		0.01	0.03	0.05	0.10	0.15	0.20
No Insurance	1.00	1%	L-	5%	10%	H-	20%
Baseline Insurance	0.05	*n.a.*	LxL	*n.a.*	*n.a.*	HxL	*n.a.*
	0.15	*n.a.*	LL	*n.a.*	*n.a.*	HL	*n.a.*
	0.20	*n.a.*	LH	*n.a.*	*n.a.*	HH	*n.a.*
Premium Discount	0.05	*n.a.*	LxL	*n.a.*	*n.a.*	HxL	*n.a.*
	0.15	*n.a.*	LL	*n.a.*	*n.a.*	HL	*n.a.*
	0.20	*n.a.*	LH	*n.a.*	*n.a.*	HH	*n.a.*
Loan	0.05	*n.a.*	LxL	*n.a.*	*n.a.*	HxL	*n.a.*
	0.15	*n.a.*	LL	*n.a.*	*n.a.*	HL	*n.a.*
	0.20	*n.a.*	LH	*n.a.*	*n.a.*	HH	*n.a.*
Loan+Discount	0.05	*n.a.*	LxL	*n.a.*	*n.a.*	HxL	*n.a.*
	0.15	*n.a.*	LL	*n.a.*	*n.a.*	HL	*n.a.*
	0.20	*n.a.*	LH	*n.a.*	*n.a.*	HH	*n.a.*

Notes: Initial wealth = 75,000; Maximum damage = 50,000; Interest rate Loan = 1%; Nr of installments in Loan treatments = 10; Premium = (1 - Deductible) × Probability × Damage; *n.a.* = not applicable.

2.2.3 Treatments

Participants were randomly distributed over five treatments: No Insurance ($n = 60$), Baseline Insurance ($n = 120$), Premium Discount ($n = 59$), Loan ($n = 60$) and Loan+Discount ($n = 58$). The relation between treatments and our hypotheses is explained in Section 2.3.2 and in more detail in Appendix 2D. Baseline Insurance included only a deductible and served therefore as the baseline mandatory insurance treatment. As we expected the highest variability in this treatment, we doubled the number of subjects allocated to it.[4] In the Premium Discount treatment, a premium discount was offered to participants if they invested in damage-reducing measures, proportional to the estimated damage-reduction. To overcome the effects of time-discounting, the Loan treatment offered the participants a loan to spread the costs of investment over multiple rounds. The final treatment, Loan+Discount, was a combination of the previous two, including both the premium discount and the mitigation loan. The advantage of this combination is that it makes the cost-effectiveness of the measures very salient when the annual premium discounts exceed the annual loan cost.

[4] As we introduced a novel design, we had no priors regarding effect sizes to perform a power analysis.

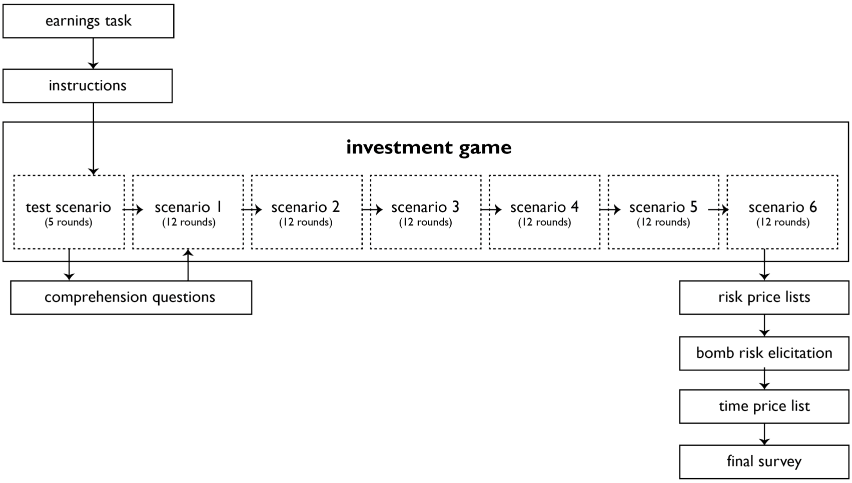


Figure 2.3: Schematic overview of the experiment

2.2.4 Extra tasks

Following the experiment, there were a set of questions and decision-tasks to gather data on risk preferences, time preferences, and other behavioral characteristics that could be related to the investment decisions. Risk preferences were measured using two price lists and the Bomb Risk Elicitation Task (BRET) (Crosetto and Filippin, 2013). Based on a recent review on risk-elicitation tasks (Csermely and Rabas, 2016), we used the new price list proposed by Drichoutis and Lusk (2016) and did not include the original (Holt and Laury, 2002). In this new iteration, probabilities are held constant at 0.50 and the payoff amounts are varied. This method seems to perform well in forecast accuracy and is relatively simple. The same price list was adapted from Drichoutis and Lusk (2016) and framed in the loss domain. In this task, subjects were first endowed with the maximum possible loss (€4.70) and the outcomes of the lotteries were negative. In both price lists, subjects were prevented by the oTree software from switching more than once between options (Holzmeister, 2017): All rows were shown on the screen simultaneously (see screenshots in the Online Supplementary Material). Finally, a *static* version of the BRET by Holzmeister and Pfurtscheller (2016) was played once. This contained 100 boxes, each worth €0.05, and one bomb. Subjects were asked to choose a total number of boxes, which were then picked at random and opened by the software. The total value of the opened boxes was earned by the subject, unless the bomb was among them, which would lead to a payoff of zero. To prevent income effects, the software selected at random one of the tasks for the payment at the end of the three risk-elicitation tasks.[5] The results of the selected task were shown on the screen and the earnings saved for payment. For the time preferences, we used the price list of the Preference Module by Falk et al. (2016), where subjects had to choose 25 times between an immediate payment of €100 and a delayed payment in 12 months. The delayed payment ranged from €100 to €185. Again, consistency was enforced by the software. After the time preferences, one task was selected for the large payment: One of the six scenarios or the result of the time preferences task. Note that the time preferences task was thus only incentivized by the large payment; both ‘immediate’ and delayed time preferences payments would be paid by bank transfer, which resulted in a front-end delay with constant transaction costs. A summary of the payments (participation fee, investment game, and risk-elicitation task) was given on the next page. At the end of the experiment, subjects were presented with risk preferences questions and some additional questions (e.g., beliefs regarding flood risk). The coding of the questions can be found in Appendix 2B.

2.2.5 Procedure

To test the instructions for the newly developed investment game, a pilot experiment was carried out with Master’s students in October 2017. Subjects were sent a link through which they could play the game online on their own laptop or desktop computer. The pilot experiment was made available on the server for one week. All participants were paid according to their performance in the game by bank transfer, one week after the pilot. To keep incentives equal for the pilot and the experiment, all pilot students were eligible for the large payment. The payment structure was explained verbally in one of the lectures and again in the invitation e-mail. In total, 20 students took part in the pilot experiment. They earned an average of approximately €12.00 in 34 minutes. We were mostly interested in testing the procedure and the average time required to finish the game. The pilot students finished faster than expected, and many invested in all scenarios. To increase heterogeneity in investment decisions across subjects, we added two scenarios to the game with an extra low deductible and two more risk levels in the No Insurance treatment. To test the length of the final procedure, a second pilot was conducted with five PhD students in our institute. No major changes were made after the second pilot.

The experiment was conducted in the CREED lab of the University of Amsterdam in November 2017. A total of 361 participants earned an average of €12.95 in 29 minutes. We conducted 11 sessions in 4 days. Note that subjects were randomly assigned to a treatment by the software; hence different treatments were played during one experimental session. Three subjects

[5] Subjects were informed about this procedure before the start of the first risk-elicitation task, which was introduced together with the others as ‘additional tasks’.

participated twice due to a minor error with the subject database. The results of their second experiment were removed from the analysis. One result was incomplete, as this subject did not finish the final survey, and the result was thus removed. This left 357 observations for analysis. All earnings - except the large payment, which included the time preferences payment - were paid out privately, in cash, immediately after the experiment. The large payment was arranged via bank transfer, after all sessions had ended.[6]

2.3 Theory and hypotheses

Based on the previous literature referred to in Section 2.1, we developed several hypotheses, which we then tested in the lab experiment. The parameters of the experiment were based on simulations of a theoretical model, as described in Appendix 2C.

2.3.1 Simulations

We used a comparative statics approach to predict best responses to the simplest hypothesis (a comparison between Baseline Insurance and No Insurance), reported in Appendix 2C. However, no clear-cut analytical solution was found for the other hypotheses. Therefore, we predicted the best response of risk-averse (versus neutral, seeking) and low (versus high) time-discounting individuals investing in self-insurance under each treatment based on simulations of the theory. We used these simulations to set our experimental parameters, such that all hypotheses could be tested with the lab experiment. The results of these simulations, which are based on Equation C2, are reported in Appendix 2D. The final set of parameters includes initial wealth $W = 75{,}000$, maximum loss $V = 50{,}000$, effectiveness of self-insurance $\beta = 0.00008$, number of installments in Loan treatment = 10 and interest rate = 1%. The following section provides the hypotheses and the intuition behind them.

2.3.2 Hypotheses

From the comparative statics in Appendix 2C, we know that investments under insurance coverage (Baseline Insurance) should be lower than without coverage (No Insurance). In general, Winter (2013) states that even though moral hazard is considered as a major issue in insurance from a theoretical perspective, empirical results are mixed. An overview of empirical studies on moral hazard has been carried out by Cohen and Siegelman (2010). The authors conclude that the existence of moral hazard is largely dependent

[6] Large earnings ranged from €86.70 to €615. The randomly selected participant earned €196.49 from one of the scenarios. The payment was thus made immediately and not delayed by 12 months, which could have happened if the time preferences payment had been selected.

on the type of insurance market. In survey studies, moral hazard has been found to play only a minor role in voluntary flood insurance markets (Hudson et al., 2017; Thieken et al., 2006). Therefore, the first hypothesis concerns the role of moral hazard in the flood risk insurance context. In simulations of the theory (Appendix 2D), damage-reduction investments in the Baseline Insurance treatment are lower than in the No Insurance treatment. Positive investments in the Baseline Insurance treatment may be optimal in high-probability scenarios, depending on the deductible level and attitude to risk.

Hypothesis 2.1 *Damage-reduction investments in the Baseline Insurance treatment are lower than in the No Insurance treatment, but greater than zero.*

In line with risk-based insurance premiums, researchers (Kunreuther, 1996; Surminski et al., 2015) and policymakers (European Commission, 2013) have suggested that a premium discount may motivate policyholders to take mitigation measures. So far, there is little empirical evidence of the effectiveness of premium discounts, except for the findings of Botzen et al. (2009b), which concern the willingness of a large sample of Dutch homeowners in floodplains to pay for low-cost flood-mitigation measures. The researchers found that the main incentive for investment was the premium discount on the flood insurance policy that was offered in the survey. The following hypothesis therefore concerns the Premium Discount treatment. The simulations in Appendix 2D show that damage-reduction investments should be higher in the Premium Discount treatment than in the Baseline Insurance treatment, under all scenarios and risk attitudes.

Hypothesis 2.2a *Damage-reduction investments are higher in the Premium Discount treatment than in the Baseline Insurance treatment.*

A second financial incentive to promote policyholder damage-reduction measures is a mitigation loan or a payment in installments (Michel-Kerjan, 2010), aimed at individuals who heavily discount the future. This treatment could overcome both high time-discounting and a moral hazard effect. The Loan+Discount treatment could be powerful, assuming that a considerable share of individuals is risk-averse and present-oriented. Therefore, we expect that the combination of incentives will lead to the largest damage-reduction investment. The simulations in Appendix 2D indicate that Loan+Discount gives the highest optimal investments for all treatments in the low-probability scenarios.

Hypothesis 2.2b *Damage-reduction investments are largest in the Loan+Discount treatment.*

Policyholder damage-reduction measures may be cost-effective under expected utility theory (Kreibich et al., 2015), but myopic individuals with high discount rates weigh the present costs much more heavily than the projected future benefits. Damage-reduction investments are lower in the Baseline Insurance and Premium Discount treatments under high time-discounting, according to our simulations. A mitigation loan may overcome this discounting effect by spreading the costs over multiple periods. The simulations indeed show that in the Loan and Loan+Discount treatments, time-discounting has no effect on damage-reduction investment.

Hypothesis 2.3a *Damage-reduction investments are lower among participants with high time discount rates. This effect is strongest in the Baseline Insurance and Premium Discount treatments, and it disappears in the Loan and Loan+Discount treatments.*

Hudson et al. (2017) argue that in natural disaster markets, decisions are mainly driven by risk attitudes, where highly risk-averse individuals take multiple precautionary measures, including flood insurance and flood damage-reduction measures. In this scenario, advantageous selection may prevail over the moral hazard effect, which may be explained by a misunderstanding of risk (Kunreuther and Pauly, 2004). However, Hudson et al. (2017) did not examine the behavioral mechanisms to back up their claim. The current experiment aims to fill that gap. The simulations show that risk-seeking individuals will not invest in the Baseline Insurance and Loan treatments, while investing 1000 or 5000 could be optimal for risk-neutral individuals and 10,000 for risk-averse individuals.

Hypothesis 2.3b *Risk-averse individuals will invest more in damage-reduction in the Baseline Insurance treatment and the Loan treatment than risk-neutral individuals will, while risk-seeking individuals will invest less.*

2.4 Results

This section reports our results, beginning with the moral hazard effect (Hypothesis 2.1) and the effect of financial incentives related to insurance (loan and premium discount, Hypotheses 2.2a and 2.2b) with non-parametric tests and a multivariate regression analysis. Subsequently, we examine the effect of time and risk preferences on investment behavior (Hypotheses 2.3a and 2.3b). Finally, we present some additional analyses, including a trend analysis and the effects of flood beliefs on investment behavior. We conclude with an overview of the predicted margins of our key findings, comparing investments in self-insurance under different loss probabilities, deductibles, and financial incentives.

Table 2.2 displays descriptive statistics of the demographic variables that should not be influenced by our experimental treatments. Demographic variables did not significantly vary between treatment groups.[7] We further analyzed the balance of the flood perception variables efficacy, worry, and regret across treatments, which were measured in a post-experimental survey and could be affected by different versions of the investment game.[8] Precise coding of the variables can be found in Appendix 2B.

Table 2.2: Descriptive statistics per treatment group

	No Insurance	Baseline	Discount	Loan	Loan+Discount	p-value
Age in years	21.05	21.89	21.39	21.17	21.48	0.593
	(2.22)	(4.82)	(2.33)	(3.24)	(3.60)	
Gender	0.58	0.52	0.49	0.50	0.67	0.264
	(0.50)	(0.50)	(0.50)	(0.50)	(0.48)	
High income	0.05	0.03	0.09	0.07	0.02	0.364
	(0.22)	(0.18)	(0.29)	(0.25)	(0.13)	
Risk averse	5.65	5.83	5.79	5.82	5.81	0.932
	(1.30)	(1.14)	(1.34)	(1.36)	(1.39)	
Present biased	13.49	14.02	12.39	13.05	12.70	0.726
	(7.80)	(8.17)	(8.19)	(8.05)	(8.62)	
Observations	59	121	57	60	60	

Note: Table displays means, SD in parentheses. Final column presents the p-value for an F-test of the null hypothesis of equal means across treatment groups. Baseline = Insurance Baseline. Gender dummy: 1 indicates female. High income dummy: 1 indicates > €5000.

2.4.1 Testing the moral hazard effect

To test Hypothesis 2.1, we compared the investment levels in the Baseline Insurance treatment with those in the No Insurance treatment. We began with an analysis of the most independent unit of observation: The first round. A one-sided t-test revealed that the average investment in the first round of Baseline Insurance was significantly higher than 0, both in the high-probability scenario ($M_{BaselineHL} = 4049.59$, $t = 9.20$, $df = 120$, $p < 0.0000$) and in the low-probability scenario ($M_{BaselineLL} = 2404.96$, $t = 6.22$, $df = 120$, $p < 0.0000$). Figure 2.4 shows the average investments in the first round

[7] Note, however, that the benefit of balancing checks after experimental randomization is debatable (see e.g. Mutz and Pemantle (2015) or Deaton and Cartwright (2018) for recent discussions).

[8] Significant differences were found for efficacy of protection ($p = 0.008$) and regret about investment ($p = 0.000$), but not for worry and regret about no investment. Participants in the Discount treatments reported higher efficacy values, which may be caused by a positive experience of mitigation measures due to the financial benefit of the premium discount. Furthermore, participants reported lower regret values in case of investment without a flood event in the game. This finding is consistent with the design of the Discount treatment, where participants received benefits (namely, premium discounts) of their self-insurance investments regardless of flood events.

in Baseline Insurance (gray boxes) and No Insurance (black boxes), split by probability and deductible levels (shade of gray). Note that the No Insurance treatment is equivalent to a 100% deductible.

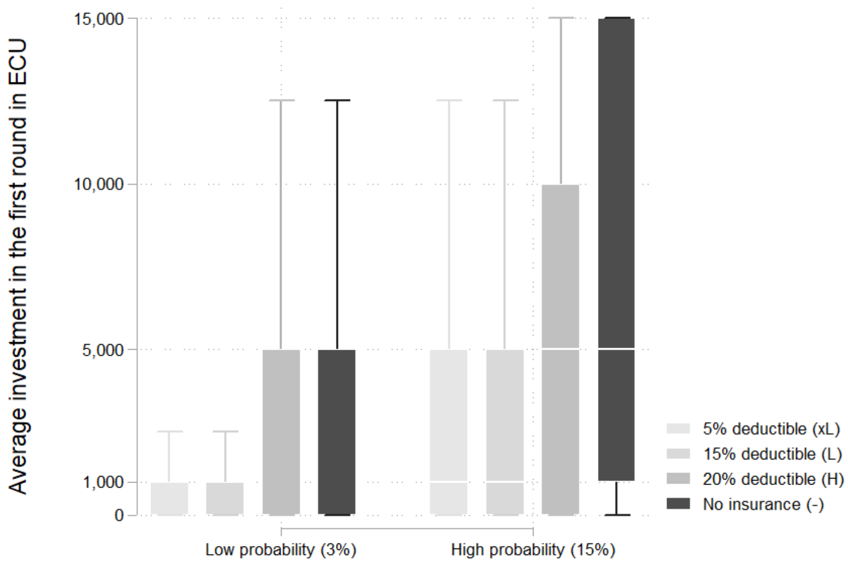


Figure 2.4: Boxplots of investments in the first round, by probability and deductible. Boxplot whiskers indicate the inter-quartile range, middle lines represent medians.

Table 2.3 shows the average investment in the first round, by treatment.[9] Significant differences between investments in Baseline Insurance and No Insurance are indicated by asterisks in the third column of the table (non-parametric Mann-Whitney-Wilcoxon (MWW) tests). The results show significant differences for the high-probability scenarios, indicating a moral hazard effect: Subjects invest less in damage-reduction when insurance is available and probabilities are high. However, we do not observe such a strong effect in the low-probability scenarios. Only in the scenario with the smallest deductible (5%) do subjects invest slightly less than in a scenario without insurance ($p < 0.1$).

To test Hypothesis 2.1 over all 12 rounds of the investment game, we ran panel regressions with scenario dummies and controls. We opted for a random effects ML specification[10] to control for subject and scenario effects.

[9] Note that both Table 2.3 and Figure 2.4 illustrate investments in the first round in ECU. However, Table 2.3 presents means, while Figure 2.4 presents medians.

[10] To control for unobservable subject-specific and scenario-specific effects, we created subject-scenario dummies and used these to cluster standard errors. The random effects

Table 2.3: Average investment in the first round in ECU

	No Insurance	Baseline	Discount	Loan	Loan+Discount
scenario H-	7,288.14				
	(5717.64)				
scenario HH		5,421.49**	9,233.33***	3,816.67	8,614.04***
		(5,431.01)	(5,732.35)	(3716.62)	(5,512.18)
scenario HL		4,049.59***	8,416.67***	3,050.00	7,807.02***
		(4,843.98)	(5,681.64)	(4,188.06)	(5,717.89)
scenario HxL		3,471.07***	8,966.67***	3,500.00	7,771.93***
		(5,010.11)	(5,971.59)	(5,000.00)	(5,840.19)
scenario L-	2,711.86				
	(4,102.36)				
scenario LH		2,727.27	3,850.00**	1,883.33	3,719.30
		(4,222.95)	(4,398.86)	(3,796.04)	(4,806.08)
scenario LL		2,404.96	3,283.33*	1,750.00	3,421.05
		(4,253.58)	(4,584.76)	(4,015.33)	(5,119.81)
scenario LxL		1,793.39*	3,550.00***	1,633.33	2,087.72
		(3,976.84)	(4,560.05)	(3,723.34)	(3,434.49)
Observations	59	121	60	60	57

Note: Table reports means, st.dev in parentheses. Asterisks in the Baseline Insurance column indicate significant differences with the No Insurance treatment. Asterisks in last three columns indicate significant differences with the Baseline Insurance treatment (MMW tests, * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$).

All explanatory variables were checked for high correlations to rule out issues of multicollinearity. As all correlation coefficients were smaller than 0.5, multicollinearity is not regarded as problematic (Field, 2009). The dependent variable is the log-transformed[11] damage-reducing investment. Table 2.4 restricts the sample to the Baseline Insurance and No Insurance treatments. The results show the same pattern as in the non-parametric tests. In the high-probability scenarios (15%), we find significantly less investment in damage-reduction when insurance is available under all deductible levels. In the low-probability scenarios, we only find lower investments when the deductible is particularly small (5%). The regression results confirm that there is no moral hazard effect in the low-probability scenarios (3%), under low (15%) or high (20%) deductible levels. The negative and significant estimates for order of scenario indicate that damage-reducing investment declines with experience. Note that the order of scenarios was determined at random by the software.

Overall, we find mixed support for Hypothesis 2.1. There is no significant difference between investments in the No Insurance and Baseline Insurance treatments in the low-probability scenario, which suggests that there is no moral hazard in an insurance market where probabilities are low and expected

ML estimates are not conditional on subject and time effects to account for clustered standard errors per subject and scenario (see e.g. Bell and Jones, 2015).

[11] We used the transformation $transformed = log(investment + 1)$ to deal with 0 investments.

damages are high, while moral hazard might occur with increasing probabilities of damage. The latter finding is in line with previous literature on moral hazard in different insurance markets (Cohen and Siegelman, 2010). Under low probabilities and high expected damages, a substantial share of the "cautious" types might decide to insure and invest in damage-reducing investments. Note, however, that the probability information in this experiment was objective information.

Table 2.4: Random effects ML panel regression estimates of investments

	Probability L: 3%	Probability H: 15%
Deductible (ref. No Insurance)		
H: 20%	-0.171	-1.089*
	(0.561)	(0.562)
L: 15%	-0.501	-1.894***
	(0.561)	(0.562)
xL: 5%	-1.611***	-3.182***
	(0.561)	(0.563)
Order of scenario	-0.562***	-0.227**
	(0.103)	(0.100)
Constant	3.083*	4.681***
	(1.780)	(1.778)
σ_u	3.332***	3.342***
	(0.121)	(0.122)
σ_e	0.989***	0.946***
	(0.011)	(0.010)
Observations	4596	4596
Nr of subjects	163	163
AIC	14,867	14,488
Log likelihood	-7,415	-7,225

Notes: Standard errors clustered by id and scenario in parentheses (* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$). Controls: age, gender, high income, understanding, perceived difficulty, flood risk perception, risk aversion, time preferences, worry, perceived efficacy, regret, 1/round. Dependent variable log-transformed.

2.4.2 Financial incentives to increase self-insurance

Next, we investigated the effect of financial incentives related to insurance on investments in damage-reduction. Hypothesis 2.2a concerns the effect of the Premium Discount treatment. Table 2.3 shows non-parametrically for round

1 that subjects invest significantly more in the Premium Discount treatment than in the Baseline Insurance treatment, regardless of risk and deductible levels. Table 2.5 presents the results of a random-effects panel regression ML estimates, which takes all rounds into account, as well as treatment dummies, scenario dummies, demographics, and various controls. We chose a panel specification to account for the correlation of decisions by the same subject and clustered standard errors by id (subject) and scenario. All models control for (1) attempts to answer understanding questions,[12] (2) perceived difficulty, (3) flood risk perception, (4) one over round to control for experience, and (5) order of scenario $\times$ probability interaction; but coefficients have been suppressed for brevity. The positive coefficients of the Premium Discount treatment confirm the results of the non-parametric analysis: A premium discount leads to larger investment. This effect is large and statistically significant under all possible controls. We can therefore confirm Hypothesis 2.2a: A premium discount leads to larger damage-reduction investment, compared to a baseline insurance situation.

The Loan treatment, however, does not encourage subjects to invest more in damage-reduction. Neither the non-parametric analysis in Table 2.3, nor the multivariate regression analysis in Table 2.5 reveal a significant effect of the Loan treatment, compared to the Baseline Insurance treatment. We expected a positive investment effect for the Loan+Discount treatment (Hypothesis 2.2b). In that case, the economic return on the loan was salient on the decision screen, because cost effective investments show lower annual costs than benefits in terms of the premium discount. Average investment in the first round in the Loan+Discount treatment, as displayed in Table 2.3 is lower than in the Premium Discount treatment in almost all scenarios. These results are confirmed by the negative insignificant estimates of the Loan $\times$ Discount dummy in Table 2.5 after controlling for Premium Discount only. Hypothesis 2.2b thus finds no support in the data.

Our findings could be explained by the dislike for the mandatory 1% interest in the Loan treatment, or a general dislike of lending among the students in our sample. Alternatively, one could argue that the operationalization of a Loan treatment in the lab lacks external validity,[13] as the investment costs are spread over 12 rounds, ranging from seconds to minutes in the lab, rather than years, as in the real world. However, incorporating true intertemporal payoffs would require a complicated experimental design, in which subjects were to return to the lab to pay back their loans. We considered this impossible to

[12] One subject attempted the comprehension questions more than 10 times. For robustness, we re-ran all analyses excluding this subject. The results do not change qualitatively.

[13] Note that lab experiments are in general low in external validity, although we did all we could to increase external validity: An engaging task explained with parameters based on real data, an incentive compatible payment scheme and a high stakes random lottery incentive mechanism to mimic the large consequences of flood risk investment decisions.

enforce. Further research on loans in the context of disaster risk reduction should therefore focus on field rather than lab experiments.

2.4.3 The effect of time and risk preferences

To examine our last two hypotheses, we use the multivariate regression analysis reported in Table 2.5. We find no effect of time-discounting on investments,[14] suggesting no support for Hypothesis 2.3a.

The risk-aversion variable is a linear combination of our four risk-elicitation methods,[15] as in Menkhoff and Sakha (2017). We find that risk-averse subjects invest more in damage-reducing investments, providing evidence for Hypothesis 2.3b. Table F1 provides additional robustness checks for each of the four risk-elicitation methods. The direction of the risk aversion effect is equal for all elicitation methods and the estimates of other variables do not change qualitatively.

2.4.4 Additional results

In addition to the evaluation of our hypotheses, some other interesting patterns emerge from our data. Model 2 in Table 2.5 includes three control variables that varied between rounds: Participant flooded in the previous round, direct neighbors (see Figure 2.5) flooded in the previous round, and decision time in seconds at the Invest screen. The positive and significant estimate for decision time shows that investments are greater when subjects spend more time on the Invest page.

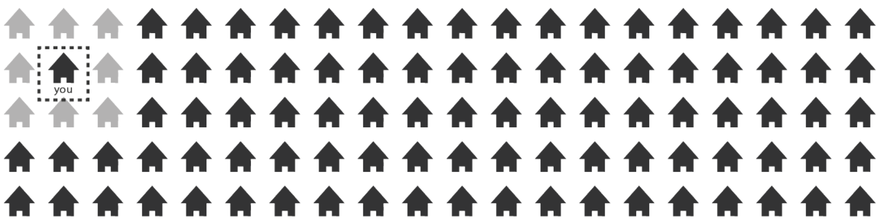


Figure 2.5: Grey color indicates direct neighbors for construction of neighbors variable

This effect may be explained by the decisions in the first round requiring some deliberation, while subjects learn to move quickly to the next page without extra investments in later rounds. The neighbor variable was constructed to control for erroneous impressions of spatial correlations between

[14] We have included an interaction term of time-discounting × Loan, but the results were not statistically significant.
[15] See Section 2.2.4 for a description of these tasks.

floods in the game. Both participant- and neighbor-flooded variables are not significant. Note that the dependent variable here is log-transformed investment, which may not differ substantially between rounds. In Appendix 2F, we specifically analyze 'extra investments' and find that subjects invest extra in damage-reduction after experiencing floods themselves, but not when a neighbor has been flooded in the game.

Model 3 includes demographic variables. All else being equal, including risk aversion, we find that investments decrease slightly with age, that women invest significantly more than men, and that subjects with a high income in real life invest less in damage-reduction in the game. In Model 4, we further include variables concerning flood beliefs. We observe significant and positive coefficients of believed efficacy of protective measures and worry about flooding. Note that the flood belief variables may be driven by some underlying characteristics that drive both beliefs and investments, which could potentially violate the assumption of strict exogeneity of explanatory variables. A significant negative estimate is seen for regret of investment, a question asked at the end of the experiment.

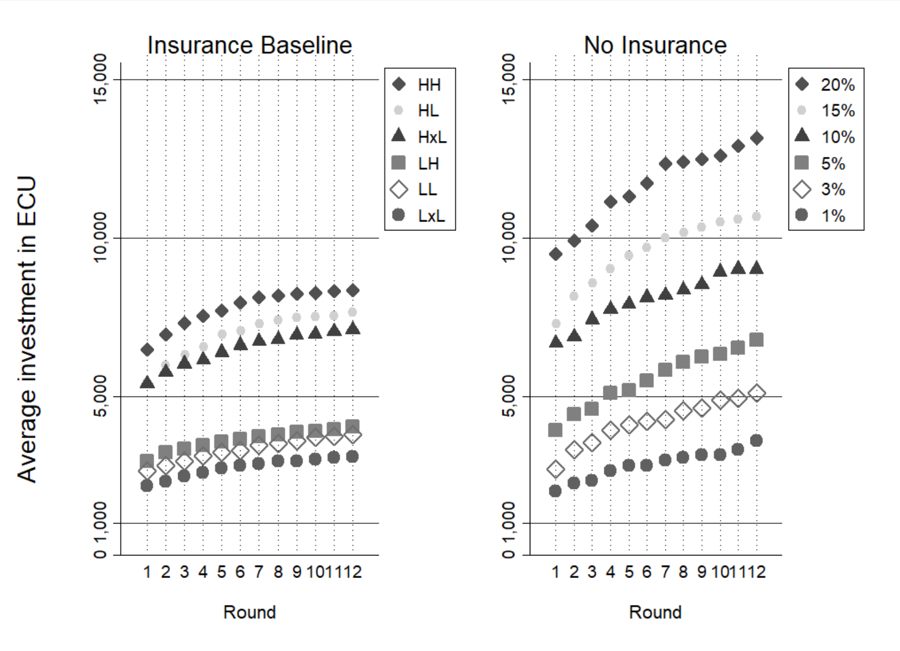


Figure 2.6: Average investment in damage-reducing measures by scenario

Figure 2.6 shows the average damage-reducing investments per round and scenario of all subjects in the Baseline Insurance and No Insurance treatments. It is no surprise that investments do not decrease, as this was not an option for subjects during a scenario. Note that investments were effective for all subsequent rounds: Investing in the first round leads to the highest expected benefits over all rounds. Still, average investments increase through the rounds, with the largest increase in the high-probability treatments of the No Insurance treatment. This can be explained by a small share of individuals who initially invest little and realize during the game that they want more protection, following the experience of a flood (see Appendix 2F). In our initial design, the No Insurance treatment contained only two scenarios (H- 15% probability and L- 3% probability), where all other treatments tested six scenarios. To keep the workload for all participants approximately equal, we added four scenarios to the No Insurance treatment to study the effect of expected value of flood losses on investments with a more refined pattern of probabilities. Figure 2.6 also shows that subjects did invest more when the expected value of a loss increased (i.e., higher deductible and/or higher probability). These extra probability scenarios in the No Insurance treatment are not included in any of the other analyses.

2.4.5 Predicted margins

Finally, Figure 2.7 summarizes our key findings with regards to the effects of probabilities, deductibles and financial incentives for self-insurance investments. It shows the adjusted predicted margins at the 95% confidence level of a log-transformed OLS regression of interactions between probabilities, deductibles, and treatments in the first round. For readability, the null-effect of the Loan treatment is not displayed. The graph further facilitates comparison of effect sizes. For example, adding a premium discount in the low-probability scenarios leads to a similar increase in self-insurance investments as that seen when increasing the probability of loss from 1% to 5%.

Loss probabilities

The black diamond markers in Figure 2.7 show that respondents invested more in self-insurance when they were confronted with a higher probability of loss, confirming the results of Figure 2.6. However, the increase in investment is not proportional to the increase in loss probabilities, which is in line with experimental work on the relationship between probabilities and self-protection investments (Shafran, 2011; Ozdemir, 2017).

Moral hazard

The graph further illustrates the mixed findings around the moral hazard problem. In the high-probability scenarios, we find evidence for moral hazard: Self-insurance investments are significantly lower in the Baseline treatment (indicated with gray triangles) than the No Insurance treatment

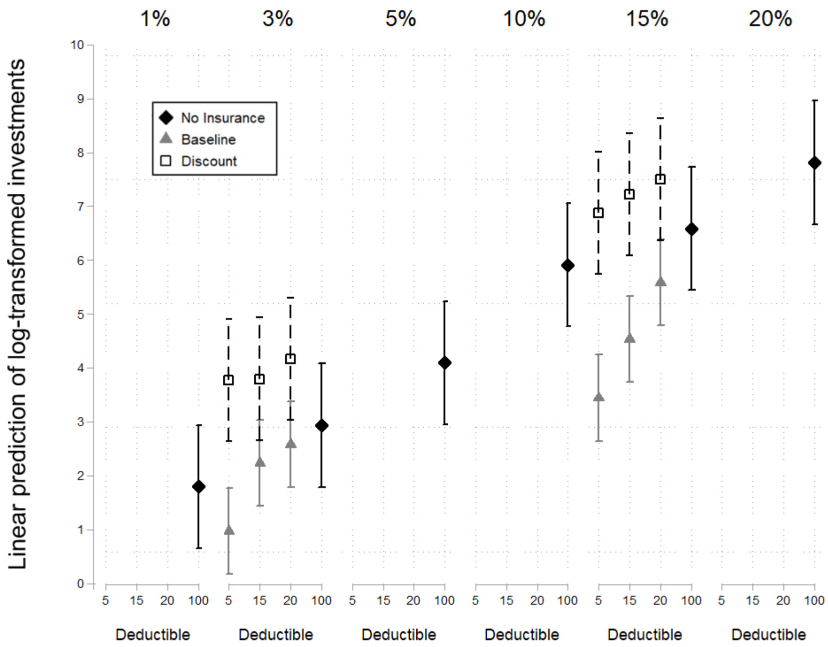


Figure 2.7: Adjusted predictions of log-transformed investments in the first round by treatment, deductible, and probability of loss. Error bars indicate 95% confidence intervals. We used the *marginsplot* command in Stata to create this figure.

(indicated with black diamonds). The only significant difference in the low-probability scenarios, however, is under the lowest deductible. In other words, a large deductible (at least 15%) may alleviate the moral hazard problem in a low-probability/high-impact context. This finding validates the empirical conjecture that moral hazard is absent in low-probability/high-impact insurance markets (Thieken et al., 2006; Hudson et al., 2017).

Deductibles

The effect of deductibles is represented in Figure 2.7 on the x-axis of each subplot. The figure shows that, in line with theoretical predictions, increasing the deductible leads to slightly higher investments in self-insurance. We thus find support for the substitution hypothesis of Carson et al. (2013), which theorizes that insurance and mitigation may be substitute goods. The deductible effect is smallest in the low-probability (3%) scenarios, which confirms previous survey research in natural disaster insurance markets (Hudson et al., 2017).

Financial incentives

Figure 2.7 shows that a premium discount (indicated with white squares) can significantly increase investment in self-insurance, although the effect is largest under high probability of loss and low levels of deductibles. Note that the premium discount is based on the expected value of damage-reduction, leading to a larger premium discount in absolute terms in the high-probability scenarios. The finding that a premium discount can be effective in increasing self-insurance investments even under low probabilities of loss, confirms previous empirical studies (Botzen et al., 2009b; Hudson et al., 2016).

2.5 Implications for disaster risk management

Both the effects of climate change and ongoing socio-economic development in floodplains are contributing to the projected increase in flood damage (Jongman et al., 2014). Floods are one of the costliest extreme weather events worldwide, with more than 26 billion US dollars in losses in 2017 (Munich RE, 2018). Flood risk insurance is often mandatory or at least heavily regulated when provided by private insurers. The implementation of mandatory insurance in our experiment closely resembles the characteristics of many natural disaster insurance markets (Paudel et al., 2012), for which it is impossible to distill moral hazard by survey and market data because a control group without insurance coverage does not exist in practice. Our experiment investigated the effect of deductibles, financial incentives, and time and risk preferences on private investments for reducing disaster risk damage. These investments can be taken by individual homeowners and are cost-effective in reducing flood risk (Poussin et al., 2015; Kreibich et al., 2011). While the estimated prevented damage can be substantial (Kreibich et al., 2015), only a small proportion of homeowners has currently taken these measures.

Our results reveal why current voluntary take-up rates of damage mitigation measures are low and how they might be improved. For example, policyholders should be well-informed about cost-effective ways of reducing damage. Furthermore, appeals to negative feelings about flooding (in terms of worry) may stimulate investment in flood damage mitigation measures. Although deductibles have a significant impact on damage-reduction, the size of this effect is not very large, which draws into question the effectiveness of high deductibles for stimulating policyholder flood risk reduction activities. Moreover, our finding that moral hazard effects are minor when probabilities of damage are low suggests that there is less need for high deductibles to limit such an effect. Premium discounts are likely to be a more effective way of stimulating policyholders to reduce flood risk.

In the face of increasing disaster risk, such as climate change, it is important to understand individual preparedness and risk-reduction activities. In our No Insurance treatment, we systematically varied the yearly probability of loss in six scenarios, from 1% to 20%. The results show that damage-reducing investment increases with loss probability, but less than proportionately. Hence, there is a need to improve individual preparedness in the face of increasing disaster risk. Experiencing a flood in the game triggers extra investment in flood damage mitigation measures. It is more beneficial for people to take such measures before a flood, rather than after, which highlights the need to explore the effectiveness of incentives that motivate people to reduce risk ex ante flood events. Future work could examine the behavior of homeowners in floodplains, who might respond differently due to their greater experience with insurance and possibly flooding than the current student sample.

2.6 Conclusion

With economic losses due to natural disasters expected to increase, it is important to study risk reduction strategies, including individual investments of homeowners in damage-reducing (mitigation) measures. Different options exist for policyholders to reduce risk, including self-insurance and self-protection. While there is an extensive literature on the empirical regularities related to insurance demand and self-protection, research on the drivers of self-insurance is limited. This chapter contributes to the discussion by investigating the relevant dimensions of heterogeneity of self-insurance under compulsory insurance coverage for low-probability/high-impact risk. These characteristics include probability levels, deductibles, and other financial incentives, which cannot be varied systematically in actual insurance markets. A new investment game was developed to study the causal relationship between financial incentives related to insurance and self-insurance investments, taking into account behavioral characteristics of individuals in an insurance market with mandatory coverage.

We found that subjects invested more when the expected value of a loss increased (higher deductible and/or higher probability of loss), although this increase in investment was not proportional to the increase in risk. Furthermore, we identified that the investments in the No Insurance treatment were significantly higher than in the Baseline Insurance treatment for the high-probability (15%) scenarios, but not significantly different in most low-probability (3%) scenarios. Mean investments in Baseline Insurance were greater than zero, confirming our conjecture that moral hazard is less of a problem in an insurance market where probabilities of damage are low and expected damages are high. Regarding financial incentives for damage-reduction, our results indicate that a premium discount can increase investment in damage-reduction, while the availability of a mitigation loan does not increase investments. Behavioral characteristics that have a positive effect on these investments are risk aversion, perceived efficacy of protective measures, and anticipated regret.

While the current research focuses on mandatory insurance, information asymmetries such as moral hazard may also emerge in insurance markets where policyholders are able to select the level of coverage. Future work could examine the interplay between financial incentives and behavioral characteristics in these voluntary insurance schemes. Another important topic for further research is uncertainty about the future. For simplicity, our participants played a fixed number of rounds in the game. An interesting possibility would be to add a random stopping rule to the game to mimic the indefinite time horizon of real-world policyholders.

Acknowledgements

We thank Fujin Zhou, Mehmet Kutluay and two anonymous referees for helpful comments on an earlier version of this chapter. This research has received financial support from the Netherlands Organization for Scientific Research (NWO) VIDI (452.14.005) grant.

Table 2.5: Random-effects ML panel regression estimates on log-transformed damage-reducing investments

	(1) Treatments	(2) Previous rounds	(3) Demographics	(4) Flood beliefs
Treatment (ref. Baseline)				
Discount	2.372***	2.370***	2.200***	1.916***
	(0.230)	(0.230)	(0.228)	(0.248)
Loan	-0.169	-0.169	-0.172	0.099
	(0.231)	(0.231)	(0.227)	(0.234)
Loan × Discount	-0.455	-0.457	-0.241	-0.285
	(0.356)	(0.356)	(0.351)	(0.367)
Probability (ref. L: 3%)				
H: 15%	1.301***	1.301***	1.374***	1.649***
	(0.386)	(0.386)	(0.379)	(0.390)
Deductible (ref. xL: 5%)				
L: 15%	0.597***	0.596***	0.597***	0.708***
	(0.207)	(0.207)	(0.203)	(0.209)
H: 20%	1.163***	1.162***	1.163***	1.223***
	(0.207)	(0.207)	(0.203)	(0.209)
Order of scenario	-0.556***	-0.554***	-0.543***	-0.493***
	(0.071)	(0.071)	(0.069)	(0.071)
Participant flooded		-0.018	-0.018	-0.024
		(0.026)	(0.026)	(0.027)
Neighbor flooded		-0.012	-0.012	0.001
		(0.026)	(0.026)	(0.027)
Decision time round		0.005***	0.005***	0.004***
		(0.001)	(0.001)	(0.001)
Age in years			-0.086***	-0.064***
			(0.022)	(0.023)
Gender (1 = female)			0.867***	0.618***
			(0.171)	(0.181)
Income (1 = above €5000)			-0.989**	-1.141***
			(0.396)	(0.407)
Risk averse			0.221***	0.262***
			(0.067)	(0.069)
Present biased			0.008	0.001
			(0.010)	(0.011)
Efficacy protection				0.275***
				(0.044)
Worried about flood				0.389***
				(0.092)
Regret no investment / flood				0.108
				(0.088)
Regret investment / no flood				-0.267***
				(0.079)
Constant	4.810***	4.808***	4.891***	1.928**
	(0.423)	(0.423)	(0.761)	(0.932)
σ_u	3.554***	3.554***	3.490***	3.416***
	(0.060)	(0.060)	(0.059)	(0.060)
σ_e	0.983***	0.983***	0.983***	0.972***
	(0.005)	(0.005)	(0.005)	(0.005)
Observations	21,456	21,456	21,456	19,440
Nr of subjects	298	298	298	270
AIC	69,251	69,227	69,172	62,245
Log likelihood	-34,610	-34,594	-34,562	-31,094

Notes: Standard errors clustered by id and scenario in parentheses (* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$). Controls: Understanding questions, perceived difficulty, flood risk perception, 1/round, scenario-order × probability. Model 4 excludes the 28 subjects of session 1 because of incomplete efficacy responses. For robustness, we ran Models 1, 2 and 3 without these subjects; the results do not change.

Appendix 2A: Literature

Table A1: Overview of experimental literature on moral hazard

Publication	Journal	Type	Treatments	Context	N
Berger and Hershey (1994)	JRU	lab	stochastic/deterministic loss	Insurance	101
Di Mauro (2002)	JSE	lab	coverage	Insurance	60
McKee et al. (2004)	SNR	lab	size of loss	Insurance	60
McKee et al. (2007)	TJLS	lab	contingency fees	Legal services	22
Deck and Reyes (2008)	TSEJ	lab	second investor	Work effort	48
Du et al. (2008)	Working paper	lab	group identity, disclosure	Group dynamics	90
Burger and Kolstad (2009)	Working paper	lab	coalitions	Group dynamics	80
Gong et al. (2009)	JRU	lab	group / individual	Public goods	202
Karlan and Zinman (2009)	Econometrica	field	contract rates	Micro finance	5028
Banerjee et al. (2011)	JQE	lab	cut-off investment	Public goods	100
Hoppe and Kusterer (2011)	EER	lab	group size, conflict	Work effort	474
Cason et al. (2012)	JEBO	lab	group/individual	Micro finance	348
Hasson et al. (2012)	SAJE	lab	stochastic/deterministic loss	Climate change	144
Nieken and Schmitz (2012)	GEB	lab	wage schemes	Work effort	358
Füllbrunn and Neugebauer (2013)	EI	lab	transparency	Public goods	112
Biener et al. (2014)	Working paper	lab	coverage, group/individual	Micro finance	992
Bixter and Luhmann (2014)	JoEP	lab	face-to-face contact	Group dynamics	40
Dhillon et al. (2014)	Working paper	lab	social networks	Work effort	136
Gong et al. (2014)	JBDM	lab	group/individual	Public goods	294
Czura (2015)	JDE	field	monitoring, punishment	Microfinance	105
Hopfensitz et al. (2016)	Working paper	litf	deterministic/stochastic loss	Public goods	110
Huck et al. (2016)	JEBO	lab	competition	Health	336
Janssens and Kramer (2016)	JEBO	field	group/individual, communication	Micro finance	355
Neuß et al. (2016)	Working paper	lab	volunteer/insurer	Public goods	162
Biener et al. (2018)	EER	field	group / individual	Insurance	1,692
Giraudet et al. (2018)	JAERE	field	insurance, quality standards	Energy	2,936
Gelade and Guirkinger (2018)	JEBO	field	internal/external monitoring	Micro finance	890
Hoppe and Schmitz (2018)	GEB	lab	observability	Work effort	754
Rud et al. (2018)	JFI	lab	competition	Work effort	79
Macera (2018)	JEBO	lab	practice	Work effort	300

Notes: lab = lab experiment, field = field experiment, litf = lab in the field experiment, JRU = Journal of Risk and Uncertainty, JSE = Journal of Socio-Economics, SNR = Society & Natural Resources, TJLS = The Journal of Legal Studies, TSEJ = The SOuthern Economic Journal, JQE = Journal of QUantitative Economics, EER = European Economic Review, SAJE = South African Journal of Economics, GEB = Games and Economic Behavior, EI = Economic Inquiry, JoEP = Journal of Economic Psychology, JBDM = Journal of Behavioral Decision Making, JDE = Journal of Development Economics, JAERE = Journal of the Association of Environmental and Resource Economists, JEBO = Journal of Economic Behavior and Organization, JFI = Journal of Financial Intermediation.

Appendix 2B: Variable coding

Table B1: Summary overview of the variables used in the statistical analysis

Age	Continuous variable, age in years
Gender	Dummy variable, 1 = participant is female
High income	Dummy variable, 1 = monthly household after-tax income is within the highest category > €5000
Worried about flood	Categorical variable (range 1-5), worried about danger of flooding at current residence, 1 = strongly disagree, 5 = strongly agree
Regret no investment	Categorical variable (range 1-5), I felt regret about not investing in protection when a flood occurred in the game, 1 = strongly disagree, 5 = strongly agree
Regret investment	Categorical variable (range 1-5), when in a certain year in the game no flood occurred, I felt regret about paying for protection, 1 = strongly disagree, 5 = strongly agree
Risk averse	Risk aversion index: weighted average of four risk elicitation methods, 1 = very risk seeking, 10 = very risk averse
Present biased	Switching row in time list[a]
Efficacy protection	Categorical variable (range 0-10), How effective do you consider investing in flood protection measures that limit flood damage[b]
Participant flooded	Dummy variable, 1 = participant flooded in previous round
Neighbor flooded	Dummy variable, 1 = one or more neighbors[c] flooded in previous round

[a] Time list parameters from Falk et al. (2016) (range 1-26), 1 = no time discounting, 26 = high time discounting. [b] This question was taken from Poussin et al. (2014), 0 = very ineffective, 10 = very effective. [c] See Figure 2.5.

Appendix 2C: Comparative statics

The following section briefly describes the model, which extends the expected utility framework on optimal loss mitigation of Kelly and Kleffner (2003) to a multiple-years framework. Note that mitigation refers to investments that reduce the size of a potential loss but not the probability, which is known as self-insurance in the original model by Ehrlich and Becker (1972).

First, consider the one-year framework. Consider an individual with initial wealth W who faces a loss V with probability p and no loss with probability $1 - p$. The individual has the possibility to reduce the size of the loss by implementing mitigation expenditures r. The effectiveness of mitigation is captured in the mitigation function $L(r)$ that denotes the maximum possible loss if r is spent on mitigation. If a consumer does not spend anything on mitigation, the size of the loss will be V. Increasing mitigation expenditures leads to a decrease of maximum possible loss such that $L(0) = V$ and $L'(r) < 0$. Finally, assume that $L''(r) \leq 0$, meaning that the marginal effectiveness of mitigation decreases with an increase in mitigation expenditures. Insurance coverage is mandatory to protect against the possible loss, with a coverage of $\alpha \in [0, 1]$. In other words, the insurance contains a deductible of $1 - \alpha$ per dollar of coverage. The term $\alpha L(r)$ denotes the compensation in case of a loss. The insurer sets the premium $\alpha\pi$, where $\pi = pL(0)$. The insurer does not observe r and, hence, does not give premium discounts for risk reduction. The individual will choose a level of r to maximize expected utility EU:

$$\max EU_r = (1 - p)U[W - \alpha\pi - r] + pU[W - \alpha\pi - (1 - \alpha)L(r) - r] \tag{C1}$$

Now consider the multi-year framework. The model is constructed such that the policyholder considers a damage reduction investment in the present based on of the net present value of utility in both the present year (in which he/she considers an investment in mitigation) and in the years to come. For simplicity, we assume that the policyholder can invest only once, namely in the first year. A parallel with reality may be that you cannot elevate your house twice. Thus, the costs of mitigation r are paid in the first year $t = 1$ only, while the benefits (a decrease in L) extend in the future up to and including the last year T. Future years are discounted with a discount factor δ (see Frederick et al., 2002). The individual will choose a level of r to maximize expected utility EU:

$$\begin{aligned}\max_r EU &= (1 - p)U[W_1 - \alpha\pi - r] + pU[W_1 - \alpha\pi - (1 - \alpha)L(r) - r] \\ &+ \sum_{t=2}^{T} \frac{1}{(1 + \delta)^{t-1}} \Big((1 - p)U[W_t - \alpha\pi] + pU[W_t - \alpha\pi - (1 - \alpha)L(r)] \Big)\end{aligned} \tag{C2}$$

We aimed to derive theoretical predictions based on comparative statics for each of our treatments. We start with the simplest case: the effect of insurance coverage, by comparing the Insurance Baseline and the No Insurance treatments (Hypothesis 2.1).

Insurance Baseline versus No Insurance

Coverage α determines the difference between the Insurance Baseline and the No Insurance treatments. We determine the optimal investment in mitigation r in relation to α. Taking the derivative of Equation C2 with respect to r leads to the first order condition:

$$F = -(1-p)U'[W_1 - \alpha\pi - r] - p((1-\alpha)L'(r)+1)U'[W_1 - \alpha\pi - (1-\alpha)L(r) - r)]$$
$$-p((1-\alpha)L'(r))\sum_{t=2}^{T}\frac{1}{(1+\delta)^{t-1}}\Big(U'[W_t - \alpha\pi - (1-\alpha)L(r)]\Big) = 0 \tag{C3}$$

Using the implicit function theorem:

$$\frac{\partial r}{\partial \alpha} = -\frac{F'_\alpha}{F'_r}$$

Fulfilled second order condition implies:

$$F'_r < 0$$

Abbreviating $W_1 - \alpha\pi - r$ as nL_1, $W_1 - \alpha\pi - (1-\alpha)L(r) - r$ as L_1 and $W_t - \alpha\pi - (1-\alpha)L(r)$ as L_t:

$$F'_\alpha = (1-p)\pi U''(nL_1) - p((1-\alpha)L'(r)+1)(L(r)-\pi)U''(L_1) + L'(r)pU'(L_1)$$
$$+L'(r)p\sum_{t=2}^{T}\frac{1}{(1+\delta)^{t-1}}U'(L_t) - p((1-\alpha)L'(r))\sum_{t=2}^{T}\frac{1}{(1+\delta)^{t-1}}(L(r)-\pi)U''(L_t) \tag{C4}$$

If we assume $1 < |(1-\alpha)L'(r)|$ and a concave utility function, F'_α is negative. Then:

$$\frac{\partial r}{\partial \alpha} < 0 \tag{C5}$$

Under more insurance coverage, optimal investment in r decreases, which is part of Hypothesis 2.1.

Loan treatment

We have found a comparative statics prediction for the simplest treatment, under the assumption that $1 < |(1-\alpha)L'(r)|$. This holds for the parameters used in our experiment, but it is not necessarily always the case. Furthermore, analytical solutions for the other hypotheses cannot be obtained. For example, consider the Loan treatment (Hypothesis 2.3a). Individuals pay part $q \in [0,1]$ of investment r for $1/q$ periods until the loan has been repaid, maximizing utility:

$$\max_r EU = (1-p)U[W_1 - \alpha\pi - qr] + pU[W_1 - \alpha\pi - (1-\alpha)L(r) - qr]$$
$$+ \sum_{t=2}^{T} \frac{1}{(1+\delta)^{t-1}} \Big((1-p)U[W_t - \alpha\pi - qr] + pU[W_t - \alpha\pi - (1-\alpha)L(r) - qr] \Big) \tag{C6}$$

Taking the derivative of Equation C6 with respect to r leads to the first order condition:

$$F = -q(1-p)U'[W_1 - \alpha\pi - qr] - p((1-\alpha)L'(r) + q)U'[W_1 - \alpha\pi - (1-\alpha)L(r) - qr)]$$
$$-p((1-\alpha)L'(r) + q) \sum_{t=2}^{T} \frac{1}{(1+\delta)^{t-1}} \Big(U'[W_t - \alpha\pi - (1-\alpha)L(r) - qr] \Big) = 0 \tag{C7}$$

Abbreviate $W_1 - \alpha\pi - qr$ as X_1, $W_1 - \alpha\pi - (1-\alpha)L(r) - qr$ as X_2 and $W_t - \alpha\pi - (1-\alpha)L(r) - qr$ as X_3:

$$F'_q = -(1-p)U'[X_1] + rq(1-p)U''[X_1] - pU'[X_2] + pr((1-\alpha)L'(r) + q)U''[X_2]$$
$$-p \sum_{t=2}^{T} \frac{1}{(1+\delta)^{t-1}} U'[X_3] + pr((1-\alpha)L'(r) + q) \sum_{t=2}^{T} \frac{1}{(1+\delta)^{t-1}} U''[X_3] \tag{C8}$$

It is not straightforward to determine the sign of F'_q without restricting some of the parameters. Similar problems occur with Hypothesis 2.2a, 2.2b and 2.3b. Therefore, we decided to use numerical simulations to generate predictions about our hypotheses (see Appendix 2D).

Appendix 2D: Parameter calculations

To determine the parameters of our investment game, we calculated the net present value (NPV) based on Expected Utility (Equation C2) for different combinations of parameters. Some parameters were chosen based on estimations from reality, such as the maximum damage (50,000 ECU) and the interest rate (1%). For the effectiveness of damage reducing investments, we used the loss function $L(r) = Ve^{-\beta r}$ proposed by Kelly and Kleffner (2003), where V denotes the maximum loss and the effectiveness of mitigation is captured by parameter β. We aimed to base our loss function on damage reduction estimates from real data: Federal Emergency Management Agency (FEMA) cost estimates and damage reduction estimates for a typical single family dwelling of flood mitigation measures. Figure D1 plots these estimates[16] against the loss function with different values of β, with V = 200,000, the average value of this type of building. The mitigation function $L(r) = Ve^{-\beta r}$ with $0.00001 \leq \beta \leq 0.00008$ seems to fit the data well.

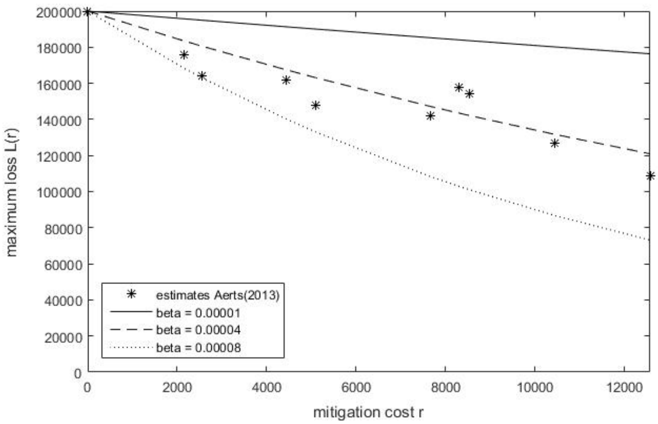


Figure D1: Parameter estimation of the mitigation function

We varied the parameters (savings account, income per round, probabilities, deductibles, β, number of installments) to find a reasonable combination[17] which allowed us to test all our hypotheses. Table D1 shows the results of these simulations with our final set of parameters.

[16] Table 2.10, Table 2.13 and Table 2.18 from Aerts et al. (2013) to be precise.
[17] For example: $0.00001 \leq \beta \leq 0.00008$, positive income.

Table D1: Normalized NPV of investment by scenario and treatment at $\delta = 0.01$

(a) Risk averse ($\theta = 0.3$)

	Insurance Baseline				Premium Discount				Loan				Loan+Discount			
	1000	5000	10,000	15,000	1000	5000	10,000	15,000	1000	5000	10,000	15,000	1000	5000	10,000	15,000
H -	**0.025**	**0.103**	**0.163**	**0.195**												
HH	**0.003**	**0.005**	-0.016	-0.055	**0.059**	**0.239**	**0.374**	**0.440**	**0.003**	**0.004**	-0.015	-0.048	**0.058**	**0.239**	**0.375**	**0.446**
HL	-0.001	-0.013	-0.046	-0.092	**0.057**	**0.231**	**0.361**	**0.425**	-0.001	-0.013	-0.042	-0.082	**0.057**	**0.232**	**0.365**	**0.434**
HxL	-0.008	-0.046	-0.099	-0.159	**0.054**	**0.219**	**0.342**	**0.401**	-0.008	-0.043	-0.091	-0.143	**0.054**	**0.221**	**0.349**	**0.416**
L -	**0.001**	-0.001	-0.012	-0.028												
LH	-0.008	-0.045	-0.097	-0.154	**0.001**	-0.003	-0.026	-0.063	-0.008	-0.043	-0.089	-0.139	**0.002**	**0.000**	-0.018	-0.049
LL	-0.009	-0.048	-0.102	-0.160	**0.001**	-0.003	-0.026	-0.064	-0.009	-0.046	-0.094	-0.145	**0.002**	-0.001	-0.019	-0.049
LxL	-0.010	-0.054	-0.111	-0.172	**0.001**	-0.004	-0.028	-0.066	-0.010	-0.051	-0.103	-0.156	**0.002**	-0.001	-0.019	-0.049

(b) Risk neutral ($\theta = 1$)

	Insurance Baseline				Premium Discount				Loan				Loan+Discount			
	1000	5000	10,000	15,000	1000	5000	10,000	15,000	1000	5000	10,000	15,000	1000	5000	10,000	15,000
H -	**0.025**	**0.103**	**0.163**	**0.195**												
HH	**0.001**	-0.003	-0.024	-0.057	**0.052**	**0.216**	**0.341**	**0.406**	**0.002**	-0.001	-0.020	-0.052	**0.053**	**0.218**	**0.345**	**0.412**
HL	-0.002	-0.017	-0.047	-0.086	**0.052**	**0.216**	**0.341**	**0.406**	-0.002	-0.015	-0.043	-0.081	**0.053**	**0.218**	**0.345**	**0.412**
HxL	-0.008	-0.044	-0.093	-0.144	**0.052**	**0.216**	**0.341**	**0.406**	-0.008	-0.042	-0.089	-0.139	**0.053**	**0.218**	**0.345**	**0.412**
L -	**0.001**	-0.001	-0.012	-0.028												
LH	-0.009	-0.047	-0.097	-0.150	**0.001**	-0.003	-0.024	-0.057	-0.009	-0.045	-0.093	-0.144	**0.002**	-0.001	-0.020	-0.052
LL	-0.010	-0.050	-0.102	-0.156	**0.001**	-0.003	-0.024	-0.057	-0.009	-0.048	-0.098	-0.150	**0.002**	-0.001	-0.020	-0.052
LxL	-0.011	-0.055	-0.111	-0.168	**0.001**	-0.003	-0.024	-0.057	-0.011	-0.053	-0.107	-0.162	**0.002**	-0.001	-0.020	-0.052

(c) Risk seeking ($\theta = 3$)

	Insurance Baseline				Premium Discount				Loan				Loan+Discount			
	1000	5000	10,000	15,000	1000	5000	10,000	15,000	1000	5000	10,000	15,000	1000	5000	10,000	15,000
H -	**0.025**	**0.103**	**0.163**	**0.195**												
HH	-0.002	-0.014	-0.029	-0.046	**0.029**	**0.123**	**0.202**	**0.249**	-0.001	-0.008	-0.022	-0.042	**0.030**	**0.129**	**0.209**	**0.253**
HL	-0.003	-0.018	-0.037	-0.055	**0.031**	**0.134**	**0.220**	**0.272**	-0.002	-0.013	-0.033	-0.057	**0.033**	**0.138**	**0.224**	**0.270**
HxL	-0.006	-0.030	-0.058	-0.083	**0.036**	**0.155**	**0.255**	**0.316**	-0.006	-0.030	-0.062	-0.096	**0.037**	**0.156**	**0.251**	**0.303**
L -	**0.001**	-0.001	-0.012	-0.028												
LH	-0.008	-0.039	-0.075	-0.107	**0.000**	-0.003	-0.015	-0.031	-0.008	-0.039	-0.079	-0.121	**0.001**	-0.003	-0.020	-0.046
LL	-0.008	-0.041	-0.077	-0.110	**0.001**	-0.002	-0.013	-0.029	-0.008	-0.041	-0.083	-0.125	**0.001**	-0.002	-0.019	-0.045
LxL	-0.009	-0.044	-0.084	-0.118	**0.001**	**0.000**	-0.011	-0.026	-0.009	-0.045	-0.090	-0.136	**0.001**	-0.001	-0.017	-0.044

Table D2: Normalized NPV of investment by scenario and treatment at $\delta = 0.1$

(a) Risk averse ($\theta = 0.3$)

	Insurance Baseline				Premium Discount				Loan				Loan+Discount			
	1000	5000	10,000	15,000	1000	5000	10,000	15,000	1000	5000	10,000	15,000	1000	5000	10,000	15,000
H -	**0.016**	**0.065**	**0.099**	**0.113**												
HH	-0.001	-0.015	-0.048	-0.095	**0.037**	**0.150**	**0.226**	**0.253**	**0.002**	**0.003**	-0.011	-0.034	**0.041**	**0.167**	**0.262**	**0.312**
HL	-0.004	-0.027	-0.069	-0.121	**0.036**	**0.145**	**0.218**	**0.244**	-0.001	-0.009	-0.030	-0.057	**0.040**	**0.162**	**0.256**	**0.304**
HxL	-0.009	-0.050	-0.106	-0.168	**0.034**	**0.137**	**0.206**	**0.230**	-0.006	-0.030	-0.064	-0.101	**0.038**	**0.156**	**0.246**	**0.293**
L -	-0.001	-0.009	-0.024	-0.044												
LH	-0.009	-0.048	-0.102	-0.161	-0.002	-0.018	-0.052	-0.097	-0.006	-0.030	-0.063	-0.098	**0.001**	**0.000**	-0.013	-0.034
LL	-0.010	-0.051	-0.106	-0.166	-0.002	-0.019	-0.052	-0.097	-0.006	-0.032	-0.066	-0.102	**0.001**	**0.000**	-0.013	-0.035
LxL	-0.011	-0.055	-0.113	-0.174	-0.002	-0.019	-0.053	-0.098	-0.007	-0.036	-0.073	-0.110	**0.001**	-0.001	-0.014	-0.035

(b) Risk neutral ($\theta = 1$)

	Insurance Baseline				Premium Discount				Loan				Loan+Discount			
	1000	5000	10,000	15,000	1000	5000	10,000	15,000	1000	5000	10,000	15,000	1000	5000	10,000	15,000
H -	**0.016**	**0.065**	**0.099**	**0.113**												
HH	-0.003	-0.019	-0.051	-0.092	**0.033**	**0.135**	**0.207**	**0.236**	**0.001**	-0.001	-0.014	-0.036	**0.037**	**0.154**	**0.244**	**0.291**
HL	-0.005	-0.029	-0.067	-0.112	**0.033**	**0.135**	**0.207**	**0.236**	-0.001	-0.010	-0.031	-0.057	**0.037**	**0.154**	**0.244**	**0.291**
HxL	-0.009	-0.048	-0.099	-0.153	**0.033**	**0.135**	**0.207**	**0.236**	-0.006	-0.030	-0.063	-0.098	**0.037**	**0.154**	**0.244**	**0.291**
L -	-0.001	-0.009	-0.024	-0.044												
LH	-0.010	-0.050	-0.103	-0.157	-0.003	-0.019	-0.051	-0.092	-0.006	-0.032	-0.066	-0.102	**0.001**	-0.001	-0.014	-0.036
LL	-0.010	-0.052	-0.106	-0.161	-0.003	-0.019	-0.051	-0.092	-0.007	-0.034	-0.069	-0.106	**0.001**	-0.001	-0.014	-0.036
LxL	-0.011	-0.056	-0.112	-0.169	-0.003	-0.019	-0.051	-0.092	-0.007	-0.038	-0.076	-0.114	**0.001**	-0.001	-0.014	-0.036

(c) Risk seeking ($\theta = 3$)

	Insurance Baseline				Premium Discount				Loan				Loan+Discount			
	1000	5000	10,000	15,000	1000	5000	10,000	15,000	1000	5000	10,000	15,000	1000	5000	10,000	15,000
H -	**0.016**	**0.065**	**0.099**	**0.113**												
HH	-0.004	-0.021	-0.042	-0.062	**0.019**	**0.078**	**0.125**	**0.149**	**0.001**	-0.005	-0.016	-0.031	**0.022**	**0.094**	**0.152**	**0.184**
HL	-0.005	-0.024	-0.047	-0.069	**0.020**	**0.085**	**0.136**	**0.163**	-0.002	-0.010	-0.024	-0.041	**0.024**	**0.100**	**0.162**	**0.195**
HxL	-0.007	-0.033	-0.063	-0.089	**0.023**	**0.098**	**0.158**	**0.190**	-0.004	-0.021	-0.044	-0.068	**0.026**	**0.111**	**0.179**	**0.215**
L -	-0.001	-0.009	-0.024	-0.044												
LH	-0.009	-0.042	-0.079	-0.111	-0.003	-0.016	-0.037	-0.059	-0.005	-0.028	-0.056	-0.086	**0.001**	-0.002	-0.014	-0.032
LL	-0.009	-0.043	-0.080	-0.114	-0.002	-0.015	-0.035	-0.057	-0.006	-0.029	-0.059	-0.089	**0.001**	-0.001	-0.013	-0.032
LxL	-0.009	-0.045	-0.085	-0.119	-0.002	-0.014	-0.034	-0.055	-0.006	-0.032	-0.064	-0.096	**0.001**	-0.001	-0.012	-0.031

The table displays the NPV of Expected Utility of investments in damage mitigation over 10 rounds[18], discounted by $\delta = 0.01$ for different degrees of risk aversion θ and normalized over the minimal and maximal possible expected values in the game. We used a power utility function of the form $u(x) = x^{\theta}$. The results are shown in comparison to zero investment. Therefore, all positive numbers are displayed in bold, as they indicate a net gain from investing a positive amount. For each combination of treatment and scenario, the largest positive number gives the optimal investment (<u>underlined</u>) for an individual. If no number is underlined the optimal investment is zero. Table D2 shows the results for high discounting, $\delta = 0.1$.

The following section repeats the hypotheses and explains briefly how each hypothesis can be tested based on the predictions in Table D1 and Table D2.

Hypothesis 1. *Damage reduction investments in the Insurance Baseline treatment are lower than in the No Insurance treatment, but greater than zero.* The NPV is higher for all investments in No Insurance (denoted as H - and L - in Table D1) compared to investments in Insurance Baseline. In the high probability scenarios, positive investments may be optimal with insurance, depending on the deductible level and attitude to risk. For example, for a risk averse individual in scenario HH (Table D1a) the optimal investment in Insurance Baseline is 5000 ECU, which leads to a positive NPV of 0.005 compared to no investment. This allows for evaluation of Hypothesis 2.1.

Hypothesis 2a. *Damage reduction investments are higher in the Premium Discount treatment than in the Insurance Baseline treatment.* Comparing the Premium Discount column with the Insurance Baseline column gives higher NPV values in each of the rows and sub-tables in Table D1. Therefore, this hypothesis can be tested under all scenarios and risk attitudes.

Hypothesis 2b. *Damage reduction investments are highest in the Loan+Discount treatment* Under low time discounting (Table D1), investments in the Premium Discount treatment were already optimal, such that they stay optimal in Loan+Discount treatment. Under high time discounting (Table D2), Loan+Discount gives the highest optimal investments of all treatments in the low probability scenarios.

[18] Note that the actual design uses a fixed number of 12 rounds, but participants are only informed that each scenario takes at least 10 rounds. The results of the simulations for 12 rounds (not shown here in detail) are very similar to the tables reported here and the corresponding hypotheses are identical.

Hypothesis 3a. *Damage reduction investments are lower for participants with high time discount rates. This effect is strongest in the Insurance Baseline and Premium Discount treatments, but disappears in the Loan and Loan+Discount treatments.* In the Insurance Baseline and Premium Discount treatments, increasing the time discount rate from low time discounting ($\delta = 0.01$ in Table D1) to high time discounting ($\delta = 0.1$ in Table D2) decreases the optimal investment level. No change is observed in the Loan and Loan+Discount treatments.

Hypothesis 3b. *Risk-averse individuals will invest more in damage reduction in the Insurance Baseline treatment and the Loan treatment than risk-neutral individuals, where risk-seeking individuals will invest less.* In the Insurance Baseline and the Loan treatment, risk-neutral ($\theta = 1$, Table D1b) individuals will invest (scenario HH and HL). A risk-averse individual ($\theta = 0.3$, Table D1a) will also get a positive NPV for investing 5000. Risk-seeking individuals ($\theta = 3$, Table D1c) will not invest in any of these scenarios.

Appendix 2E: Comprehension questions

Correct answers are marked in **bold**.

Questions asked in all treatments

- What was the flood risk in the test scenario?

  a) 1% b) 3% c) 5% d) 10% e) 15% f) 20%

  The answer depends on the risk in the test scenario (randomly determined).
- If you are flooded in year 1, what is the flood risk in year 2?
  (a) Less than in year 1
  (b) **Flood risk does not change**
  (c) More than in year 1
- How long are protective investments effective?
  (a) From the moment you implement to the end of the experiment
  (b) **From the moment you implement to the end of the scenario**
  (c) From the start of the scenario to the end of the scenario

Extra question in the No Insurance treatment

- What happens if you are flooded and you did not take protective investments?
  (a) **I have to pay the full damage: 50.000 ECU**
  (b) I have to pay a small fee
  (c) I will cry

Extra question in all Insurance treatments

- What was your deductible (eigen risico) in the test scenario?

  a) 5 percent b) 15 percent c) 20 percent d) 50 percent

  The answer depends on the deductible in the test scenario (randomly determined).

Extra question in the Loan and Loan+Discount treatments

- Should you always repay your loan?
  (a) No, I can refuse to pay the loan cost
  (b) No, if the loan is not fully repaid in the last year, I am lucky
  (c) Yes, I will pay the loan cost in the first 5 years
  (d) **Yes, if the loan is not fully repaid in the last year, I will pay the remainder**

Extra question in the Premium Discount and Loan+Discount treatments

- What is the benefit of a protective investment?
  (a) A reduced damage in case of a flood
  (b) A lower premium
  (c) **Both reduced damage and a lower premium**
  (d) None of the above

Appendix 2F: Additional analyses

Risk aversion index Our risk aversion index was a linear combination of the four risk aversion measures, following Menkhoff and Sakha (2017). Table F1 shows the results of our random-effects ML panel regressions for each of the four measures separately, in comparison to the combined measure (Model 5). All risk aversion measures except the price list in the loss domain have positive and significant estimates.

Extra investors As investments in damage reduction lasted for all rounds of the game, it was optimal to invest in the first round. However, a substantial number of subjects increased their existing investment after the first round, or started investing after the first round. The number of these 'extra investors' and the average extra investment, pooled by the appearance of each scenario, are plotted in Figure F1. The number of subjects that invests extra drops by half from the first to the last scenario. Even though all subjects started with 5 rounds of the test scenario, a substantial number of subjects invests extra in the experimental scenarios. Interestingly, extra investments are rather stable over the scenarios at about 7000 ECU.

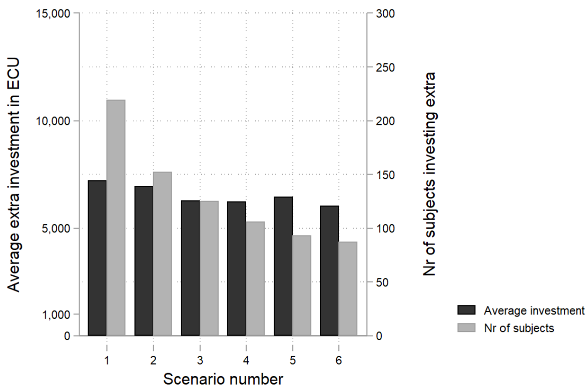


Figure F1: Extra investments after first round

To analyze the extra investors in more detail, we ran our random-effects ML panel regressions with log-transformed extra investments as the dependent variable. This variable was constructed to capture a change in investment from the previous round, starting from round 2. For example, if a subject invests 1000 ECU in round 1, nothing more in round 2 and increases to 5000 ECU in round 3, the extra investment variable takes the values 0, 0, 4000. Table F2 shows that extra investments increase after a flood in the game that hit the subject's house, but not after hitting the neighbors. The non-significant estimates of probability and deductibles suggest that extra investments do not differ per scenario. In contrast to the analysis of investments in all rounds, we find no effect of risk aversion and efficacy of protection on extra investments; these seem to be primary motivators to invest at the start of the game. Extra investors seem to be primarily motivated by firsthand experience of flood in the game and anticipated regret.

Table F1: Random-effects ML panel regressions of log of investments

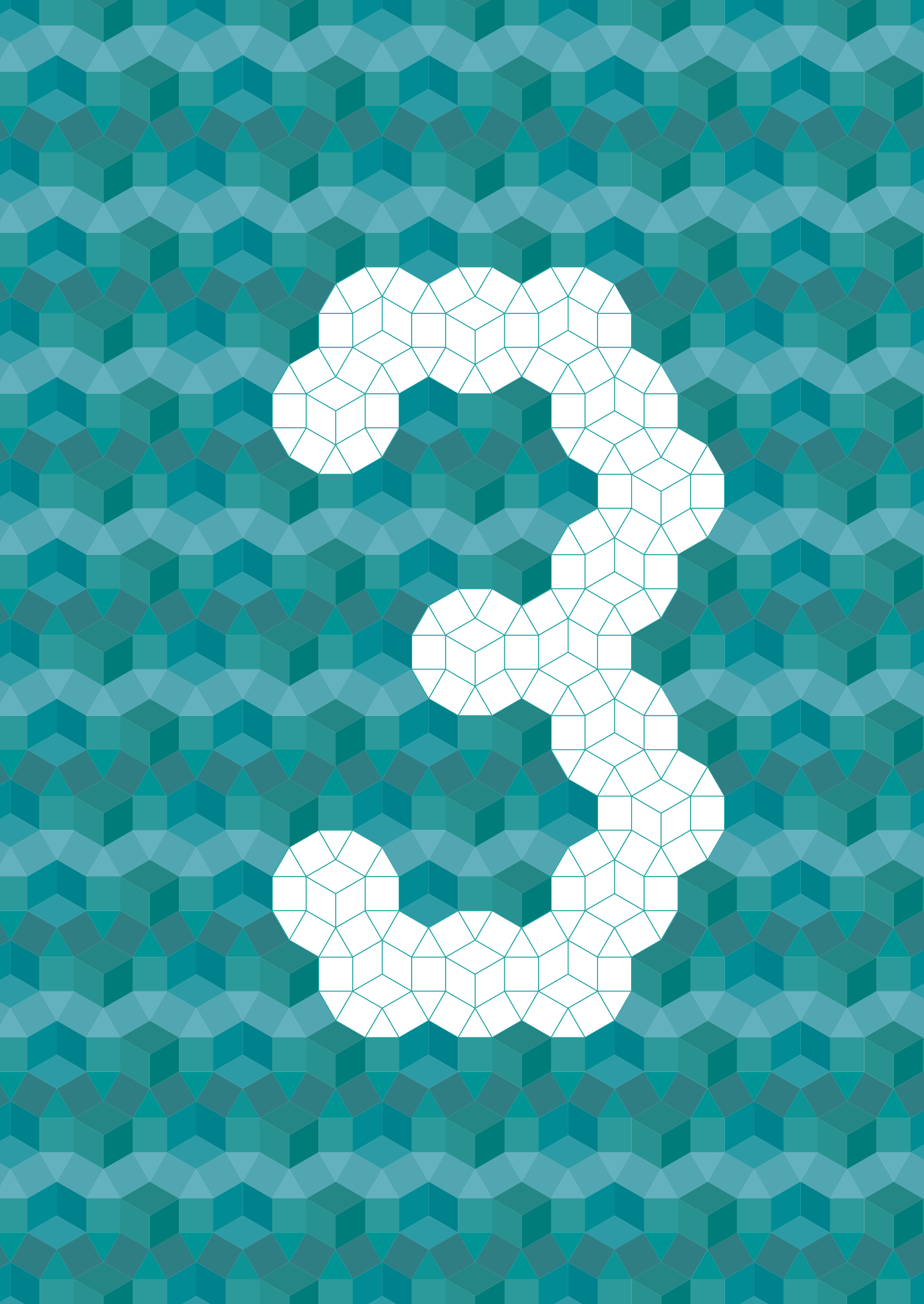
	(1) qualitative	(2) list gain	(3) list loss	(4) BRET	(5) combined
Treatment (ref. Baseline Insurance)					
Premium Discount	1.886***	1.927***	1.909***	1.892***	1.916***
	(0.249)	(0.249)	(0.249)	(0.249)	(0.248)
Loan	0.137	0.139	0.115	0.048	0.099
	(0.235)	(0.235)	(0.236)	(0.235)	(0.234)
Loan $\times$ Discount	-0.217	-0.302	-0.243	-0.228	-0.285
	(0.368)	(0.369)	(0.369)	(0.367)	(0.367)
Probability (ref. L: 3%)					
H: 15%	1.656***	1.639***	1.623***	1.640***	1.649***
	(0.391)	(0.392)	(0.392)	(0.390)	(0.390)
Deductible (ref. xL: 5%)					
L: 15%	0.708***	0.708***	0.708***	0.708***	0.708***
	(0.209)	(0.209)	(0.210)	(0.209)	(0.209)
H: 20%	1.223***	1.223***	1.223***	1.223***	1.223***
	(0.209)	(0.209)	(0.210)	(0.209)	(0.209)
Order of scenario	-0.492***	-0.494***	-0.497***	-0.494***	-0.493***
	(0.071)	(0.071)	(0.071)	(0.071)	(0.071)
Participant flooded	-0.024	-0.024	-0.024	-0.024	-0.024
	(0.027)	(0.027)	(0.027)	(0.027)	(0.027)
Neighbor flooded	0.001	0.001	0.001	0.001	0.001
	(0.027)	(0.027)	(0.027)	(0.027)	(0.027)
Decision time round	0.004***	0.004***	0.004***	0.004***	0.004***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Risk averse self reported	0.145***				
	(0.046)				
Risk averse in gain domain		0.062**			
		(0.030)			
Risk averse in loss domain			-0.007		
			(0.036)		
Risk averse in BRET on 1-10 scale				0.153***	
				(0.042)	
Risk averse					0.262***
					(0.069)
Constant	2.785***	3.084***	3.541***	2.856***	1.928**
	(0.867)	(0.861)	(0.894)	(0.855)	(0.932)
σ_u	3.421***	3.426***	3.431***	3.417***	3.416***
	(0.060)	(0.061)	(0.061)	(0.060)	(0.060)
σ_e	0.972***	0.972***	0.972***	0.972***	0.972***
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
Observations	19440	19440	19440	19440	19440
Nr of subjects	270	270	270	270	270
AIC	62,249	62,255	62,259	62,246	62,245
Log likelihood	-31,097	-31,099	-31,101	-31,095	-31,094

Notes: Standard errors clustered by id and scenario in parentheses (* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$). Controls: Understanding questions, perceived difficulty, flood risk perception, scenario-order $\times$ probability, high income, gender, age, efficacy, worry, regret and 1/round.

Table F2: Random-effects ML panel regressions of extra investments

	(1) treatments	(2) previous rounds	(3) demographics
Treatment (ref. Baseline Insurance)			
Premium Discount	0.132***	0.118***	0.145***
	(0.028)	(0.028)	(0.030)
Loan	0.016	0.012	0.003
	(0.028)	(0.028)	(0.028)
Loan × Discount	-0.184***	-0.166***	-0.149***
	(0.044)	(0.044)	(0.044)
Probability (ref. L: 3%)			
H: 15%	-0.032	-0.026	0.043
	(0.048)	(0.047)	(0.047)
Deductible (ref. xL: 5%)			
L: 15%	0.039	0.039	0.038
	(0.025)	(0.025)	(0.025)
H: 20%	0.021	0.021	0.028
	(0.025)	(0.025)	(0.025)
Order of scenario	-0.051***	-0.050***	-0.043***
	(0.009)	(0.009)	(0.009)
Participant flooded	0.197***	0.198***	0.197***
	(0.034)	(0.034)	(0.035)
Neighbor flooded	0.016	0.016	0.022
	(0.034)	(0.034)	(0.035)
Decision time round	0.009***	0.009***	0.008***
	(0.001)	(0.001)	(0.001)
Age in years		-0.011***	-0.003
		(0.003)	(0.003)
Gender (1 = female)		0.070***	0.050**
		(0.021)	(0.022)
Income (1 = above €5,000)		-0.025	-0.017
		(0.049)	(0.049)
Risk averse		0.007	0.005
		(0.008)	(0.008)
Present biased		0.001	0.001
		(0.001)	(0.001)
Efficacy protection			0.008
			(0.005)
Worried about flood			0.012
			(0.011)
Regret no investment / flood			0.046***
			(0.011)
Regret investment / no flood			0.043***
			(0.010)
Constant	0.261***	0.419***	-0.115
	(0.053)	(0.095)	(0.112)
σ_u	0.223***	0.216***	0.183***
	(0.015)	(0.015)	(0.017)
σ_e	1.312***	1.312***	1.278***
	(0.007)	(0.007)	(0.007)
Observations	21456	21456	19440
Nr of subjects	298	298	270
AIC	73,104	73,085	65,112
Log likelihood	-36,533	-36,518	-32,528

Notes: Standard errors clustered by id and scenario in parentheses (* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$). Controls: Understanding questions, perceived difficulty, flood risk perception, order × probability and 1/round.


3

Behavioral motivations for self-insurance under different disaster risk insurance schemes

This paper presents a lab-in-the-field experiment with 2111 Dutch homeowners in floodplain areas to examine the impacts of financial incentives and behavioral motivations for self-insurance under different flood insurance schemes. We experimentally varied the insurance type (mandatory public versus voluntary private) and the availability of a premium discount incentive for investing in flood damage mitigation measures. This set-up allowed us to examine the existence of moral hazard, advantageous selection and the behavioral motivations of individual agents who face these different insurance types, without the selection bias that makes a causal inference from survey studies problematic. The main results show that a premium discount can increase investments in self-insurance under both private and public insurance. Moreover, we find no support for moral hazard in our natural disaster insurance market, but we do find a substantial share of cautious people who invest both in private insurance as well as in self-insurance, indicating advantageous selection. The results have implications for the design of insurance schemes to cope with increasing natural disaster risks.

This chapter is published as:
Mol, J. M., Botzen, W. J. W., & Blasch, J. E. (2020). Behavioral motivations for self-insurance under different disaster risk insurance schemes. *Journal of Economic Behavior & Organization,* 180, 967–991.

3.1 Introduction

The impacts of natural hazards on society have increased in the past decades and are expected to increase further in the future, as a result of climate change as well as population and economic growth in disaster prone areas (IPCC, 2012; Munich RE, 2018). Of all weather-related disasters, flooding is considered to have the largest consequences both in number of people affected and in total economic cost (UNISDR, 2015). As a response to this problem, researchers have investigated potential risk reduction strategies, such as flood protection infrastructure like dikes (Kreibich et al., 2015) and disaster risk insurance schemes (Michel-Kerjan, 2010; Kunreuther, 2015; Hudson et al., 2016). In the EU, a variety of arrangements exist in member states for compensating flood losses, including public insurance which is often mandatory and private market insurance which is often voluntary (Schwarze et al., 2011; Paudel et al., 2012). In various countries it has been debated whether these arrangements should be reformed to provide policyholders stronger incentives to limit the risk. Stimulating individuals to invest in self-insurance - defined as measures that reduce the size but not the probability of a loss (Ehrlich and Becker, 1972) - is an additional promising approach in the attempt to decrease expected damages from natural disasters (Den et al., 2017).

In the case of flood risk, various cost-effective measures can be taken by private homeowners to prevent flood damage. These measures fall into three broad categories: dry flood proofing (shielding a house to prevent water from entering), wet flood proofing (minimizing damage once water has invaded a house), and the elevation of structures. However, investments in self-insurance by individual homeowners are still rare, even though these measures can be cost-effective (Bubeck et al., 2012; Poussin et al., 2015).

There are three main explanations for low investments in self-insurance: namely, the availability of insurance, the features of insurance, and the behavioral characteristics of individual agents. A behavioral explanation for low investments in self-insurance in the context of flood risk is that individuals have difficulties understanding low-probability high-impact (LPHI) risks, such as flood risk, and underestimate these risks when they do not personally experience such disasters (Kunreuther and Pauly, 2004). They might only respond to the risk when a certain threshold level of concern is reached (McClelland et al., 1993) or generally underweight the probability in their insurance decision. Such underweighting of risk can be accommodated by Prospect Theory (Kahneman and Tversky, 1979), which is a frequently used model for decisions under risk that has been used to explain behavior related to natural disasters (Page et al., 2014; Koetse and Brouwer, 2016). Under Prospect Theory, risk attitudes are defined by a combination of utility curvature, loss aversion and probability weighting. While a large existing literature has examined probability weighting of LPHI risks and loss aversion (see e.g. Barberis, 2013), the current chapter focuses on the influence of insurance and financial incentives on individual investments in self-insurance.

Other behavioral explanations include incorrectly high perceived costs of implementing self-insurance measures, a present bias that leads to procrastination of long-term investments or a potential moral hazard effect arising from insurance (Michel-Kerjan, 2010). Economic theory predicts that individuals invest less in self-insurance under insurance coverage, unless they are incentivized to make such investments through premium discounts (Ehrlich and Becker, 1972). However, individuals may respond differently to insurance features, such as a premium discount, when the insurance offered is mandatory (public insurance), rather than voluntary (market insurance), which is nearly impossible to study with non-experimental data. Furthermore, the results provided in Cutler et al. (2008) suggest that less risk reducing behavior is associated with lower insurance take-up, which could be due to low risk aversion. Similarly, de Meza and Webb (2001) showed that highly risk averse individuals tend to purchase insurance and also take other measures to limit risks. The importance of the role of risk preferences is also recognized by Corcos et al. (2017) who conduct a lab experiment on the conditional demand for insurance under premium variations, while controlling for risk preferences. In our study, we investigate the influence of financial incentives and behavioral motivations on the level of self-insurance against LPHI risk.

The current chapter focuses on incentives for self-insurance in the domain of flood risk, both in the presence and absence of flood risk insurance, to offer insights into all three categories of explanations. A large online experiment with homeowners in floodplain areas was conducted. The homeowners were randomly assigned to face either a public or private insurance scheme, which rules out potential endogeneity bias. In the field, different types (e.g. with regards to risk attitudes and self-insurance) may have access to different types of insurance schemes, which makes it difficult to make correct causal claims about the effect of a typical insurance scheme on investments in self-insurance. Homeowners in the river delta in the Netherlands with relatively high flood probabilities seem to be a suitable sample to study flood risk mitigation of households. To the best of our knowledge we are the first to study self-insurance behavior experimentally under both public and private insurance schemes, accounting for both insurance features and behavioral characteristics of the decision-makers. Furthermore, we use a large sample size such that the group of respondents who self-select into insurance coverage will be large enough to make valid comparisons with the publicly insured experimental subjects.

The main results show no difference in self-insurance investments between respondents with public (mandatory) versus private (market) insurance. With regards to the features of insurance, we find that a premium discount increases investments in self-insurance under both private and public insurance. Moreover, we find no support for moral hazard in our natural disaster insurance market, but we do find a substantial share of cautious people who invest both in private insurance as well as in self-insurance, indicating advantageous selection. These cautious people take their investment decision consciously and are primarily motivated by the efficacy of mitigation measures, social norms and risk aversion, as well as by a lower trust in dike maintenance.

The remainder of this chapter is organized as follows: Section 3.2 gives a short overview of the literature on behavioral insurance in the low probability context, Section 3.3 describes the experimental design, Section 3.4 derives hypotheses for each of the treatments. Finally, Section 3.5 presents results, Section 3.6 discusses the results and their implications and concludes.

3.2 Literature review

A growing body of empirical research has examined factors contributing to private self-insurance decisions and preventive behavior in the context of natural disaster risk. In this section, we briefly review the papers most relevant to our study (for more detailed literature reviews see Bubeck et al., 2012; Koerth et al., 2017).

3.2.1 Presence of insurance

Insurance companies generally do not expect policyholders to self-insure, due to the existence of information asymmetries between the insurer and the insured. This implies risk reduction behavior of policyholders is not observed by insurers and hence not reflected in premiums (Arrow, 1963; Arnott and Stiglitz, 1988). In theory, such a moral hazard effect removes individuals' motivation to self-insure if they have insurance coverage, as they expect to be compensated in case of damage irrespective of their risk reduction efforts. In this case the expected benefits of self-insurance remain at the insurer level. The moral hazard effect has been studied empirically in different insurance markets and appears to vary with the type of insurance product and the magnitude of asymmetric information (Cohen and Siegelman, 2010). For example in the health insurance market, Einav et al. (2013) identified an ex-post moral hazard effect in data of insurance coverage and medical spending of a large U.S. company. In contrast, Chiappori and Salanié (2000) found no evidence for moral hazard in the automobile insurance industry. If the asymmetric information involves private information on the side of the policyholder about the probability of loss, it is essential that this information is correctly understood to be of any advantage to the policyholder.[1] Moreover, as shown by de Meza and Webb (2001) behavioral characteristics can explain why a moral hazard effect may not

[1] In the automobile insurance example, drivers have private information about their personal driving skills. However, if a large majority (mistakenly) thinks their driving is extraordinarily safe compared to others, the private information about risk is less accurate. This inaccurate private information may explain why the correlation between coverage and risk is not universally present across insurance markets (Cohen and Siegelman, 2010).

occur, for example, when people who are highly risk averse purchase insurance and also take other measures to limit risks. This has been demonstrated in the U.S. long term care insurance market, where individuals with more insurance coverage were on average not higher risk (Finkelstein and McGarry, 2006).

Previous work in the domain of natural hazards found no moral hazard effect using statistical methods to analyze survey data of flood insurance coverage and the implementation of flood risk reduction measures in Germany and the United States (Hudson et al., 2017). The empirical analyses by Carson et al. (2013) have found no evidence for a substitution effect between self-insurance and market insurance to protect homes in Florida against storms. Petrolia et al. (2015) surveyed homeowners in coastal areas of the United States and found no moral hazard effect either: the same respondents who buy wind insurance also invest more in wind risk mitigation. Likewise, Osberghaus (2015) showed that German individuals who think they have flood insurance coverage are also more likely to invest in flood risk mitigation measures. While the high external validity of field survey data is very valuable, the disadvantage of this type of empirical research is that it is hard to find causal relationships, as different insurance plans are not allocated randomly to homeowners. Moreover, these survey studies were not able to identify the behavioral mechanisms that may explain why a moral hazard effect was absent (Hudson et al., 2017).

3.2.2 Features of insurance

The moral hazard problem is often dealt with by shifting part of the risk to the policyholder, for instance by introducing a deductible. A deductible is thus an example of a financial incentive to stimulate self-insurance. Furthermore, it has been proposed that risk-based premiums could encourage investments in self-insurance by offering premium discounts to policyholders who limit flood risk to their property (Kunreuther, 1996; Botzen and van den Bergh, 2008; Kleindorfer et al., 2012). Such incentives towards self-insurance are common in health insurance, for example when policyholders are stimulated by financial incentives to increase physical activity or quit smoking (see Tambor et al. (2016) for a review). Previous empirical research based on hypothetical stated preference survey data suggests that a premium discount may affect homeowners' decisions to invest in low cost flood mitigation measures (Botzen et al., 2009b).

A higher level of control can be accomplished by a (quasi-)experimental design. So far, little experimental research has been conducted on incentives for individual damage reduction in a flood insurance context, which is characterized by low probabilities and high expected damages. An exception is Mol et al. (2020a), who studied the impact of different mandatorily public flood insurance schemes and related financial incentives on risk reduction behavior in a controlled lab experiment with mainly students as participants (N=357). The results showed that investments in damage reduction increased with higher probabilities of loss, higher deductibles and a premium discount. Interestingly, moral hazard was found to be less of a problem in the scenarios with low probabilities of loss. Although this design had a high degree of control, one drawback is that students are not representative of the decision makers in the flood insurance context. For example, students are inexperienced with the purchase of homeowners insurance and their individual characteristics (such as risk attitudes and time preferences) may differ from the population. Moreover, Mol et al. (2020a) did not examine self-insurance in the context of voluntary private insurance, like we do here.

3.2.3 Behavioral motivations for self-insurance

A commonly examined behavioral motivation to decide upon precautionary action in general is risk attitude. For example, Cutler et al. (2008) analyzed the relationship between risk aversion, risk reducing activities and insurance purchases in five different types of insurance markets. The authors demonstrated that less risk reducing behavior was associated with lower insurance take-up and argue that this is due to low risk aversion. More recently, Corcos et al. (2017) examined the premium sensitivities in demand for insurance, both theoretically and experimentally. They found that an increase in premiums causes risk loving subjects to leave the market, while the conditional demand (the level of coverage demanded) does not change. Their careful examination of the risk loving types indicated that this behavior is related to gambling and opportunism. In the context of natural disaster insurance markets, Hudson et al. (2017) provided evidence that individuals with insurance-coverage in these markets were more likely to have undertaken disaster preparations, although the role of risk aversion was not examined directly in that study.

Considering that self-insurance in our flood risk context is often a large lump-sum investment with expected benefits spread over a time-span of about 25 years into the future, time preferences might also influence the decision to self-insure (see e.g. Michel-Kerjan, 2010; Kunreuther and Michel-Kerjan, 2015). Other behavioral motivations are more focused on the self-insurance measures themselves, such as response efficacy, response cost and self-efficacy of these measures, where the latter refers to the subjective feeling of being able to install the measures in practice. Grothmann and Reusswig (2006) showed that coping appraisal, and in particular a combination of high response efficacy, low response costs and high self-efficacy, positively influences precautionary action against flooding.

An interesting behavioral motivation for preventive behavior is the psychological construct internal locus of control, which refers to the trade-off between one's own efforts and external factors (e.g., fate) in determining life outcomes. Individuals with an internal locus of control feel more inclined to take protection in their own hands. Locus of control has been shown to impact hurricane preparedness in the U.S. (Sattler et al., 2000), but also in preventive health behaviors (Conell-Price and Jamison, 2015). Furthermore, investments in self-insurance could be motivated by emotional factors, such as high worry of flooding (Bubeck et al., 2012) and anticipated regret about not prevented or uninsured losses (Krantz and Kunreuther, 2007).

Finally, the behavior of others may be an important behavioral motivation to take action against flood risk (Van der Linden, 2015). Social norms concern expectations of what others think one should do ('prescriptive social norms'), what others would approve ('injunctive social norms') or what is typically done ('descriptive social norms') (Cialdini and Goldstein, 2004). Social norms have been shown to have a positive influence on behavioral intentions across domains (Doran and Larsen, 2016; Nyborg et al., 2016) and with the visible construction works to flood-proof a house, individuals might well be influenced by their personal environment (e.g. family, friends, neighbors) to invest in self-insurance themselves.

3.3 Experimental design

Individual flood preparedness decisions may be largely influenced by individual risk attitudes and perceptions that are specifically related to flooding, such as worry about floods, flood risk perceptions due to climate change, social norms regarding mitigation measures and response efficacy of these measures. Such perceptions may differ substantially between student samples and homeowners in flood-prone areas. To measure the effects of these behavioral motivations for flood risk reduction, a large sample of inhabitants of flood-prone areas is needed. A large sample size also allows for an analysis of investments in risk reduction in (voluntary) market insurance, as it is expected that a small fraction of participants are willing to pay the premium for insurance against low probability flood risk. Selection into private market insurance might be affected by the anticipated behavioral response to insurance; risk and/or loss averse individuals with a high risk perception who expect to claim more under insurance coverage might be willing to pay more for insurance coverage (Einav et al., 2013). Such individuals may also invest more in risk reduction measures, even if they have insurance coverage. A treatment with voluntary insurance would allow for a comparison between self-insurance decisions of mandatory insured individuals and voluntarily insured individuals. Preferences for insurance, risk tolerance and private information about risk could contribute independently to the decision to self-select into insurance (Cutler et al., 2008). Relating the individual characteristics of these voluntarily insured people helps to understand why some cautious people insure and perhaps also take other measures to reduce risks, while others do not insure nor reduce risk at all.

We intended to do a large experiment to examine homeowners' investments in damage reduction under different insurance conditions (exogenous variation) and behavioral characteristics (endogenous variation). However, due to large travel costs and higher incentives to convince individuals to participate, it would be very costly to invite large groups of homeowners to the lab. Moreover, a selection effect might be unavoidable with such a lab experiment, when the type of participants (those willing to travel) is related to one of the individual variables of interest. To address these concerns, a short experiment was embedded in a survey and conducted online. The survey consisted of 30 questions that examined flood experience, flood risk perception, response efficacy of mitigation measures, social norms with regards to flood protection, related insurance purchases and demographic data.

The survey questions were based on surveys about flood risk perceptions and flood preparedness decisions in Canada, Germany, the U.S. and the Netherlands (Thistlethwaite et al., 2018; Bubeck et al., 2013; Botzen et al., 2015, 2009b). While risk and time attitudes may be measured with incentive-compatible experimental tasks, these tasks are often too costly and complex to perform in surveys among a large, representative sample. Recent studies have addressed this problem by investigating the predictive power of qualitative survey items that elicit risk and time attitudes on behavior in paid real-stakes lotteries in representative and cross-cultural samples (Dohmen et al., 2011; Vieider et al., 2015). These studies found that the (non-incentivized) survey measures have approximately similar descriptive power in explaining risk and time preferences compared with the incentive-compatible experimental tasks. Furthermore, recent evidence indicates that the survey measure of risk attitudes correlates with risky behavior outside the lab, such as geographical mobility and occupational choice (see e.g. Fouarge et al., 2014; Bauernschuster et al., 2014). As we faced similar time and complexity constraints as other surveys, we adopted the qualitative survey instruments of Falk et al. (2018) to assess risk and time preferences in our survey. The survey question used to elicit risk attitudes was “In general, are you a person who is willing to take risks?” and the answers ranged from 0 (= completely willing) to 10 (= completely unwilling). The question used to assess present biased time preferences was “In general, are you willing to give up something now in order to profit from that in the future?” where the answers ranged from 0 (= completely willing) to 10 (= completely unwilling). In addition, we used the number of insurances held by a respondent[2] as a proxy variable for risk aversion in the insurance domain. For instance, Botzen and van den Bergh (2012) find that the number of insurance held by Dutch homeowners positively relates to their demand for flood insurance. The self-reported voluntary health insurance deductible is included as a proxy variable for risk seeking attitudes in the insurance domain. In the Netherlands, citizens have a mandatory deductible of €385 per year for their health insurance. Beyond this mandatory deductible, individuals may opt for an additional voluntary deductible of €100, €200, €300, €400 or €500

[2] Continuous variable. Total number of boxes checked in the question “Which insurance(s) do you hold at the moment?” (Appendix 3C, question 17).

in exchange for a premium discount. A voluntary health insurance deductible might indicate risk seeking in the insurance domain (Dillingh et al., 2016).

A clear advantage of these revealed preferences questions is that they involve real life outcomes with high stakes. A potential drawback is that these insurance decisions may be affected by other factors, which may lead to unobserved heterogeneity in preferences. A detailed overview of all other questions used in the statistical analysis (including their coding) can be found in Appendix 3A.

The investment game was a simplified and translated version of the lab experiment from Chapter 2 (Mol et al., 2020a) and was embedded in the middle of the survey questions. The currency used in the investment game was ECU (Experimental Currency Units). All respondents were paid a fixed participation fee of 62,000 points[3] (equivalent to approximately €1), while one participant was randomly selected for a large payment. This payment corresponded to the participant's bank balance at the end of the main scenario at a conversion rate of 100 ECU = €1, which could be up to €650. The online experiment was preregistered.[4]

3.3.1 Investment game

In the investment game, respondents were asked to imagine owning a house in a floodplain for the next 25 years[5] and a savings balance of 65000 ECU. All payments in the game were subtracted from this balance. A scenario started with instructions (see Online Supplementary Material) and the introduction of the parameters: the yearly flood probability (1%), the maximum damage to the participant's house in case of flooding (50,000 ECU), the savings balance (65000 ECU) and whether flood risk insurance was available ("No"/"Yes, with 5% deductible"). The parameters were based on net present value (NPV) calculations similar to the simulations in Chapter 2, by adjusting the 12-round investment game to a one-shot version. Some parameters remained the same (e.g. maximum damage 50,000 ECU) and others changed slightly (e.g. savings balance 65,000 ECU instead of 75,000 ECU).

Table 3.1: Investment options in ECU.

Investment	0	1000	5000	10,000	15,000
Reduced damage	50,000	45,242	30,327	18,394	11,157
Discount on yearly premium[6]	0	49	190	304	373

[3] These points refer to the currency of the survey company and they are not related to our experimental currency units.
[4] See the AEA RCT Registry entry: https://www.socialscienceregistry.org/trials/2966/.
[5] In 25 years, most flood damage mitigation measures are cost-effective, see Poussin et al. (2015)

Figure 3.1 shows the first page (*Investment*) of a scenario: respondents could choose to invest in damage reduction measures with accompanying benefits in terms of a reduced damage from flooding and a premium discount in case they are in the Premium Discount treatment (see Table 3.1). Next, the *Pay premium* page was shown to individuals in the Insurance treatments: here the actuarially fair premium was (automatically) paid from their savings balance for all 25 years at once. The *Flood risk result* page showed a grid with 100 houses, where the house of the participant was indicated with a square. All houses flooded (according to the yearly 1% flood probability) at least once in the 25 years of the scenario were highlighted in blue. In case a participant's house was one of these, the deductible (or damage in the No Insurance treatment) was paid from the savings balance. Finally, the *Overview of results* page showed the history of the savings balance (65000 ECU - premiums - deductible/damage - investment). The scenario covered 25 years, but decisions were made only once to facilitate a short and simple version of the investment game, suitable for our consumer panel participants.[7] An additional advantage of this setup is that it corresponds to the long lifetime of many flood risk mitigation measures, which has been estimated to be between 10 up to 50 years (Poussin et al., 2015). This lifetime of about 25 years means that once the measure is taken by a homeowner, it would be present in their house and reduce the flood risk over this lifetime, which is consistent with the setup of our experiment. We acknowledge that the current design does not capture learning over time, while in practice decision makers are able to observe peers and experience potential losses. The instructions were supported by graphics and were always available as a pop-up screen throughout the experiment.

The investment game started with a test scenario to allow participants to be come more familiar with the decision screens. To ensure the participants' understanding of the game and the savings balance, the test scenario was followed by a few comprehension questions, conditional on the treatment (see Appendix 3B). The answers were available in the pop-up instructions. The number of times these pop-up instructions were opened was stored by the software, as well as the number of attempts to answer the comprehension questions correctly. These counts were used as experimental control variables in the regression analysis. After answering all comprehension questions correctly, subjects could start with the main scenario.

---

[6] Only in the Premium Discount and Voluntary + Discount treatments.

[7] In the lab experiment in Chapter 2 (Mol et al., 2020a), participants played the investment game for multiple years (experimental rounds). While this design allowed us to study the effect of flood damage experience on mitigation investments, it was rather complex and repetitive for participants. We anticipated that the consumer panel participants in the current study might be irritated or get bored when being asked to make their choice repeatedly, which could lead to lower completion rates and erratic choices.

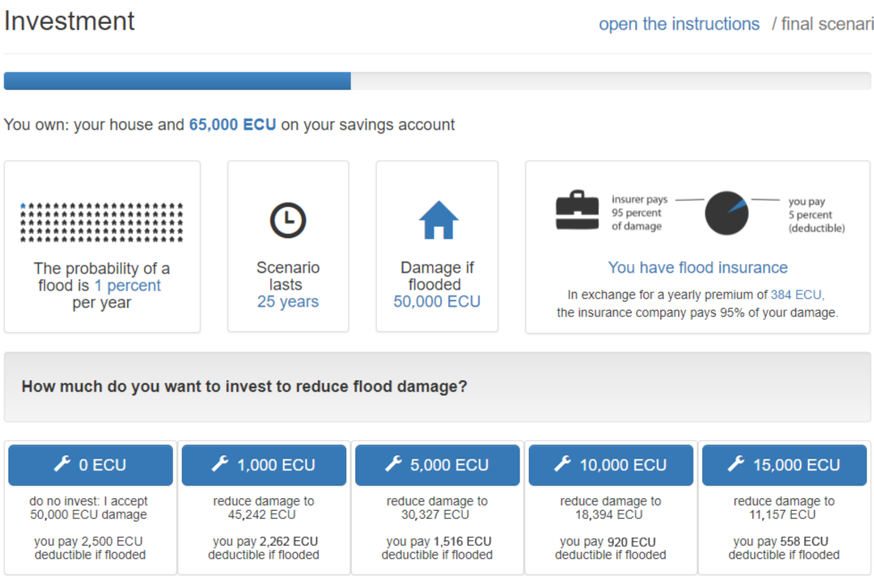


Figure 3.1: Screen shot of the investment page.

3.3.2 Treatments

Each respondent was randomly selected by the software into one of the five treatment groups: No Insurance, Mandatory Insurance, Premium Discount, Voluntary Insurance and Voluntary + Discount (see Table 3.2 for details). Respondents in the No Insurance treatment played the Investment game without insurance. Respondents in the Mandatory Insurance and Premium Discount treatments played the Investment game with mandatory insurance coverage at a premium of 384 ECU per year.[8] Respondents in the Voluntary and Voluntary + Discount treatments were asked whether they would be willing to buy flood insurance (deductible: 5%) at the actuarially fair premium of 480 ECU per year (40 ECU per month). The *Willingness to pay* page showed the yearly costs, as well as the monthly costs and the total costs for 25 years of insurance (see Online Supplementary Material for screen shots). The willingness to pay (WTP) was not restricted. Subjects gave answers between 0 and 150 ECU per month (see Figure 4). The scenario lasted for 25 years: total costs to spend on insurance were 25 × 12 × WTP. Participants were informed that monthly insurance costs were constant over the 25 years. For the example

[8] The actuarially fair premium of 480 ECU was slightly subsidized to increase the sample of voluntarily insured respondents. Besides, subsidizing the premium is a realistic assumption under a mandatory insurance scheme which are often public insurance systems, such as the National Flood Insurance Program in the United States.

of 32 ECU (the subsidized premium) the total costs would be $25 \times 12 \times 32 = 9600$ ECU, which would equal 96 euro. Those who agreed to the actuarially fair premium were insured for the rest of the game, while those who refused were asked again at a subsidized premium of 384 ECU per year (32 ECU per month). Respondents who agreed to the subsidized premium were insured for the rest of the game. Individuals who rejected the insurance offer again were forwarded to the No Insurance treatment of the investment game. After the binary insurance take-up question(s), an open-ended question followed to ask for the exact maximum willingness to pay. To facilitate comparisons across treatments, all respondents insured in the investment game (Mandatory, Voluntary agreed to actuarially fair premium, Voluntary agreed to subsidized premium) were confronted with the same - subsidized - premium of 384 ECU per year. In the Premium Discount treatment, respondents were offered a premium discount that equals the expected value of the damage reduction (probability $\times$ damage) of their self-insurance investment. The optimal investment in self-insurance based on simple expected value calculations was 0 ECU in the Insurance treatments, 1000 ECU in the No Insurance treatment and 5000 ECU in treatments with Premium Discount.

A sample size analysis assuming a significance level of 0.05 and a power of 80% indicated that we would need a sample size of at least 252 participants in the Mandatory Insurance and No Insurance treatments. This sample size would allow us to detect the effect sizes found in the scenario of the lab experiment in Chapter 2 (Mol et al., 2020a), closest to our current parameters, with a Wilcoxon Mann Whitney test. All treatment groups received versions of the survey that shared the same structure, starting with socioeconomic questions and flood perception questions (see Appendix 3C). The investment game was followed by a final set of questions (see Appendix 3D) to gather data on risk preferences, time preferences and other behavioral factors that could be important characteristics related to flood risk, such as flood experience and trust in dike maintenance. Figure 3.2 gives an overview of the flow of the experiment, starting from each of the five treatments.

3.3.3 Procedure

The experimental part of the survey was a simplified version of the lab experiment in Chapter 2 (Mol et al., 2020a) which was extensively pretested by 25 participants and completed by 357 participants in November 2017. The current set-up was pretested with flood hazard experts at the Institute for Environmental Studies (IVM) and a sample of 10 Dutch homeowners. After the pretest, a few minor adjustments were made in the formulation of the survey questions and the instructions of the investment game. The response rate of the final survey was 25.3%. To determine the optimal sample size for each of the treatments, we ran a pilot with a sample of 100 respondents in the Voluntary Insurance treatment to determine voluntary insurance take-up

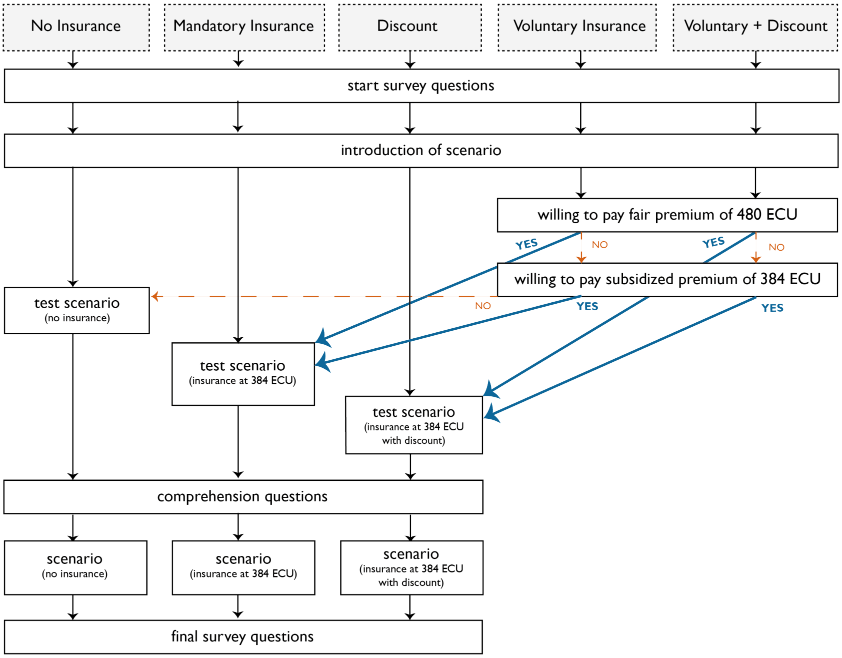


Figure 3.2: Overview of the experiment, by treatment.

rates. 74 out of 100 individuals indicated they were unwilling to pay at least 384 ECU for insurance; they played the No Insurance version of the game. The residual 26 individuals selected into insurance.

The Dutch online experiment was distributed by the survey company Panelinzicht in May and June 2018 and was completed by 2122 unique respondents. Eight responses were deleted because of missing answers in the final survey. Three responses were excluded because of unreasonable outliers in WTP value: monthly premiums above 216 ECU could not be paid from the bank balance. This left 2111 responses for analysis (see Table 3.2 for details). The sample specifically targeted homeowners who were located in the river delta areas of the Netherlands with a flood probability standard of 1 in 1250.[9] The survey was administered over the Internet using the experimental software oTree (Chen et al., 2016) and started with a selection question to ensure that only respondents who owned a house in the river delta zip-code areas could

[9] We could sample 1846 responses in the dike rings corresponding to the 1:1250 protection standard. We sampled the remaining 265 responses from the zip-codes of the 1 in 2000 flood probability standard. We ran additional analyses without these 265 responses. The results do not change qualitatively. A dummy for sample area has been included in the regression analyses.

Table 3.2: Implementation of treatments.

	No Insurance	Mandatory Insurance	Discount	Voluntary	Voluntary + Discount	Total
Mandatory No Insurance	261	0	0	0	0	261
Mandatory Insurance	0	300	0	0	0	300
Mandatory Discount	0	0	351	0	0	351
Self-selected No Insurance	0	0	0	439	411	850
Self-selected Insurance	0	0	0	159	0	159
Self-selected Discount	0	0	0	0	190	190
Total	261	300	351	598	601	2111

Notes: This table shows the distribution of treatments and the number of observations.

continue answering the rest of the survey. The investment game was optimized for tablets and desktop computers.[10]

3.4 Hypotheses

We first consider the case where the availability of flood insurance is publicly determined. A government offers public insurance which has to be mandatorily purchased by its citizens who then face only the (in our case 5%) deductible as expected damage. Alternatively, if no flood insurance is available, citizens face the expected damage of the full loss. Clearly, in this case the uninsured have a higher incentive to invest in self-insurance than the insured.

From a cost benefit analysis perspective[11], the investments of publicly insured individuals in self-insurance should approach zero. However, the combination of very small probabilities of loss and very high potential damages in a natural disaster insurance situation may still lead to investments by individuals with specific behavioral motivations, like high risk aversion or high loss aversion and probability overweighting in Prospect Theory. Previous survey studies in the context of low probability disaster risks have found no evidence for a moral hazard effect (Thieken et al., 2006; Osberghaus, 2015). Therefore, our first hypothesis concerns the non-existence of moral hazard:

Hypothesis 3.1a *Investments in self-insurance in the Mandatory treatments do not differ between individuals with insurance coverage and without insurance coverage.*

Hudson et al. (2017) suggest that natural disaster insurance markets may give rise to advantageous selection; some individuals both purchase insurance

[10] A warning was given to all participants attempting to start the survey from a mobile device. Mobile device users were not excluded from taking the survey, but the software saved browser details of each respondent to control for mobile devices in the analyses.

[11] With the benefits being the expected value of avoided flood damage.

coverage and take available protective measures. However, advantageous selection is very hard to test empirically as it is often not possible to control for behavioral characteristics between the self-selected and the mandatory insured policyholders. The current large-scale online experiment intended to fill this gap with different between-subject treatments with mandatory (public) insurance and voluntary (private market) insurance. We hypothesize that advantageous selection leads to higher investments in self-insurance in the voluntary insurance treatments in comparison to respondents with mandatory insurance, and no insurance.

Hypothesis 3.1b *Self-insurance investments from individuals self-selected into Insurance are higher than those from individuals in the Mandatory Insurance treatment.*

Hypothesis 3.1c *Investments in self-insurance are higher for people who select into purchasing voluntary private insurance than for people who choose not to insure.*

In order to design an affordable insurance scheme for natural disasters and encourage the taking of cost-effective risk reduction measures, researchers and policymakers have suggested premium discounts to promote individual investments in protective measures (Kunreuther, 1996; European Commission, 2013; Surminski et al., 2015). Some empirical evidence suggests that premium discounts might be effective in convincing homeowners to invest in flood mitigation measures of low cost (Botzen et al., 2009b). These initial findings were supported by Chapter 2 Mol et al. (2020a) with a student sample.

Hypothesis 3.2a *Average self-insurance investments are higher in the Discount treatments compared to investments in the Insurance Baseline treatments.*

The current design also allows to test for an interaction effect between voluntary insurance and premium discounts. We expect that because of behavioral characteristics of individuals selecting into voluntary flood insurance (e.g. high risk aversion, high risk perception), those individuals are already more motivated to invest in flood risk reduction measures. Hence, the additional positive effect of the insurance premium discount in terms of stimulating risk reduction measures is less strong for this sub-group compared with the mandatory insured group: we hypothesize a larger effect of the premium discount in the Mandatory Insurance treatment.

Hypothesis 3.2b *The effect of a premium discount on investments in self-insurance is larger for respondents with mandatory insurance than for respondents who self-selected into insurance.*

We now turn to several behavioral motivations to invest in self-insurance. An important motivation to invest in self-insurance is risk aversion. Following the literature summarized in Section 3.2, we expect that respondents with a high willingness to pay (WTP) for flood insurance as proxy for risk aversion in the flood risk domain are more likely to invest in self-insurance.

Hypothesis 3.3a *Risk-averse individuals will invest more in self-insurance than risk-neutral individuals, while risk-seeking individuals will invest less.*

As the expected benefits of a large self-insurance investment may spread over a time-span of 25 years or more, time preferences might be an important factor in the decision process (Michel-Kerjan, 2010; Kunreuther and Michel-Kerjan, 2015). When individuals place too much value on current costs, they might neglect the future benefits of self-insurance investments.

Hypothesis 3.3b *Individuals with present-biased time preferences will invest less in self-insurance than individuals who report neutral time preferences.*

Furthermore, a vast body of literature in both psychology and economics has shown that emotions can influence economic decisions (see e.g. Lerner et al., 2004; Lin et al., 2006; Hanley et al., 2017). A relevant emotion in the context of protective behavior is worry (see e.g. Slovic, 2010; Peters et al., 2006). Schade et al. (2012) conclude from a large insurance experiment with LPHI risks that worry explains more variation in WTP for insurance than the subjective probability of loss. Meyer et al. (2013) also study the role of worry in a computer-mediated environment with a simulated storm. They find that those subjects with the highest levels of worry are the fastest to gather information and indicate the intention to take protective action. Previous survey studies have shown that positive relationships exist between worry about flooding and perceived flood probabilities and damages (Botzen et al., 2015) as well as flood risk mitigation activities (Bubeck et al., 2012).

Hypothesis 3.3c *Individuals with high levels of worry about flooding will invest on average more in self-insurance than individuals who do not worry.*

Some researchers have argued that social norms are positively related with flood insurance purchases (Lo, 2013). Moreover, both descriptive and prescriptive norms have been found to influence risk perceptions of climate change such that individuals with peers who recognize climate change, have higher climate risk perceptions (Van der Linden, 2015). Others have found no support for the impact of social networks and social norms on risk mitigation decisions and flood insurance demand (Harries, 2012; Poussin et al., 2014). The final survey of the current study contains a question about investments in the social network, prescriptive norms as well as injunctive norms.

Hypothesis 3.3d *A higher level of approval concerning self-insurance investments by peers increases self-insurance investments.*

A different emotion that has been shown to affect preventive behavior is anticipated regret about facing a large loss that could have been prevented (Braun and Muermann, 2004). Anticipated regret could increase all types of protective investments (Krantz and Kunreuther, 2007), including investments in self-insurance.

Hypothesis 3.3e *Individuals who anticipate regret about not preventing flood losses will invest on average more in self-insurance than individuals who do not anticipate regret.*

Our large sample size and extensive final questionnaire allows us to take a closer look at the individuals who drive this potential advantageous selection effect. Traditionally, a combination of insurance and preventive behavior - defined here as cautious types- has been explained by risk tolerance preferences. In their seminal paper, de Meza and Webb (2001) argued that people do not have identical (risk) preferences with regards to the risks they are exposed to. Cautious people may prefer both insurance coverage and self-insurance, while 'bold' types prefer less of both. Talberth et al. (2006) found advantageous selection in an experiment in the context of wildfire risks. One other influential factor in their findings was response efficacy of mitigation measures. Fang et al. (2008) have examined the origins of advantageous selection in the context of health insurance, where they found no effect of risk preferences. They do find that education level, cognitive ability and financial numeracy are important predictors of advantageous selection.

Hypothesis 3.3f *Cautious types express higher levels of risk aversion, are more highly educated, and perceive self-insurance measures as more effective than non-cautious types.*

3.5 Results

In this section we present the experimental findings. The main outcome of interest is the discrete level of investment in self-insurance. In addition, we analyze willingness to pay (WTP) for flood insurance by participants in the Voluntary treatments. We first present descriptive statistics and aggregated treatment effects of insurance and insurance features. This is followed by an Ordered Probit estimation to analyze the effects of behavioral motivations and the interactions with incentives on self-insurance investments.

Table 3.3 presents some descriptive statistics of our sample. Demographic variables are largely identical in each treatment group, except for small

Table 3.3: Descriptive statistics per treatment group.

	No Insurance	Mandatory Insurance	Discount	Voluntary	Voluntary + Discount	Total
Gender	0.53	0.43	0.48	0.50	0.49	0.49
	(0.50)	(0.50)	(0.50)	(0.50)	(0.50)	(0.50)
Age in years	52.88	54.49	53.50	55.34	53.96	54.22
	(14.97)	(15.24)	(14.66)	(14.43)	(14.05)	(14.56)
Education	0.10	0.11	0.10	0.10	0.09	0.10
	(0.30)	(0.31)	(0.30)	(0.29)	(0.29)	(0.30)
Expensive home	0.04	0.06	0.06	0.07	0.05	0.06
	(0.20)	(0.23)	(0.24)	(0.26)	(0.22)	(0.24)
Nr of insurances	5.44	5.49	5.44	5.51	5.53	5.49
	(2.19)	(2.06)	(1.87)	(1.98)	(2.05)	(2.02)
Browser	0.15	0.10	0.11	0.10	0.16	0.13
	(0.36)	(0.30)	(0.31)	(0.31)	(0.37)	(0.33)
Observations	261	300	351	598	601	2111

Note: Table displays means, SD in parentheses. Gender dummy: 1 indicates female. Education dummy: 1 indicates Master degree. Home value dummy: 1 indicates > €500,000. Browser dummy: 1 indicates smartphone.

differences in age and browser type.[12] On average, respondents are 54 years old and approximately 49% are female. The average after-tax household income is the answer category "between €2500 and €2999 per month", which would include the average after-tax household income of homeowners in the Netherlands, namely €2933 per month (Netherlands Statistics, 2018a). The average home value is the answer category "between €250,000 and €299,000", which is close to the average home value in the Netherlands, namely €216,000 (Netherlands Statistics, 2018b).

3.5.1 Presence of insurance

To investigate Hypothesis 3.1a we compared the investment levels in the mandatory treatment without insurance with investments in the mandatory insurance treatment. The results are illustrated by Figure 3.3. A one-sided t-test revealed that the average investment in the Mandatory Insurance treatment was significantly higher than 0 (t = 14.89, df = 299, $p <$ 0.000). In other words, self-insurance and mandatory insurance are not complete substitutes. A Mann-Whitney-Wilcoxon (MWW) test showed that the investments in No Insurance (M_{M-no} = 5099.62) are not significantly

[12] This may be caused by the distribution of respondents into treatments per session, where some sessions were larger than others. We therefore cluster standard errors in the regressions at session level. Note that by 'session' we do not mean a typical laboratory session, but we refer to a wave of participation invitations sent out by the survey company. Most sessions held approximately 100 subjects.

different from Mandatory Insurance (M_{M-ins} = 4743.33, z = 1.137, p = 0.256), indicating no moral hazard effect. Therefore, we find support for 3.1a: investments in self-insurance in the Mandatory treatments do not differ between individuals with insurance coverage and individuals without coverage. To examine Hypothesis 3.1b we compared the investments in self-insurance in

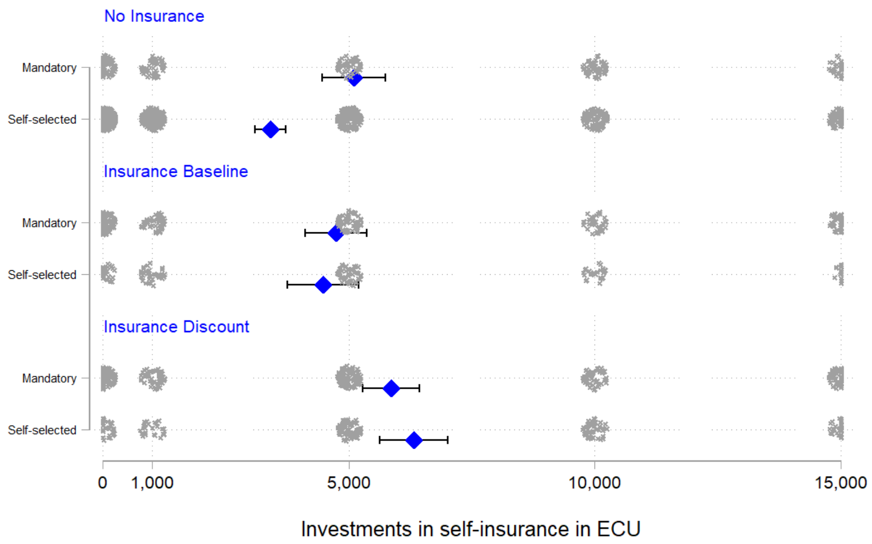


Figure 3.3: Investments in self-insurance over treatments. *Notes:* Each observation is indicated with a gray cross with 4% random jitter to facilitate readability. Diamonds indicate means with confidence intervals.

the Mandatory Insurance treatment with the investments of respondents who self-selected into Insurance (M_{S-ins} = 4477.99) with a MWW test. The results indicate that no significant difference is supported by the data ($z = 0.837$, p = 0.403). The difference between the Mandatory Discount (M_{M-dis} = 5857.55) and self-selected Discount (M_{S-dis} = 6321.05) is not significant at the 5% level either ($z = 1.667$, $p = 0.096$). We do not find support for Hypothesis 3.1b: self-insurance investments from individuals self-selected into Insurance are not significantly higher (nor lower) than individuals in the Mandatory Insurance treatment. In contrast, Hypothesis 3.1c is clearly supported by the data: investments in self-insurance in the self-selected Insurance treatment are significantly higher than in the self-selected No Insurance treatment (M_{S-ins} = 4477.99, $M_{S-no-ins}$ = 3405.88, $z = -4.386$, $p < 0.000$). Note that the probability of loss was equal for all respondents in our experiment. If we consider that risk = probability × damage, individuals with high investments in self-insurance lowered their risk, while individuals with low investments in self-insurance can be classified as high risk. Following this argument, the effect of lower self-insurance (= high risk) by individuals who selected no insurance coverage, indicates advantageous selection.

3.5.2 Features of insurance

We examined the effect of a premium discount both in the Mandatory treatments and in the Voluntary treatments, as well as pooled data across these treatments. We find that a premium discount increases investments under Mandatory insurance ($M_{M-ins} = 4743.33$, $M_{M-ins-disc} = 5857.55$, $z = -3.072$, $p = 0.002$), as well as under Voluntary insurance ($M_{S-ins} = 4477.99$, $M_{S-disc} = 6321.05$, $z = -3.715$, $p < 0.000$). This pattern is confirmed when the investments in both discount treatments are pooled ($z = -5.109$, $p < 0.000$). We can confirm Hypothesis 3.2a: a premium discount increases investments in self-insurance.

Figure 3.3 shows that the effect of a discount is slightly larger for individuals with self-selected insurance coverage than for the mandatorily insured respondents. To analyze this result more formally, we ran regressions with treatment dummies and other explanatory variables.[13], such as demographics and behavioral motivations for investment in self-insurance in Table 3.4. The models have an Ordered Probit specification to account for the discrete investment options. Model 1 restricts the analysis to the subsample of respondents who were insured during the investment game: i.e. respondents in the Mandatory Insurance and Mandatory Discount treatments, as well as respondents who self-selected into the Voluntary Insurance and treatments. This model confirms our findings from the non-parametric tests concerning Hypothesis 3.2a: the premium discount is effective in increasing self-insurance investments, both in the Mandatory insurance treatment, as well as among respondents who self-selected into insurance.

We ran a Wald test for equality of estimates to test the interaction[14] between the discount and insurance type and found no significant difference ($F(1, 965) = 0.79$, $p = 0.373$). Because the increase in self-insurance by a premium discount does not differ between mandatorily and voluntarily insured individuals, we cannot confirm Hypothesis 3.2b, i.e. there is no evidence that the effect of a premium discount on investments in self-insurance is larger for respondents with mandatory insurance than for respondents who self-selected into insurance.

[13] To rule out issues of multicollinearity, we checked all explanatory variables for high correlations; most were smaller than 0.5, indicating no problematic variables (Field, 2009) For the pair level of worry vs. threshold of concern ($\rho = 0.537$) we included only worry in the model, as this question was directly related to Hypothesis 3.3c.

[14] Null hypothesis: Mandatory Discount - Mandatory Insurance = Self-selected Discount - Self-selected Insurance

3.5.3 Behavioral motivations for self-insurance

Next, we investigate the behavioral motivations to invest in self-insurance against flood risk. Hypothesis 3.3a concerned the risk attitude of respondents as measured by their willingness to pay (WTP) for flood insurance. Respondents in the Voluntary and Voluntary + Discount treatments were asked to specify their monthly WTP for flood insurance. Figure 3.4 shows that a majority (71% of the sample) is not willing to pay at least the subsidized premium, which according to Prospect Theory suggests that many people underweight the flood probability in their insurance decision.

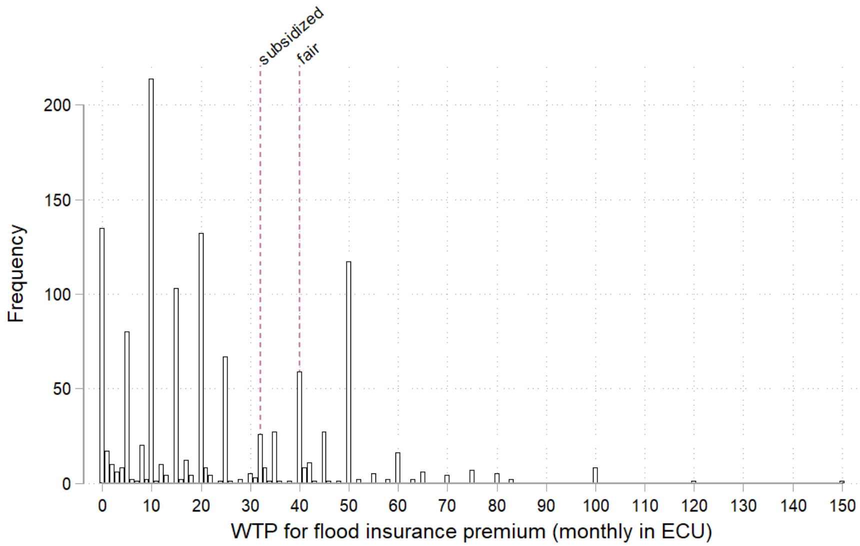


Figure 3.4: Frequency distribution of monthly WTP for flood insurance. *Note:* Dotted lines indicate subsidized and fair premium.

Table 3.4 shows the results of a regression analysis on the effects of behavioral motivations on investment in self-insurance. The table presents treatment dummies, demographics and variables related to our hypotheses (worry, anticipated regret and social norms, risk and time preferences). For the risk and time preferences, we did not classify subjects into 'risk averse', 'risk neutral' or 'risk seeking', but used the reported values for the proxies as predictors in the regression analysis (see Appendix 3A). We suppress coefficients of other flood beliefs and control variables for brevity. Note that we report McFaddens Pseudo R^2, because the R^2 statistic is not defined for our nonlinear (probit) model. In general, Pseudo R^2 statistics of models of flood preparedness decisions are low (Botzen et al., 2009a), indicating large individual differences in factors of influence on these decisions.

The pseudo R^2 values reported in Table 3.4 and Table 3.5 are in a typical range for models with binary dependent variables of flood preparedness decisions reported in other studies (e.g. Hudson et al., 2017; Osberghaus, 2017; Peacock et al., 2005; Botzen et al., 2009b). While Model 1 restricts the analysis to respondents with insurance during the investment game, Model 2 includes only respondents without insurance coverage: i.e. respondents in the Mandatory No Insurance treatment, pooled with respondents who self-selected into Voluntary No Insurance. In this regression we include a dummy variable for the respondents who self-selected to have no insurance coverage.

The significantly negative estimate for this dummy confirms that self-insurance investments by respondents without coverage in a voluntary (market) insurance scheme are lower than in a situation where no flood insurance is available. The third model examines the full sample, but includes WTP for flood insurance as an explanatory variable, which restricts the sample to the subjects who were offered voluntary (market) insurance. The WTP coefficient indicates that investment in self-insurance is positively related to higher WTP values for flood insurance. This WTP variable reflects individual risk aversion for flood risk, but can also capture some other behavioral motivations for reducing flood risk, like anticipated regret for flood damage, like a subsequent analysis reported in Table 3.5 shows. We therefore base our assessment on several indicators for risk aversion. The coefficients of the self-reported general risk aversion question are positive and significant at least at the 5% level across models of investments in self-insurance. The coefficient of number of insurances points into the same direction; for every additional insurance policy in real life, subjects invest more in self-insurance in the game, although the effect is not always significant. Overall, these results suggest that individuals who show a higher level of risk aversion are likely to invest more in self-insurance, which is in line with Hypothesis 3.3a. The self-reported measure regarding time preferences shows that present biased individuals are significantly less willing to invest in self-insurance in the game, as in Hypothesis 3.3b. This may seem obvious, but note that although the time horizon of the investment game describes 25 years, the results are realized within a couple of minutes.

The last model in Table 3.4 includes the full sample, with dummies for each of the treatments, where Mandatory No Insurance is the reference category. We do not find support for Hypothesis 3.3c: no significant coefficient of worry about flood on the average investment in self-insurance is found in either of the four models. We find a positive effect of social norms on investments in self-insurance, confirming Hypothesis 3.3d. However, we need to acknowledge the possibility that subjects answer consistently with their chosen investment level in the experiment[15] as the social norms question was part of the final

[15] See Appendix 3D for the final survey and Appendix 3C for the questions asked before the start of the investment game.

survey. The social norms estimate is not significant in the Insurance only sample (Model 1). For anticipated regret, the regression results indicate that a strong feeling of anticipated regret leads to higher investments, as predicted by Hypothesis 3.3e. Nonetheless, the effect is only significant in the pooled model. Note that as in Chapter 2, we also elicited the regret of investing in case of no flooding. We found no effect of such regret in our analyses in the current chapter. Given that it was not in our hypotheses, we have supressed the results. One reason for the difference in results between the two chapters could be that the multi-year setup in Chapter 2 allows for learning effects, leading to a more pronounced effect of regretting investment when no flood occured (which occurs more often) compared to the one-shot game in this chapter.

Other behavioral motivations In addition to the behavioral motivations which we expected to affect investments in self-insurance, we observe some other important factors in our models. The demographic variables indicate that there is no gender effect, but that more highly educated respondents invest more in self-insurance. Note that this is different from the findings in Chapter 2, where we found that women invest significantly more in the multi-year investment game. All else equal, we find that both older individuals and those who own an expensive home (> €500,000) invest less in self-insurance, although this seems to be mainly the case if no insurance coverage is available. The low investment behavior of older individuals could be explained by the time horizon of 25 years that was presented in the game. As one participant mentioned in the feedback field at the end of the questionnaire: *"If you are 30 years old, the 25 years are within your scope, but I am 71 and that makes me think I will not outlive those investments."*

To understand the determinants of self-selection into insurance coverage, we ran an additional Tobit[16] model with WTP as the dependent variable (Model 1) and a Probit model to predict self-selected insurance coverage (Model 2), which are presented in Table 3.5. To facilitate comparison of coefficient estimates, we used the same set of variables in all four models, even though some variables (such as response efficacy of mitigation measures) mainly intended to explain cautious and uncautious types in Model 3 and 4. We find that risk averse individuals have a higher willingness to pay for flood insurance, as indicated by the self-reported measure. Respondents who decreased their health insurance coverage by raising the deductible in exchange for a lower premium, have a lower likelihood to select flood insurance coverage in the investment game. This may indicate their general dislike of insurance, although there does not seem to be any effect of additional insurance policies. Present biased respondents not only invest less in self-insurance, they also have a lower WTP for flood insurance.

[16] The Tobit model accounts for possible censoring at zero, as respondents were not allowed to enter negative WTP values.

Table 3.4: Ordered Probit regression of investments in self-insurance.

	Dependent variable: Discrete investment in self-insurance			
	(1) Insurance	(2) No insurance	(3) Voluntary	(4) Pooled
Treatments				
Mandatory No Insurance		0		0
		(.)		(.)
Mandatory Insurance	0			-0.088
	(.)			(0.099)
Mandatory Discount	0.234***			0.142
	(0.076)			(0.129)
Self-selected No Insurance		-0.213***	0	-0.262***
		(0.083)	(.)	(0.081)
Self-selected Insurance	-0.043		-0.760***	-0.169
	(0.063)		(0.076)	(0.111)
Self-selected Discount	0.314***		-0.431***	0.168*
	(0.091)		(0.120)	(0.101)
Risk and time preferences				
Willingness to pay for flood insurance			0.024***	
			(0.003)	
Risk averse self-reported	0.040**	0.072***	0.038**	0.056***
	(0.018)	(0.017)	(0.015)	(0.012)
Nr of extra insurances	0.027**	0.013	0.033*	0.020*
	(0.013)	(0.017)	(0.019)	(0.012)
Raised health insurance deductible	0.172	0.076	0.163	0.121
	(0.117)	(0.104)	(0.137)	(0.086)
Present biased self-reported	-0.044***	-0.046**	-0.045***	-0.047***
	(0.011)	(0.021)	(0.016)	(0.014)
Demographics				
Gender (1 = female)	-0.047	0.134	0.040	0.049
	(0.049)	(0.108)	(0.064)	(0.065)
Age in years	0.001	-0.007***	-0.005*	-0.002
	(0.002)	(0.002)	(0.003)	(0.002)
Master's degree	0.279**	0.202	0.272**	0.257***
	(0.109)	(0.149)	(0.119)	(0.084)
Home > €500,000	-0.027	-0.353**	-0.249	-0.175
	(0.162)	(0.165)	(0.161)	(0.116)
Hypothesized flood beliefs				
Worried about floods	0.018	0.036	0.026	0.035
	(0.052)	(0.056)	(0.055)	(0.037)
Social norm approve	0.070	0.110**	0.128***	0.095**
	(0.065)	(0.044)	(0.044)	(0.042)
Anticipated regret	0.043	0.048	0.033	0.048***
	(0.042)	(0.038)	(0.026)	(0.017)
Observations	1000	1111	1199	2111
AIC	3056.1	2999.9	3233.0	6079.3
Log likelihood	-1508.0	-1479.0	-1597.5	-3018.7
Pseudo R^2 (McFadden)	0.044	0.089	0.106	0.069
Controls	✓	✓	✓	✓

Notes: Standard errors clustered at session level in parentheses (* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$). Controls: mobile device, reopened instructions, wrong attempts understanding questions, dummy very difficult, time in minutes, sample area, property includes ground floor. Suppressed coefficients: high income, availability, response efficacy, response cost, self-efficacy, climate risk will increase, subjective flood probability, locus of control, neighbors measures, nr of measures implemented, trust in dikes, high expected damage, house damaged in past.

While we find no gender effect in the previous analyses, men have a higher WTP for flood insurance and are more likely to select coverage in the game. Older respondents have a lower WTP for flood insurance and are less likely to self-select into flood insurance. No significant coefficient estimates were found for education level and home value.

Social norms and anticipated regret increase both WTP and coverage, while worry about floods only increases coverage. Both efficacy variables show there is a positive relation between WTP for flood insurance and response efficacy of mitigation measures, but a negative effect with self-efficacy. These findings suggest that individuals who think that it is effective to invest in flood risk mitigation measures, also have a high demand for flood insurance, but that those who think that implementing mitigation measures is an easy way for coping with floods only mitigate risk. The coefficient sizes show the former effect dominates the latter. WTP for flood insurance is positively related with the number of implemented flood risk mitigation measures, which is consistent with the positive relation between insurance demand and self-insurance observed in the experiment.

Trust in the maintenance of Dutch dikes decreases WTP for flood insurance, but not to such an extent that it decreases coverage in the experiment. The feeling of having control over one's life (locus of control) increases WTP and flood insurance coverage, while the statement that flood risk will increase due to climate change does not have any effect. Interestingly, respondents who are certain that they live in a floodplain area, select significantly less often into insurance coverage than respondents who think they live outside a floodplain area. Note that all respondents do live in a floodplain and that we have controlled for the "real" floodplain where respondents live ("sample area") as well as for past flood experience ("availability"). The fact that respondents' neighbors have implemented damage reducing measures increases WTP for flood insurance slightly, although the coefficient is insignificant in the Probit models. When asked about their strategy in the investment game, many respondents' answers included words like "analyze", "budget", "calculation" and "compare". The answers could be roughly categorized into those who used words related to calculations and those who did not. This dummy variable is strongly significant, indicating that the calculating types have a higher WTP for flood insurance and subsequently select more often into insurance coverage. Interestingly, Figure A1 shows that calculating types did not select the optimal (i.e. maximizing expected value) investment in self-insurance more often than respondents with other strategies. However, calculating types over-invest more and under-invest less than the other types and vice versa.

Cautious and uncautious types Finally, we examine the sources of advantageous selection by a classification of extremely cautious and uncautious types. Out of 1199 subjects who were offered voluntary insurance, 349 selected insurance coverage, of which 287 also invested at least 1000 ECU

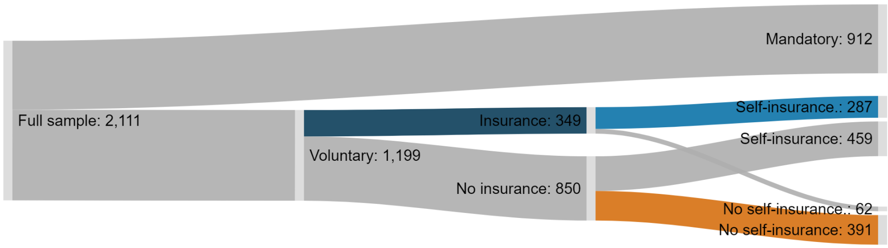


Figure 3.5: Self-selection into insurance and self-insurance.

(the lowest possible non-zero investment) in self-insurance. These respondents are classified as the cautious type. Out of the 850 self-selected non-insured respondents, 391 decided to invest 0 ECU in self-insurance, so we classify this sub-group as uncautious. Figure 3.5 illustrates the proportion of cautious (light blue) and uncautious (red) types.

We analyzed the behavioral motivations of these types through a Probit model of cautious types and uncautious types (Model 3 and Model 4 respectively in Table 3.5). The estimates changing from column 2 to column 3, indicate the difference between only purchasing insurance coverage (dark blue sample in Figure 3.5) and additional investments in self-insurance. Recall that we hypothesized that cautious types are more risk averse, higher educated and perceive self-insurance measures as more effective than non-cautious types. Comparing columns 2 and 3 in Table 3.5), we observe that the estimates of self-reported risk aversion, response efficacy and education level indeed change in the expected direction. Cautious types have higher coefficients for risk aversion, response efficacy and Master's degree as compared to the estimates of respondents with only coverage.

Table 3.5: Regressions on WTP, coverage and types.

	(1) Tobit WTP	(2) Probit coverage	(3) Probit cautious	(4) Probit uncautious
Risk and time preferences				
Risk averse self-reported	2.648***	0.045***	0.043***	-0.036***
	(0.342)	(0.006)	(0.006)	(0.006)
Nr of extra insurances	-0.363	-0.008*	-0.002	-0.001
	(0.253)	(0.005)	(0.004)	(0.006)
Raised health insurance deductible	-0.360	-0.067**	-0.017	0.008
	(1.800)	(0.032)	(0.036)	(0.051)
Present biased self-reported	-1.148***	-0.029***	-0.027***	0.017**
	(0.264)	(0.005)	(0.003)	(0.007)
Demographics				
Gender (1=female)	-2.355***	-0.029*	-0.053***	0.004
	(0.833)	(0.016)	(0.015)	(0.028)
Age in years	-0.223***	-0.004***	-0.004***	0.004***
	(0.034)	(0.001)	(0.001)	(0.001)
Master's degree	-0.341	-0.043	-0.029	-0.073*
	(1.871)	(0.039)	(0.039)	(0.041)
Home > €500,000	-1.218	-0.037	-0.052	0.103
	(2.523)	(0.047)	(0.052)	(0.064)
Hypothesized flood beliefs				
Worried about floods	1.051	0.036***	0.024**	-0.011
	(0.805)	(0.014)	(0.011)	(0.019)
Social norm approve	2.418***	0.030**	0.043***	-0.043***
	(0.757)	(0.012)	(0.013)	(0.014)
Anticipated regret	1.692***	0.036***	0.022	-0.026***
	(0.608)	(0.012)	(0.014)	(0.009)
Other behavioral motivations				
Response efficacy	2.036***	0.020*	0.029***	-0.096***
	(0.564)	(0.011)	(0.011)	(0.011)
Self efficacy	-1.151**	-0.019	-0.0109	0.036***
	(0.467)	(0.013)	(0.012)	(0.012)
Nr of measures implemented	0.851***	0.011**	0.009*	-0.017***
	(0.209)	(0.005)	(0.005)	(0.006)
Trust in dikes	-0.891**	-0.000	-0.017	-0.0219**
	(0.422)	(0.010)	(0.012)	(0.010)
Locus of control	0.581*	0.014*	0.009	-0.026***
	(0.330)	(0.008)	(0.006)	(0.007)
Climate risk will increase	1.050	-0.001	-0.014	-0.052**
	(1.375)	(0.027)	(0.020)	(0.025)
Sure live in flood plain	-2.235	-0.105***	-0.090**	0.039
	(1.533)	(0.038)	(0.041)	(0.030)
Neighbors measures	2.964*	0.065	0.040	0.027
	(1.781)	(0.044)	(0.049)	(0.055)
Calculating strategy	3.650***	0.090***	0.059***	-0.127***
	(1.080)	(0.019)	(0.020)	(0.026)
Observations	1199	1199	1199	1199
AIC	9534.5	1215.3	1112.2	1183.2
Log likelihood	-4748.2	-588.6	-537.1	-572.6
Pseudo R^2 (McFadden)	0.037	0.186	0.186	0.244
Controls	✓	✓	✓	✓

Notes: Marginal effects; Standard errors clustered at session level in parentheses (* $p < 0.1$, ** $p < 0.05$, $p < 0.01$). Additional controls: mobile device, dummy very difficult, sample area, understanding questions. Cautious type defined as: selected both coverage and self-insurance. Uncautious type defined as: selected no coverage and no self-insurance. Suppressed coefficients ($p > 0.1$): response cost, house damaged in past, high expected damage, subjective flood probability, high income, availability.

Additionally, we find that the estimates of injunctive social norms and trust in dike maintenance also change across models. The differences in scores of those five variables are illustrated by Figure 3.6. A lower trust in the Dutch dike maintenance might motivate respondents to take all possible measures to protect their house against water. Education level does not seem to affect cautious behavior. We conclude that cautious types are more motivated by social approval, have higher response efficacy regarding mitigation measures, higher risk aversion and lower trust in dike maintenance than their single' cautious counterparts (who only select insurance coverage), partially validating Hypothesis 3.3f.

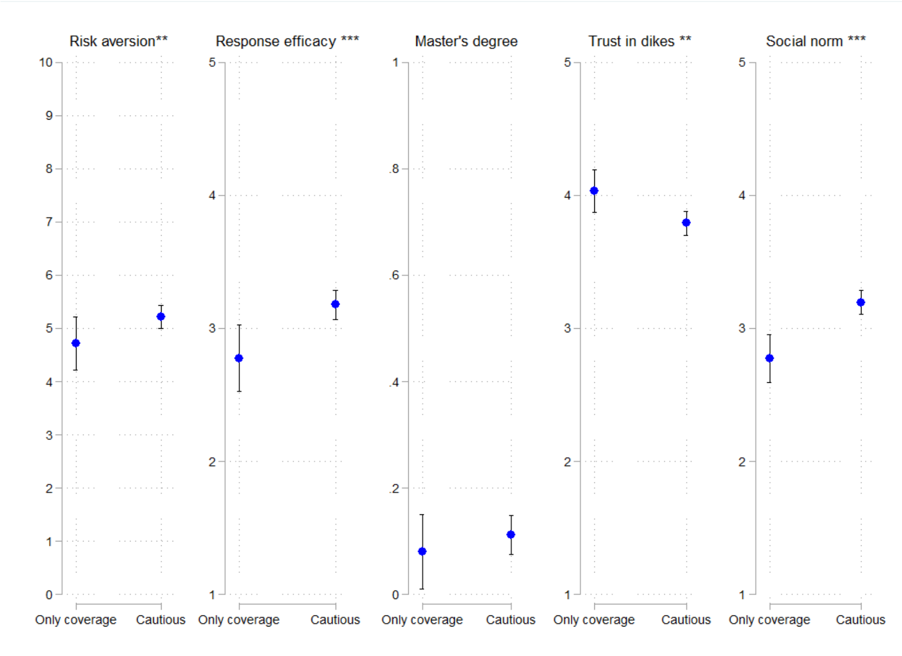


Figure 3.6: Variable (top) means by cautious type (bottom). *Notes:* Stars indicate significant differences by MWW tests ** $p < 0.05$, *** $p < 0.01$.

Following our result of low investments in self-insurance by individuals who self-selected no insurance coverage, we analyzed the uncautious types in Model 4. Although we did not construct hypotheses about this type, we observe some reassuring results: almost all estimates have opposite signs when compared to the cautious types in Model 3. Additionally, we find that uncautious types score significantly lower on trust in Dutch dike maintenance and internal locus of control. They are also significantly less likely to think that flood risk due to climate change is likely to increase. The uncautious types regard damage reducing measures as significantly less effective but also easier to implement (self-efficacy). The dummy for calculating strategy has a strongly significant negative value for uncautious types, while it is significantly positive for cautious types and subjects who select insurance coverage. This suggests that the uncautious types do not make their decision based on calculations, but have more emotional motivations, such as an external locus of control and the feeling that flood risk will not increase due to climate change.

3.6 Conclusion

In response to the growing expected damages of flooding, academics and flood risk managers have recently started to examine different flood risk reduction strategies and cost-effective self-insurance measures in particular. Previous studies have indicated that individual flood preparedness decisions may be largely influenced by individual flood risk perceptions and behavioral motivations (Kunreuther and Pauly, 2004). Empirical research in health insurance markets has indicated that heterogeneity in preferences may explain the appearance of either adverse or advantageous selection (Cutler et al., 2008). This study offered a careful examination of the interplay between financial incentives and behavioral motivations for investing in self-insurance on a group of relevant decision makers (homeowners in floodplains). To the best of our knowledge we are the first to study self-insurance behavior experimentally under both public and private insurance schemes, accounting for insurance features and behavioral characteristics of the decision-makers. Furthermore, our large sample size allowed for an in-depth analysis of heterogeneous behavioral motivations among respondents.

Our analysis started with the impacts of the presence or absence of insurance: we find no support for moral hazard in our data. As expected, we find that a premium discount can increase investments in self-insurance, although it does not matter whether this insurance is provided in a public or private market. A small majority of individuals in the voluntary insurance treatments are not willing to pay the subsidized insurance premium, but we do find a substantial share of cautious types, indirectly indicating advantageous selection. Important behavioral motivations stimulating investments in self-insurance are response efficacy, social norms and risk aversion. When we examine the sources of advantageous selection by a classification of extremely cautious and uncautious types, we find that cautious types tend to take their decision based on some sort of calculation, although the calculating respondents are more inclined to invest more than optimal amounts. These individuals are particularly motivated by response efficacy, social approval by their peers and risk aversion, as well as by a lower trust in dike maintenance. In contrast, uncautious types have opposite motivations and can be characterized by a lower locus of control and the belief that flood risk will not increase due to climate change. Even though all our respondents were floodplain inhabitants, only a minority of subjects stated confidently that their house was located in a floodplain and many did not consider damage reducing measures as cost-effective. Although our design differs in some key points[17] from the experiment of Corcos et al. (2017), it is interesting to compare the results. Our split between cautious and uncautious types suggests that the cautious types make decisions based on calculation, while the split between risk lovers and risk averters of Corcos et al. (2017) indicated strategic gambling rather than a lack of interest in insurance by the risk lovers. A careful examination of the strategic motivations such as opportunism and strategic ignorance of the uncautious types requires further research. The limited length of our survey restricted the explanatory variables to simple survey questions, while it would have been interesting to take a closer look at risk attitudes, by differentiating between utility curvature, probability weighting and loss aversion as in Prospect Theory (Kahneman and Tversky, 1979). Previous research indeed indicates that many individuals underweight the low probability of flooding and that this behavior may be explained by Prospect Theory (Botzen and van den Bergh, 2012; Barberis, 2013). Nevertheless, probability weighting seems to be different for precautionary decisions about real life hazards compared to simple monetary gambles (Kusev et al., 2009). An interesting topic for future research is to examine how loss aversion, utility curvature, and probability weighting can explain individual investments for self-insurance against flood risk.

Regarding policy implications, these results may justify the strengthening of purchase requirements for flood insurance as we found no support for moral hazard and voluntary take-up rates in our experiment are low. Furthermore, the result that the uncautious types (who do not believe that flood risk will increase, nor that they should take action) select less insurance coverage could lead to substantial claims for government support which may drain public resources. These could be important topics for informational campaigns aimed at improving flood preparedness, which should be focused on explaining possible cost-effective measures, rather than on increasing awareness about flood risk in general. Our analysis also indicated that individuals who used calculations in the decision-making process were more inclined to select insurance coverage and (over-)invest in self-insurance. The fact that reporting a calculating strategy does not increase optimal investments may indicate either miscalculation or preferences for over-investment. Further research will have to show whether calculation tools could help to increase investments in cost-effective self-insurance measures among cautious as well as uncautious types. As our results suggest, changing the social norm for self-insurance by means of information and communication measures may be another policy lever to stimulate a wider uptake of these cost-effective measures. Finally, our finding

[17] Due to the simple online set-up in order to achieve a high sample size, we were not able to measure risk preferences with an incentive compatible task, but rely on a general self-reported measure and two insurance related questions instead. The WTP task was part of the incentivized investment game, but it was only present in the Voluntary treatment sample.

that there is no moral hazard in this LPHI insurance market suggests that high deductibles may not be necessary to limit such an effect. This is in line with previous survey results of Hudson et al. (2017) who found that a majority of (hurricane insurance) policyholders are not even aware of having a deductible and that deductibles played a minor role in hurricane preparedness activities. Using premium discounts is likely to be a more effective way for insurers to stimulate policyholders to reduce natural disaster risk in general and flood risk in particular. These results support the ongoing debates and reforms aimed at linking flood insurance coverage with risk reduction in the European Union (Surminski et al., 2015; Hochrainer-Stigler et al., 2017) and the United States (Tullos, 2018).

Acknowledgments

We would like to thank conference participants at ESA World Meeting 2018 and TIBER 2018 for valuable comments. This research has received financial support from the Netherlands Organisation for Scientific Research (NWO) VIDI (452.14.005) grant.

Appendix 3A: Explanatory variables

Table A1: Summary overview of the variables used in the statistical analysis

Risk and time preferences	
Risk averse self reported	Categorical variable (range 0-10) *In general, are you a person who is willing to take risks?*, 0 = completely willing, 10 = completely unwilling
Nr of extra insurances	Continuous variable. Total number of boxes checked in the question 'which insurance(s) do you hold at the moment' (Appendix 3C question 17). Used as a proxy for risk aversion in the insurance domain.
Voluntary deductible	Dummy voluntary health insurance deductible (1 = yes). In the Netherlands, citizens have a mandatory deductible of €385 per year for their health insurance. Beyond this mandatory deductible, individuals may opt for an additional voluntary deductible of €100, €200, €300, €400 or €500 in exchange for a premium discount. A voluntary health insurance deductible might indicate risk seeking in the insurance domain (Dillingh et al., 2016).
Present biased self reported	Categorical variable (range 0-10) *In general, are you willing to give up something now in order to profit from that in the future?* (0 = completely willing, 10 = completely unwilling)
Demographics	
Gender (1=female)	Dummy variable gender (1 = respondent is female)
Age in years	Continuous variable, age in years
Master's degree	Dummy variable education level (1 = holds Master's degree)
High income	Dummy variable income (1 = monthly household after-tax income is > €5,000)
Expensive house	Dummy variable house value (1 = house value is within the highest category > €400,000)
Flood beliefs	
Worried about floods	Categorical variable (range 1-5), *Worried about danger of flooding at current residence* (1 = strongly disagree, 5 = strongly agree)
Social norm approve	Categorical variable (range 1-5), *People in my direct environment would approve an investment in damage reducing measures* (1 = strongly disagree, 5 = strongly agree)
Anticipated regret	Categorical variable, Response to statement *I would feel regret if my house flooded and I had not taken any measures* (1 = strongly disagree, 5 = strongly agree)
Response efficacy	Categorical variable (range 1-5), *How effective do you consider investing in flood protection measures that limit flood damage* (1 = very ineffective, 5 = very effective)
Response cost	Categorical variable (range 1-5), *How costly do you think it is to take flood protection measures?* (1 = very cheap, 5 = very expensive)
Self-efficacy	Categorical variable (range 1-5), *How difficult do you think it is to take flood protection measures?* (1 = very difficult, 5 = very easy)

Nr of measures implemented	Continuous variable, number of flood protection measures already implemented at home
Trust in dikes	Categorical variable (range 1-5), *Dikes in Netherlands are well maintained* (1 = strongly disagree, 5 = strongly agree)
Locus of control	Categorical variable (range 4-20) combined 4 locus of control questions (4 = extremely external LOC, 20 = extremely internal LOC)
Climate risk will increase	Dummy consequences for flood risk at your current residence (1 = flood risk will increase)
Sure live in flood plain	Dummy flood-prone (1 = *I am certain that I live in a flood-prone area*)
Neighbors measures	Dummy respondent knows people who have invested in damage reducing measures (1 = yes)
Calculating strategy	Dummy respondent used words such as 'analyze', 'budget', 'calculation' and 'compare' in answer to open question strategy in the investment game, indicating a calculating strategy (1 = yes)
House damaged in past	Dummy property damaged due to floods in the past (1 = yes)
High expected damage	Dummy high expected damage (1 = respondent expects damage > €50,000 in case of flooding at residence)
Subjective flood probability	Continuous variable, log of estimated flood probability by respondent
Availability	Dummy availability (1 = Yes, I can recall high water levels)
Controls	
Time	Time from the first to the last page in the experiment in minutes
Mobile device	Dummy browser dimensions of respondent (1 = mobile device)
Dummy difficult	Dummy difficult (1 if respondent answered 'difficult' or 'very difficult' to the question *How easy or difficult did you find it to make a choice in the investment game presented to you?*)
Sample area	Dummy sample area (0 = 1:1250 floodplain, 1 = 1:2000 floodplain)
Understanding questions	Continuous variable, number of wrong attempts to answer understanding questions
Property ground floor	Dummy property of respondent includes ground floor (1 = yes)
Reopened instructions	Continuous variable, number of times respondent reopened pop-up screen with instructions

Appendix 3B: Comprehension questions

Correct answers are marked in **bold**.

Question asked in all treatments

- What was the flood risk in the test scenario?
  (a) **1% per year**
  (b) 3% per year
  (c) 5% per year
  (d) 10% per year
  (e) 15% per year
  (f) 20% per year

Extra question in the No Insurance treatment

- What happens if you are flooded and you did not take protective investments?
  (a) **I have to pay the full damage: 50.000 ECU**
  (b) I have to pay a small fee
  (c) The government will compensate me

Extra question in all Insurance treatments

- What was your deductible (eigen risico) in the test scenario?
  a) **5 percent** b) 15 percent c) 20 percent d) 50 percent
- What is the benefit of a protective investment?
  (a) A reduced damage in case of a flood
  (b) A lower premium
  (c) Both reduced damage and a lower premium
  (d) None of the above

The correct answer is:
(a) in Insurance Baseline
and
(c) in Insurance Discount

Appendix 3C: Start survey (translated from Dutch)

1. Are you male or female?
   • *Male* • *Female*

2. What is your age?

3. What is the highest level of education you have completed?
   - *No diploma*
   - *Primary school*
   - *Lower vocational education (VBO, LBO)*
   - *Lower general secondary education (ULO, MULO, VMBO, MAVO)*
   - *Lower vocational secondary education (MBO)*
   - *Higher general secondary education or pre-university education (HAVO, VWO, HBS)*
   - *Higher vocational and university education (HBO, WO Bachelor)*
   - *Master's degree (WO Master)*
   - *Doctorate, PhD (Promotie-onderzoek)*
   - *Other: [text box for open answer]*

4. Do you live in a flood-prone area at the moment?
   - *I am certain that I live in a flood-prone area*
   - *I think that I live in a flood-prone area, but I am not sure*
   - *No, I am certain that I do not live in a flood-prone area*
   - *Don't know*

5. Have you ever been evacuated due to a threat of flooding?
   • *Yes* • *No*

   In case subject answered *Yes* in question 5:

5.a Do you think your experience with evacuation makes it easier to imagine a flood in the nearby future?
   - *Yes, I can now imagine that a flood is very likely*
   - *No, I cannot imagine that a flood is very likely*
   - *I do not think that this experience has changed my thoughts on the likelihood of a flood*

6. Have you ever experienced damage to your house due to a flood?
   • *Yes* • *No*

7. How large or small do you think the probability is that your house will be flooded?
   - *The probability is zero*
   - *Very low*

- *Low*
- *Not low/not high*
- *High*
- *Very high*
- *Do not know*

8. What consequences of climate change for flood risk do you expect at your current residence?
   - *Flood risk will increase*
   - *Flood risk will remain constant*
   - *Flood risk will decrease*
   - *Don't know*

9. Do you recall any situations of exceptionally high water levels in rivers close to your residence?

   - *Yes, I can recall high water levels*
   - *I cannot recall high water levels*

10. Imagine your neighborhood is flooded, how what height do you think the water would reach in your house?
    - *The water would not reach my house*
    - *Low (1-10 cm)*
    - *Pretty high (11-50 cm)*
    - *Fairly high (50-100 cm)*
    - *High (1-2 m)*
    - *Very high (whole floor flooded)*

11. To what extent do you agree with the following statement?
    "I would feel regret if my house flooded and I had not taken measures"

12. What is your household monthly income (after taxes)?
    - *Less than €499*
    - *Between €500 and €999*
    - *Between €1,000 and €1,499*
    - *Between €1,500 and €1,999*
    - *Between €2,000 and €2,499*
    - *Between €2,500 and €2,999*
    - *Between €3,000 and €3,499*
    - *Between €3,500 and €3,999*
    - *Between €4,000 and €4,499*
    - *Between €4,500 and €4,999*
    - *€5,000 or more*
    - *Don't know*
    - *Rather not say*

13. What is approximately the market value of your home?
    - *Less than €100,000*
    - *Between €100,000 and €149,000*
    - *Between €150,000 and €199,999*
    - *Between €200,000 and €249,000*
    - *Between €250,000 and €299,999*
    - *Between €300,000 and €349,000*
    - *Between €350,000 and €399,999*
    - *Between €400,000 and €449,000*
    - *Between €450,000 and €499,999*
    - *Between €500,000 and €549,000*
    - *Between €550,000 and €599,999*
    - *Between €600,000 and €649,000*
    - *Between €650,000 and €699,999*
    - *Between €700,000 and €749,000*
    - *Between €750,000 and €799,999*
    - *€800,000 or more*
    - *Don't know*
    - *Rather not say*
14. What is your postcode in numbers and letters? [18]
15. Please indicate in what kind of property you live.
    - *House*
    - *Ground floor apartment*
    - *Apartment on 1st floor or higher*
    - *Other*
16. How much damage do you expect to your house and contents in case you would be flooded?
    - *Less than €1,000*
    - *Between €1,000 and €4,499*
    - *Between €5,000 and €9,999*
    - *Between €10,000 and €49,999*
    - *Between €50,000 and €99,999*
    - *Between €100,000 and €499,999*
    - *€500,000 or more*
    - *Don't know*
    - *Rather not say*
17. Could you indicate which insurance(s) you hold at the moment?

    ☐ *Dentist insurance*

[18] This answer was not required for privacy reasons.

☐ *Other extra option in health insurance (e.g. physiotherapy, glasses)*

☐ *Home contents insurance*

☐ *House insurance*

☐ *All risk car insurance*

☐ *Continuous travel insurance*

☐ *Life insurance*

☐ *Legal counsel insurance*

☐ *Bike insurance*

☐ *Occupational disability insurance*

☐ *Other: [text box for open answer]*

☐ *None*

18. In your Dutch health insurance, what do you think was your deductible in 2018?
    - *385 euro, the minimum set by the Dutch government*
    - *485 euro, I raised it by 100 euro*
    - *585 euro, I raised it by 200 euro*
    - *685 euro, I raised it by 300 euro*
    - *785 euro, I raised it by 400 euro*
    - *885 euro, I raised it by 500 euro (the maximum)*
    - *I do not know*
    - *I do not have Dutch health insurance*

Appendix 3D: Final survey (translated from Dutch)

1. Can you indicate which measures you have taken to protect your house against flood damage?

   ☐ *No valuables in basement*

   ☐ *Water-resistant furniture on ground floor*

   ☐ *Elevated ground floor*

   ☐ *Strengthened foundation*

   ☐ *Walls made of water-resistant materials*

   ☐ *Floor of ground floor made of water-resistant materials (e.g. tile floor)*

   ☐ *Raised power sockets on ground floor*

   ☐ *Anti-backflow valves*

   ☐ *(Empty) sand bags or flood barriers*

   ☐ *Elevated electrical appliances*

   ☐ *Elevated boiler*

   ☐ *Raised electricity meter*

   ☐ *Bought separate flood insurance*

   ☐ *Other: [box for open answer]*

   ☐ *None*

2. Do you know anyone in your close environment who has taken one or more of these measures?

   • *Yes* • *No*

   In case subject answered *Yes* in question 2:

2.a Could you indicate your relationship to the person who invested in one or more damage reducing measures?

   - *Partner*
   - *Friend*
   - *Parent*
   - *Aunt/Uncle*
   - *Son/Daughter*
   - *Cousin*
   - *Neighbor*
   - *Acquaintance*
   - *Other: [Text box for open answer]*

3. How effective to you consider investing in flood protection measures that limit flood damage?[19]

4. How costly do you think it is to take flood protection measures?

5. How difficult do you think it is to take flood protection measures that limit flood damage?

6. Please tell me, how willing or unwilling you are to take risks if it concerns floods?

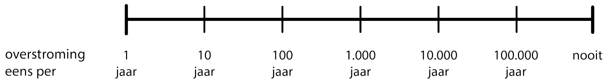
7. How willing are you to give up money today in order to benefit more from that in the future?

8. To what extent do you agree with the following statements?
   (a) I am worried about the danger of flooding at my current residence
   (b) I am confident that the dikes in the Netherlands are maintained well
   (c) I felt regret about not investing in protection when a flood occurred in the game[20]
   (d) People in my direct environment would approve an investment in damage reducing measures
   (e) People in my direct environment think that I should invest in damage reducing measures
   (f) When I get what I want, it is usually because I am lucky[21]
   (g) It is not always wise for me to plan too far ahead because many things turn out to be a matter of good or bad fortune
   (h) I believe that there are a number of measures that people can take to reduce their risk
   (i) I can pretty much determine what will happen in my life
   (j) The probability of flooding at my current residence is too low to be concerned about

[19] This question was taken from Poussin et al. (2014)
[20] If the subject did not experience a flood during the experimental phase, this question was phrased as "When in the scenario no flood occurred, I felt regret about paying for protectio"
[21] These four questions are developed to measure locus of control (see Sattler et al., 2000)

9. The government is responsible for the maintenance of dikes. A dike in your neighborhood should be strong enough such that a flood does not happen more than once each 1250 years. The scale below shows different flood probabilities.

   What is according to you the probability of a flood in your



   neighborhood?

   - *Flood on average once every ... years*
   - *Never*

10. How easy or difficult did you find it to make a choice in the investment game presented to you?

    - *Very easy*
    - *Easy*
    - *Not easy/not difficult*
    - *Difficult*
    - *Very difficult*

    In case subject answered *Difficult* or *Very difficult* in question 10:

10.a Could you describe what made the investment game difficult for you?

11. What is according to you the probability of a cloudy sky in your residence tomorrow?

12. What is according to you the probability of a cloudy sky and rain in your residence tomorrow?

13. Could you briefly explain how you made your decisions in the investment game?

14. This is the end of the survey. If you have comments, you can write them below.

Appendix 3E: Additional analysis

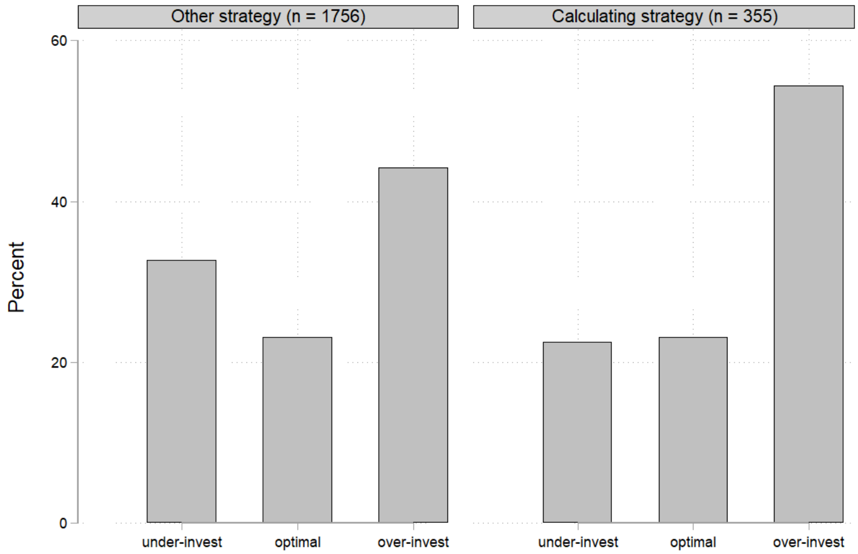
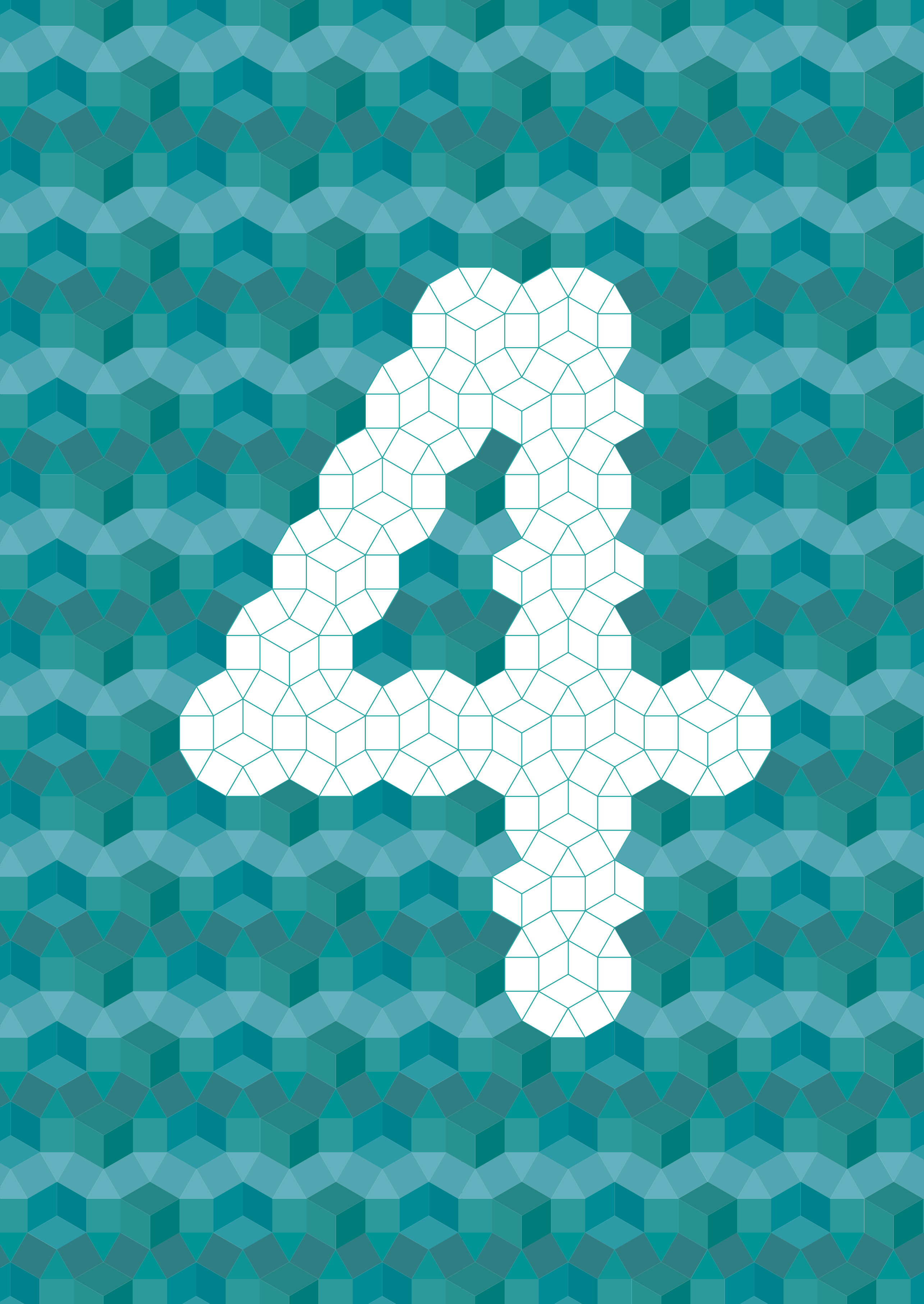


Figure A1: Proportion of optimal and sub-optimal investments, by self-reported strategy.


4

Insights into flood risk misperceptions of homeowners in the Dutch river delta

Flooding is one of the most significant natural disasters worldwide. Nevertheless, voluntary take-up of individual damage reduction measures is low. A potential explanation is that flood risk perceptions of individual homeowners are below objective estimates of flood risk, which may imply that they underestimate the flood risk and the damage that can be avoided by damage reduction measures. The aim of this paper is to assess possible flood risk misperceptions of floodplain residents in the Netherlands, and to offer insights into factors that are related with under- or overestimation of perceived flood risk. We analyzed survey data of 1848 homeowners in the Dutch river delta and examine how perceptions of flood probability and damage relate to objective risk assessments, such as safety standards of dikes, as well as heuristics, including the availability heuristic and the affect heuristic. Results show that many Dutch floodplain inhabitants significantly overestimate the probability, but underestimate the maximum expected water level of a flood. We further observe that many respondents apply the availability heuristic.

This chapter is published as:
Mol, J. M., Botzen, W. J. W., Blasch, J. E., & de Moel, H. (2020). Insights into flood risk misperceptions of homeowners in the Dutch river delta. *Risk Analysis*, 40(7), 1450–1468.

4.1 Introduction

Flooding is one of the most significant natural disasters worldwide in terms of number of people evacuated and total economic damages (UNISDR, 2015). With both sea levels as well as population increasing in flood-prone areas, the impacts of flooding are expected to increase further in the future (IPCC, 2012; Munich RE, 2018). Hence, it is becoming more important to implement flood damage reduction strategies. Recent evidence shows that damage reduction measures taken by private homeowners are cost-effective and can substantially limit the expected damages from flooding (Kreibich et al., 2015). However, current voluntary investments in private flood damage reducing measures are low. A potential explanation is that flood risk perceptions of homeowners differ considerably from objective estimates, which may skew their assessment of the damage that can be avoided by risk reduction measures (Siegrist and Gutscher, 2008; Bubeck et al., 2013). Flood risk perceptions further affect support for public investments in flood protection infrastructure (Ripberger et al., 2018). This leads to a growing interest in risk perception research, which is important for the design of effective risk communication campaigns that stimulate people to better prepare for increasing natural disaster risks (Botzen, 2013; Kellens et al., 2013).

The aim of this chapter is to assess possible flood risk misperceptions of floodplain residents in the Netherlands, and to offer insights into factors that are related with the under- or overestimation of perceived flood risk. We build upon previous studies which have examined flood risk perception in relation to knowledge of the causes of flood events (Botzen et al., 2009a), distance to a perceived flood zone (O'Neill et al., 2016) and climate change information (de Boer et al., 2016). However, a systematic assessment of flood risk misperceptions is lacking for the Netherlands, as well as more generally, as becomes evident from a comprehensive literature review on the topic of flood risk perception by Lechowska (2018). This study takes the analysis of flood risk misperceptions one step further by relating the type of misperception (over- versus under-estimation) to objective risk assessments, heuristics, and personal characteristics. Risk perceptions are an important component of theories of decision making under risk in both economics and psychology. The current chapter examines drivers of risk perceptions from both domains to arrive at a comprehensive assessment of flood risk perceptions.

The Netherlands, with its long history of protection against potentially severe flooding, lends itself as a relevant case to study these relationships. Moreover, the Dutch government has released several informational campaigns,[1] but flood risk perceptions have since not been evaluated. While respondents in our sample have not experienced a flood recently, we examine whether we find similar patterns of risk perception as in the sample of Botzen et al. (2015), where respondents recently survived a major hurricane.

[1] see e.g. www.onswater.nl and www.overstroomik.nl

The chapter is structured as follows: Section 4.2 presents our theoretical framework and hypotheses, Section 4.3 describes the methodology, Section 4.4 presents results and Section 4.5 discusses these results in relation to the literature. Finally, Section 4.6 gives policy implications and concludes.

4.2 Theory and hypotheses

In this section, we discuss several theories and motivate our hypotheses about specific relations between risk perceptions and explanatory variables. Risk perceptions are an important component of theories of decision under risk from both economics and psychology. In economics, theories of decision-making under risk have taken rationality as a starting point. In psychology, the importance of intuitive thinking (System 1) has been stressed, which is defined as fast, automatic and directed by emotional reactions, as compared to deliberative thinking (System 2) which requires more effort to undertake trade-offs. Generally, individuals combine both modes of thinking and they may apply simple rules of thumb (heuristics) whenever the cost of deliberative thinking are perceived too high. Heuristics are quick and straightforward decision rules that can be used to deal with complex decision environments (such as flood preparedness decisions) without draining an individual's cognitive capacities (Tversky and Kahneman, 1973).

4.2.1 Objective risk assessment

An important economic model of individual decision making under risk is expected utility theory (EUT), which assumes that individuals assess the likelihood and consequences of several choice alternatives, and subsequently choose the alternative that gives the highest expected utility (von Neumann and Morgenstern, 1947). When the objective likelihood is uncertain or unavailable, individuals may still maximize expected utility by using their own subjective estimates of probabilities and losses (Savage, 1954), which in our applications are the perceived flood probability and damage.

Kunreuther and Pauly (2004) postulated based on the expected utility framework that individuals facing low-probability/high-impact risks expect a low return from searching for information about their risk, and hence are unlikely to be fully informed about the risk they face. This implies that perceptions of low-probability/ high-impact risks are likely to be biased, but would still be related to the objective risk faced by individuals (Kunreuther and Pauly, 2004). This means that risk perceptions would at least partially relate to objective risk and, hence, the latter may relate to the degree to which people under- or overestimate their risk. Such a heterogeneity in risks is applicable to the Dutch flood risk context, because although flood probabilities are generally low, expected flood inundation depths vary considerable between areas. In line with expected utility theory, we predict that individuals under higher actual flood risk have higher flood risk perceptions.

Hypothesis 4.1a *Respondents who live in an area with a larger flood probability have higher flood risk perceptions than respondents living in an area with lower flood probability.*

Ruin et al. (2007) found that flash flood risk perception (expected damage) among French motorists was higher among those who lived close to the place of impact. In a similar study among Dutch homeowners, Botzen et al. (2009a) found that individuals living close to a river have higher flood risk perceptions. Recent studies have confirmed these findings, both for expected probability (Miceli et al., 2008; Lindell and Hwang, 2008) as for expected damage (Zhang et al., 2010; O'Neill et al., 2016).

Hypothesis 4.1b *Respondents who live closer to dikes have higher flood risk perceptions than respondents who live further away from dikes.*

Generally, we expect that respondents who live in low-lying areas have higher flood risk perceptions than those who live on higher grounds, simply because the houses of the latter cannot be reached by floods and because they will experience lower inundation depths if they are flooded.

Hypothesis 4.1c *Respondents who live in low-lying areas (as indicated by higher maximum water levels) have higher flood risk perceptions than respondents who live on higher grounds.*

4.2.2 Heuristics

A growing body of evidence shows that individuals often do not behave as if they were following expected utility theory; they rather engage in intuitive thinking, using heuristics or simple rules of thumb to evaluate a certain situation (Kahneman, 2003; Slovic et al., 2004). These heuristics are potentially helpful in many situations in daily life, but systematic biases may occur when they are applied to low-probability/high-impact events, causing errors in risk judgments. This may lead to completely ignoring the risk as well as overreacting to a recent disaster (Kunreuther and Michel-Kerjan, 2015). Several systematic biases have been documented in the flood risk domain: in particular, the affect heuristic (Slovic et al., 2004; Keller et al., 2006) and the availability heuristic (Siegrist and Gutscher, 2006).

Loewenstein et al. (2001) noted that affective feelings toward risk, such as worry, are important determinants of risk perception (affect heuristic). However, Sjöberg (2000) argued that it is crucial to distinguish between worry and hazard properties when analyzing risk perception. Sjöberg (2007) showed in three Swedish survey data sets (each $n > 400$) that negative emotions are the strongest predictors of perceived risk. Botzen et al. (2015) surveyed 1035 floodplain residents in New York City and found that high levels of worry were related to a higher perceived flood probability.

Hypothesis 4.2a *High degrees of worry about flooding are related to higher perceptions of the flood probability.*

When people lack objective information about a certain hazard, they might rely on local risk management. Previous research has found that individuals who distrust local risk management have higher risk perceptions of hazardous facilities, such as nuclear waste repositories (Slovic et al., 1991). Terpstra (2011) conducted three Internet surveys among 1071 Dutch households vulnerable to flooding and found that individuals who trust local risk management, expect the probability of a flood to be lower. Also the survey by Botzen et al. (2015) revealed that high trust in flood risk management officials is related to lower anticipated flood damage. We thus expect that trust in flood risk management lowers perceptions of flood probability and damage.

Hypothesis 4.2b *Individuals with a high level of trust in local flood risk management have lower perceptions of the flood probability and damage.*

A related cognitive bias is the availability heuristic, where the probability or frequency of events is judged to be higher when the event is easier to recall (Tversky and Kahneman, 1973). Generally, individuals overestimate the probability of an event if they have experienced it, and underestimate the probability of events they have not experienced before (Siegrist and Gutscher, 2006; Viscusi and Zeckhauser, 2006). A first-hand flood experience may make the flood risk more salient and easier to recall, leading to higher subjective flood probabilities, which is reflected in lower housing prices (Bin and Landry, 2013) and higher insurance take-up (Shao et al., 2017). Most empirical studies indeed find a positive relationship between flood experience and flood risk perception (Reynaud et al., 2013; Richert et al., 2017; Royal and Walls, 2019), which gives us a rationale for the next hypothesis.

Hypothesis 4.2c *Individuals with flood experience have higher perceptions of the flood probability.*

With the last severe coastal flood in the Netherlands dating back to 1953, we expect few respondents in our sample who personally experienced a flood in their homes. However, a larger group of respondents might recall high water levels in their neighborhood, for example during the 1995 river floods, which could be an alternative indicator of the availability heuristic in the flood context. Działek et al. (2019) demonstrated that memory of flood events tends to decrease quickly over time, with individuals recalling significantly smaller flood surface areas two years after the initial survey. Media exposure could play a role in memorizing flood events, which could increase recall. Siegrist and Gutscher (2006) showed that media coverage can increase risk perceptions for individuals lacking personal experience with flooding. A recent empirical study confirmed that risk perception increases following media exposure of the 2013 tornado in Moore, Oklahoma (Zhao et al., 2019). Therefore, we expect a similar effect of recalling high water levels on flood risk perceptions as with the previous hypothesis concerning flood experience.

Hypothesis 4.2d *Individuals who recall high water levels have higher perceptions of the flood probability.*

All in all, heuristics in the flood risk domain may lead to serious misperceptions. While there is a growing body of literature on flood risk perceptions (cf. Kellens et al., 2013; Lechowska, 2018), few studies have examined the difference between individual risk perceptions and objective risk estimates with regards to natural hazards. One notable example is O'Neill et al. (2016), who examined the difference between real and perceived distance to a hazard source. They found that respondents who live in a flood zone but indicate that they are outside, are generally higher educated and less worried about flooding. To the best of our knowledge, the only paper that examined the deviation between objective and subjective flood risk estimates with respect to both probability and damage is Botzen et al. (2015). The authors report substantial under-estimations and over-estimations for both aspects of flood risk, but in general respondents overestimate the flood probability and underestimate potential damage.

Hypothesis 4.3 *Individuals will overestimate the probability of a flood and underestimate the consequences (damage and water levels).*

While Botzen et al. (2015) quantify flood risk misperceptions, and examine which variables relate to perceptions of the absolute level of the perceived flood probability and damage, they do not examine which variables contribute to under- versus over-estimations of flood risk in particular. Therefore, we cannot motivate hypotheses about the variables related to misperceptions. Nevertheless, we will examine whether the variables we expect to influence flood risk perceptions also influence over- or under-estimations of probability, damage and water levels.

4.3 Methodology

4.3.1 Survey method

We conducted a survey with a sample of 2122 Dutch homeowners living in floodplains in May and June 2018. The Netherlands is a relevant geographical area for flood risk perception research, as it has a long history of protection against flooding. Approximately half of the country is located behind dikes, including the metropolitan area where the main business districts and the government are situated. These low-lying areas (dike-rings) are protected from flooding by large dike infrastructures, leading to one of the highest flood safety standards across the globe. For example, some dike-rings at the coast have safety standards of 1:10000, which means that the dikes are designed to withstand an extreme flood event that may occur once in 10,000 years. The consequences of flooding in this area could be catastrophic, with maximum potential damages of 100 billion Euros (Aerts et al., 2008). Nevertheless, floodplain inhabitants might not be aware of the possibility of flooding, as the most recent severe river floods in the Netherlands occurred in 1993 and 1995 (even though none of the dikes breached), while the most recent coastal flood dates back to 1953.

We targeted homeowners in particular, as they bear the full costs of flood damage to their house, in contrast to tenants. We opted for an online survey instrument to reach a large sample of homeowners in flood-prone areas. The invitation email did not specify the topic of the survey, to prevent selection bias. The survey was distributed online and started with a selection question to ensure that only homeowners in pre-defined zip code areas could participate. Figure 4.1 shows that respondents were located in the areas with relatively low dike-ring safety standards (1:1250 and 1:2000 years, as opposed to 1:4000 and 1:10000 years in the coastal areas), in close proximity of the main rivers (Rhine and Meuse with their respective branches). The final response rate was 25.3%. We excluded 269 respondents who indicated that their home did not include the ground floor, which would give invalid results with respect to objective maximum water levels. From the 1856 valid responses, 8 were incomplete, leaving 1848 responses for analysis.

4.3.2 Elicitation of dependent and explanatory variables

This section describes the questions of our dependent and explanatory variables, which were based on previous surveys about disaster risk perceptions (Bubeck et al., 2013; Botzen et al., 2015). An extensive description of the survey, including a complete English translation of the questions can be found in Appendix 3C and Appendix 3D.

Two questions were used to elicit respondents' perception of the flood probability. Eliciting perceived flood probability estimates is a challenge, because individuals generally have difficulties with probabilistic concepts. In the context of influenza vaccination, which is a low-probability/high-impact event, analogous to flooding, Weinstein et al. (2007) showed that a qualitative question may better predict behavior under risk than a quantitative question on a percentage scale. Accordingly, we asked respondents about their perceived flood probability (*How large or small do you think the probability is that your house will be flooded?*) on a scale with seven answer categories. The drawback of such a question format is that people may attach

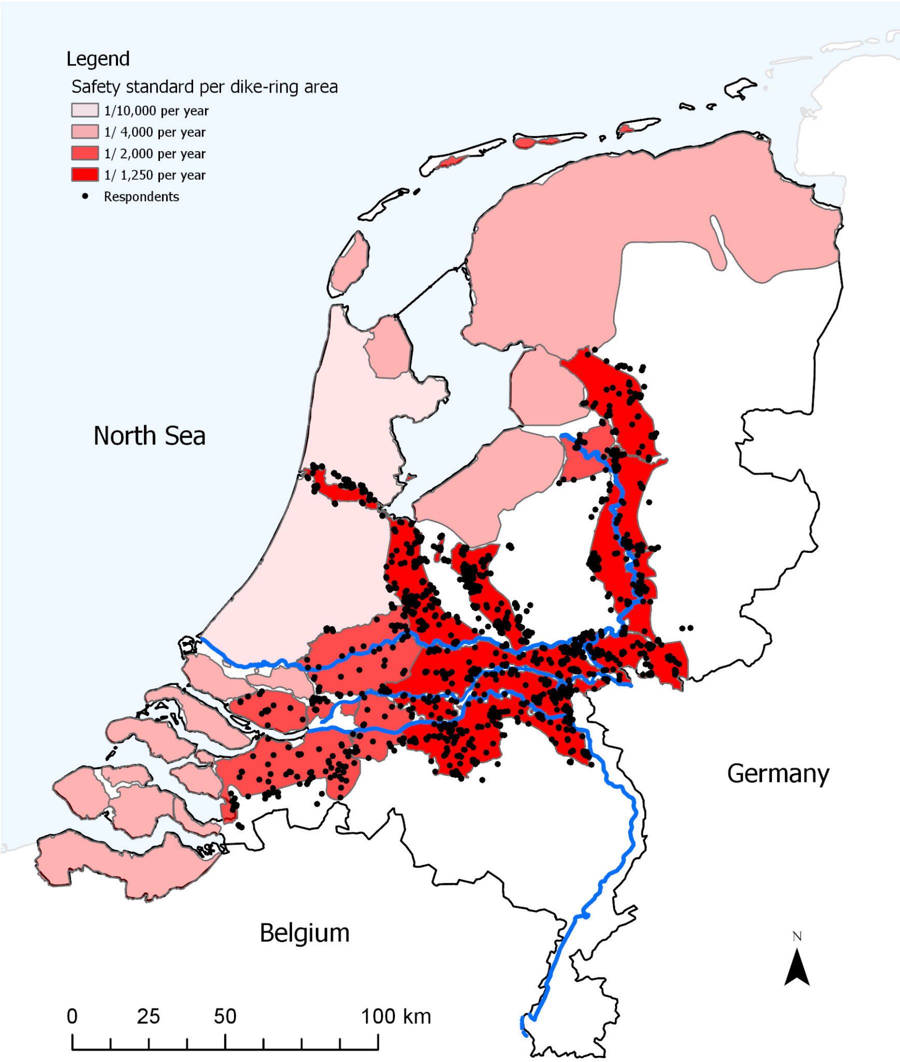


Figure 4.1: Locations of respondents to the survey on a map with safety standards of dike-ring areas in the Netherlands. Every dot represents a respondent. Main rivers are indicated in blue.

different meanings to probability phrases, which complicates a comparison with objective, quantitative estimates.

To be able to quantify over- and under-estimation among our respondents, we were interested in a more precise estimate of respondents' perceived flood probability. Recent evidence shows that compared to percentage and frequency scales, a logarithmic scale performs best in eliciting low-probability (< 1%) perceptions in terms of validity, usability and reliability (Woloshin et al., 2000; de Bruin et al., 2011). Therefore, we introduced a logarithmic scale with different return periods of flooding as a visual aid. Since our main interest is in flood probability misperceptions, we did not provide any anchor (compared to e.g. Botzen et al., 2009a, who used the legal safety norm as an anchor) with the scale. Figure 4.2 shows the decision screen of this question. Respondents could either enter their best estimate of the flood probability or express their belief in a zero-flood probability with the tick box on the right.

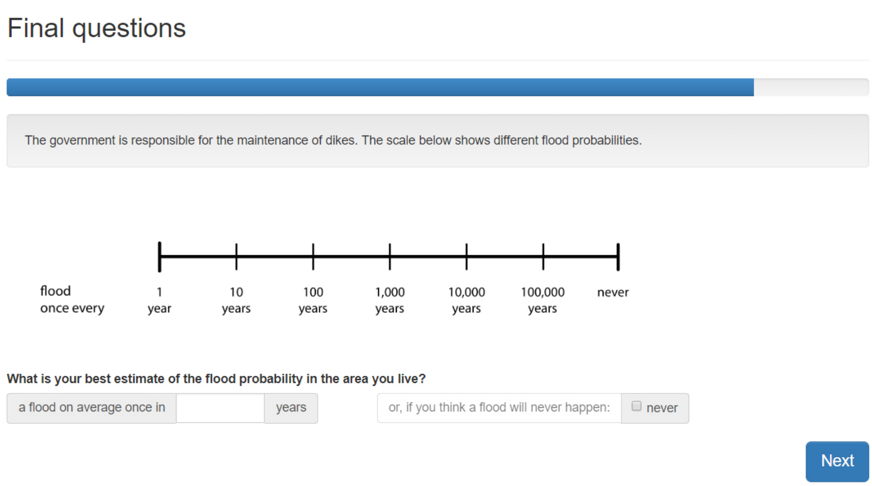


Figure 4.2: Decision screen of the subjective probability question, translated from Dutch. Respondents could either fill in an estimate on the left or tick the 'never' box on the right, but not both.

With regards to damage, we asked respondents to estimate potential flood damage to their house (*How much damage do you expect to your house and contents in case you would be flooded?*) on a scale with nine answer categories, as our pretest indicated that an open-ended question would lead to substantial participant dropouts. An alternative indicator for perceived flood risk is the expected water level in a home once a flood occurs. We asked respondents about the water level during a flood, which might be easier to imagine and is, therefore, potentially less prone to errors. We used the following question: *Imagine your neighborhood is flooded, what height do you think the water would reach in your house?*, on a scale with six answer categories. We acknowledge that we asked for the *expected* water level in case of a flood, which is not identical to the *maximum* water level used as an objective indicator of flood risk. However, we believe that respondents who imagine a flood reaching their neighborhood will think of an extreme event, which may lead to answers corresponding to the maximum water level. In flood risk communication research, depicting maximum water level or inundation is standard (see e.g. Lindner et al., 2018). Moreover, communication about water levels by the Dutch government presents exactly these maximum water levels.[2] Lastly, there is little variance in flood water levels expected in the Netherlands due to the high safety standards, which result in either no flood (i.e. the dikes hold) or a large catastrophic flood (Vergouwe, 2015) with maximum or close to maximum water levels.

Objective flood risk indicators

The objective flood probability is equal to the legal return period of flooding as described in the 2009[3] Dutch water law, which was 1:1250 for the majority (87%) of respondents, and 1:2000 otherwise. Spatial information about objective flood risk was gathered with detailed geographical information system (GIS) maps of respondents' zip codes (PC6).[4] From these GIS maps we calculated the distance to the nearest dike and the maximum objective water level for each respondent. The maximum objective water level was based on recent scenario estimates[5] provided by the Dutch government (Kok and Doef, 2008). Potential flood damage is typically estimated with depth-damage curves, which provide the proportion of value at risk for a specific inundation depth (Merz et al., 2010). To obtain the approximate rebuilding value of the home, rather than the market value, we applied a standardization[6] to the continuous home values derived from the survey answers. We applied the damage curves of the Dutch SSM-2017[7] of residential buildings to the rebuilding values, a fixed home content value of €70,000 and the maximum water level to calculate the objective damage per respondent (De Moel et al., 2014a).

[2] See www.overstroomik.nl.

[3] See https://wetten.overheid.nl/BWBR0025458/2016-07-01#BijlageI. Although a new water law was passed in 2017, the new law articulates that the flood protection infrastructure should meet the new norms only by 2050: https://www.helpdeskwater.nl/onderwerpen/waterveiligheid/primaire/nieuwe-normering/.

[4] 16 respondents entered invalid letters in the zip code input field. We calculated their location based on the four digit zip code (PC4).

[5] https://basisinformatie-overstromingen.nl/liwo/#/viewer/23

[6] Each home value was multiplied by the ratio of the average market price of the respective region and the average market price of the region with the lowest prices (Groningen). Data were obtained from: https://bit.ly/3pii1Kl

[7] https://www.helpdeskwater.nl/onderwerpen/applicaties-modellen/applicaties-per/aanleg-onderhoud/aanleg-onderhoud/schade-slachtoffer/

Heuristics

We asked several questions to elicit flood beliefs, based on the extensive reviews by Kellens et al. (2013) and Lechowska (2018). Kellens et al. (2013) classify frequently used variables in risk perception research into four main indicators: affect, awareness, likelihood and impact. Note that the likelihood and impact (expected damage) variables have been discussed above in the dependent variables subsection. To measure affect (worry), we asked subjects to respond to a statement (*I am worried about the danger of flooding at my current residence.*) on a 5-point Likert scale. We used the same linear coding for the statement on trust, (*I am confident that the dikes in the Netherlands are maintained well.*), which was almost an exact reproduction of the question in the original paper by Terpstra (2011). To assess previous flood risk experience, we asked respondents about damage (*Have you ever experienced damage to your house due to a flood?*). Furthermore, a Yes/No question was asked to examine recall of flood events (*Do you recall any situations of exceptionally high water levels in rivers close to your residence?*).

Personal characteristics (control variables)

Finally, personal characteristics such as gender, age and numeracy may play a role in determining risk perceptions. We asked two questions about the probability of a certain weather in a respondent's residence, following Dillingh et al. (2016), to get a proxy for probability numeracy. Respondents who gave a larger estimate for 'cloudy sky' than for 'cloudy sky and rain' were coded as probability innumerate. Besides, risk preferences may be important when individuals evaluate risks (Loewenstein and Prelec, 1992): risk-seeking individuals often foresee a lower probability of flooding (Botzen et al., 2009a; Mills et al., 2016). We used a qualitative question to elicit risk preferences (*How willing or unwilling are you to take risks?*), as in Falk et al. (2018).

In the domain of natural hazards, socio-demographic variables such as education, income and home value often explain little of the variance in risk perception (Peacock et al., 2005; Van der Linden, 2015). Considering the inconsistent effects of personal characteristics on risk perception in previous literature (Kellens et al., 2013; Lechowska, 2018), we will adopt these variables as control variables in our analysis (see Table 4.1 for coding).

4.3.3 Statistical analysis

Flood risk perceptions

We estimate various regression models where flood risk perception Y of individual i depends on a vector of objective risk variables ($\boldsymbol{O}$), heuristics ($\boldsymbol{H}$) and personal characteristics of the individual ($\boldsymbol{P}$).

The general specification takes the following form:

$$Y(flood\ risk\ perception)_i = \beta_1 + \beta_2 O_i + \beta_3 H_i + \beta_4 P_i + \epsilon_i \qquad \text{(C1)}$$

where ϵ_i is the error term. In Model 1, the dependent variable Y_i is a binary variable indicating whether respondents answered "Zero" to the categorical flood probability question, which is why a probit model is employed as an estimation method. In Model 2, we use an ordered probit specification to estimate flood probability perceptions: the dependent variable Y_i in this model is an ordinal variable that captures the categorical answer structure of the qualitative flood perception question. The dependent variable in Model 3 is the log-transformed estimated flood probability (return period) and this model was estimated by ordinary least squares (OLS). Note that positive coefficient estimates indicate a high perceived flood probability in all three models. In Model 4, we estimate the perceived flood damage Y_i with an ordered probit specification, to account for the categorical answer structure of the perceived flood damage question. We tested our data for multicollinearity, but this was not a concern: the correlation between the independent variables was small ($r < 0.4$).

Flood risk misperceptions

To classify our respondents into those that underestimate, those that correctly estimate and those that overestimate risk, we compared the perceived estimate (PE) of each respondent with the objective estimates (OE), allowing for different error margins (EM). The perceived risk estimate was considered correct if $OE(1 - EM) \leq PE \leq OE(1 + EM)$. As an illustration, if the objective return period is 1:2000 years and we allow for a 50% error margin, we consider estimates under 1:3000 years as under-estimation and estimates above 1:1000 years as over-estimation, while estimates within that interval are correct. Since respondents were presented with fixed answer categories for the perceived damage and water level questions, we applied the error margins to the upper- and lower bound of those intervals. For example, if a respondent answered "*10-50 cm*" for the perceived maximum water level, we considered this as correct if the objective estimate was within the 5-75 cm interval (50% error margin).

To understand the determinants of flood risk misperception in more detail, we estimated probit regressions where the dependent variable Y_i is a dummy indicating under-estimation (excluding over-estimation) or over-estimation (excluding under-estimation) of individual i. The reference category in all models is the correct estimation.

4.3.4 Sample characteristics

Our sample has equal proportions of male and female (49%) respondents. The average age of respondents is 54 years old and the distribution of age groups is very similar to that of homeowners in the general Dutch population.[8] 10% have at least a Master's degree as highest education level, which is equal to the general population.[9] The average after-tax income category is €2500-€2999 per month, which corresponds to the average after-tax income of the actual Dutch population (€2933 per month, Netherlands Statistics, 2018a). The average home value of our respondents is €250,000-€299,000, which is slightly higher than the actual average home value in the Netherlands (€216,000 Netherlands Statistics, 2018b). Summary statistics of all explanatory variables used in the analyses are presented in Table 4.1.

Table 4.1: Summary statistics

	N	Mean	St. Dev.	Min	Max
Objective risk assessment					
Sample area (0 = 1:1250, 1 = 1:2000)	1,848	0.13	0.34	0	1
Distance to nearest dike in km[a]	1,848	1.66	1.41	0.003	6.81
Maximum water level in m	1,848	1.34	1.37	0.00	8.29
Heuristics					
Worry about flooding[b]	1,848	2.08	0.96	1	5
Trust in dike maintenance[b]	1,848	3.88	0.83	1	5
Experienced flood damage (dummy)	1,848	0.06	0.24	0	1
Recall high water levels (dummy)	1,848	0.63	0.48	0	1
Personal characteristics (control)					
Gender (1 = female)	1,848	0.49	0.50	0	1
Age	1,848	53.76	14.49	18	90
Probability innumerate[c] (dummy)	1,848	0.07	0.25	0	1
Risk aversion index[d]	1,848	4.49	2.04	0	10
Education[e]	1,848	5.86	1.43	1	9
Ln income[f]	1,389	7.95	0.42	5.52	8.57
Ln home value[g]	1,680	12.53	0.38	10.82	13.62

[a] Euclidian distance from center of zipcode area to nearest dike, based on GIS maps. [b] Categorical answers, coded from 1 (strongly disagree) to 5 (strongly agree). [c] Respondents were asked to estimate the probability of (1) a cloudy sky tomorrow and (2) a cloudy sky and rain. Respondents who gave a larger estimate for event (2) were counted as probability innumerate. [d] How willing or unwilling you are to take risks? Categorical answers, coded from 1 (very unwilling) to 7 (very willing) [e] Education in 9 categories were 1 indicates no diploma and 9 indicates a PhD. [f] Respondents could indicate their after-tax income category, starting at €0-€499, increasing in steps of €500. Continuous values of income variables were constructed by setting the income value of each respondent to the midpoint of the interval. €5,250 was used for the highest income category (>€5,000). The results were log-transformed. Respondents who answered "Rather not say" or "Don't know" were excluded from this measure. [g] Question format similar to income. Starting category <€100,000, increasing in steps of €50,000. €825,000 was used for the highest category (>€800,000).

[8] See CBS details at: https://bit.ly/2YcJrpe
[9] See CBS details at: https://bit.ly/2LUyqq6

4.4 Results

Flood risk is generally defined as the product of flood probability and flood damage. We first report respondents' answers to the perceived probability, damage and water level questions and relate them to the objective flood risk estimates. We analyze the drivers of flood risk perceptions in detail with a regression analysis to evaluate our hypotheses. Subsequently, we examine the direction of flood risk misperceptions by inspecting the predictors of under and over-estimations.

4.4.1 Flood risk perceptions

Few respondents (< 5%) consider the probability of a flood as high or very high, which confirms that a large majority of Dutch citizens is aware of the high flood protection standards in the country. Almost 15% of respondents mark a perceived flood probability of zero in the categorical flood probability question (see Figure A1 for the full distribution of answers). When asked to give a more precise estimate of the flood probability in the form of an estimated return period, more respondents report that a flood will never reach their current residence. Figure 4.3 shows a histogram of the perceived return period of flooding, with dashed reference lines to indicate the objective return period. A large fraction of respondents (28%) expects that a flood will never occur at their present address, which is a serious misperception as the sample was drawn from the zip code areas that are at risk of flooding in the Netherlands (within dike-ring areas with relatively protection norms). While these individuals may be unaware that they live in a flood-prone area, other individuals largely overestimate the probability of a flood reaching their house. Approximately 10% of respondents estimate that the return period of a flood at their present address is 10 years or less, indicating a very high flood risk perception. Note that a return period of 100 years is considered a relatively high flood probability in the Netherlands, where most areas are protected up to 4000 and even 10,000 years. Overall, we find a bi-modal pattern of risk perception, with a large group of respondents reporting high risk perceptions (return periods of 100 years and below) and a slightly smaller group who neglects the flood probability altogether. Very few responses were collected in between those two extremes.[10] When it comes to expected damage, the majority of respondents (70%) estimated that flood damage would cost up to €50,000.

[10] To account for these different flood risk perception 'types' in our data, we constructed a dummy variable to indicate the 'never types'. We re-ran our regressions (not reported here) for this subgroup of 'never types'. The sign and significance of the coefficients do not differ from the main regressions.

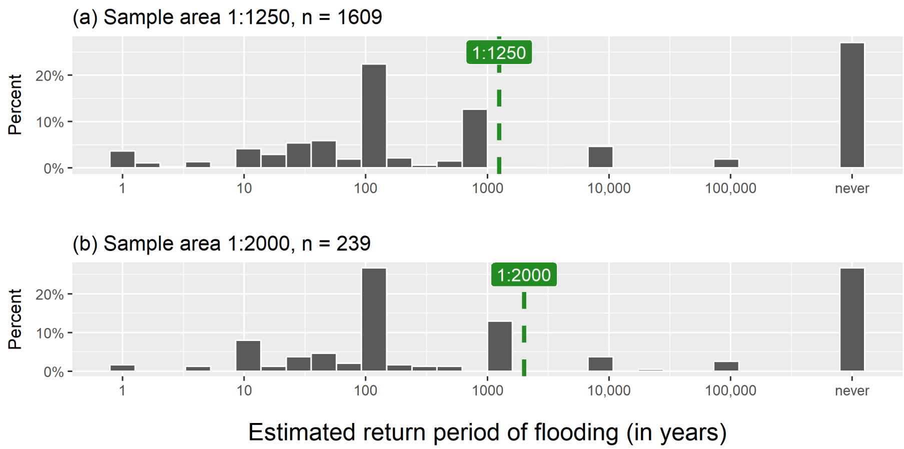


Figure 4.3: Histogram of respondents' estimated return period of flooding. Green dashed reference lines indicate actual return periods.

Objective risk assessment

Table 4.2 reports the results of our regression analyses. To examine the relationship between perceived and objective risk, consider the coefficients of the geographical characteristics in the first block of the table. We find no effect of objective return periods (as indicated by sample area) on the perceived probability of flooding, nor on perceived damage. In other words, the data do not support Hypothesis 4.1a. With regards to Hypothesis 4.1b, we find partial support. In Model 1, 3 and 4 we find no significant effect of dike-distance on flood risk perceptions. The significantly negative coefficients of Model 2 indicate that respondents who live further away from dikes, expect a lower probability of flooding than respondents who live closer to dikes, as hypothesized. We find, however, a significant, strong and positive effect of the objective maximum water level on risk perceptions across all four models, confirming Hypothesis 4.1c.

Heuristics

We find a strong effect of worry on flood risk perceptions across models. The significantly positive estimates for worry confirm Hypothesis 4.2a: individuals with high levels of worry about flooding estimate the likelihood of flooding to be higher. Moreover, the coefficient of Model 4 implies that those who worry a lot about flooding expect significantly higher damage to their house in case of a flood. We find no effect for trust in dike maintenance on flood risk perceptions: Hypothesis 4.2b cannot be confirmed. Individuals who have previous flooding experience, indicated by the dummy variable of 'experienced flood damage' generally perceive a higher likelihood of flooding, as predicted by Hypothesis 4.2c. However, the results are not statistically significant in Model 1. Interestingly, individuals who have had their home damaged due to flooding in the past, have lower damage expectations for future floods. One explanation for this effect is that flood events in the Netherlands in the last decades have been relatively small, which may have led to minor damages. Finally, we find strong support for the use of the availability heuristic (Hypothesis 4.2d) in the data: individuals who remember high water levels, have significantly higher flood probability perceptions for all three models.

Personal characteristics (control variables)

In addition to the explanatory variables related to our hypotheses, we observe some other interesting patterns with regards to our control variables. We find that respondents with a higher income generally expect higher damages. The significantly positive estimates for education indicate that more highly educated respondents perceive a higher likelihood of flooding, while the significantly negative estimate in Model 4 indicates that they expect a lower level of flood damage. Moreover, risk-averse and younger respondents seem to have higher flood risk perceptions. We find no effect of gender and probability innumeracy on risk perceptions.

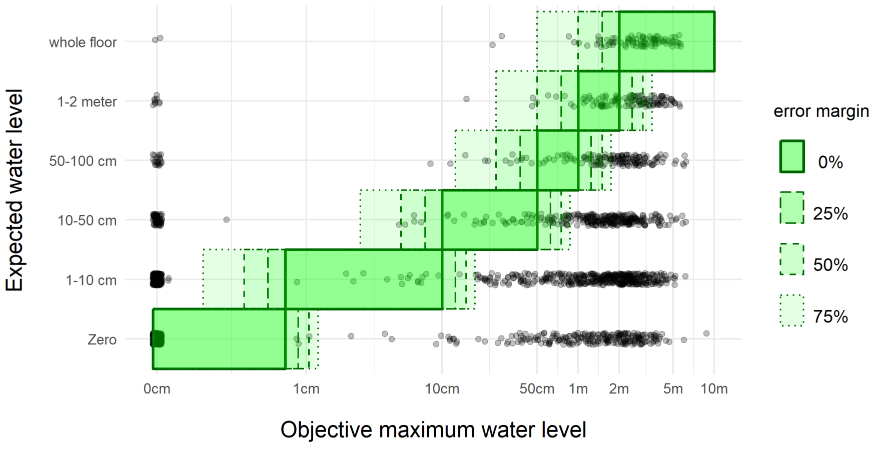


Figure 4.4: Perceived versus objective water levels; green shaded bars indicate correct estimates.

4.4.2 Flood risk misperceptions

In this section we examine the direction of flood risk misperceptions: over- versus under- estimation. Figure 4.4 shows a scatter plot of the perceived and

Table 4.2: Regression results of flood risk perceptions

	Probability *probit* (1)	Probability *oprobit* (2)	Probability *OLS* (3)	Damage *oprobit* (4)
Constant	−1.486		−10.926**	
	(1.561)		(4.214)	
Objective risk assessment				
Sample area	0.111	0.133	0.217	0.153
	(0.111)	(0.087)	(0.309)	(0.092)
Distance to nearest dike in km	0.014	−0.042*	−0.004	0.0003
	(0.029)	(0.021)	(0.075)	(0.023)
Maximum water level in m	0.115***	0.155***	0.288***	0.112***
	(0.032)	(0.023)	(0.079)	(0.025)
Heuristics				
Worry about flooding	0.444***	0.623***	1.443***	0.181***
	(0.055)	(0.043)	(0.121)	(0.034)
Trust in dike maintenance	0.048	−0.021	0.005	0.053
	(0.050)	(0.041)	(0.139)	(0.042)
Experienced flood damage	0.268	0.675***	1.476***	−0.266*
	(0.254)	(0.140)	(0.431)	(0.119)
Recall high water levels	0.408***	0.293***	1.183***	0.164*
	(0.083)	(0.064)	(0.245)	(0.074)
Personal characteristics				
Gender (1 = female)	−0.060	0.121	0.099	0.088
	(0.085)	(0.064)	(0.232)	(0.072)
Age	−0.001	−0.009***	−0.014	−0.004
	(0.003)	(0.002)	(0.008)	(0.003)
Probability innumerate	0.006	0.007	0.256	0.058
	(0.183)	(0.135)	(0.441)	(0.123)
Risk aversion index	0.068***	0.043**	0.131*	0.013
	(0.020)	(0.015)	(0.054)	(0.017)
Education	0.148***	0.073**	0.288***	−0.068*
	(0.033)	(0.023)	(0.087)	(0.027)
Ln income	−0.088	−0.157	−0.533	0.253**
	(0.100)	(0.083)	(0.275)	(0.096)
Ln home value	0.021	−0.040	0.168	0.506***
	(0.126)	(0.090)	(0.339)	(0.107)
Log likelihood	-668.8	-1628.5	-3816.9	-1669.1
Pseudo R^2 (McFadden)	0.379	0.374		0.208
Observations	1,370	1,332	1,370	1,083
R^2			0.199	

Notes: Dependent variable Model 1: dummy estimated flood probability not zero; Model 2: categorical flood probability, higher numbers indicate higher flood probability; Model 3: log-transformed estimated flood probability; Model 4: categorical damage estimate. Robust standard errors in parentheses (${}^{*}p < 0.05$; ${}^{**}p < 0.01$; ${}^{***}\ p < 0.001$). Dummy sample area: 0 indicates 1:1250; 1 indicates 1:2000. Other dummy variables: experienced flood damage, recall high water levels and probability innumerate (1 indicates yes).

the objective maximum water level. Each observation (respondent) is indicated with a gray dot with 1% random jitter to facilitate readability. The graph reveals a small subset of respondents who have zero as their objective maximum water level.[11] Green shaded bars indicate the range where perceived and objective water level estimates match. To acknowledge that flood risk involves large uncertainties and is therefore difficult to estimate for respondents, we allow for different error margins around the objective estimate. All data points above the green diagonal represent respondents who overestimate maximum water levels, while data points below the diagonal represent those who underestimate. The graph shows that most Dutch homeowners seriously underestimate the maximum water level in their home in case of flooding, even when we allow for a 75% margin of error.[12] A similar pattern emerges for the relationship between perceived and objective damage (see Figure A2).

Figure 4.5 gives an overview of the ratio of under-, correct and over-estimations under different error margins for the three different aspects of flood risk perception (probability, water level and damage). The majority of respondents overestimates the flood probability and underestimates the maximum water level, under all error margin specifications, which is in line with Hypothesis 4.3. Figure 4.5 also shows that respondents have more correct estimates when it comes to anticipated damage, rather than the maximum water level in case of flood.

Table 4.3 reports regression results of probit regressions on a dummy of under-estimation versus correct estimation (excluding over-estimation) or over-estimation (excluding under-estimation). The significantly positive constant term in Model 3 confirms that individuals generally underestimate the maximum water level during a flood, while the non-significant constant terms in Model 5 and 6 verify that most respondents correctly identify the expected flood damage.

Objective risk assessment

The positive coefficients for the variable sample area indicate that respondents in the safer dike-ring area are more likely to overestimate the maximum water level and less likely to underestimate the potential damage of a flood. The coefficients for dike distance indicate that individuals who live far away from dike protection significantly underestimate the maximum water level and the potential damage of a flood: “out of sight, out of mind”. The pattern of

[11] We have tested this subset on coding errors but none were found: these individuals simply live close to the border of a dike-ring or on slightly higher grounds. For robustness, we re-ran our analysis on flood risk perceptions excluding this sample. The results do not change qualitatively.

[12] We use error margins following Botzen et al. (2015) and checked with experts whether the 25%, 50% and 75% margins could be applied to the Dutch context. The reader is referred to De Moel et al. (2014b) and Huizinga et al. (2017) for a detailed discussion of uncertainty and sensitivity in flood risk modeling.

Table 4.3: Probit regressions of flood risk misperceptions

	Probability		Water level		Damage	
	under-estimate	over-estimate	under-estimate	over-estimate	under-estimate	over-estimate
	(1)	(2)	(3)	(4)	(5)	(6)
Constant	3.586	2.918	4.127^{**}	1.513	-0.653	-1.328
	(2.094)	(1.752)	(1.556)	(2.164)	(1.526)	(1.746)
Objective risk assessment						
Sample area	-0.169	-0.007	-0.104	0.620^{**}	-0.273^{*}	0.205
	(0.174)	(0.146)	(0.139)	(0.195)	(0.128)	(0.157)
Distance to nearest dike	0.012	0.029	0.240^{***}	0.027	0.153^{***}	-0.008
	(0.044)	(0.034)	(0.034)	(0.049)	(0.029)	(0.038)
Maximum water level in m	-0.090	0.050	0.448^{***}	-2.498^{***}	0.389^{***}	-0.762^{***}
	(0.046)	(0.039)	(0.050)	(0.560)	(0.037)	(0.200)
Heuristics						
Worry about flooding	-0.230^{***}	0.259^{***}	-0.054	0.506^{***}	-0.020	0.033
	(0.070)	(0.057)	(0.047)	(0.081)	(0.045)	(0.060)
Trust in dike maintenance	0.027	0.017	-0.102	-0.050	-0.121^{*}	-0.079
	(0.074)	(0.065)	(0.054)	(0.074)	(0.049)	(0.059)
Experienced flood damage	0.191	0.582^{*}	0.136	0.199	0.150	0.581^{*}
	(0.451)	(0.253)	(0.243)	(0.292)	(0.211)	(0.253)
Recall high water levels	-0.408^{***}	-0.027	-0.174^{*}	0.022	0.065	0.073
	(0.119)	(0.114)	(0.085)	(0.125)	(0.089)	(0.096)
Personal characteristics						
Gender (1 = female)	-0.057	0.031	0.097	0.182	-0.232^{**}	-0.252^{*}
	(0.119)	(0.106)	(0.093)	(0.136)	(0.088)	(0.111)
Age	-0.003	-0.008^{*}	-0.011^{**}	-0.013^{**}	-0.008^{*}	-0.011^{**}
	(0.004)	(0.004)	(0.003)	(0.005)	(0.003)	(0.004)
Probability innumerate	0.129	0.239	0.090	-0.025	-0.212	0.065
	(0.289)	(0.209)	(0.201)	(0.241)	(0.185)	(0.190)
Risk aversion index	-0.040	-0.006	-0.018	-0.032	0.033	0.006
	(0.026)	(0.023)	(0.021)	(0.029)	(0.019)	(0.023)
Education	-0.121^{**}	-0.036	0.025	0.096	0.027	-0.062
	(0.044)	(0.042)	(0.034)	(0.050)	(0.032)	(0.038)
Ln income	0.126	-0.130	0.128	0.168	-0.026	0.180
	(0.153)	(0.153)	(0.108)	(0.182)	(0.112)	(0.133)
Ln home value	-0.183	-0.091	-0.371^{**}	-0.258	0.043	0.092
	(0.169)	(0.145)	(0.130)	(0.178)	(0.129)	(0.153)
Log likelihood	-346.6	-427.2	-573.2	-278.5	-639.4	-439.5
Pseudo R^2 (McFadden)	0.355	0.417	0.399	0.516	0.315	0.288
Observations	621	926	1,104	631	1,064	890

Notes: Probit regression estimates of misperception (over- and under-) versus correct estimation (at 50% error margin) for three indicators of flood risk. Robust standard errors in parentheses ($^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}\ p < 0.001$). Dummy sample area: 0 indicates 1:1250; 1 indicates 1:2000. Other dummy variables: experienced flood damage, recall high water levels and probability innumerate (1 indicates yes).

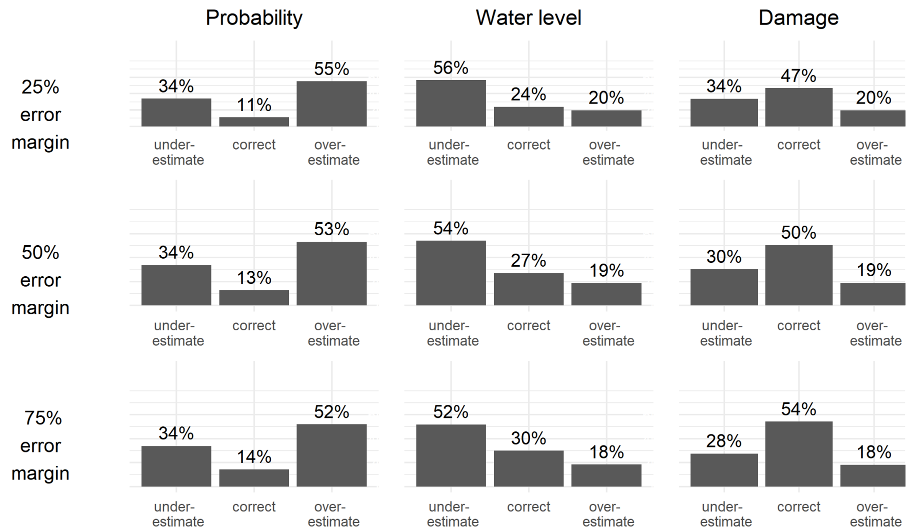


Figure 4.5: Distribution of flood risk perceptions at different error margins.

coefficients of maximum water level demonstrates that high-risk individuals with high maximum water levels are more likely to underestimate water levels and damage. The pattern is consistent: these high-risk individuals are also less likely to overestimate water levels and damage. We find no significant misperceptions of flood probability based on objective risk variables.

Heuristics

Respondents with high levels of worry, have serious over-estimations of probability and water levels, but not of damage. High trust in dike maintenance makes it less likely that respondents will underestimate potential flood damage. This suggests that trust in dike maintenance does not activate a false sense of safety, which has raised concerns by previous researchers (see e.g. Tobin, 1995, on the 'levee effect'). Experience with flood damage increases the likelihood of over-estimating flood probability and potential flood damage. Finally, we find that respondents who recalled high water levels are less likely to underestimate flood probability and maximum water levels.

Personal characteristics (control variables)

With regards to our control variables, we find that older individuals are less likely to have misperceptions (both under- and over-estimations) on all three risk factors. The significantly negative estimate for education indicates that more highly educated individuals are less likely to underestimate the flood probability.[13] However, education seems not to affect misperceptions about maximum water level and damage. Respondents with more expensive homes are significantly less likely to underestimate the maximum water level. We find no effects of risk aversion, income and probability innumeracy on flood risk misperceptions.

4.5 Discussion

This section discusses our main results in relation to our hypotheses and places these findings in the context of the existing literature. Starting with the indicators of objective flood risk, we find no support for the effect of flood probability (Hypothesis 4.1a) and dike-distance (Hypothesis 4.1b) on flood risk perceptions. However, we sampled from two different protection standards, which were rather similar. This lack of initial variation could explain why our results do not show the hypothesized effect of flood probability on flood risk perceptions. We do find strong support for Hypothesis 4.1c: individuals living in low-lying areas as indicated by maximum water level, have higher subjective flood probability estimates, as well as higher potential flood damage estimates. The same individuals are more likely to underestimate water levels and damage. In other words, individuals living in low-lying areas know that they face flood risks, but they underestimate them. One reason for the lack of effect of dike-distance and the strong effect of maximum water levels, is visibility. Respondents cannot easily observe the distance to the nearest dike, while maximum water level (which corresponds to the height of the land) may be easier to observe, for example during periods of rainfall.

With regards to heuristics, we examined the affect heuristic, trust in dike maintenance, flood risk experience and the availability heuristic. We find support for Hypothesis 4.2a: individuals with high levels of worry about flooding estimate the likelihood of a flood to be higher. These findings are consistent with Botzen et al. (2015), who find that low perceptions of flood probability are related to low worry and high trust in local flood risk management. However, the current analysis finds no support for Hypothesis 4.2b about the effect of trust in local flood risk management on flood risk perceptions. The lack of support for the trust hypothesis is in contrast to some previous work (Sjöberg, 2007; Terpstra, 2011) but not all (Carlton and Jacobson, 2013; Verlynde et al., 2019). Moreover, trust in local flood risk management was rather high (less than 5% disagreed or strongly disagreed

[13] We conjectured that older participants would have more flood experience. Instead, we found a small but negative Pearson correlation between age and the experienced flood damage dummy ($\rho = -0.081$, $p < 0.001$) and that higher educated participants have more flood damage experience ($\rho = 0.067$, $p = 0.004$). We further found that younger people are more likely to feel worried about flooding ($\rho = -0.160$, $p < 0.000$), which may be one of the reasons why younger people have more misperceptions about flooding.

with the statement) in our sample.[14] Future studies could examine the effect of trust on risk perception in a sample with more variability in trust ratings. Regarding Hypothesis 4.2c, note that only a small fraction of our sample has first-hand flood experience (6%) and that we cannot exclude the possibility of reversed causality: individuals with higher risk perceptions are more likely to remember high water levels (c.f. Spence et al., 2011; Osberghaus, 2017). Indeed, we find ample support for Hypothesis 4.2d, which operationalized the availability heuristic as being able to recall a flood event. These findings are consistent with the previous findings on the effect of the availability heuristic on risk perceptions (Kellens et al., 2013; O'Neill et al., 2016).

Some limitations of our study should be addressed. First, the study uses an individualistic approach to risk perception, whereas homeowners might share their homes with family and discuss home-related issues within their neighborhood. Van der Linden (2015) demonstrated that the behavior of others can be an important motivation to take action against flood risk. Future studies could examine the impact of social norms, an additional heuristic, on flood risk perception. Another limitation is that we used validated, but single-item scales due to time constraints for respondents in completing the online survey. Some studies show that multiple-item risk measures perform better in predicting risky behavior (Menkhoff and Sakha, 2017), but not all studies confirm this finding (Mol et al., 2020a). Numeracy and trust measures could be improved in future research by implementing a numeracy (McNaughton et al., 2015) and trust (Grimmelikhuijsen and Knies, 2017) scale. When interpreting the results, it should be noted that the current dataset contains survey data collected at one particular point in time. The assumption of exogeneity of explanatory variables may therefore be violated, when risk perception and worry about flooding are both driven by an underlying and unmeasured characteristic. To be able to draw causal conclusions, further research should use experiments or longitudinal surveys (Hudson et al., 2019; Bubeck et al., 2020).

Our typology of flood risk misperceptions revealed that a majority of Dutch floodplain inhabitants overestimates the probability of a flood event, while underestimating the potential water level in case of a flood, supporting Hypothesis 4.3. Most damage estimates appear to be correct, although up to 34% of our sample underestimates potential flood damage. One explanation for this finding is that the maximum flood damage is bounded by the value of a home. Even without knowledge about depth-damage curves and water levels, respondents who opted for a certain fraction of the home value would have picked the right range quite often. These findings largely confirm the results of Botzen et al. (2015), who found that most New York City floodplain inhabitants overestimate flood probability, while underestimating the potential damage. A major difference between the two studies is that our sample has no

[14] We constructed a dummy variable for those who agreed or strongly agreed with the statement. We re-ran our analyses with this dummy variable. The sign and significance of the coefficients do not change.

recent flood experience, while the New York City sample was surveyed within one year after a major hurricane.

4.6 Conclusion

Flooding is one of the most significant natural disasters worldwide and its impacts are expected to increase further in the future. The implementation of damage reduction strategies is therefore of increasing importance. Damage reduction measures taken by private homeowners can be cost-effective, but current take-up is low. A potential explanation is that flood risk perceptions of individual homeowners differ considerably from objective estimates, which may alter their assessment of the cost-effectiveness of damage reduction measures. Flood risk perceptions further affect support for public investments in flood protection infrastructure. While the literature on flood risk perceptions is extensive, so far a systematic assessment of the determinants of flood risk misperceptions was lacking. This chapter aimed to understand and quantify the flood risk misperceptions of Dutch floodplain residents, which is important for the design of effective risk communication campaigns and insurance schemes to cope with increasing natural disaster risks.

The main contribution of this chapter to the literature lies in the detailed analysis of factors that are related with flood risk misperceptions. For instance, this analysis revealed that individuals who recall high water levels are less likely to have misperceptions of flood risk. It further shows that affective feelings about risk, in this case worry, may lead to over-estimations of probability and water level. Experience of a flood and trust in dike maintenance seem to decrease flood risk misperceptions.

The following policy recommendations can be drawn from our results. The observation that a majority of respondents underestimates the water level of a flood implies that many Dutch homeowners may underestimate the cost-effectiveness of damage reduction measures. It may hence be worthwhile for the Dutch government to proceed with information campaigns for homeowners in the river delta. The government could target homes which can be improved with cost-effective measures. Moreover, these campaigns could specifically target homeowners in low-lying areas as they are currently over-represented in the share of under-estimators of flood risk. A second implication of this study is that worry about flooding may increase flood risk perceptions, but it may lead to over-estimations. Hence a promising approach could be to focus on communicating consequential factors of risk, such as damage estimates and the maximum water level, as they are salient and rather easy to imagine, rather than communicating difficult to interpret probabilities or return periods. Future research could focus on the effectiveness of these informational campaigns, considering the absence of recent flood experience among Dutch floodplain inhabitants.

Acknowledgments

We would like to thank conference participants at the European Conference on Risk Perception, Behavior, Management and Response in Paris and SPUDM 2019 in Amsterdam for valuable comments. This research has received financial support from the Netherlands Organization for Scientific Research (NWO) VIDI (452.14.005) grant.

Appendix 4A: Additional figures and tables

Figure A1 shows a histogram of the given answers in the categorical question on perceived flood probability.

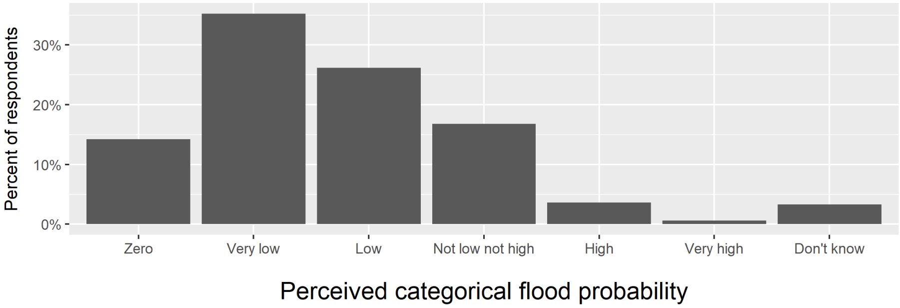


Figure A1: Histogram of respondents' answers to the categorical flood probability question.

Figure A2 shows a scatter plot of perceived flood damage and the objective flood damage. The figure confirms the pattern of Figure 4.4; a large majority of respondents underestimates the damage that a flood can potentially cause.

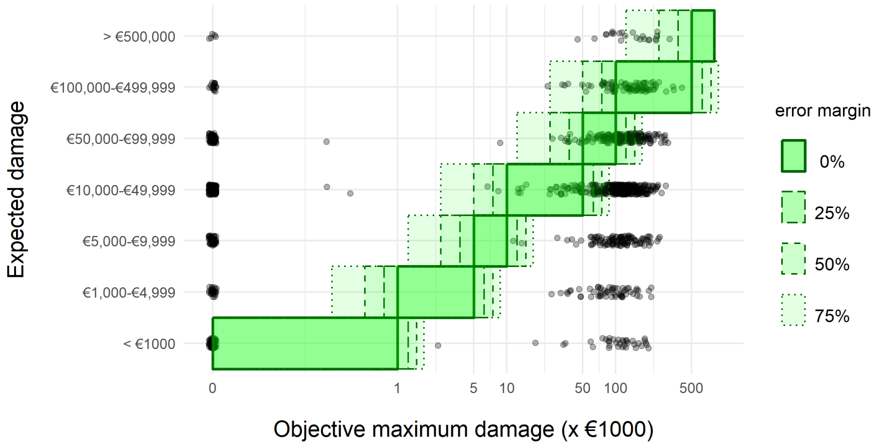


Figure A2: Perceived versus objective flood damage; green bars indicate correct estimates.

Appendix 4B: Inconsistent types

Since we used two different questions to elicit the perceived probability of a flood, we could examine respondents' consistency. Figure B1 shows a scatter plot of the categorical perceived flood probability versus the numerical estimate. We find a large variation in numerical estimates for the different probability phrases, which is in line with previous research on interpretation of probability phrases (c.f. Visschers et al., 2009; Willems et al., 2019). One could argue that the probability phrase *Very low* is inconsistent with a return period of 10 years. Our focus is on the most extreme answer categories, indicated with red bars in the figure: *Zero* on the categorical scale is clearly inconsistent with all numerical estimates <100,000 years and the explicit *never* answer to the estimated flood probability is inconsistent with all categorical estimates larger than *Low*. As a robustness check, we reran our analyses (not reported here) excluding inconsistent respondents ($n = 97$). All main effects and interactions remained unchanged.

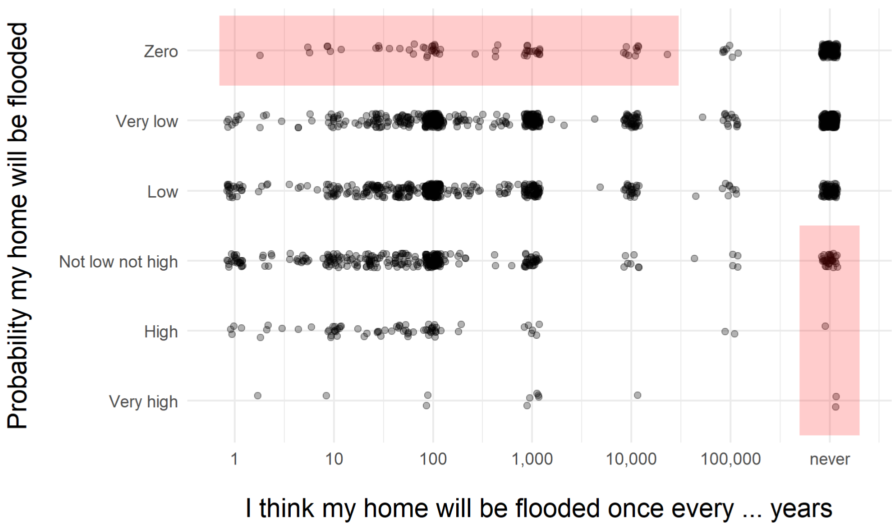
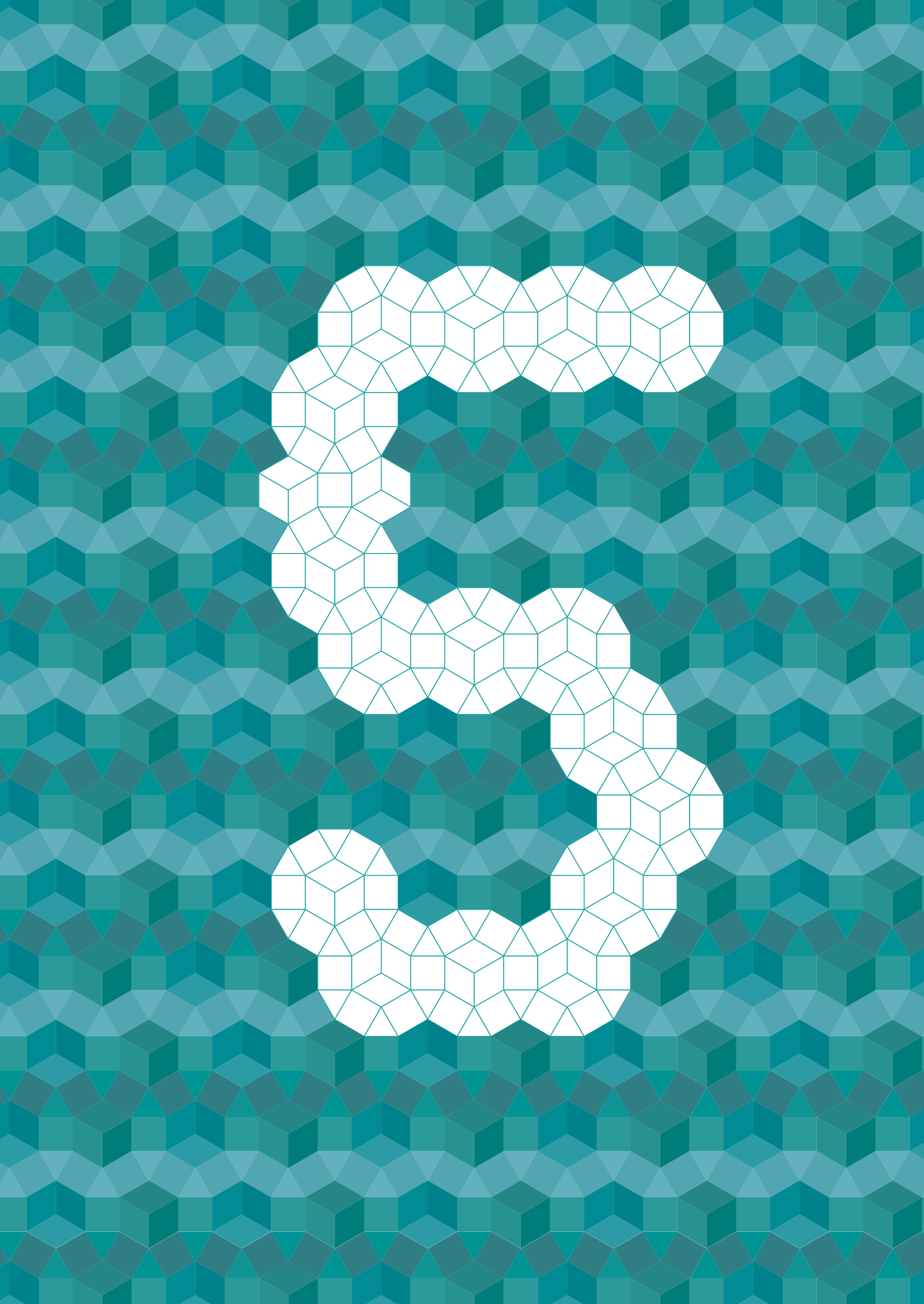


Figure B1: Categorical versus numerical flood risk perception; red shaded bars indicate respondents classified as inconsistent.


5

Goggles in the lab: Economic experiments in immersive virtual environments

This review outlines the potential of virtual reality for creating naturalistic and interactive high-immersive environments in experimental economics. After explanation of essential terminology and technical equipment, the advantages are discussed by describing the available high-immersive VR experiments concerning economic topics to give an idea of the possibilities of VR for economic experiments. Furthermore, possible drawbacks are examined, including simulator sickness, the costs of VR equipment and specialist skills. By carefully controlling a naturalistic experimental context, virtual reality brings some field into the lab. Besides, it allows for testing contexts that would otherwise be unethical or impossible. It is a promising new tool in the experimental economics toolkit.

This chapter is published as:
Mol, J. M. (2019). Goggles in the lab: Economic experiments in immersive virtual environments. *Journal of Behavioral and Experimental Economics,* 79(C), 155–164.

5.1 Introduction and terminology

Virtual Reality (VR) is a popular new technology by which almost any environment can be simulated and projected in 3D to the user. The rapid growth of VR is in large part driven by technological innovations and a sharp decline in the costs of VR devices. While VR as a research tool is now commonly applied in psychotherapy (Dibbets and Schulte-Ostermann, 2015), engineering (Freeman et al., 2016), spatial planning (Natapov and Fisher-Gewirtzman, 2016) and social psychology (Bombari et al., 2015), to date there are very few VR experiments in economics. Yet, the possibilities are promising: VR could add crucial realism to lab experiments and more control to field experiments. A recent review by Innocenti (2017) discussed how VR experiments may contribute to the field of economics by offering context to check the external validity of economic theories, with a focus on low-immersive virtual environments such as online virtual worlds. The current review does not address these low-immersive virtual worlds, but focuses on high-immersive virtual reality.

Recent reviews have highlighted the potential of VR for marketing (Barnes, 2016) and business research (Meißner et al., 2017). The current review complements by offering a critical overview of the possibilities and challenges for experimental economics in high-immersive virtual environments. The remainder of this chapter is organized as follows: Section 5.2 explains the essential terminology and technical equipment. Section 5.3 discusses the main advantages by describing the available VR experiments concerning economic topics to give an idea of the possibilities for economists, including an overview of relevant VR experiments in Table 1. In Section 5.4 possible drawbacks are discussed, including simulator sickness, the demand for physical equipment and specialist skills. Finally Section 5.5 provides some practical advice and Section 5.6 concludes.

5.2 Terminology

The possibility to escape the world by virtually going elsewhere has always triggered human imagination. In the 1990s, this idea of creating a virtual world was first introduced in science, when communication researchers started to study virtual reality as a medium (Biocca and Levy, 1995). Virtual reality includes a computer generated environment and an interaction aspect. The Oxford English Dictionary defines VR as “the computer-generated simulation of a three-dimensional image or environment that can be interacted with in a seemingly real or physical way by a person using special electronic equipment, such as a helmet with a screen inside or gloves fitted with sensors” (Oxford Dictionaries, 2018).

Several definitions describe how ‘real’ participants experience virtual reality. Following Bombari et al. (2015), in this review the term “presence” is used to describe the "subjective feeling of 'being there' and interacting with one's body in a virtual world projected by VR technology". As technology improved, the possibility of having more than one person in a VR environment was created in many modern labs. Consequently, the term "copresence" was coined: "the feeling of presence together with other virtual humans" (Bombari et al., 2015, p.33). Two classes of virtual humans can be defined: those controlled by algorithms (agents) and those controlled by other humans (avatars) (Bailenson and Blascovich, 2004). Sometimes, participants respond differently to these two types of virtual humans, for example by keeping more distance to agents than to avatars (Bailenson et al., 2003).

"Immersion" is defined by Bombari et al. (2015) as "the objective amount and quality of the perceptual input provided to the participant through technology" (p. 3). Immersion can be increased by showing a participant's own limbs in the virtual environment, while movements are projected in real time. Thus, by varying the amount of perceptual input or technological capabilities of the VR system (immersion), participants will experience the environment either as more or less 'real' (presence). A more thorough discussion of the concepts immersion and presence can be found in the survey of Slater and Sanchez-Vives (2016). Innocenti (2017) defines two classes of virtual reality environments by level of immersion, where low-immersive virtual environments (LIVE) represent desktop renderings and (online) virtual worlds, such as Second Life and World of Warcraft. The focus of this review is on the other class: high-immersive virtual environments (HIVE), where a virtual environment is projected in 3D to the user at the cost of more complex and expensive equipment.

VR equipment for HIVE falls into two broad categories: head-mounted displays (HMD) and projection screens, where the latter type is sometimes called a CAVE activated virtual environment (Cruz-Neira et al., 1993). Figure 5.1 depicts the two categories in the DAF Technology Lab at Tilburg University. An HMD brings the virtual environment close to the eyes of the participant, leading to a wide-angle view, including the virtual ground and ceiling. A set-up with projection screens in combination with stereoscopic glasses (CAVE), gives participants the freedom to walk around in the virtual environment and to enter the environment with multiple users. The downside to this setup is that the floor and the ceiling are often not used as projection screens, such that the borders of these areas are clearly visible, creating a less immersive environment.

In addition to virtual reality, two frequently used terms in both industry and academia are augmented reality (AR) and mixed reality (MR). Where VR excludes the real world almost completely from the (mainly visual) senses, in AR the physical environment is visible but overlaid with extra (computer graphic) information. MR adds interaction to the computer graphic objects

[1] Pictures taken at the DAF Technology Lab at Tilburg University, retrieved from: https://www.tilburguniversity.edu/campus/experiencing-virtual-reality/.

(a) HMD

(b) CAVE

Figure 5.1: Different categories of VR equipment[1]

projected by AR. Examples of modern-day AR/MR devices are the Google-glass[2] and the Microsoft Hololens.[3] This review focuses on high immersive virtual reality.

An important concept in VR is (virtual) embodiment, which refers to substitution of the real body by a virtual body (see Slater and Sanchez-Vives, 2016, for a survey of work on embodiment). Under the right technical conditions (perfect visuomotor synchrony, among others) embodiment can lead to the illusion of body ownership. Even though a person's own body might look very different from the virtual projection, the illusion can lead to a strong feeling that the virtual body is the real one. Embodiment allows for changing the virtual body, for example as an avatar that is taller (Yee and Bailenson, 2007), skinnier (Fox et al., 2009) or with a different skin color (Peck et al., 2013) than subjects' appearance in reality. A related term is the 'Proteus effect' of Yee and Bailenson (2007), meaning that self-representation is modified in a meaningful way, which leads the user to conform to the modified self-representation regardless of the physical self. Fox et al. (2009) found that participants exercised more when they saw a virtual representation of the self that changed in body weight in accordance to exercise efforts, than participants without a responsive representation.

Transformed social interaction refers to interpersonal communication in VR, where the appearance or ability of a participant has been changed. This includes possibilities that do not exist in the real world, such as changed perceptual abilities, forced perspective taking and controlled self-representation (Bailenson et al., 2005). For instance, Yee and Bailenson (2007) examined the effect of the height of avatars on negotiation behavior in an ultimatum game and found that participants with taller avatars behaved more confidently and proposed more unfair allocations than participants with shorter avatars. One

[2] https://developers.google.com/glass/.
[3] https://www.microsoft.com/microsoft-hololens/en-us/.

could also change the appearance (e.g. height) of all other avatars in the virtual environment. Changing the communication itself can be achieved by manipulating the gaze of avatars, for example by shorter or longer eye contact (Bombari et al., 2015).

5.3 Advantages

Virtual reality experiments offer unique advantages to experimental economists, including the combination of experimental control and increased naturalistic context. Some of the most recent VR publications concern topics relevant in economics, such as helping behavior, cheating behavior and real-effort tasks. This section discusses these advantages by describing the available VR experiments concerning economic topics to give an idea of the possibilities of VR for economic experiments. A more complete overview of recent virtual reality experiments can be found in Table 1. The table shows only high-immersive VR experiments, although some desktop experiments are described in the paragraphs below for their innovative research design and their possibility to be extended to more immersive VR equipment. Another possible direction of experimental economic research is the execution of field experiments in on-line virtual worlds, such as World of Warcraft and Second Life. The present review does not concern these low-immersive virtual worlds, but a recent discussion can be found in Innocenti (2017), who argues that VR experiments (both low and high immersive) can be classified as framed field experiments.

5.3.1 Experimental control

One of the important advantages of virtual reality is its high level of experimental control. Outdoor environments can be tested without problematic interference of unintended contextual cues such as sound, smell and weather. Moreover, as Fox et al. (2009) phrase it: “VR can be used to create stimuli that are unavailable or difficult to manage in the real world, such as large crowds, snakes, or children” (p.101). Using VR in addition to traditional lab or field experiments could solve the lack of exact replication in the social sciences that some researchers consider problematic (Blascovich et al., 2002; Rebelo et al., 2012). Furthermore, VR can offer high standardization in contexts that traditionally lacked it, such as social interaction. For example, Slater et al. (2013) used the standardization possibilities of VR to examine in-group versus out-group behavior. In particular, the authors studied the beliefs of 40 Arsenal[4] supporters about the relationship between victim and perpetrator in a violent pub situation. An argument was simulated between a victim wearing a football-shirt/Arsenal-

[4] I.e. the football club.

Table 5.1: Overview of papers using high-immersive virtual reality experiments

Publication	Research question	Dependent variable	Tool	N	Field
Bailenson et al. (2003)	What interpersonal distance do participants keep towards virtual humans?	distance	HMD	160	soc psy
Bailenson et al. (2005)	Do listeners show more agreement with a presenter who is gazing at them?	gauged social presence	HMD	72	comm
Slater et al. (2006)	To which extent do participants respond to an extreme social situation (Milgram) as if it were real, even though it is VR?	shocks administered, skin conductance, hr	CAVE	38	soc psy
Yee and Bailenson (2007)	Does behavior conform to a digital self-representation independent of how others perceive them?	ultimatum game	HMD	50	comm
Gillath et al. (2008)	What is the effect of context on helping? (businessman / beggar)	helping, empathy scale	HMD	107	psy
Fox et al. (2009)	Can real-time vicarious reinforcement (avatar losing/gaining weight) improve exercise behavior?	exercise repetitions	HMD	189	clin psy
Hershfield et al. (2011)	What is the effect of age-processed renderings of future self on saving behavior?	choice task	HMD	103	eco
Latu et al. (2013)	Do successful female role models empower women's behavior in a leadership task?	speech length & quality	HMD	149	soc psy
Peck et al. (2013)	Can embodiment in a different skin color change racial bias?	IAT	HMD	60	soc psy
Rosenberg et al. (2013)	Does giving people superpowers in VR lead them to behave more prosocial in reality?	number and speed of pens picked up	HMD	60	soc psy
Slater et al. (2013)	Under what conditions will a bystander intervene to try to stop a violent attack by one person on another?	number of verbal and physical interventions	CAVE	38	soc psy
van Gelder et al. (2013)	Can exposure to a VR age-progressed self predict delinquency?	cheating (quiz)	HMD	67	crime psy
Dixit et al. (2014)	What is the impact of subjective beliefs of risk on driver safety?	virtual crashes	CAVE*	132	eco
Hadley et al. (2014)	What is the effect of risky cued VREs on physiological arousal?	hr, arousal	HMD	42	clin psy
Kinateder et al. (2014)	What is the influence of a peers on emergency route choice?	movement trajectories	CAVE	42	safety

Gamberini et al. (2015)	What is the effect of time and race on helping in VR emergency?	helping (binary)	HMD	96	psy
Kinateder et al. (2015)	What is the effect of dangerous goods transporters on hazard perception?	movement trajectories	CAVE	40	safety
McCall and Singer (2015)	Do physical movements (or interpersonal distances) in VR predict (financial) behavior outside VR?	distance, gaze direction	HMD	56	soc psy
Murray et al. (2015)	What is the impact of present others on exercise behavior?	distance rowed	CAVE	60	psy
Qu et al. (2015)	Can bystanders' judgments influence a person's beliefs, self-efficacy and emotions?	speech length, arousal, beliefs	HMD	26	edu
Toppenberg et al. (2015)	To what extent are diagnosis (HIV, cancer or broken leg) and sexual orientation related to approach behavior?	distance, speed, head orientation, IAT	HMD	49	soc psy
van Herpen et al. (2016)	Can real-life shopping behavior in a supermarket be captured in VR?	products selected	CAVE	100	marketing
Puschmann et al. (2016)	Can VR-based risk assessments offer an alternative to document-based or CAD-based approaches?	machine operation	CAVE	27	safety
Hale et al. (2017)	Can specific trust towards strangers be measured in a virtual maze task?	directions, advice	HMD	24	soc psy
Schutte and Stilinović (2017)	Can a virtual reality experience increase empathy?	empathy scale	HMD	24	psy
Chittaro et al. (2017)	What are the effects of a VR experience on risk attitudes?	hr, (risk) surveys	HMD	108	psy
DeHoratius et al. (2018)	Quantify the role of product similarity in execution failures	sorting errors	CAVE	87	eco
Gürerk and Kasulke (2018)	Does virtual reality increase charitable giving?	donations, empathy	HMD	61	eco
Kugler et al. (2018)	What is the effect of disgust emotions on trust behavior?	trust game	HMD	104	eco
Graff et al. (2018)	How do tournament incentives and peer effects interact in a dynamic setting?	real effort	CAVE	131	eco
Gürerk et al. (2019)	What is the effect of the presence of a virtual co-worker on real effort?	speed, accuracy	CAVE	108	eco
Mol et al. (2020c)	Can cheating be affected by the presence of a virtual observer?	cheating (mind game)	CAVE	121	eco

Notes: Abbreviations used: comm = communication research, soc = social, clin = clinical, psy = psychology, env = environmental, eco = economics, edu = education science, hr = heart rate. * multi-screen driving simulator.

shirt and the perpetrator. The victim was programmed to look at the participant in some of the conditions. The results show that in-group participants (i.e. Arsenal supporters watching an Arsenal fan being attacked) were more likely to intervene in the conflict than out-group participants. From this in-group, those who believed that the victim was looking at them, intervened more than those who did not believe they were looked at.

Qu et al. (2015) studied a different aspect of social interaction with the help of virtual standardized humans: the effect of bystanders in a classroom setting with a within-subject design. 26 participants were asked to take part in a virtual language lesson where their virtual classmates where whispering either approvingly or skeptically. As a result, participants' self-reported beliefs, self-efficacy and anxiety levels shifted. Furthermore, beliefs about the teacher (whose behavior was in fact always neutral) varied as well, leading participants in the negative-comments condition to think that the teacher disapproved too. On the other hand, participants gave longer answers in the case of positive whispering classmates, which correlated with a lower self-reported level of anxiety.

Recently, Mol et al. (2020c) studied the effects of a virtual observer on cheating in a VR version of the mind game, which is a variation of the die-under-the-cup paradigm. In this game, subjects had the incentive to be dishonest by reporting the highest payoff, without the chance of getting caught. A VR agent as observer allowed for a more naturalistic variation of observability than the typical images of 'watching eyes' in the literature on social control. They found similar levels of cheating as in the conventional lab equivalent of the mind game. The presence of the virtual observer did not affect cheating, compared to the same VR environment without a virtual observer. However, participants cheated significantly more when the virtual observer was passively seated in a corner, rather than actively staring at the participant. The authors discuss the impact of human-like virtual observers on cheating behavior, which involves more than simple cues of social control. Note that using VR experiments eliminate the need of confederates, an experimental practice using deception, which is generally disapproved by economists (cf. Hertwig and Ortmann, 2001; Ortmann and Hertwig, 2002).

5.3.2 Experimental realism

In the past decades, economic experiments were not only used to test theories, but also to motivate and develop new theories, which makes the external validity of experiments more essential (cf. Schram, 2005). The highly naturalistic situations participants experience in a VR experiment can generate more natural responses than traditional lab experiments (Fox et al., 2009). By visualizing life-like situations, emotional arousal can be elicited to the extent that post-traumatic symptoms may be reported. Dibbets and Schulte-Ostermann (2015) used VR to induce a mild trauma (a scene about physical abuse) upon participants and found a large degree of presence and immersion, as well as traumatic symptoms in the week after the view.

Participants can thus be confronted with decisions in a more natural way (naturalistic cues) than via conventional ways such as vignettes, scenarios and self-report questions. The scenario-approach is typically low in ecological validity: asking participants what-if questions requires them to imagine the situation, where the quality of imagination can never be controlled. Virtual reality allows for the careful controlling of perception confounds, by showing participants the context of the question. This way, participants have no need to 'bring' their own frames or life experiences to the game (see Harrison et al., 2011, for a careful discussion on this topic). For example, DeHoratius et al. (2018) used a virtual conveyor belt as an environment similar to the work environment of many retail employees to study the effect of packaging and similarity on sorting errors. Their results have clear implications for retailers who wish to improve employee productivity, for example by adding visual cues. Haruvy et al. (2017) also take advantage of rich contextual cues to study the effect of communication and visibility on contributions in a public goods game. The authors contrast an abstract zTree environment with a 3D avatar-based virtual world and find that communication improves contributions in both environments, but that communication and visibility are complements in the virtual world.

Besides, the high degree of experimental control in VR allows for repeated viewing of the same or slightly different environments, which is one of the reasons that VR is applied in the treatment of phobias (Wiederhold and Bouchard, 2014). In economics, this gradual change of environments can be used to study preferences that are hard to imagine, for example in the domain of risky and dangerous decisions. The outcomes of hypothetical risky decisions, such as damage due to (natural) disasters and accidents might be visualized. Research from psychology shows that VR exposure might change participants' risk perception, depending on the VR environment (Chittaro et al., 2017). Furthermore, VR allows for detailed studies on subjective probability formation based on simulated environments, in contrast to abstract lab experiments based on simple objective probabilities that are not so common in the field (Harrison et al., 2015). As there is considerable heterogeneity in risk attitudes across elicitation methods and domains (Csermely and Rabas, 2016; Pedroni et al., 2017), such rich visualizations of (compound) risk and uncertainty might be of interest to economists. Using an environment that is very close to the natural environment in which people make decisions, while controlling for perception confounds, is a new type of experiment that could add valuable contributions to experimental economics.

Furthermore, the higher level of presence that can be achieved by VR, in comparison to mainstream photos or videos, may enhance emotions, empathy or altruism. 360° VR videos can be used to induce stronger emotions in participants than conventional methods such as images or 2D video (Diemer et al., 2015; Schutte and Stilinović, 2017). Many researchers have shown that emotions can alter decisions in economic contexts (Fiala and Noussair, 2017; Martinez et al., 2011; Lin et al., 2006). In a recent experiment, Gürerk and Kasulke (2018) presented participants with a real effort task to earn their endowment, which they could donate later to a local refugee organization. Before donating, participants viewed a 360 $^\circ$ video of the destroyed city of Aleppo in Syria on a computer screen or a VR version in a HMD. A control group watched no video at all. Besides the donation decision, the researchers measured empathy with the Interpersonal Reactivity Index questionnaire. They found the highest scores in the VR treatment, both for empathic concern as for donations. These results are in line with the findings of Schutte and Stilinović (2017); greater engagement and higher reported empathy by participants in the VR condition compared to the control condition where a documentary on a refugee camp was presented in 2D format. Another illustrative example is provided by Kugler et al. (2018), who used HMDs to induce disgust emotions in participants, to study the effect on trust in an economic trust game. They find that disgusted participants are less trusting, presumably because they misattribute their emotions to the course of the game.

It should be noted that a rich and natural set of stimuli or context that can be provided by VR is not useful for all domains of economics. Many economic experiments are mainly abstract and neutrally framed and it is not the aim of this review to change such good practice. However, in some domains VR could help to generate more stable decisions in complex environments.

5.3.3 Logging of responses

Another interesting feature of VR devices is the automatic logging of response data such as movement and rotation (Parsons, 2015), which can be captured in detail depending on the hardware used. Gillath et al. (2008) for example, measured individual differences in helping behavior of a virtual person in need. In a first experiment, participants encountered a blind man in need (he lost his walking cane) on an urban side walk. Apart from self-report empathy measures, physical helping (approaching) responses were recorded and coded. The results showed that 30% of participants expressed their concern (either verbally or by offering help) when approaching the man, which is a similar measure as has been found in field experiments outside VR (Guéguen and De Gail, 2003). In a second experiment the blind man was replaced by either a beggar or a businessman. Gaze direction of the participant and distance to the man were measured by the HMD and the results from the first experiment were replicated. A different example of a VR study using detailed logging of response data, is Gürerk et al. (2019), who simulated a virtual conveyor belt and asked participants to sort pieces according to the color on one side of the blocks. The controllers used by the participants to rotate the blocks in the virtual environment allowed the authors to rate performance both on speed and accuracy, while manipulating the performance of a virtual co-worker in the background. The authors were able to "evaluate how subjects make the trade-off between quantity and quality as a function of the economic incentives provided" (p. 4). They found that competitive subjects perform better when working with a highly productive peer compared to when they work in the presence of a low-productive co-worker.

McCall and Singer (2015) also took advantage of the detailed logging of interpersonal distances by studying approach and avoidance behavior in a virtual environment. First, participants were asked to play a trust game twelve times on a desktop computer with two confederates (players A and C) as opponents (one fair and one unfair player). In the next stage, participants were immersed in a VR with two agents: players A and C. Participants were led to believe that these agents were avatars, controlled by the actual humans that they played the trust game with in the first stage. The task performed in VR was a memory task, while the dependent variable of interest was the distance between participants and the other players. In the last stage (outside VR) participants could punish the other player(s) by paying to remove tokens from another player. Participants came significantly closer to the fair agent than to the unfair agent. Interestingly, those participants who chose to punish considerably, spent more time in front of the unfair agent, which was interpreted as mildly aggressive behavior.

Overall, the potential of VR in the automatic logging of responses is considerable, as it offers new objective variables, such as gaze direction and hand rotation. It should be acknowledged that detailed movement tracking in itself does not provide added value to all economic experiments. Yet some topics, such as principal agent paradigms using real effort tasks, may benefit from the detailed analysis of time, position and visibility (DeHoratius et al., 2018). Note that the greatest precision in the measurement of human movement can be accomplished by the use of motion trackers, while an HMD or controller will yield only data on the head or the hand movement of the participant. Besides, eye trackers may be combined with VR hardware, which enables researchers to track precisely which information participants are viewing (Meißner et al., 2017). Future developments may improve automated interactivity, for example by simulating a corresponding responsive negotiator in front of the participant. Evidently, the recommended hardware selection depends on the specific research question at hand.

5.3.4 Visualizing complex questions

Virtual reality is frequently used to visualize complex problems in environmental science, as well as in landscape architecture (Patterson et al., 2017) and construction business (Portman et al., 2015; Pérez Fernández and Alonso, 2015). For example Patterson et al. (2017) used low-immersive VR to refine the coefficients of discrete choice experiments on neighborhood choice. Another complex environment that can benefit from VR experiments is transportation. Dixit et al. (2014) used virtual reality driving simulators to study the subjective beliefs of participants under different risky traffic scenarios, while controlling for experience and risk attitudes. They found that participants who crashed were generally more optimistic about their success in the task, although this was unrelated to risk attitudes.

Virtual reality allows for naturalistic exploration of large areas with multiple users simultaneously, which is useful for environmental scientists to study wild fire prevention (Fiore et al., 2009), land use change (Bateman et al., 2009) and coastal erosion management (Matthews et al., 2017). Bateman et al. (2009) performed a choice experiment on coastal land use both with and without a virtual reality visualization, while keeping the objective information presented constant. The VR visualization showed a smaller variability in elicited preferences and a smaller the willingness to pay (WTP) - willingness to accept (WTA) gap. Matthews et al. (2017) used virtual environments in a desktop choice experiment about coastal erosion managment. In line with the results of Bateman et al. (2009) the authors found a significant decrease in choice error and a different WTP in the virtual reality group as compared to the static images control group. Fiore et al. (2009) showed a VR visualization of forest fire consequences to study individuals' assessment of risks of prescribed burns, in comparison to a multi-image visualization of the consequences. A multiple price list was used to determine subjective beliefs of the subjects with regard to the risk of the simulated forest fire. The results showed that the subjective beliefs in the VR visualization treatment were closer to the actual risks than the subjective beliefs in the image treatment. The authors conclude that the primary benefit of VR is the naturalistic way in which counter factual scenarios can be generated. This is particularly important in environmental issues, where individuals often have difficulty with the comprehension of possible consequences in the long run, for example in assessing the effects of global warming.

In a follow up study, Harrison et al. (2015) studied the relationship between prior experiences and perception formation in natural risky decision settings by forest ranger experts and non-expert residents. They found that experts are focused too much on prior beliefs and therefore do not outperform non-experts in estimating compound risks.

5.3.5 Conducting "impossible" experiments

One of the unique advantages of virtual reality is that it gives the experimenter the freedom to test situations that would never be possible in the real world. For example, Rosenberg et al. (2013) offered participants the ability to fly over a virtual environment, after which they measured the degree of helping. They found that participants who were able to actively fly in the VR environment (as opposed to the control group, who were seated in a virtual helicopter) picked up more pens in a subsequent helping task. Gamberini et al. (2015) manipulated the ethnicity of the victim in different emergency situations (*None* versus *Time pressure* versus *Fire*). The experimenters sent participants into a virtual building with the assignment to leave the building after exploring it. Suddenly, a screaming voice asked for help from the cafeteria inside the virtual building. In addition to the binary variable helping (defined as moving back to the cafeteria before moving to the emergency exit), the researchers registered promptness, number of collisions with the walls and number of backward movements. They found that 68% of the participants helped, but a significant racial bias was found (black victims were helped less often than white victims).

Other possibilities include experiments that would be unethical in the real world, such as showing the (fatal) results of a choice in a moral dilemma (see e.g. Navarrete et al., 2012, for the trolley problem in VR) or replicating the classic Milgram obedience experiment (Slater et al., 2006). Responses in risky situations can be trained repeatedly without exposing participants to unethical situations. Evacuation behavior can be tested experimentally with non-expert participants, for example in a virtual tunnel-fire (Kinateder et al., 2014) or during an earthquake (Lovreglio et al., 2017). Zaalberg and Midden (2013) exposed participants to a (desktop) VR simulation of a dike breach to test how flood awareness can be improved. The results showed that information search, evacuation motivation, and stated preference to buy flood insurance increased after the VR simulation compared to a film and slide show version of the dike breach.

A further promising approach is to use VR to visualize the future, thereby confronting participants with consequences of their behavior. This approach was tested successfully in the domain of exercise behavior, where participants were encouraged to exercise in response to a virtual future self who either gained or lost weight (Fox and Bailenson, 2009). The results showed that participants exercised more when they saw a virtual representation of the self that changed in body weight in accordance to exercise efforts, than without a responsive virtual representation. The same idea can be applied to intertemporal choice to increase saving behavior, by showing participants a virtual construction of their elderly self. Hershfield et al. (2011) embodied participants in a virtual construction of an elderly self and let them through a mirror with their (visually) elderly body. After a short walk to get familiar with their body in the virtual environment, participants could watch their virtual body in a virtual mirror, which lead to increased saving behavior in a subsequent task. Interestingly, embodiment in another elderly person did not increase saving behavior. In a related experiment, van Gelder et al. (2013) used the same method to construct projections of participants (present self) and age-processed these (future self). The authors compared cheating behavior after exposure to either their present self or their age-processed future self and found that interaction with the future self significantly decreased cheating.

5.4 Challenges

While VR experiments as a research tool has many advantages, a number of challenges need to be addressed. The following section discusses the current state of affairs with regards to costs, specialist skills, simulator sickness, familiarity, naturalistic avatars and lab time.

5.4.1 Costs

The costs of a virtual reality lab can be divided into two categories: hardware and software. As mentioned before, different possibilities exist for the hardware set-up. In addition to the headset and controllers, many HMDs require a platform (e.g. desktop computer, smartphone) to render the virtual environment, although "standalone HMDs" are a recent addition to the VR hardware market[5]. The costs of an HMD set-up range from €10 (excluding smartphone) for the Google Cardboard[6] to the more expensive displays with a higher resolution and a larger field of view, such as the Samsung Gear VR[7] (€115, including one controller, excluding smartphone), the Oculus Rift[8] (€450, including two controllers) and the HTC Vive[9] (€600, including two controllers.). The most expensive VR headset at the time of writing is the Pimax 8K[10]. This headset can be purchased from €900 (excluding controllers) and offers a 200-degree field of view which comes closest to the 220-degree field of view of the human eye. Note that all devices try to strike a balance between costs, wearability and screen quality. Recent releases of new VR products have focused on improving screen resolution and field of view. A larger field of view could decrease simulator sickness susceptibility as it would require less head movement (Serge and Fragomeni, 2017). A larger screen resolution is desirable to increase immersion and thus presence, especially when it is detailed enough to remove the pixelated view known as *screen door effect*[11] that arises when the display is magnified in front of the eyes of the user. Solutions to the screen door effect are in development (Cho et al., 2017; Sitter et al., 2017) and might be implemented in the newest (business) releases of VR hardware. A recent discussion of screen latencies for both CAVE and HMD can be found in Meißner et al. (2017). Note that these technological advancements are costly and might increase hardware prices. Researchers who

[5] For example Oculus Go (€250, https://www.oculus.com/go/) or HTC Vive Focus.
[6] https://vr.google.com/cardboard/get-cardboard/.
[7] https://www.oculus.com/gear-vr/.
[8] https://www.oculus.com/rift/.
[9] https://www.vive.com/eu/product/.
[10] https://pimaxvr.com/products/pimax-8k-vr-headset/.
[11] The term originates from the comparison to a view through a fine mesh as in anti-insect screen doors

wish to purchase VR HMD equipment could compare the current HMD devices on computer magazine websites.[12]

The hardware set-up costs of a CAVE are considerably higher. Prices range from €5.000 for a 3D projection screen to €20.000 for a simple CAVE to €1.5 million for a complete CAVE including stereoscopic glasses, motion capture and sensing technology (Pérez Fernández and Alonso, 2015). Note that these prices are an indication and the VR technology market is constantly developing. Different hardware set-ups require different software. Most 3D scripting languages are interchangeable but caution is required when avatars are used in combination with motion capture: using the right skeleton[13] is crucial. Many of these programming applications are open-source software and therefore free to use while others are commercial, but academic subscriptions are available. Different software is necessary for each step in the process: from constructing the 3D environment (e.g. Autodesk 3DS Max, Maya, Sketchup) to texturing (e.g. Adobe Photoshop) and scripting (e.g. Unity, Unreal, Vizard). For a comprehensive overview of the process of developing a virtual environment, see Chapter 11.4 in Wiederhold and Bouchard (2014).

5.4.2 Specialist skills

One might fear that the construction of a VR environment requires specialist programming skills. In essence this is true but the accessibility of software (e.g. Vizard, Unity 3D) and assets has been greatly improved over the past decades. In the words of Fox et al. (2009): "a computer science degree is no longer necessary to understand and implement them (VE environments)" (p. 106). In addition, graphic simulations, avatars and 3D renderings can be found and bought on the Internet, where a specialist marketplace has been created in parallel to the developments in the gaming industry.

5.4.3 Simulator sickness

Probably the best documented negative side-effect of the use of VR equipment is simulator sickness, a type of motion sickness. During or after exposure to a virtual environment, a mismatch between vision and input of the vestibular system can cause symptoms such as nausea, blurred vision and instability (Rebelo et al., 2012). Simulator sickness seems to get worse in the case of a large display delay: a temporal delay between the physical movement of the participant and the updated screen. However, due to increased computational power, recent VR equipment is constructed to reduce the display delay to the minimum by maximizing the field of view and the refresh rate (Parsons, 2015). A larger field of view inside a HMD would require less head movement (Serge and Fragomeni, 2017), decreasing the likelihood of simulator sickness.

[12] See e.g. https://www.slant.co/topics/1668/~best-vr-headsets/.
[13] The basic joints structure to which different avatars and animations can be added.

Unsurprisingly, these technological advancements are a costly part of the VR hardware price. The severity of simulator sickness symptoms is further connected to the type of VR equipment, where HMDs may lead to more severe symptoms than projection screens (CAVEs) and desktop computers (Sharples et al., 2007). Practical experience from the DAF Technology lab at Tilburg University demonstrates that control over the navigation in the virtual environment decreases simulation sickness, while passive participants experience more simulation sickness. A recent test with 24 participants using the HTC Vive found no uncomfortably high sickness ratings on average (Serge and Fragomeni, 2017). Another recent study with the Oculus Rift found that some participants experience simulator sickness, but much depends on the type of game (Munafo et al., 2017). Particularly movements in the game that are not synchronized with real (bodily) movements are likely to cause simulator sickness, such as riding a virtual roller-coaster while sitting in a fixed (non-moving) chair."

Another parameter in the context of simulator sickness is exposure duration. Longer exposure tends to produce more symptoms (Stanney et al., 2003), although after approximately 60 minutes habituation can occur: participants will adapt to the new environment, leading to a decrease in symptoms. Habituation will increase by offering repeated (short) exposure periods. The availability of breaks can decrease the severity of simulation sickness (Rebelo et al., 2012) but it may have a negative effect on presence.

5.4.4 Familiarity

Some participants are more familiar than others with the usage of VR equipment, for example because they play 3D video games frequently. In rare cases this may cause a confounding factor in the analysis of the results. A few researchers have argued that individual differences in computer familiarity can indeed moderate the effect of VR interventions (Turner and Casey, 2014). However, little research has been performed to back up this claim. A self-report question about familiarity with video games and VR equipment may be asked in the post-experimental questionnaire to control for this effect.

5.4.5 Naturalistic avatars

Social interaction in virtual reality requires avatars. While naturalistic avatars are not crucial to induce a feeling of interaction or embodiment, they have a powerful impact on presence. Detailed and naturalistic avatars demand computational power to render and more time to animate. VR software often comes with some free stock avatars (see Figure 5.2a) and extra avatars can easily be bought on-line. The quality of these avatars has improved over the past decade, although the face is difficult to model and each muscle should be animated. To circumvent this problem, one could consider to use avatars who do not face the participant, for example because they perform a task at the next conveyor belt (DeHoratius et al., 2018; Gürerk et al., 2019). Animations are available on-line, including many free ones (see Figure 5.2a). However, joining these animations and adding a certain movement path requires software skills. Alternatively, a motion tracker suit could record the animations, which gives very natural results but adds another hardware item to the VR startup costs[14]. Recent developments in the domain of motion tracking combine the data of several trackers (e.g. €120 HTC Vive Tracker) with motion capturing software[15] to track and model real-time full body avatars.

Note also that the focus of the gaming industry is mainly on fantasy characters, which leads to a large supply of monsters, soldiers and animé characters, while "normal" people are harder to find. A solution would be to create your own character[16], which gives the opportunity to confront participants with subtle variations in avatars, but comes at the expense of programming time and requires software skills. A recent technique is to make 3D scans of real humans, which results in a detailed and naturalistic avatar (see Figure 5.2b). Achenbach et al. (2017) present a 3D-scanning setup which takes less than ten minutes to complete, enabling researchers to scan each experimental subject prior to VR exposure.

(a) Avatars of Adobe Mixamo

(b) Avatars created with 3D scans[17]

Figure 5.2: Examples of naturalistic avatars. (Animated version online: https://bit.ly/2VoxysL, https://bit.ly/2VsIbei)

5.4.6 Lab time

In comparison to experiments in traditional labs with multiple workstations, VR experiments will require more time to conduct because there is often only

[14] See https://www.rokoko.com/ (from €2.500) or https://neuronmocap.com/ (from €1.000).

[15] https://www.ikinema.com/full-body-ik-for-vr

[16] For example with Adobe Fuse: https://www.adobe.com/products/fuse.html/.

[17] Reprint courtesy of Latoschik et al. (2017).

one CAVE or HMD available, taking about 10 to 30 minutes per participant, sequentially. However, the costs (especially of HMDs) may decrease in the future and the set-up is not time-consuming, as it is with invasive biometric tools such as heart rate, fMRI and EEG.

5.5 VR in practice

Even though VR experiments offer the opportunity to increase external validity, that does not mean that it happens by design or without effort. Harrison et al. (2011) discuss some issues on both external and internal validity in the design of VR experiments, including perception confounds and sample selection. Some practical suggestions with regards to conducting a VR experiment are discussed below.

5.5.1 Ethical use of VR

As with any new technology, the use of virtual reality might pose risks that are yet unknown to its users. VR might not seem as invasive as several biometric methods, but it has the potential to have lasting effects (cf. Dibbets and Schulte-Ostermann, 2015). It is therefore strongly recommended to adhere to the *VERE code of conduct for the ethical use of VR in research* by Madary and Metzinger (2016) and to exclude vulnerable participants from the experiment. These at-risk participants include epileptic patients and patients with psychosis or personality disorders as they could possibly mix up reality with the virtual environment (Wiederhold and Bouchard, 2014). Most economists might not be handling a clinical population, but the recommendations on non-maleficence and informed consent are important for all disciplines.

5.5.2 Minimizing simulator sickness

Even though simulator sickness is not commonly reported with modern-day VR facilities, researchers take measures to mimimize and track potential sickness. Sharples et al. (2007) report several guidelines for VR researchers to minimize the negative effects of simulator sickness, such as giving participants control over their movement in the virtual environment (cf. Wiederhold and Bouchard, 2014). A further recommendation is to be aware of physiological signs of participants suffering from simulator sickness (sweating, pallor, fidgeting with HMD, closing eyes). VR researchers have developed different measures in order to track simulator sickness, including physiological measures such as EEG, blood pressure and heart rate. Still, the most widely used measure is a self-reported questionnaire, such as the simulator sickness questionnaire (SSQ, Kennedy et al., 1993). To prevent an experimenter demand effect, one might consider conducting only the post experimental SSQ (see Young et al., 2006, for a discussion on this issue). The SSQ has recently been revised by Balk et al.

(2013) to update the factors with current technology and to examine dropout predictability. They conclude that the SSQ is "still relevant today" (Balk et al., 2013, p.263), and is therefore recommended for future VR research.

5.5.3 Measuring presence

Without a substantial level of presence, the benefits of a VR experiment compared to a conventional lab experiment could be neutralized. When a certain condition is clearly more engaging for participants than another, treatment effects might be confounded by presence levels. Thus, researchers may want to control for presence levels of participants. The traditional method to measure presence is with a self-reported questionnaire (c.f. Witmer and Singer, 1998; Schubert et al., 2001), although questionnaires are known to have limited stability (Slater, 2004). Most presence questionnaires use seven-point Likert Scales on questions such as *How aware were you of events occurring in the real world around you*, *How natural did your interactions with the environment seem* and *Somehow I felt that the virtual world surrounded me*. Slater (2009) distinguishes two types of presence: place illusion and plausibility. Place illusion refers to the physical feeling of being in the virtual environment, where plausibility captures the idea that whatever happens in the virtual environment is real, regardless of the knowledge that the virtual environment was constructed by technology. Subjects with strong feelings of plausibility would respond similarly in reality as in the real world. Considering that conventional presence questionnaires focus mostly on place illusion, Qu et al. (2015) developed a presence response scale to capture plausibility scores. Recently, Diemer et al. (2015) suggested that participants might judge their presence level based on immersion, as well as on emotional arousal. Thus, in certain emotional (e.g. fearful) situations, one might measure presence by physiological measures, such as galvanic skin response. A detailed discussion of measuring presence can be found in Sanchez-Vives and Slater (2005).

5.6 Conclusion

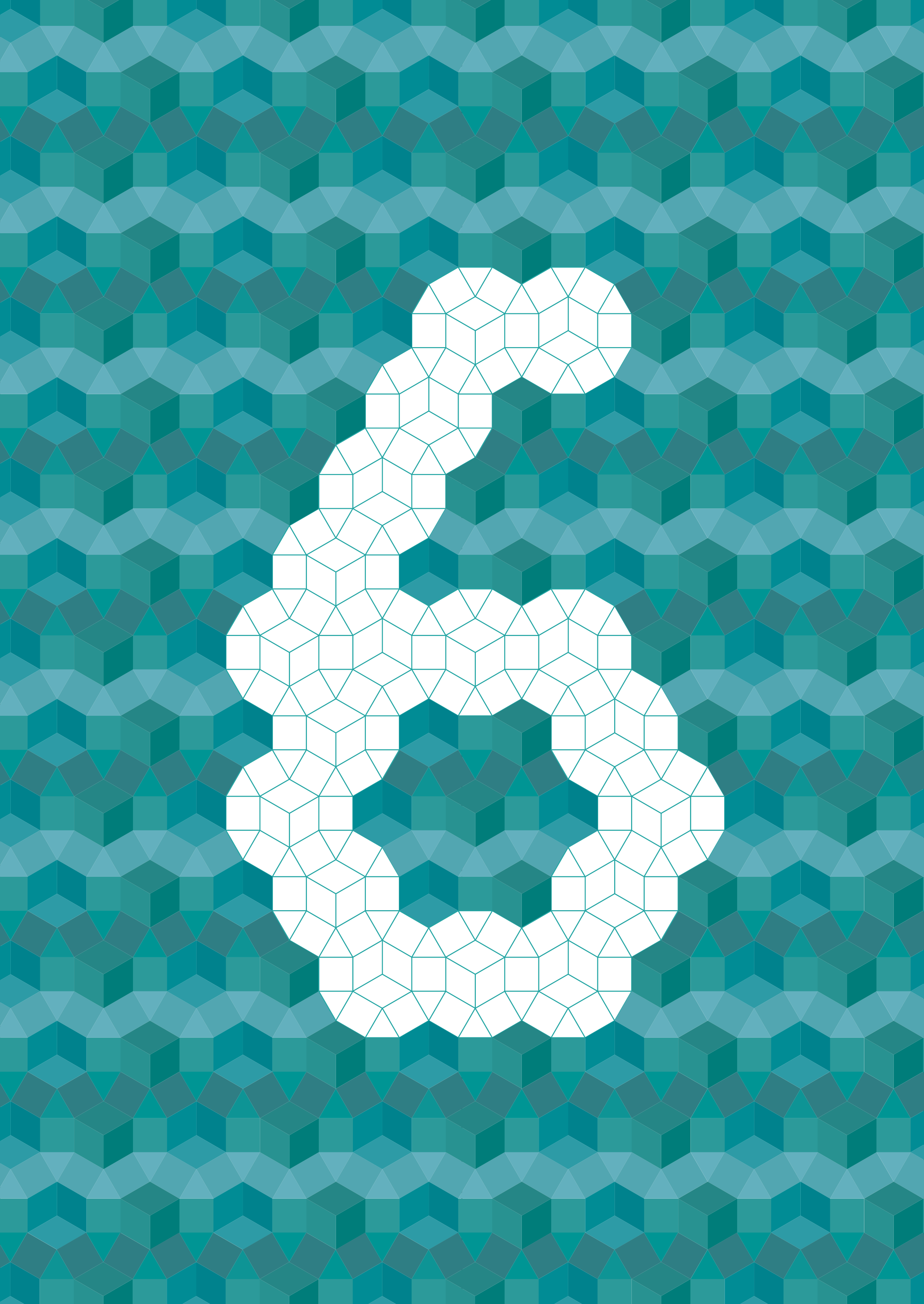
This review aimed to give a critical overview of the possibilities and challenges for experimental economics in high-immersive virtual environments. While VR is becoming more mainstream in disciplines such as engineering, psychology and spatial planning, VR experimental economics is still in its infancy. Some domains of economics could benefit from visualizing a rich and natural context that can be provided by VR.

One of the key advantages of VR above conventional field experiments is that it is relatively easy to control for confounding factors such as weather, gender and non-verbal cues. Many economic field experiments could be improved by this technology, leading to more robust findings and helping to exclude alternative explanations. Thanks to the improved technologies in the past decade, perceived realism (presence) now allows for VR research to move from methodological publications to experiments with respect to content and the objective measurement of human movement may offer new insights. Furthermore, experiences in VR seem to extend to real life and a close parallel has been found between behavior in VR experiments and conventional labs. By carefully controlling the context of an experiment, virtual reality could bring a bit of the field into the laboratory. VR experiments can be considered framed field experiments, as the context they provide to subjects is completely controlled by the experimenter (Innocenti, 2017). VR is a promising new research tool when it comes to visualizing complex economic questions. Future research with virtual reality could help to visualize those questions, such as belief elicitation, risk perception and preference, gain-loss asymmetry in environmental planning and inter-temporal choice. By helping people to visualize these situations, they might be better able to form stable beliefs and preferences. Other suitable topics include social interactions that are not easily controlled in field experiments and a detailed logging of responses. Social dilemmas may be presented much more naturally than in a conventional computerized experiment and games may be played with multiple players in the same VR environment. Alternatively, consider a VR physical real effort task (e.g. where subjects have to physically move many objects) to examine a response to incentives, where current real effort tasks may be insufficiently elastic (Araujo et al., 2016). Nevertheless, caution is required to prevent that subjects simply enjoying the virtual environment show an even more inelastic response to incentives.

The main drawbacks of VR experiments are the costs of equipment and the required programming skills, although developments in the game industry might lead to cheaper devices and straightforward software, as well as improved specifications to minimize simulator sickness. At any rate, researchers should adhere to the conduct for the ethical use of VR, be aware of signs of simulator sickness and pay careful attention to the measurement of presence. Note that as technology advances, VR experiments have the potential to increase both in the realism and the control dimension. At the moment, the costs of starting a simple economic VR experiment are decreasing and the possibilities for testing and developing behavioral models are promising. Many university campuses around the globe already have a VR lab, for example in a psychology or computer science department. Collaborating with someone familiar with VR equipment and programming is an affordable way to conduct an economic experiment in VR. It might be the right time to consider using it.

Acknowledgements

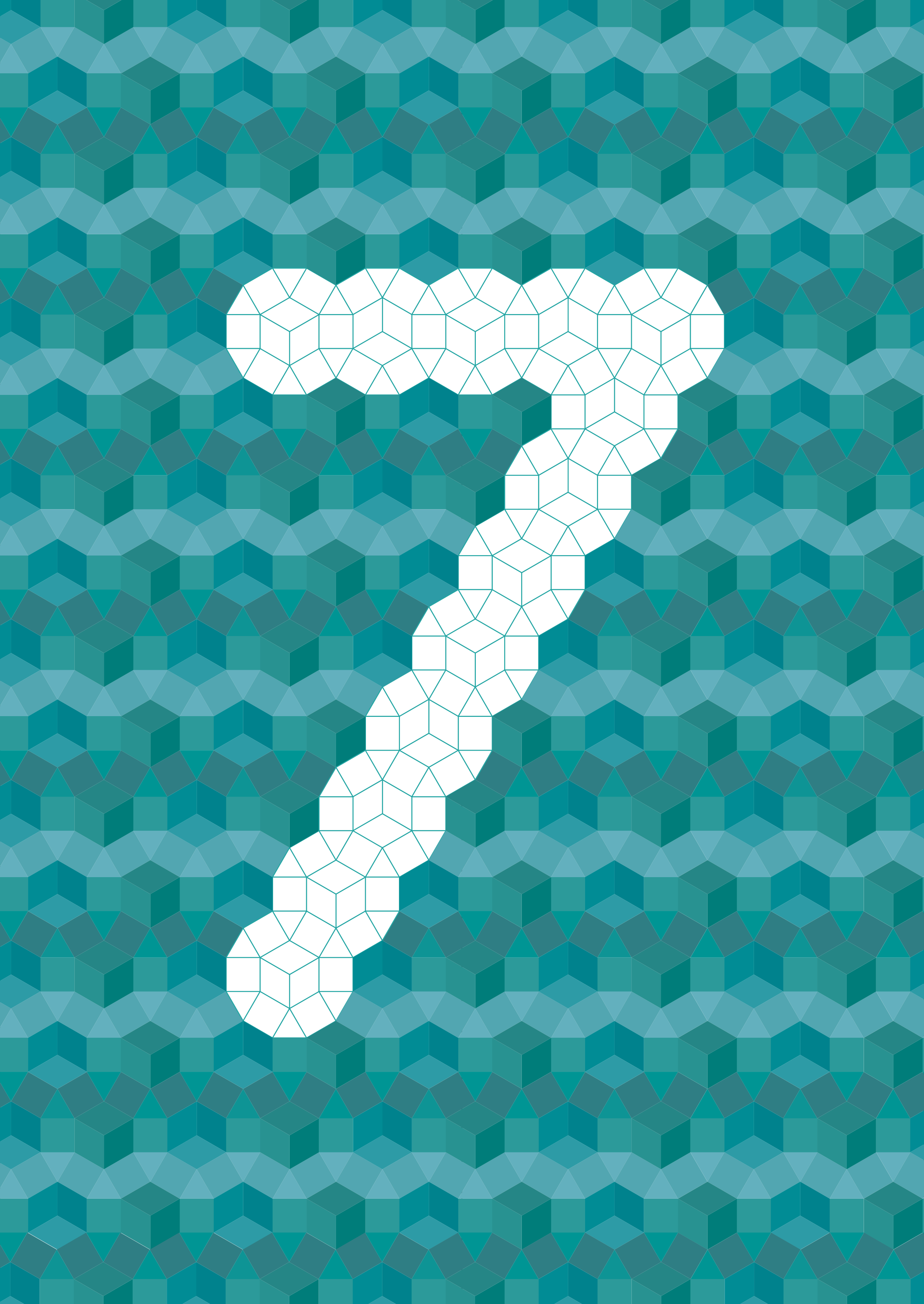
I wish to thank Jan Potters and Eline van der Heijden for their invaluable comments and suggestions. The first version of this review was written during my stay at CentER and the DAF Technology Lab at Tilburg University. I acknowledge with thanks the assistance given by Daniel Roth in permitting the reproduction of Figure 5.2b, reprinted courtesy of Latoschik et al. (2017), ACM Digital Library.



After the virtual flood: risk perceptions and flood preparedness after virtual reality risk

Many individuals experience problems understanding and preparing for low-probability/high-impact risk, like natural disasters and pandemics – unless they experience these events, yet then it is often too late to avoid damages. Individuals with recent disaster risk experience are, on average, better prepared. This seems to be mediated through emotions and a better understanding of the consequences. In this study, we use immersive virtual reality (VR) technology to examine whether a simulated disaster can stimulate people to invest in risk reducing measures in the context of flooding, which is one of the deadliest and most damaging natural disasters in the world. We investigate the possibility to boost risk perception, coping appraisal, negative emotions and damage-reducing behavior through a simulated flooding experience. We find that participants who experienced the virtual flood invest significantly more in the flood risk investment game than those in the control group. These effects are persistent up to four weeks after the VR intervention.

This chapter is published as:
Mol, J. M., Botzen, W. J. W., & Blasch, J. E. (2021). After the virtual flood: risk perceptions and flood preparedness after virtual reality risk communication. *In review with Judgment and Decision Making.*


7

All by myself? Testing descriptive social norm-nudges to increase flood preparedness among homeowners

Nudges based on social norms (norm-nudges) can be compelling behavioral interventions compared to traditional interventions such as taxes and regulations, but they do not work in all circumstances. We tested two empirical norm-nudge frames in an online experiment on taking measures for flood preparedness with large samples of homeowners (N = 1805) in two European countries, to evaluate the possible interactions between norm-nudge effectiveness, individual characteristics and intercultural differences. We contrasted these norm-nudge treatments with a control and norm focusing treatment by asking respondents to express their beliefs about what other respondents would do before making a decision relevant to their own payoff. We find no evidence of a treatment effect, suggesting that our social norm-nudges do not affect flood preparedness in the context of a flood risk investment game.

This chapter is published as:
Mol, J. M., Botzen, W. J. W., Blasch, J. E., Kranzler, E. C., & Kunreuther, H. C. (2021). All by myself? Testing descriptive social norm-nudges to increase flood preparedness among homeowners. *Behavioural Public Policy,* 1-33.

7.1 Introduction

Social norms are rules of behavior that are commonly approved by society while personal norms represent what people believe to be appropriate behavior for themselves (Bicchieri, 2006). If deviations from a norm are likely to be sanctioned by society, individuals are inclined to follow the norm. A popular behavioral intervention based on social norms is a norm-nudge (Bicchieri and Dimant, 2019), which encourages certain behavior by informing individuals about the actions of others, for example by showing energy conservation behavior of neighbors (Allcott, 2011) or tax compliance rates of fellow citizens (Hallsworth et al., 2017). Norm-nudges may prompt people to act the way others are acting, because humans are inclined to model behavior on what others do, or what they believe others do (Bicchieri and Dimant, 2019).

Norm-nudges are compelling interventions because they are cheap, easy to implement and less prone to political resistance, compared to traditional interventions such as taxes or regulations (Benartzi et al., 2017). Nevertheless, norm-nudges do not work in all circumstances and their effectiveness depends on the design of the norm-nudge (Hummel and Maedche, 2019). Moreover, there is a risk that a norm-nudge will be ineffective (see e.g. Mackay et al., 2020; Chabé-Ferret et al., 2019) or even backfire, if not properly tailored to the population and context of interest (Hauser et al., 2018). For example, norm-nudges may backfire when they provide information about norm-violating behavior (e.g. tax evasion), which may lower motivations for compliance (Richter et al., 2018). Thus, it is relevant to test different kinds of norm-nudges and empirically assess their effectiveness across contexts.

The aim of this chapter is to examine the effectiveness of different norm-nudge messages with varying information in increasing individual investments in flood damage mitigation measures. Moreover, this study aims to examine heterogeneity in individual responses to these nudges as well as in the individual investment amounts, including individual characteristics and intercultural differences. We test two empirical norm-nudge frames with a large sample in Spain and the Netherlands and contrast these with a control treatment and a norm focusing treatment. In the latter, respondents are asked to guess what other respondents would do before making an investment decision relevant to their own payoff. This task has been shown to influence behavior in past work, namely by increasing donations to charity (Bartke et al., 2017) and encouraging pro-social behavior, such as sharing funds.

Many studies on norm-nudges have focused on applications for health, finances, the environment and energy (Hummel and Maedche, 2019; Abrahamse and Steg, 2013). To our knowledge, previous research has not explored the effect of norm-nudges in the context of natural disaster risk reduction measures such as investment in flood damage mitigation. Over the last decades, natural hazards such as floods have increasingly impacted society, and this trend is expected to continue in the coming years due to climate change and population and economic growth in disaster-prone areas (IPCC, 2012; Munich RE, 2018). Floods are among the most costly natural disasters (UNISDR, 2015). Despite the availability of cost-effective measures that limit flood damage to buildings (Aerts et al., 2013), few people in flood-prone areas invest in or implement such measures (Botzen et al., 2019a). This highlights the importance of studying whether norm-nudges can incentivize individuals to invest in cost-effective mitigation measures. Examples of cost-effective individual damage-reducing investments include installing dry flood proofing measures which keep water out of the building during a flood (e.g. flood shields) or wet flood proofing measures that minimize damage when water enters a building (e.g. by applying water-resistant building materials). A recent review showed that flood risk management strategies will be much more cost-effective when including individual-level damage reducing measures in addition to structural measures from traditional flood risk management, especially under an increased frequency and severity of floods as a result of climate change (Kreibich et al., 2015).

Investments in individual damage-reducing measures can be considered a public good. For instance, in countries where the government can provide compensation for flood damage, such as the Netherlands, individual investments in reducing flood damage saves tax money for compensated victims after flood events. In a previous survey in Chapter 3 with Dutch homeowners, we elicited social norms with regard to individual flood damage-reducing measures (Mol et al., 2020b). We found that 25% of Dutch homeowners think that their peers would approve if they invested in damage-reducing measures, 50% are indifferent, and the remaining 25% think that their peers would disapprove. Therefore, we believe that investments in flood preparedness are subject to social norms and provide an opportune case for testing social norm-nudges. We focus on descriptive social norm-nudges in this chapter, because our previous elicitation of injunctive social norms showed that only 10% of Dutch homeowners indicate that their peers think that they should invest in damage-reducing measures. In our experiment, the norm nudges refer to the flood protection investment behavior of participants in the flood risk game in Chapter 3 (Mol et al., 2020b), and not to the behavior of peers in real life.

Previous research indicates that flood preparedness behavior is driven by the risk-reduction behaviors of others (Poussin et al., 2014; Grothmann and Reusswig, 2006). For example, a survey of households in Australia found that perceived social norms had a greater influence on flood insurance purchases than homeowners' perceptions of flood risk (Lo, 2013). In a separate survey, Bubeck et al. (2013) showed a positive relationship between mitigation behavior and having neighbors and friends who implemented flood mitigation measures. However, these studies have not examined the effectiveness of different social norm messages in stimulating individual investments in flood damage mitigation measures.

In this study, we examine the efficacy of several different messages to stimulate flood preparedness measures in a controlled experimental study. As an additional innovation, we compare the impact of social norms on preventive behavior across two countries characterized by different flood risk management regimes. In addition, the populations of these countries differ in the average scores of individualism-collectivism (Pineda et al., 2015), a measure that indicates to what extent people conceptualize themselves in relation to others (Triandis, 1989). Both characteristics may influence the effectiveness of social norm-nudges. We hence assess whether differences in current flood risk management between those countries with the Netherlands more focused on public flood protection through dikes, and Spain on individual protection measures influence risk attitudes and personal norms for protecting ones home.

Ideally, social norm-nudges are examined in a large-scale field experiment, such as the classical examples on energy conservation and water conservation (Allcott, 2011; Price, 2014). Such a large scale field experiment was practically infeasible for the case of flood preparedness, because (1) making substantial investments to make a home flood-proof is a more costly behavior than habitual behavior like energy saving or recycling and (2) there is no obvious field partner, such as a utility company, to measure and stimulate flood preparedness. An online lab experiment is a feasible and less costly alternative which can give directions for future field experiments, for example by identifying the most promising interventions to be tested in the field. Even though the results of lab experiments often correlate well with self-reported behavior in the field (Dohmen et al., 2011; Dai et al., 2018), the results should be interpreted with caution (Levitt and List, 2007; Lades et al., 2020).

7.2 Literature review

A growing body of scientific research has identified important aspects of norm-nudge designs (see e.g. Bicchieri and Dimant, 2019). One line of research suggests that norm-nudges are only effective if the targeted behavior is interdependent; that is, when individual preferences are conditional on the empirical expectations of the behavior of others (Bicchieri, 2016). In contrast, when individuals are primarily motivated by their own basic needs or by their beliefs about what is morally right (i.e. targeted behavior is independent), individuals may expect others to behave in one way while behaving in a different way themselves (Bicchieri, 2010). Note that expectations may be normative (what other people would approve of), empirical (what other people do), or both normative and empirical (Bicchieri, 2016). In this chapter, I use the terminology of Bicchieri et al. (2014). I focus on interdependent behavior under empirical expectations, or *descriptive norms* - a preference to do X following the expectation that others do X as well (see Bicchieri and Dimant, 2019). Note that the term descriptive norm is used slightly differently in the psychological literature, namely as the perception of what is common behavior

(Bicchieri and Dimant, 2019). Alternatively, norm-nudges may be based on injunctive norms, or expectations of what others find appropriate behavior (Cialdini et al., 1990), such as *Most people think you should not litter* (Farrow et al., 2017).

Another important component of norm-nudge design is choosing the appropriate reference network. According to social identity theory, individuals are much more strongly affected by the actions of others if they share a certain group membership, such as gender, neighborhood or ethnicity (Tajfel, 1982). For instance, Goldstein et al. (2008) found that referring to a specific reference network in the norm-nudge message "other hotel guests who stayed in the same room" more effectively promoted towel reuse than a generic message about other hotel guests. Some research suggests that the credibility of the message or message source may alter the effectiveness of norm-nudges. For example, Gifford et al. (2018) claimed that mistrust in messages from government officials could prevent citizens from taking action to combat climate change. However, recent experimental evidence on feedback frames to increase pro-environmental behavior did not demonstrate any evidence of a messenger effect (Hafner et al., 2017). Note that citizens who believe climate change is real may also mistrust government messages if they think the problem is underestimated. Conversely, individuals may feel threatened in their freedom of choice by the nudge, which may prompt them to act in opposition to the desired behavior. For example, Arad and Rubinstein (2018) provided respondents with a nudge to increase savings, which increased the number of respondents selecting the savings arrangement. However, when respondents were told the government used a nudge, some respondents opted out of the savings arrangement. A strategy for overcoming this effect is to be transparent about the aim of the nudge, for example by informing respondents that the default option may encourage higher contributions to charity. Recent evidence shows that transparent nudges are judged as more trustworthy (Osman et al., 2018) and might be equally effective as traditional nudges that conceal their aim (Bruns et al., 2018).

Finally, the exact framing of empirical norm-nudge messages may improve their efficacy. For example, Stoffel et al. (2019) studied the effect of different quantifiers (*'a large number'* and *'nearly half'*) on intentions to participate in cancer screening. They found that both verbal quantifiers increased intentions compared with an exact numerical norms message (*43%*). While most norm-nudge messages are binary (e.g., people pay or do not pay their tax on time, Hallsworth et al., 2017), some contexts allow for a continuous approach (e.g. *'Neighbors used 1,092 kWh on average,'* Allcott, 2011). However, many cases of norm-nudges use a binary message, even when the exact distribution of this variable is known. Contributing to the literature on the transparency of nudges (Bruns et al., 2018), this study tests whether transparently showing the full distribution of choices by previous respondents (i.e. providing the exact percentage of those who chose each option) is more effective than summarizing these choices as a binary message. Furthermore, we measure several individual characteristics that previous research has identified as influencing the effectiveness of social norms messages, such as identification with the reference network (Liu et al., 2019), political identities (Chang et al., 2019), and a concern for social comparison (Buunk and Gibbons, 2006; Garcia et al., 2013). An additional possible moderator of norm-nudge messages is the extent to which people perceive their relationships to others, which can be measured on a scale ranging from individualist (people conceptualize themselves as individuals) to collectivist (people conceptualize themselves as members of a group) (Triandis, 1989). With regard to social norms, collectivists may be more motivated to follow the behavior of others implying that people who demonstrate a more collectivist worldview are more inclined to respond to norm-nudge messages (Baldwin and Mussweiler, 2018; Oh, 2013).

7.3 Methodology

We used an experimental study to examine the impact of different norm-nudge messages on individual flood preparedness in two European countries. Following Hafner et al. (2019), who argued that the effect of norm-nudge messages on behavioral intentions in real life may only apply to respondents who are in the position to execute the intention, we restricted the sample to homeowners. The design included an incentivized investment game in which respondents were asked to make decisions about investing in cost-effective measures to prevent damage of low-probability floods. To mimic the large consequences of real flood investment decisions, we implemented a random lottery incentive mechanism with high monetary stakes (see Camerer and Hogarth, 1999). Specifically, at the end of the experiment, the software randomly selected one respondent who had the chance to earn up to 650 euro, based on his/her decisions and luck in the game. The payment mechanism was explained at the start of the game, to motivate subjects to consider their decisions carefully under high stakes.

Investment game

The investment game was a simplified version of a the online experiment from Chapter 3 (Mol et al., 2020b). We used identical parameters to facilitate comparison of the results. In this game, respondents were asked to imagine owning a house with a flood risk for 25 years. With the hypothetical house comes a savings balance that could be used to make payments in the game, such as purchasing flood damage reduction measures. The currency used in the investment game was ECU (Experimental Currency Units). The game started with the introduction of the parameters: a yearly flood probability of 1%, the maximum damage of 50,000 ECU in case of a flood and the savings balance of 65,000 ECU. The next page offered five discrete investments with accompanying benefits in terms of reduced damage from flooding. The investment decision was made once, at the start of the game, and damage-reducing measures were effective for the full 25-year period of the game. This one-shot set-up of the game was designed to be suitable for the online sample of respondents, accounting for the recommended time span for online surveys.

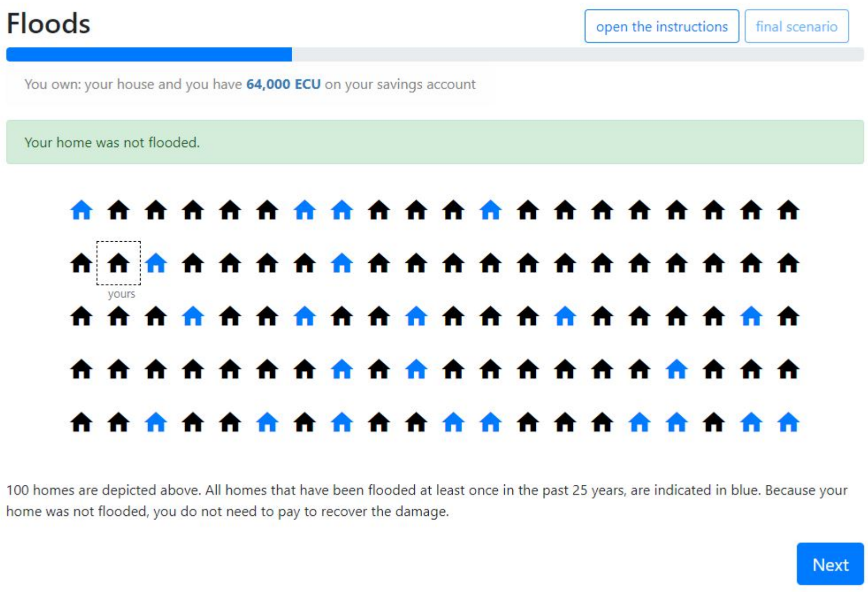
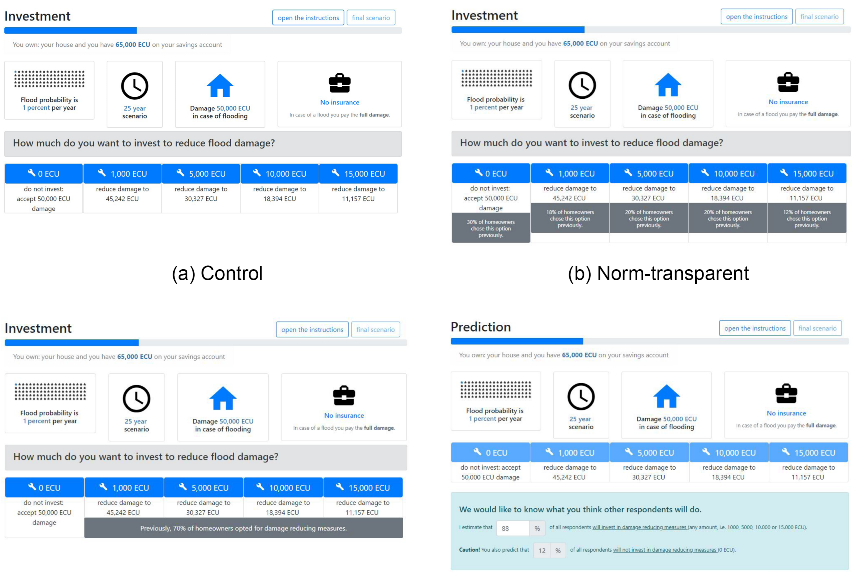


Figure 7.1: Screen shot of flood risk page.

Figure 7.1 provides a screen shot of the page in the investment game where the flood risk was realized. This page showed a grid with 100 houses, with the house of the respondent enclosed in a square. The software randomly selected (based on the 1% flood probability) a number of houses that were flooded in the 25 years of the game and highlighted these in blue. In case the house of the respondent was flooded, the 50,000 ECU damage was subtracted from the savings balance. The optimal investment based on expected value calculations was 1000 ECU for a risk neutral ($\theta = 1$) subject with low time discounting rate ($\delta = 0.01$).

Experimental treatments

Each respondent was randomly assigned by the software into the control group or one of three treatment groups. Based on an a priori sample size analysis with a significance level of 0.05 and power of 80%, we decided to sample 250-300 participants in each treatment group. A sample size of 252 participants



(a) Control (b) Norm-transparent

(c) Norm-high (d) Norm-focusing (belief elicitation)

Figure 7.2: Screen shots of the four treatments.

per group could detect an effect size ($d = 0.227$) equal to the impact of having insurance coverage in a previous experiment with the flood risk investment game in Chapter 3 (Mol et al., 2020b). Our budget restricted the number of treatment groups to six. We decided to run all three treatments and the control group in one country ($n = 4 \times 300 = 1200$, the Netherlands), and the most promising treatment plus the control group in the other country ($n = 2 \times 300 = 600$, Spain). This approach was preregistered ((https://aspredicted.org/q37kj.pdf)). In two treatment groups we displayed an empirical norm-nudge message with information about decisions of previous respondents. We used the percentages of previous investments in the Control treatment[1] of Chapter 3 (Mol et al., 2020b) to construct these messages. Note that these percentages were based on decisions of a sample of Dutch homeowners. To prevent confounding effects of in-group/out-group preferences, we did not mention the nationality of the reference group, but simply characterized them as homeowners. A third treatment group faced a focusing norm treatment, by eliciting beliefs about others' investment choices before participating in the investment game (Krupka and Weber, 2009). We did not include an

[1] This treatment was called 'Mandatory No Insurance' and had exactly the same parameters as the Control treatment in the current experiment.

injunctive norm message because results from our previous experiment showed that 90% of Dutch homeowners do not think their peers should invest in flood risk reduction. A message highlighting this information has the potential to backfire such that people are less motivated to invest, particularly if they would otherwise expect that a larger proportion of their peers think they should invest, leading to downward-adjustment of beliefs (c.f. Bicchieri and Dimant, 2019). The game was constructed such that individual decisions could not be observed by other respondents, to focus on the effect of the norm-nudge messages in isolation from observability effects. In practice, many flood preparedness measures are taken inside a house and cannot be observed by neighbors either, except for the most extreme case of elevating a house. Figure 7.2a shows the investment screen in the Control treatment.

Norm-transparent

This treatment showed the full distribution of previous flood preparedness decisions in terms of the percentage of investments of previous respondents in five small text boxes below the five investment options (see Figure 7.2b). We expected that respondents would be unfamiliar with the flood preparedness decision environment. In other words, few respondents are confronted with flood preparedness decisions on a daily basis[2], in contrast to other contexts which have been successfully related to social norms, such as energy conservation. Therefore, we expected respondents to have no (strong) beliefs about others' behavior in the investment game. The Norm-transparent treatment provided new information on mitigation decisions by others, illustrating the informational effect of a descriptive norm (Krupka and Weber, 2009). All boxes showed flood preparedness decisions by others in percentages. We presented this information as a percentage in the empirical norm-nudge message following Hallsworth et al. (2017), who found in a large natural field experiment that percentage norm messages are more effective than norm messages presented as a fraction (*nine out of ten*) or a general statement (*the majority*) in increasing tax compliance. Note that the online study by Stoffel et al. (2019) found opposite effects: verbal statements (*a large number*) were most effective in increasing cancer screening intentions. We believed that the findings of Hallsworth et al. (2017) are more relevant for our flood preparedness context, as they also focus on financial behavior, rather than intentions in the health domain.

Norm-high

In the Norm-high treatment, an empirical norm-nudge message was displayed directly below the positive investment options (see Figure 7.2c). The message

[2] To illustrate, only 9% of our respondents indicated having purchased private flood insurance coverage and even those respondents probably think about flood insurance only when they pay their premium or renew coverage.

emphasized that a large majority of previous respondents had invested in damage-reducing measures, again by showing a percentage (*"70% of respondents in previous research opted for damage-reducing measures."*). While previous social norms research has mostly focused on binary outcome variables (e.g. whether or not to donate to charity or play a risky lottery), our set-up requires respondents to choose from five discrete investment options. The Norm-high treatment highlights the binary first step of the decision (invest versus not invest) and is an intuitive way to describe the distribution.

Norm-focusing

The final treatment was designed to manipulate the strength of the norm focus (Cialdini et al., 1990; Kallgren et al., 2000). Krupka and Weber (2009) showed in a large lab experiment that the norm focus intervention is effective in increasing pro-social behavior even when respondents believe that others do not behave according to the norm. Recently, Bolton et al. (2020) showed that the mere thought of what other people might do, operationalized with incentivized belief elicitation, leads to the same desired increase of donations compared to a more costly intervention (i.e. with monetary consequences). In line with Krupka and Weber (2009) and Bolton et al. (2020), the dependent variable in our experiment cannot be confounded by strategic concerns. We used an incentive-compatible method to elicit beliefs about others' behavior, before requesting that respondents make a decision about their personal investment in the investment game. We asked respondents to estimate the percentage of other respondents investing in damage-reducing measures (1000, 5000, 10.000 or 15.000 ECU). An interactive screen emphasized that the remainder of respondents would not invest (see Figure 7.2d). We opted for such an explanation to facilitate comparison of answers with the Norm-high treatment, which also showed the percentage of people investing versus not investing. Furthermore, eliciting the full distribution would be a rather complicated task to explain, which could lead to undesirable attrition effects. Belief elicitation was incentivized with an €20 payment on top of the respondent fee for one randomly selected respondent. A large yellow alert marked the transition from the own investment decision to the belief elicitation decision screen. In the control treatment, belief elicitation was conducted after respondents completed the investment game. Figure 7.3 provides an overview of the experimental treatments.

The experiment started with a short set of socio-demographic questions. Subsequently, the investment game was introduced through several pages of instructions supported with graphics to facilitate the understanding of respondents with different education levels. As in Chapter 3 (Mol et al., 2020b), the investment game was preceded by a test scenario to familiarize respondents with the decision screens. Before finishing the test scenario stage, respondents were requested to answer a few comprehension questions. The instructions

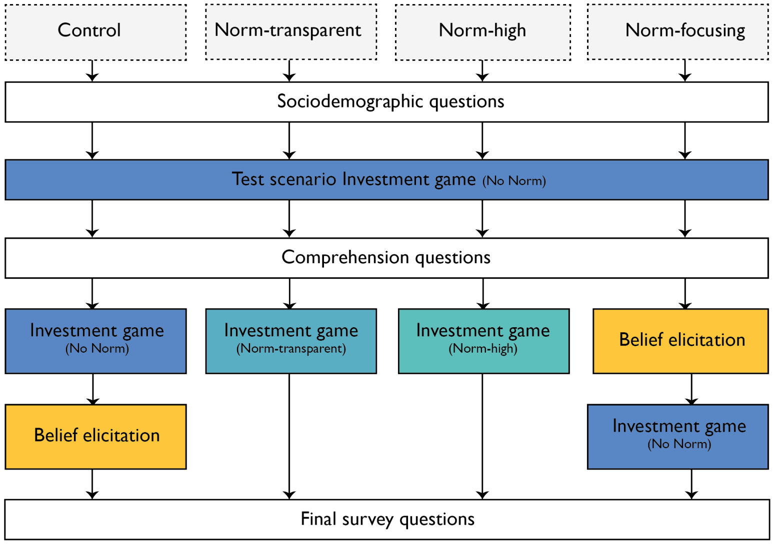


Figure 7.3: Overview of experiment per treatment.

were always accessible to respondents throughout the game. Additionally, the experimental software tracked attempts to answer these comprehension questions and the reopening of instructions. These were included in all regression analyses to control for understanding of the experimental design. All respondents were paid a fixed participation fee of approximately €1, and one participant was randomly selected for a large payment corresponding to the bank balance at the end of the experiment at a conversion rate of 100 ECU = €1, which could range from €100 to €650. Participants were informed at the beginning of the experiment that this random selection would take place after all responses had been recorded. The total number of participants was not communicated upfront. All participants agreed to the informed consent page that explained the payment mechanism and the data storage policy. An email address was provided for anyone that would request further information, but no emails were received. We further paid one participant €20 out of the 57 participants who correctly estimated the share of participants who invested in the game (75%). The average duration of the experiment was 28 minutes and the median duration was 12 minutes. The duration distribution is rather skewed with some extreme values regarding survey length, because the software did not prevent breaks, which allowed subjects to start the experiment and continue later.

The online experiment was distributed by the survey company Panelinzicht in two rounds: in August 2019 to a sample of Dutch homeowners and in

October 2019 to a sample of Spanish homeowners. The panel is representative of the population of each country with respect to education, income and gender. The experiment was translated into the local language of the respondents (Dutch and Spanish) and administered over the Internet using the experimental software oTree (Chen et al., 2016), which allowed for a similar procedure across countries. The experiment started with a selection question to ensure that only homeowners were eligible to participate. The experimental interface was optimized for the screen size of tablets and desktop computers, which was communicated in the invitation email of the panel company. Nevertheless, it was possible to complete the experiment on smaller screens such as smartphones, but this required more effort through zooming and scrolling. The final data set contains 1200 Dutch responses and 605 Spanish responses.

Variables that were part of our hypotheses included trust in the presented information, susceptibility to peer influence and individualism-collectivism personality scores. These variables were evaluated with survey questions at the end of the investment game. Moreover, we elicited variables that are likely to influence demand for flood damage mitigation investments independent of social norms, such as personal norms, response efficacy of mitigation measures and risk perception. Table 7.1 provides an overview of all survey questions administered and the order in which they appeared.

Table 7.1: Summary overview of the survey questions

Demographics	
Gender (f32)	Dummy variable gender (1 = respondent is female)
Age in years (f33)	Continuous variable, age in years
Master's degree (f35)	Dummy variable education level (1 = holds Master's degree)
High income (f36)	Dummy variable income (1 = monthly household after-tax income > €5,000)
Expensive house (f37)	Dummy variable house value (1 = house value > €400,000)
Hypotheses variables	
Trust in messenger (f12)	Categorical variable (range 1-5), following Hafner et al. (2017), only displayed in Norm-high and Norm-transparent treatments
Independence (f27)	Reverse of susceptibility to peer influence. Scale of three categorical variables (range 1-5), following Eckel et al. (2011).
Self-responsibility (f25)	Categorical variable (range 1-5), following Maidl and Buchecker (2015)
Collectivism (f30)	Short 11-item scale (range 1-7) (Cai and Fink, 2002), revision of INDCOL scale (Hui and Triandis, 1986). Scores averaged: higher numbers indicate more collectivism.
Nationality (from wave)	Dummy nationality (0 = the Netherlands, 1 = Spain)
Control variables	
Awareness (s1)	Dummy sure live in flood-prone area (1 = Yes), adapted from Botzen et al. (2015)
Evacuated (s2)	Dummy ever evacuated due to threat of flooding (1 = Yes)

Damaged (s3)	Dummy property damaged due to floods in the past (1 = Yes)
High damaged amount (s4)	Dummy variable damaged amount (1 = amount > €50,000)
Flood probability (s5)	Categorical variable (*Zero, Very low, Low, Not low/not high, High, Very high, Do not know*), adapted from Mol et al. (2020b)
Expected water level (s6)	Expected water level in case of a flood. Categorical variable (*0 cm, 1-10 cm, 11-50 cm, 50-100 cm, 1-2 meters, > 2 meter*), adapted from Mol et al. (2020b))
High expected damage (s7)	Dummy high expected damage (1 = respondent expects flood damage > €50,000)
Worry about floods (s8)	Categorical variable (range 1-5), adapted from Botzen et al. (2015)
Threshold of concern (s9)	Categorical variable (range 1-5), adapted from Botzen et al. (2015)
Trust in dikes (s10)	Categorical variable (range 1-5), adapted from Mol et al. (2020b)
Flood probability (s11)	Continuous variable, log of estimate, adapted from Mol et al. (2020b)
Anticipated regret (f13-15)	Categorical variable (range 1-5), adapted from Kunreuther and Pauly (2018)
Difficult (f16)	Categorical variable (range 1-5) on difficulty of investment game
Strategy (f17)	Open answer to assess strategy in investment game
Measures implemented (f18)	Continuous variable, number of measures, adapted from Mol et al. (2020b)
Measures neighbors (f19)	Dummy respondent knows person who has installed measures (1 = Yes), adapted from Mol et al. (2020b)
Measures self (f20)	Categorical variable, Person responsible for installing measures in question f3 (*Me, Previous owner, Homeowners association, Other*)
Neighbors relation (f21)	Categorical variable, Relationship to person in f19 (*Partner, Friend, Parent, Aunt/Uncle, Son/Daughter, Cousin, Neighbor, Acquaintance, Other*), adapted from Mol et al. (2020b)
Response efficacy (f22)	Categorical variable (range 1-5), adapted from Poussin et al. (2014)
Response cost (f23)	Categorical variable (range 1-5), adapted from Poussin et al. (2014)
Self-efficacy (f24)	Categorical variable (range 1-5), adapted from Poussin et al. (2014)
Personal norm (f26)	Categorical variable (range 1-5), adapted from Doran and Larsen (2016)
Risk aversion (f28)	Categorical variable (range 0-10), adapted from Falk et al. (2018)
Present bias (f29)	Categorical variable (range 0-10), adapted from Falk et al. (2018)
Numeracy (f31)	Short numeracy scale by McNaughton et al. (2015)
House type (s34)	Dummy house includes ground floor (1 = Yes)
House size (f38)	Continuous variable, size of ground floor in m^2, for calculating objective risk

Notes: Order of variable in parentheses: 's' indicates start survey, 'f' indicates final survey. For example, 's7' indicates it was the seventh question and appeared in the start survey.

7.4 Hypotheses

We formulated hypotheses based on results from previous literature. All hypotheses were formally preregistered prior to data collection.[3]

Our main hypothesis concerns the effect of empirical norm-nudge messages on investments in damage-reducing measures in the investment game. Norm-nudges may help individual homeowners to act the way others are acting, as humans are inclined to model behavior on what others do. A large body of literature has shown that norm-nudges may be effective to increase environmental-friendly behavior (Allcott, 2011; Abrahamse and Steg, 2013). Furthermore, survey research indicates that flood preparedness behavior is driven by the risk-reduction behavior of others (Poussin et al., 2014; Grothmann and Reusswig, 2006) and perceived social norms[4] (Lo, 2013). When the information presented as a social norm-nudge differs from of respondents' a priori expectation as to what others will do, this may lead them to correct their beliefs. We expect that most respondents perceive few others will invest in flood damage mitigation measures and predict that an empirical norm-nudge message will lead to an increase of investments in flood risk protection measures compared to not including a norm-nudge message.

Hypothesis 7.1 *Respondents confronted with an empirical norm-nudge (Norm-high and Norm-transparent) will invest more in damage-reducing measures than respondents in the Control treatment.*

To our knowledge, we are the first to test an empirical norm-nudge showing the percentages of previous decisions for each of the five discrete investment options (Norm-transparent), as compared with an empirical norm-nudge highlighting the percentage of previous respondents who either have or have not invested (Norm-high). Hence, we have no empirical information to hypothesize whether investment in damage-reducing measures will differ between these two norms. We expect that respondents in the Norm-transparent treatment will have greater trust in the norm-nudge message than respondents in the Norm-high treatment, due to the provision of more transparent information in the former condition (see e.g. Osman et al., 2018).

Hypothesis 7.2a *Respondents in the Norm-transparent treatment have greater trust in the norm-nudge message than respondents in the Norm-high treatment.*

Accordingly, we expect that respondents in the Norm-transparent treatment will be more likely to follow the majority of highlighted responses

[3] (https://aspredicted.org/q37kj.pdf). Preregistered hypothesis 6 was unrelated to the treatments and is suppressed here for brevity.

[4] Measured as the level in which the respondents believed that their family or friends want them to purchase flood insurance.

(investing some amount), relative to those in the Norm-high treatment, due to greater levels of trust in the message.

Hypothesis 7.2b *Respondents in the Norm-transparent treatment are more likely to follow the majority of the highlighted responses and invest than respondents in the Norm-high treatment.*

In the Norm-focusing treatment, respondents are confronted with a belief elicitation page before they are asked to make an investment decision for their own payoff. The question, "What proportion of other respondents would invest in damage-reducing measures?" may encourage respondents to think about the norm (Kallgren et al., 2000), which could increase investments even when respondents do not believe many others will invest (Bolton et al., 2020). Based on the large effect-sizes in previous research on norm-focusing and *norm beliefs* (Krupka and Weber, 2009; Bartke et al., 2017), we expect that the Norm-focusing treatment leads to the highest investments of all treatments.

Hypothesis 7.3 *Respondents in the Norm-focusing treatment will invest the most in damage-reducing measures, relative to all other treatments.*

Susceptibility to peer influence (Dielman et al., 1987; Bearden et al., 1989) may be another important driver of norm-nudge effects. Susceptibility to peer pressure is commonly studied in adolescents and young adults, where it has been found to drive gambling (Langhinrichsen-Rohling et al., 2004) and delinquent (Prinstein et al., 2011) behavior. Eckel et al. (2011) found that high school students who report being highly independent are less likely to conform to normative behavior in decisions about sharing funds. Recent empirical evidence among adults shows that individuals are more likely to follow provided peer information if they are susceptible to informational influence, across several domains, including investment decisions (Hoffmann and Broekhuizen, 2009), consumer choice (Orth and Kahle, 2008), vaccination behavior (FitzSimons et al., 2014) and retirement decisions (Verhallen et al., 2018). In a recent paper, Stöckli and Hofer (2020) examined susceptibility to social influence among a large sample of adult online social media users. The authors found that susceptible individuals are more likely to buy what other users post about, and to obtain information about political issues following other uses. Thus, we expect to find the same pattern of results.

Hypothesis 7.4 *The effect of the empirical norm-nudge messages is greater for respondents who demonstrate higher levels of susceptibility to peer influence.*

The degree to which empirical norm-nudges affect individuals may further be subject to differences in individualism-collectivism. These differences in self-concept can influence engagement in protective behaviors. For example, Parboteeah et al. (2012) found that collectivists are more likely to support sustainability initiatives. Recent evidence shows that people from individualist cultures respond differently to nudges in the context of vaccination behavior (Betsch et al., 2017), in that they are more willing to be vaccinated. Individuals in southern European countries (such as Spain) generally score higher on collectivism than individuals in the Netherlands (Hofstede, 2001; De Raad et al., 2016). Therefore, we expect relevant variation in this variable within our sample that may explain heterogeneity in responses to the social norms message. To investigate the cultural differences of empirical norm-nudges on flood preparedness, we will examine respondents scores on an 11-item individualism-collectivism scale (Cai and Fink, 2002), a revised version of the original INDCOL scale (Triandis, 1989). We expect that respondents with a more collectivist worldview are more strongly influenced by an empirical norm-nudge message as they are more inclined to follow group behavior.

Hypothesis 7.5 *The effect of an empirical norm-nudge is larger for individuals with high collectivism scores on the individualism-collectivism scale, relative to those with high individualism scores.*

7.5 Results

In this section we report the experimental results, beginning with descriptive statistics for each pre-registered hypothesis. We then turn our attention to a secondary treatment in the Spanish dataset with regard to intention to search for more information about flood risks. Finally, we report a number of observations from an exploratory analysis of the data.

Table 7.2 presents descriptive statistics of our sample. Demographic variables are very similar across countries, except for income[5] and home value,[6] which are in line with the population statistics of each respective country.

Results by hypotheses

Our main hypothesis concerned the effect of empirical norm-nudge messages on investments in damage-reducing measures in the investment game. Figure 7.4 shows the proportions of each investment level chosen by our respondents, split per treatment. The shaded areas indicate positive investments (1,000 ECU; 5,000 ECU; 10,000 ECU or 15,000 ECU) and the remaining white area indicates the proportion of respondents who did not invest anything. We observe almost identical investment levels across treatments and countries.

[5] The average after-tax income is €2368 per month in Spain (Institute Nacional de Estadística, 2020) and €3042 per month in The Netherlands (Netherlands Statistics, 2020b).

[6] The average home value is €151,084 in Spain (Gobierno de España, 2020) and €307,978 in The Netherlands (Netherlands Statistics, 2020a).

Table 7.2: Descriptive statistics by country

	Spain ($n = 605$)	**Netherlands** ($n = 1200$)	**Total** ($n = 1805$)
Gender			
Male	302 (50%)	633 (53%)	935 (52%)
Female	303 (50%)	567 (47%)	870 (48%)
Age (years)			
Mean (SD)	45 (±13)	52 (±17)	49 (±16)
Education level			
Low	99 (16%)	190 (16%)	289 (16%)
Medium	201 (33%)	488 (41%)	689 (38%)
High	305 (50%)	522 (44%)	827 (46%)
Income (per month)			
Mean (SD)	2100 (±1200)	3100 (±1300)	2800 (±1300)
Missing	64 (10.6%)	230 (19.2%)	294 (16.3%)
Home value (× €1,000)			
Mean (SD)	200 (±140)	290 (±130)	260 (±140)
Missing	91 (15.0%)	112 (9.3%)	203 (11.2%)

This result is unexpected, given the experimental research on the effectiveness of norm-nudges in the environmental domain (Abrahamse and Steg, 2013; Farrow et al., 2017; van Valkengoed and Steg, 2019) and the survey research on importance of perceived social norms in the flood risk domain (Bubeck et al., 2013; Lo, 2013).

We examined this result more formally with Mann Whitney Wilcoxon (MWW) tests and probit regressions and discuss the results in detail below. Table 7.3 provides an overview of all hypotheses and reports two-tailed MWW tests to analyze the differences in frequencies of investments with respect to treatments and independent variables of interest. For robustness we also applied chi-squared tests but the results do not change. Figure 7.4 seems to suggest that all investment options were chosen equally often. A two-sided chi-squared test rejected this conjecture ($\chi^2 = 37.09$, $p < 0.001$). In other words, respondents did not select their investment option at random.

Hypothesis 7.4 and 7.5 predicted interaction effects of susceptibility to peer influence and collectivism on the relationship between norm treatments and damage-reducing investments. Therefore, Table 7.3 reports Z-statistics of the interaction term in a probit regression with binary investment as the dependent variable and susceptibility peer influence and collectivism as independent variables.

The first two rows of Table 7.3 show no support for a main treatment effect as predicted by Hypothesis 7.1; investments do not differ between respondents

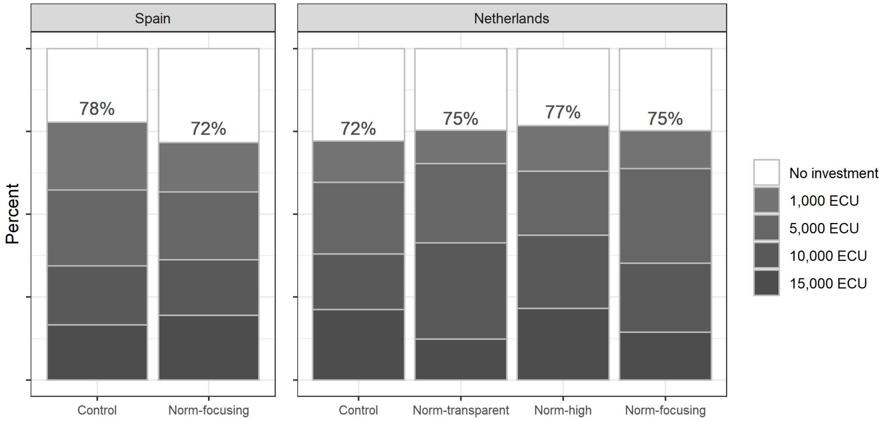


Figure 7.4: Investments in damage-reducing measures by treatment

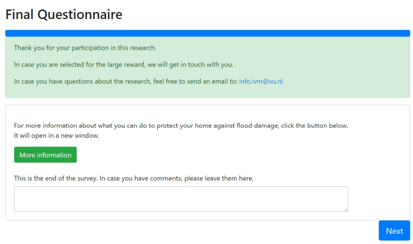
in the Control group and the Norm-high group ($W = 85493.5$, $p = 0.11$), nor between respondents in the Control group and the Norm-transparent group ($W = 87618.5$, $p = 0.53$). Our Norm-transparent treatment did not lead to higher trust in messenger ($W = 47932$, $p = 0.27$) (Hypothesis 7.2a), and investments were identical in the Norm-transparent and the Norm-high treatment ($W = 43589.5$, $p = 0.3$) (Hypothesis 7.2b). With regard to Hypothesis 7.3, we find no differences in investments between the Norm-focusing group and the Control group ($W = 181483$, $p = 0.8$), the Norm-high group ($W = 98444$, $p = 0.06$), and the Norm-transparent group ($W = 93252.5$, $p = 0.39$). The difference in investments between the Control group and the Norm-transparent treatment is significantly stronger ($z = -2.298$, $p = 0.02$) for respondents with high levels of susceptibility to peer influence, which is in line with Hypothesis 7.4. However, this result is not found when comparing the Control group with the Norm-high ($z = -0.662$, $p = 0.51$) and Norm-focusing ($z = 0.707$, $p = 0.48$) treatments. Finally, we find no support for Hypothesis 7.5; the coefficients of the interaction terms between collectivism and the treatment conditions on investments in damage-reducing measures are not significant (Norm-focusing: $p = 0.48$; Norm-transparent: $p = 0.16$; Norm-high: $p = 0.5$).

Table 7.3: Results by hypotheses

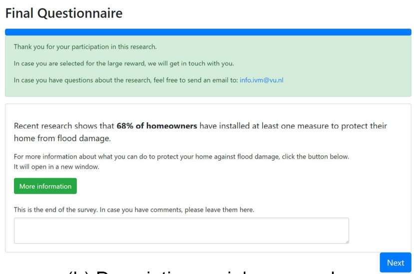
Hyp	Prediction	Variable	Test	Support
H1	Norm-high > Control	Investments	$W = 85493.5$, $p = 0.11$	✗
H1	Norm-transparent > Control	Investments	$W = 87618.5$, $p = 0.53$	✗
H2a	Norm-transparent > Norm-high	Trust	$W = 47932$, $p = 0.27$	✗
H2b	Norm-transparent > Norm-high	Investments	$W = 43589.5$, $p = 0.3$	✗
H3	Norm-focusing > Control	Investments	$W = 181483$, $p = 0.8$	✗
H3	Norm-focusing > Norm-high	Investments	$W = 98444$, $p = 0.06$	✗
H3	Norm-focusing > Norm-transparent	Investments	$W = 93252.5$, $p = 0.39$	✗
H4	(Susceptible: Norm-focusing > Control) > (Not susceptible: Norm-focusing > Control)	Investments	z = -0.654, $p = 0.51$	✗
H4	(Susceptible: Norm-transparent > Control) > (Not susceptible: Norm-transparent > Control)	Investments	z = -2.298, $p = 0.02$	✓
H4	(Susceptible: Norm-high > Control) > (Not susceptible: Norm-high > Control)	Investments	z = -0.662, $p = 0.51$	✗
H5	(Collectivist: Norm-focusing > Control) > (Individualist: Norm-focusing > Control)	Investments	z = -0.707, $p = 0.48$	✗
H5	(Collectivist: Norm-transparent > Control) > (Individualist: Norm-transparent > Control)	Investments	z = -1.39, $p = 0.16$	✗
H5	(Collectivist: Norm-high > Control) > (Individualist: Norm-high > Control)	Investments	z = -0.679, $p = 0.5$	✗

Notes: We report the W-statistic of MWW-tests for main effects and z-scores of probit regressions for hypotheses predicting an interaction. Support indicated for $p < 0.05$. H2a Trust indicates Trust in messenger.

The fact that our statistical tests do not support any differences between experimental treatments should not automatically lead to accepting the null. To examine the possibility of a null result in more detail, we conducted a power-analysis. We find that we achieve power in excess of 85% (H2a, Norm-transparent versus Norm-high) and up to 93% (H1, Norm-transparent versus Control) to find an effect size of 0.227, which was the effect of insurance policy in Chapter 3 (Mol et al., 2020b). The effect sizes in the current experiment range from 0.01 to 0.12 for pairwise comparisons, and a Kruskall-Wallis test confirms that there are no significant differences in average investments across all treatments ($p = 0.26$). These effect sizes are so small that they are not meaningful: the smallest effect size of interest based on a previous experiment with the same investment game was 0.227, which is substantially larger than what we find here.



(a) Control



(b) Descriptive social norm-nudge

Figure 7.5: Screen shots of secondary treatment.

Secondary treatment

As soon as the data collection for the Dutch respondents was completed, we conducted a preliminary analysis to determine the most promising treatment condition for the Spanish respondents, as indicated in the preregistration. We hypothesized that the non-significant effects of Norm treatments on investment decisions, as described above, might be attributed to the cost of this investment decision. In other words, changing intentions following the Norm treatments is a first step that many are willing to make, while the next step of changing behavior is more difficult to achieve, especially if it requires a costly investment (e.g., time, money, effort). In our application, the behavioral change of interest is costly by definition: to make or increase a financial investment in flood preparedness measures. In line with previous research (Dur et al., 2019), we speculated that norm-nudge messages are more apt to influence behavioral outcomes for which there is no monetary cost, such as clicking a link to retrieve more information, but less effective at producing changes in behaviors for which there is a tangible cost, such as investing money or increasing savings. To test this alternative explanation, we constructed a button to open a page[7] with more information about flood risk and mitigation possibilities in Spain. We randomly distributed a descriptive social norm-nudge message based on results from a previous survey in Chapter 3 (Mol et al., 2020b) to half of the respondents: *Recent research shows that* ***68% of homeowners*** *have installed at least one measure to protect their home from flood damage* (see Figure 7.5). We expected more clicks on the link for information from respondents who received the norm-nudge message, compared to respondents in the control group. Results demonstrate that only a very small proportion of our sample clicked the link (63 respondents), and we find no differences in information search between these groups ($\chi^2 = 0$, $p = 0.989$).

Beliefs

Next, we investigate the results of our belief elicitation question. On average, 75% of our respondents (Spanish and Dutch respondents combined) invested at least 1,000 ECU (the minimum amount). The correct answer to the belief elicitation question, which asked respondents to indicate the percentage of other respondents investing a positive amount, was thus 75%. Figure 7.6 shows the distribution of beliefs about other respondents' investment behavior in our sample, ranging from 0 to 100. The average belief was 46% and the median belief was 50%. A majority of respondents underestimated the correct answer, both in the Control treatment (72%) and in the Norm-focusing treatment (70%). Furthermore, belief responses indicate a preference for round numbers, such as 10, 50 and 80.

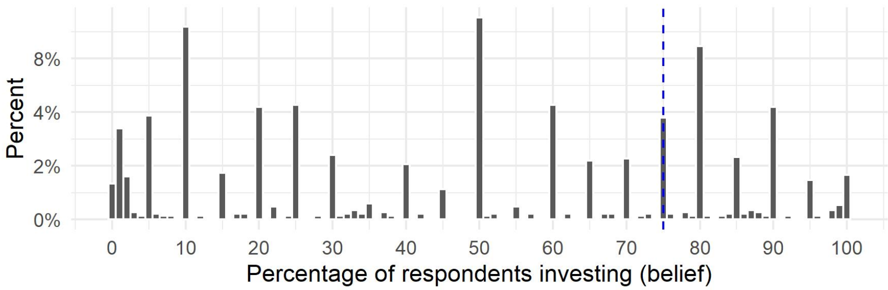


Figure 7.6: Histogram of beliefs. *Note:* The blue dotted line indicates the correct answer.

[7] The following link leads to this document, which is on the website of the Spanish government: https://www.miteco.gob.es/es/agua/formacion/guia-reduccion-vulnerabilidad-edificios_tcm30-379148.pdf

If we assume that beliefs in the Norm-transparent and Norm-focusing treatments are generally equal to the beliefs elicited in the Control treatment, this implies that the information presented in those treatments (namely, that in previous research 70% of homeowners invested a positive amount) would correct beliefs upward for approximately 72% of respondents. This illustrates there was a gap between information and beliefs, which would allow for an upward correction of these beliefs. Yet, an upward adjustment of beliefs did in this case not result in an adjustment of behaviour. The absence of a treatment effect can hence not be explained by the absence of a gap between beliefs and the descriptive norm. However, the opposite effect may be possible: that the information-belief gap is so large that it actually signals that respondents are not aware of any norm with regards to flood-preparedness measures, resulting in overall low norm-sensitivity. This would be in line with the argumentation in Bicchieri (Bicchieri, 2006).

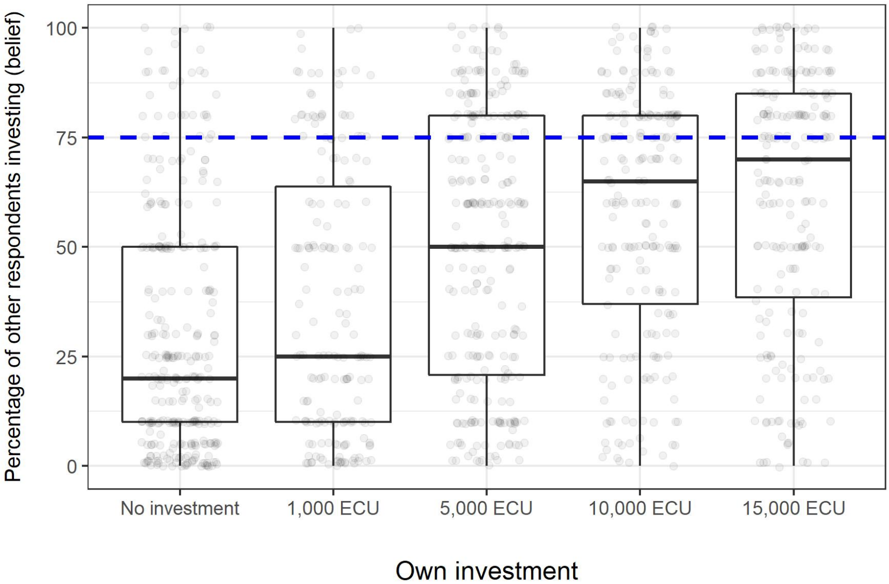


Figure 7.7: Beliefs of other respondents investing by own investment. *Notes:* The blue dotted line indicates the percentage of respondents who invested in our sample. Each individual observation is indicated with a gray dot, to which a small arbitrary noise has been added to the x coordinate to facilitate readability. Boxplot whiskers indicate the inter-quartile range, middle lines represent medians.

The belief elicitation question was asked in two of our four treatments: Control and Norm-focusing. The only difference between these treatments was whether the belief elicitation question was asked before (Norm-focusing) or after (Control) respondents' own investment decisions were taken. We found no difference in belief distributions across treatments (Kolmogorov-Smirnov test, $p = 0.982$). The lack of a treatment effect with regard to elicited beliefs suggests that beliefs about others' behavior and investment decisions are made concurrently - it does not matter which question is posed first. Figure 7.7 shows the relationship between beliefs and one's own investment decisions. The figure demonstrates a positive relationship between investments and elicited beliefs, implying that investments increase with the belief that more people are investing.

Other correlates of investments

Prior to our next set of analyses, we explored which variables have the most predictive power when it comes to investments in damage-reducing measures. We estimated simple binary correlations between all hypothesized predictor variables, control variables and the dependent variable. Figure 7.8 shows the distribution of the ten variables with the strongest correlations with decisions in the investment game in order of correlation strength.

Type of respondents

An alternative explanation for the absence of norm-nudge message effects is that such messages are not effective for respondents who have already decided they want to invest. In contrast, those respondents with no clear preferences with regard to investing, for example those lacking a strong personal norm to invest, or those without positive experiences with measures already installed at home, could be more sensitive to information about other respondents' behavior. As Sunstein (2017) has noted, though nudges may appear to be ineffective at the aggregate level, they may demonstrate effects in distinct sub-populations.

To test this alternative explanation, we constructed two dummy variables to indicate a type of respondent who may be more susceptible to the treatments based on the most important predictors in Figure 7.8. The strongest predictor of investment in damage reduction is the number of measures already installed at home (see question $f18$ in Table 7.1). Therefore, we constructed a dummy of 'No-measures individuals' (1 = zero measures installed at home, 0 = at least one measure installed at home). Other important predictors from Figure 7.8 included personal norm, present bias, response efficacy and expected water level[8] from Figure 7.8. We constructed a 'No-investors' dummy ($n = 369$)

[8] We conducted an additional regression analysis (not reported here) on the decision to mitigate (probit and ordered probit) with the top 5 predictors. We find that expected water level, response efficacy and personal norms are robust and significant predictors in either specification.

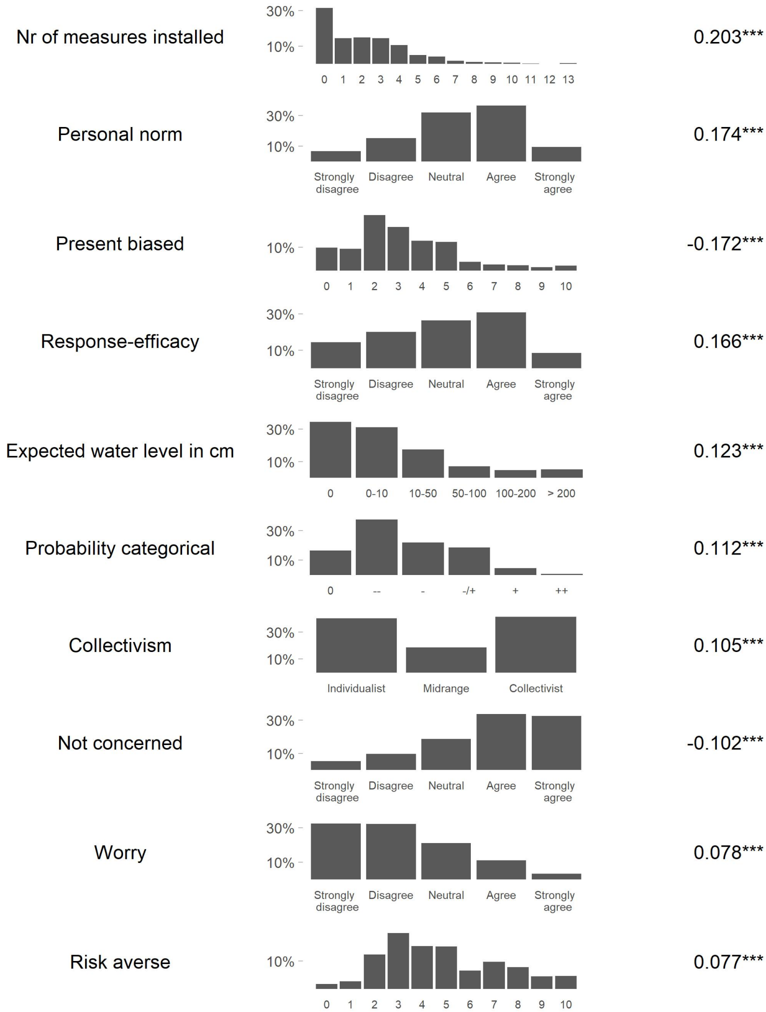


Figure 7.8: Histograms of flood belief variables and correlations with investment decisions. *Note:* Stars indicate significant Spearman correlations (*** $p < 0.001$).

to indicate individuals who do not expect high water levels (expected water level in cm = 0), have low response efficacy (strongly disagree or disagree) and do not have a strong personal norm (strongly disagree or disagree).[9] After

[9] Although present biased is the third most important predictor, constructing a dummy based on present bias values would require an arbitrary split, which is why we did not use present bias in the construction of the 'No-investors' dummy.

constructing the two dummies to indicate the 'No-measures-individuals' and the 'Non-investors' type of respondents, we conducted probit regressions to assess whether the norm-nudge treatments worked differently for these sub-samples. The dependent variable in these regressions was binary investment in protection (in the flood risk investment game) and the treatment dummies were included as explanatory variables. The model was estimated separately for each of the different sub-samples ('No-measures-individuals', 'Measures-individuals', 'Non-investors' and 'Investors'). We expected that the treatment is more effective for the non-investors or those respondents who have not yet installed any measures at home than for the other samples. We expected that the 'Non-investors' and the 'No-measures-individuals' would not be intrinsically motivated, based on the observation that these people do not have a strong personal norm or have not installed any measures at home. Therefore we expected a larger effect of the treatments for those sub-samples, as the treatments are external (they provide information).

Table 7.4: Probit regressions of treatment by type of respondents

	Dependent variable: investment in protection			
	Investors	Non-investors	Measures	No measures
	(1)	(2)	(3)	(4)
Constant	0.954^{***}	0.061	0.847^{***}	0.451^{**}
	(0.115)	(0.207)	(0.114)	(0.198)
Treatment (ref = Control)				
Norm-transparent	0.088	−0.125	−0.135	0.201
	(0.147)	(0.200)	(0.149)	(0.199)
Norm-high	0.182	−0.157	−0.0003	0.037
	(0.152)	(0.191)	(0.153)	(0.192)
Norm-focusing	0.039	−0.101	−0.023	−0.103
	(0.123)	(0.180)	(0.123)	(0.178)
Country dummy	Yes	Yes	Yes	Yes
Log likelihood	-442	-243.9	-445.4	-240.7
Pseudo R^2 (McFadden)	0.003	0.008	0.002	0.005
Observations	920	369	927	362

Notes: Robust standard errors in parentheses ($^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}\ p < 0.01$).

Table 7.4 reports the results of the probit regressions of treatment by type of respondents. Model 1 restricts the sample to 'Investors', whereas Model 2 restricts the sample to the opposite set ('Non-investors'). Model 3 restricts the sample to respondents who installed at least one measure at home ('Measures'), while Model 4 restricts the sample to respondents who did not install any measure at home ('No measures'). Across all models, we find no effect of treatment on investment in protection for any of the sub-samples. As a robustness check, we ran a probit regression analysis (not reported here) on the full sample with interaction terms, all of which were non-significant. In sum, we find no support for the alternative explanation that norm-nudge treatments are more effective for a sub-sample of the respondents, such as those lacking a strong personal norm to invest, or those without positive experiences with measures already installed at home.

Personal norms

As a next step in our analyses, we explored the differences between personal norms and social norms. As a complement to social norms, personal norms represent what people believe to be appropriate behavior for themselves (Schwartz, 1977), or what they feel morally obliged to do (Harland et al., 1999). Previous research has shown that personal norms can be powerful determinants of pro-environmental behavior (Bamberg and Möser, 2007; Yazdanpanah and Forouzani, 2015; Farrow et al., 2017). For example, Doran et al. (2019) showed that personal norms (moral concerns) are a stronger predictor of policy support to mitigate climate change than consequence evaluations. Huber et al. (2020) examined five years of longitudinal US household data and found that that personal norms are strongly related to recycling behavior. We measured personal norm as a response on a 5-point scale to the statement *"I am morally obligated to take measures to reduce flood risk to my home"*, adapted from Doran and Larsen (2016). We find that personal norm is significantly correlated with investment decisions (Spearman correlation $\rho = 0.174$, $p < 0.001$), such that stronger personal norms correspond to higher investments. Note that the results on personal norms are correlational, thus providing limited information about causality.

7.6 Discussion

We conducted a high-powered preregistered experiment with homeowners to assess the effectiveness of norm-nudge frames on flood preparedness across countries. We found no evidence of a treatment effect: investments in damage-reducing measures of respondents in the Norm-transparent, Norm-focusing and Norm-high treatment groups did not differ from investments in the Control group. We examined the alternative hypothesis that social norms affect intentions rather than costly behavioral change with a secondary treatment in the Spanish sample, but the results show no difference between the two treatment groups. Furthermore, we analyzed a subset of respondents who were not motivated by individual flood beliefs and personal norms and replicated the null effect of our full sample in this subset.[10] Several recent examples of

[10] A possible explanation for the lack of effect of susceptibility to peer influence on our treatments, is that we sampled adults from 18 to 90 years old, while most research on susceptibility to peer influence has been conducted on adolescents (Prinstein et al., 2011;

studies that do not identify treatment effects of social norm-nudges are in line with our results: in the environmental domain (Mackay et al., 2020; Chabé-Ferret et al., 2019) and the financial domain (Franklin et al., 2019). Generally, it has been noted for various domains, including corruption (Köbis et al., 2019) and obesity (Oliver and Ubel, 2014), that behavioral approaches such as norm-nudges should not be taken as substitutes but rather as supplements to traditional policies.

A recent paper by one of the founding fathers of nudging, outlined the main reasons for ineffective nudges and three possible responses (Sunstein, 2017). We can rule out one of the two main reasons for failing nudges, namely counternudges, which are nudges aiming to promote the opposite behavior from the original nudge, as they were not at stake in our experiment. The second reason would be that decision-makers have strong antecedent preferences, which would be hard to change regardless of the strength of the nudge. We assumed that most respondents were unfamiliar with the flood damage-reducing investment decision, which would argue against strong preferences. Nevertheless, strong preferences with regard to risk aversion, for example, could explain our results.

Comparing our results with prior findings on the effects of social norm-nudges, we find that our results differ from a recent meta-analysis of field experiments using social norms to promote pro-environmental behavior that reveals a medium-sized main effect of social norms compared to control conditions (Bergquist et al., 2019). Nevertheless, the authors find that social norms are less effective when communicated explicitly (i.e. by computerized messages) rather than implicitly (i.e. by cues in the environment), and that the influence of social norms is weaker in non-student samples than in student samples. Our design used explicit social norms in a non-student sample, which, according to Bergquist et al. (2019), could explain the weak effects.

We are not the first to report a null-effect of social norm-nudges: recent publications have revealed similar findings and in some empirical studies there was a backfiring effects of norm-nudge interventions (see e.g. Fellner et al., 2013; Cranor et al., 2020; Dimant et al., 2020). For example, Fellner et al. (2013) and Cranor et al. (2020) examine the effects of a social norm-nudge on compliance in payments of TV license fees and taxes and find zero effects on compliance. Tyers (2018) uses a social norm-nudge to try and convince consumers to buy carbon offsets for flight tickets and find no effect. The author argues that the main problem is that carbon offsetting is an unfamiliar concept to most participants, in contrast to other domains where norm-nudging has been effective, such as recycling, organ donation or charitable giving. This explanation appears to also reflect our findings. Even though we have no focus group results to back up the claim that respondents are unfamiliar to

Eckel et al., 2011). To control for this explanation, we reran our analysis (not reported here) for Hypothesis 7.4 on a subset of respondents younger than 25 years old and we found the same pattern of results as in the full sample.

the flood preparedness decision, several respondents answered the final open answer feedback question by indicating that they had never thought about the topic before.

An alternative explanation for null effects of social norm-nudges by Dur et al. (2019) states that norm-nudges work well on changing intentions but may ultimately fail to change (costly) behavior. In a field experiment in a retail bank, Dur et al. (2019) find that a social norm-nudge increases intended savings and information search about saving plans, but not actual savings. We find a null-effect for both behavior (investments in the flood risk investment game) as well as intentions (information search in the secondary treatment), rejecting this alternative explanation for our results. It should be noted that our study was an online lab experiment, rather than a field experiment that comprises the majority of the literature on social norm interventions. In a study similar to ours, Dimant et al. (2020) examine the effect of social norm-nudge messages on cheating in an online laboratory experiment and find that their simple norm-nudges are unsuccessful at shifting norms, presumably because a behavioral norm is already in place. Another relevant online lab experiment using norm-nudge messages is Capraro et al. (2019) that examines simple messages promoting personal norms ('what do you think is the morally right thing to do?'). They show that these messages can effectively increase donation behavior immediately after the nudge, as well as in subsequent choices. These findings are in line with the high correlations between investments in our game and answers to the personal norms statements.

A further potential explanation for the absence of an effect could be that an online lab experiment is an artificial setting, that does not perfectly translate to decisions in the field. Nevertheless, findings in Chapter 3 (Mol et al., 2020b) demonstrate that the flood risk game is generally well suited to illustrate the behavior of homeowners regarding flood risk preparedness. For example, the behavior of the players of the flood risk game under different flood risk probabilities and flood risk insurance schemes (mandatory vs. voluntary, high vs. low deductible) is very much in line with theoretical predictions. This gives us confidence that the general mechanisms leading to higher or lower investments into flood risk preparedness in the game can be transferred at least qualitatively - to the real world. One additional result that points in this direction is that the Spearman correlation between investments in the game and the number of flood damage reduction measures implemented by respondents (survey question *f18*) is positive and significant ($\rho = 0.203$, $p < 0.001$). Furthermore, our finding that flood risk perceptions and perceived efficacy of flood damage mitigation measures in a real world context are similar to behavior in the game (see Figure 7.6), suggests that the lab game may translate to behavior and attitudes in the field. Nevertheless, when interpreting the results one should keep in mind that the translation of behavior in the lab to relevant domains in the field is never perfect. One way to counter this limitation in future research is to develop a large-scale experiment in cooperation with homeowners associations and local governments to test the effect of interventions on real-world behavior.

One alternative explanation of the absence of an effect is that people consider investments in flood risk protection as a purely private good. As a result, people could be less interested in information about others' behavior regarding this private good. Future research could examine to what extent people consider flood protection to be a public good and how this belief could interact with social norm nudge effectiveness.

Finally, Czajkowski et al. (2019) compared different variations of descriptive social norm-nudges in a field experiment for household waste sorting with a stated preference approach and found that the willingness to pay for recycling increases with the size of the norm, but that the effect is not monotonic. In other words, high absolute levels of the norm are less effective as they are 'out of reach'. This finding relates to the difference between beliefs (of the current norm) and the information presented in the nudge.

One limitation of this study is that we did not communicate upfront to participants the probability of being paid, as this was dependent on the total number of participants. A recent theoretical study has shown that to be fully incentive-compatible, the pay-one mechanism should be transparent about the chances of being paid before the start of the experiment (Azrieli et al., 2018). Therefore, we recommend that further research should communicate the chances of being paid clearly and early. Another limitation is that we did not elicit beliefs in the Norm-transparent and Norm-high treatments. We did not do this because we did not expect to find independent beliefs about investments in the current sample after providing respondents with the percentages of previous investors. Moreover, we argued that investment data from a previous sample only provides an indication about the nature of investment behavior in the current sample.

In retrospect it would have been interesting to elicit beliefs in the Norm-transparent and Norm-high treatments to check the consistency of beliefs across treatment groups. Had we found the same distributions of beliefs in the Norm-transparent and Norm-high treatments as in the Control group, we would have inferred that respondents simply ignored our norm-nudge messages. Furthermore, we could have used the beliefs in the Norm-transparent and Norm-high treatments to test for the 'norm distance effect' (Bergquist and Nilsson, 2018) that suggests that the power of social norms (messages) is larger when behavior is closer to the (perceived) norm. A second limitation is that the cultural differences between Spain and the Netherlands are not extremely large (Pineda et al., 2015). To obtain a more heterogeneous sample with regard to the individualism-collectivism scale, researchers should consider surveying homeowners in more culturally diverse countries, such as Japan and the U.S. (Hofstede, 2001).

This study suggests three main takeaways for flood risk communication policies. First, communication to raise risk awareness should take risk related emotions into account. This recommendation follows from our finding that worry and concern are significant predictors of investments in damage-reduction (see Figure 7.8), which is in line with previous literature (Kunreuther, 2018). Second, informing homeowners about the effectiveness of damage mitigation measures may enhance flood preparedness. This recommendation follows from the strong positive correlation between response efficacy and investments in damage-reduction in our study (see Figure 7.8), confirming previous findings on this topic (Poussin et al., 2014; Mol et al., 2020a). Third, policy makers should pay particular attention to activating personal norms that which were found to be associated with flood risk preparedness (as indicated by the strong correlation of personal norms with investments in flood preparedness in Figure 7.8). These results are in line with Wenzig and Gruchmann (2018), who showed that personal norms generally have a much larger influence on pro-environmental behavior than social norms, and with Botzen et al. (2019b), who showed that personal norms matter more than social norms in a flood risk context. Schwartz (2012) argued that norms need to be activated to be able to influence intention or behavior by being aware of the consequences of actions and feeling responsible. This explanation complements our results in the context of flood risk preparedness, which was framed on the individual level - respondents who feel responsible for their home (i.e. personal norms) invest more in damage-reducing measures than those who do not feel morally obligated to protect their homes. The actions of neighbors and other fellow homeowners (i.e. social norms) may be of lower importance for mitigation decisions in the context of flood preparedness.

We further found a significant negative correlation between present bias and investments in damage-reducing measures. This finding is in line with previous literature about myopia in the context of preparedness for low-probability/high-impact events, such as floods (Royal and Walls, 2019; Botzen et al., 2019a). Homeowners perceive the high upfront costs of investing in damage-reduction to be much higher than the expected benefits and when they are present biased, they care more about costs now than about benefits later. One way to overcome this bias is through offering low-interest loans that spread the investment costs over time (Meyer and Kunreuther, 2017; Kunreuther and Pauly, 2018), which could stimulate flood preparedness. If a nudge in the environmental domain proves ineffective, as we show in the current chapter, this may warrant the use of stronger measures, such as incentives, regulations and bans, to influence preparedness (Sunstein and Reisch, 2013; Carlsson et al., 2019).

7.7 Conclusion

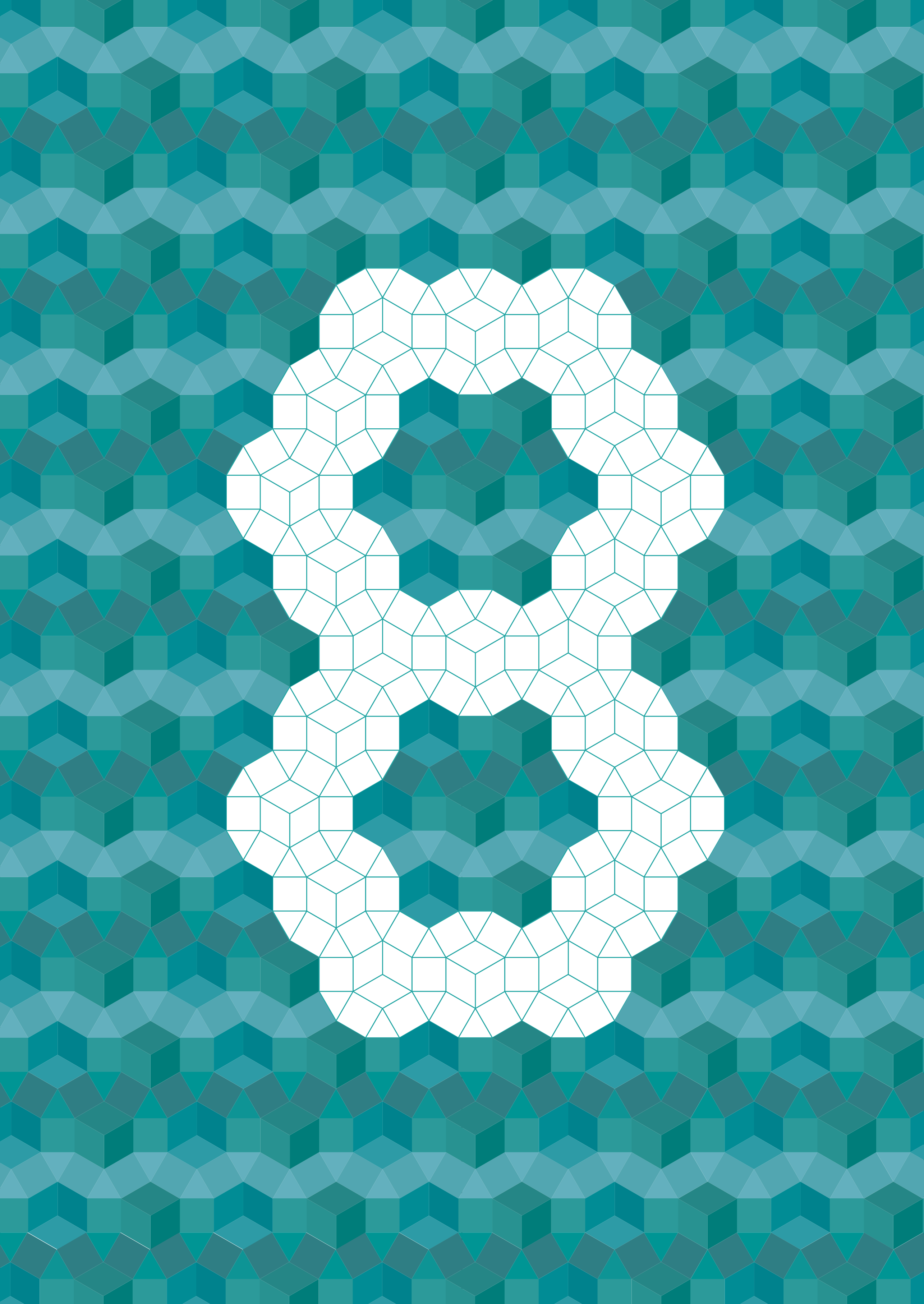
Floods are one of the deadliest and costliest natural disasters worldwide. Fortunately, individual homeowners can invest in several cost-effective measures to prepare their homes for flooding. We attempted to increase investments in flood risk reduction measures in a controlled experiment by subtly nudging respondents (homeowners in the Netherlands and Spain) to consider the social norm of fellow homeowners. In particular, we created different norm-nudge messages by either providing percentages of the population that previously invested in different flood-reduction options (Norm-transparent), or the percentage of previous respondents who invested in flood reduction (Norm-high). These treatments were contrasted with a Control treatment and a Norm-focusing treatment, in which respondents' beliefs about normative patterns of flood-reduction investments were elicited. We did not find any evidence of a treatment effect, suggesting that our social norm-nudges do not affect flood preparedness of respondents in a flood risk investment game.

Our results suggest that the problem of under-preparedness for natural disasters cannot be addressed by providing these social norm-nudges, a finding that differs from previous studies that concluded social norm-nudge messages can be effective ways to facilitate behavioral change in the environmental domain. Our exploratory analyses reveal that there is a strong correlation between beliefs of others' behavior and one's own investments; however our treatments did not influence either.

A crucial difference between this chapter and other successful social norm-nudges from the literature is that the flood-preparedness context is not a very familiar one for most respondents. Such unfamiliarity may lead to a large information-belief gap; respondents may not be aware of any norm with regards to flood-preparedness measures, which could result in overall low norm-sensitivity. Further research could look into this relationship between familiarity and norm-nudge effectiveness. A positive recommendation for further research would be to develop personal or moral norm nudge-messages.

Acknowledgments

We would like to thank conference participants at NoBeC 2019 in Philadelphia, BEEMA5 in Villanova and WINK Nudging and Beyond in Utrecht for valuable comments. This research has received financial support from the Alfred P. Sloan Foundation Grant (No. G-2018-11100), the Netherlands Organization for Scientific Research (VIDI 452.14.005) and Wharton Risk Management and Decision Processes Center.


8

Conclusion

Flooding is the most consequential natural disaster worldwide in terms of monetary losses and the number of people affected. These impacts are expected to increase under higher economic activity and population growth in flood-prone areas, as well as ongoing climate change. One particular way to deal with the risk of flooding is to invest in public flood protection infrastructure. Nevertheless, flood defenses can never offer complete protection and a residual flood risk will remain. This trend calls for careful research into risk reduction strategies, which can be used to manage the financial risk for individuals and institutions (disaster risk insurance) or mitigate potential damage once flood events occur (disaster risk reduction). The economically optimal damage reduction strategy for flooding is a combination of different elements, including investments by individual homeowners to prepare their homes for flooding. Current take-up of damage reduction measures by private homeowners is low, despite the fact that it is cost-effective in areas with high flood risk. A theoretical explanation for the lack of investments in damage reduction measures by private homeowners is moral hazard. This problem may occur in the absence of incentives to take care and limit risk, when losses are shared. The current thesis is one of the first to examine moral hazard empirically in a disaster insurance context, and to study the effects of different probability levels, deductibles, and other incentives, including social norm-nudge messages and virtual reality risk communication, on self-insurance investments. Another potential explanation for low self-insurance investments that was examined in this thesis, is that flood risk perceptions of individual homeowners differ considerably from objective estimates, which may alter their assessment of the cost-effectiveness of damage reduction measures.

Different options exist for private homeowners to reduce risk, including self-insurance (reducing the damage in case of a loss) and self-protection (reducing the probability of a loss occurring). Traditionally, scientists have used field survey data and insurance market data to examine risk reduction behavior. While field survey data has high external validity, it is less suited to identify causal relationships, as different insurance plans are not allocated randomly to homeowners. Insurance market data for natural hazard insurance markets is often not available and if it is, crucial data on preparedness behavior may not be documented. To address the potential confounds in field survey data and the lack of available market data, this thesis used experimental economics methods to examine behavioral motivations and financial incentives to stimulate individual investments in self-insurance. The geographical focus of this thesis lies on Europe. Most experiments were conducted with participants in the Netherlands, which is a relevant case-study because of its low-lying lands that are vulnerable to flooding. Moreover, although the Netherlands has a long history of public flood protection, there is an increased interest in integrated flood risk management strategies in which individuals also take measures that limit flood damage once flood protection infrastructure fails. In addition, Chapter 6.4 examined responses of homeowners in Spain, another European country which is more focused on individual protection measures rather than public flood protection.

The introductory chapter identified several research gaps in the literature. These issues were addressed in Chapter 2 to 7 by assessing the following research questions:

1. To what extent are investments in damage-reducing measures determined by loss probabilities, deductibles and a moral hazard effect? (Chapter)
2. Are financial incentives from insurance effective in increasing investments in damage-reducing measures and does effectiveness vary with insurance scheme (public or private)? (Chapter & 3)
3. Do households generally under- or overestimate flood risk and what factors explain these misperceptions? (Chapter 3.6)
4. What are the possibilities and challenges for experimental economics in high-immersive virtual environments? (Chapter 4.6)
5. Is it possible to increase flood preparedness with the experience of a flood in high-immersive virtual environments? (Chapter 5.6)
6. Could social norm-nudges help people in better preparing for flood risk and do they interact with individual characteristics and intercultural differences? (Chapter 6.4)

Key findings on explaining flood preparedness

While there is an extensive literature on the empirical regularities related to disaster insurance demand and self-protection, research on the drivers of self-insurance is fairly limited. This thesis contributes to the discussion by investigating the relevant dimensions of heterogeneity of self-insurance under compulsory insurance coverage for low-probability/high-impact risk. For example, Chapter examined different probability levels, deductibles, and other financial incentives, which cannot be varied systematically in actual insurance markets. As a starting point, I developed a new investment game to study the causal relationship between financial incentives related to insurance and self-insurance investments, taking into account behavioral characteristics of individuals in an insurance market with mandatory coverage. Chapter shows that subjects invested more in this investment game when the expected value of a loss increased (that is, under higher deductibles and/or higher probabilities of loss). Nevertheless, the increase in investment was not proportional to the increase in risk. Chapter 2.6 extends the investment game to an insurance market with voluntary coverage.

Moral hazard

An important theoretical concept that may explain investments in self-insurance is moral hazard. Under moral hazard, individuals have no incentives to limit risk when losses are shared (for example in a group, such as an insurance contract). Moral hazard theory predicts that without these incentives, policyholders will act careless. This process has indeed been documented in many contexts of asymmetric information, including insurance contracts, court settlements, tax evasion and work effort. Moral hazard may further play a role in natural disaster insurance markets, which are a type of low-probability/high-impact insurance markets. However, few studies have investigated moral hazard under low probabilities and none of these studies used an experimental method, which makes it difficult to draw causal conclusions (Thieken et al., 2006; Hudson et al., 2017). Chapter addressed this research gap by examining the moral hazard problem under low probabilities and high expected damages in a laboratory experiment with students as subjects. The results showed that investments in a treatment without insurance were significantly higher than in an insurance treatment for the high-probability (15%) scenarios, but not significantly different in most low-probability (3%) scenarios. Furthermore, mean investments in the treatment with basic insurance were greater than zero. This means that moral hazard not much of a problem in low-probability/high-impact insurance markets, confirming previous survey research findings (Thieken et al., 2006; Hudson et al., 2017) and answering research question 1.

Chapter 2.6 offered a careful examination of the interplay between financial incentives and behavioral motivations for investing in self-insurance on a group of relevant decision-makers (homeowners in floodplains). This chapter presented the first experimental study of self-insurance behavior under both mandatory and voluntary insurance schemes (answering research question 2), accounting for insurance features and behavioral characteristics of the decision-makers. Furthermore, the large sample size of this study allowed for an in-depth analysis of heterogeneous behavioral motivations among respondents. The chapter also investigated the impacts of the presence or absence of insurance and confirmed the absence of moral hazard under low probabilities and high expected damages, which was initially identified in Chapter . The setup of the investment game under a voluntary insurance scheme allowed to test for advantageous selection. Indeed, a substantial share of respondents was willing to pay for insurance, as well as for self-insurance in the form of damage-reducing investments. I found that respondents who demanded voluntary flood insurance coverage invested approximately 1,000 ECU more in self-insurance than those under mandatory insurance coverage. These 'cautious' types tended to make their decisions based on calculations, and were particularly motivated by the degree to which a measure was deemed effective (response efficacy), social approval by their peers and risk aversion, as well as by a lower trust in dike maintenance.

Risk misperceptions

One other important element of the decision to invest in self-insurance is risk perception. However, flood risk perceptions of individual homeowners may differ considerably from objective estimates. As a result, the assessment of the cost-effectiveness of self-insurance measures may be skewed: homeowners who underestimate the flood probability may not consider self-insurance measures as cost-effective. While the literature on flood risk perceptions is extensive, so far a systematic assessment of the determinants of flood risk misperceptions was lacking. Chapter 3.6 addressed this research gap by quantifying the flood risk misperceptions of Dutch floodplain residents. The analysis revealed that 53% of households overestimate the flood probability and 54% underestimate the maximum water level of a flood. Many respondents correctly estimate the maximum damage in case of a flood. These findings largely confirm the results of Botzen et al. (2015), who found that most New York City floodplain inhabitants overestimate flood probability, while underestimating the potential damage. Furthermore, individuals living in low-lying areas know that they face flood risks, but they underestimate them. One reason for the lack of effect of dike-distance and the strong effect of maximum water levels, is visibility. Respondents cannot easily observe the distance to the nearest dike, while maximum water level (which corresponds to the height of the land) may be easier to observe, for example during periods of rainfall. The main addition of this chapter to the literature lies in the detailed analysis of factors that are related with flood risk misperceptions. This chapter further showed that affective feelings about risk, in this case worry, may lead to over-estimations of probability and water level. Experience of a flood, age, recalling high water levels and trust in dike maintenance seem to decrease flood risk misperceptions. Interestingly, education, risk aversion, income and probability innumeracy do not affect misperceptions. These results answer research question 3 and may be used by policymakers to design effective risk communication campaigns and insurance schemes to cope with increasing natural disaster risks.

Decision heuristics

Low investments in self-insurance may further be explained by decision heuristics, as decision-making in the context of natural disasters is not often fully rational. This bounded rationality may lead to different predictions about human behavior, which may be of interest to insurers and policy makers. For example, people may only respond to a certain threat when this threat has reached a certain threshold level of concern (McClelland et al., 1993). The analysis of Chapter 2.6 showed that threshold level of concern is strongly related to worry about flooding. Furthermore, in Chapter 6.4 I demonstrated that both worry and threshold of concern are important predictors of individual flood-preparedness.

Another heuristic or rule of thumb that people often apply in the context of natural disaster risks is myopia, or present bias. Myopic individuals are generally more oriented towards short-term bank balances rather than potential reduced losses in the future. Present bias can be assessed through a price list, where subjects choose several times between an immediate payment and a (larger) delayed payment. A different and very simple method to estimate present bias is through a qualitative survey question: "In general, are you willing to give up something now in order to profit from that in the future?". In the pretests of Chapter , I used both the complicated but incentivized price list method, as well as the simple survey question to elicit present bias. The results of both methods were highly correlated, which is why only the simple survey question was used in all subsequent chapters. Neither present bias elicited through a price-list nor self-reported present bias was significantly related to investments in self-insurance in Chapter . In contrast, Chapter 2.6 showed that present-biased individuals are significantly less willing to invest in self-insurance in the flood risk investment game than respondents who reported less present biased time preferences. Chapter 6.4 confirmed this positive relationship between present-biased preferences and lower investments in self-insurance.

Chapter 3 examined the interplay of financial incentives and behavioral motivations for investing in self-insurance among homeowners in floodplains. One of the most important predictors of investments in self-insurance were social norms or compliance with the statement "People in my direct environment would approve an investment in damage reducing measures." The results in Chapter 2.6 suggested that changing the social norm for self-insurance by means of information and communication measures could be a potential policy lever to stimulate a wider uptake of these cost-effective measures. Chapter 7 examined this premise by constructing empirical norm-nudge messages to increase investments in damage-reducing measures in the investment game. Contrary to our expectations and the large body of experimental research on the effectiveness of norm-nudges in stimulating pro-environmental behavior (see e.g. van Valkengoed and Steg, 2019), I found no evidence for a treatment effect of social norm-nudges on investments in self-insurance. Instead, the results showed that personal norms, rather than social norms, were positively associated with flood risk preparedness. These results are in line with Wenzig and Gruchmann (2018), who showed that personal norms generally have a much larger influence on pro-environmental behavior than social norms, and with Botzen et al. (2019b), who showed that personal norms matter more than social norms in a flood risk context. Schwartz (2012) argued that norms need to be activated to be able to influence intention or behavior by being aware of the consequences of actions and feeling responsible. This explanation complements results of Chapter 6.4 in the context of flood risk preparedness, which was framed on the individual level. In other words, respondents who feel responsible to protect their home (personal norms) invest more in damage-reducing measures than those who do not feel morally obligated to protect their homes. The actions of neighbors and other fellow homeowners (social norms) may be of lower importance for mitigation decisions in the context of flood preparedness.

Other behavioral explanations

A final explanation for individual risk reduction behavior comes from protection motivation theory (PMT), which originates from psychological theories on preventive behavior in the health domain (Rogers, 1975). PMT states that people experience both coping appraisal and threat appraisal when facing a threat.

Threat appraisal is defined as the subjective evaluation of risk and has been measured throughout this thesis. Chapter 3.6 examined different components of flood risk perception among Dutch floodplain inhabitants. The results show that a majority of respondents overestimates the probability of a flood event and underestimates the potential water level in case of a flood. The chapter further showed that worry about flooding may lead to over-estimations of probability and water level. Coping appraisal refers to the evaluation of responses to a threat, including the perceived ability to install mitigation measures (self-efficacy) and the perceived effectiveness of these measures (response-efficacy) (Floyd et al., 2000). Self-efficacy and response efficacy are often referred to as coping values and they are among the most important determinants of disaster preparedness. Three chapters in this thesis (Chapter , Chapter 2.6 and Chapter 6.4) confirmed that response efficacy and self-efficacy are among the dominant behavioral motivations stimulating investments in self-insurance. Particularly individuals who are cautious in their actions, seem to be motivated by response efficacy.

Key findings on stimulating flood preparedness

Insurance features and related incentives

Chapter examined financial incentives for damage reduction. First, I considered increasing the traditional financial incentive from the insurance industry, which is increasing the deductible to decrease a policyholder's coverage. In the lab experiment in Chapter , deductible levels varied between 5%, 15% and 20%. The results are in line with theoretical predictions: increasing the deductible leads to slightly higher investments in self-insurance. This finding supports the substitution hypothesis of Carson et al. (2013), who theorized that insurance and mitigation may be substitute goods. The deductible effect is smallest in the low-probability (3%) scenarios, which confirms previous survey research in natural disaster insurance markets (Hudson et al., 2017).

The results further indicate that a premium discount can increase investment in damage-reduction, while the availability of a mitigation loan does not increase investments. Since the premium discount is based on the expected value of damage-reduction, a larger premium discount in absolute terms is given under low levels of deductibles and high probabilities of damages. The results show that the effect of the premium discount is actually larger under low levels of deductibles. Chapter 2.6 confirms that a premium discount can increase investments in self-insurance in an online sample of Dutch homeowners, although it does not matter whether this insurance is provided in a public or private market. A premium discount can increase investments in self-insurance to the same extent as an increase in probability of loss from 1% to 5%. The finding that a premium discount can be effective in increasing self-insurance investments even under low probabilities of loss, confirms previous empirical studies (Botzen et al., 2009b; Hudson et al., 2016).

Virtual reality technology

A novel approach to stimulate flood preparedness is to use virtual reality (VR) technology to show people the consequences of a flood in their home. Chapter 4.6 aimed to give a critical overview of the possibilities and challenges for experimental economics in high-immersive virtual environments (answering research question 4). First and foremost, experiences in VR seem to extend to real life and a close parallel has been found between behavior in VR experiments and conventional labs. One of the key advantages of VR above conventional field experiments is that it is relatively easy to control for confounding factors such as weather, gender and non-verbal cues. Many economic field experiments could be improved by this technology, leading to more robust findings and helping to exclude alternative explanations. Thanks to the improved technologies in the past decade, perceived realism (presence) now allows for VR research to move from methodological publications to experiments with respect to content. VR allows experimenters to safely expose participants to various virtual risks and to investigate their responses. High-immersive VR environments can be used to have participants experience the impact of a disaster, as well as the effectiveness of preventive measures (see e.g. Jansen et al., 2020). The main drawbacks of VR experiments are the costs of equipment and the required programming skills, although developments in the game industry might lead to cheaper devices and straightforward software, as well as improved specifications to minimize simulator sickness. Note that as technology advances, VR experiments have the potential to increase both in the realism and the control dimension.

In Chapter 5.6, I applied virtual reality technology to the context of individual flood preparedness. Participants could experience the impacts of a flood inside a typical Dutch home, through visualization of high water levels and the associated damages, e.g. in the form of floating objects swept along by the water. The results showed that participants who experienced the virtual flood invest significantly more in the flood risk investment game than those in the control group. These effects are persistent up to four weeks after the VR intervention.

Social norm-nudges

Chapter 6.4 attempted to increase investments in flood risk reduction measures in a controlled experiment by subtly nudging respondents (homeowners in the Netherlands and Spain) to consider the social norm of what fellow homeowners are doing. In particular, I presented different norm-nudge messages showing percentages of the population that previously invested in different flood-reduction options (Norm-transparent), or the percentage of previous respondents who invested anything in flood reduction (Norm-high). These treatments were contrasted with a Control treatment and a Norm-focusing treatment, in which respondents' beliefs about normative patterns of flood-reduction investments were elicited. I did not find any evidence of a treatment effect, suggesting that social norm-nudges do not affect flood preparedness of respondents in a flood risk investment game, answering research question 6. These results contrast with the existing evidence that social norm-nudge messages can be effective ways to facilitate behavioral change in the environmental domain. The exploratory additional results show that there is a strong correlation between beliefs of others' behavior and one's own investments, however our treatments did not influence either. If a nudge in the environmental domain proves ineffective, as in the current thesis, this may warrant the use of stronger measures, such as incentives, regulations and bans, to influence preparedness. To conclude, our results suggest that the problem of under-preparedness for natural disasters cannot (even partly) be solved by social norm-nudges. Investments in flood risk reduction could be largely explained by the number of measures installed at home, present bias, personal norms and response efficacy.

Policy recommendations

The following policy recommendations can be drawn from this thesis. First, the finding that there is no moral hazard in this LPHI insurance market suggests that high deductibles may not be necessary to limit such an effect. This is in line with previous survey results of Hudson et al. (2017) who found that a majority of (hurricane insurance) policyholders are not even aware of having a deductible and that deductibles played a minor role in hurricane preparedness activities. Using premium discounts is likely to be a more effective way for insurers to stimulate policyholders to reduce natural disaster risk in general and flood risk in particular. Moreover, advantageous selection was observed in the data, which implies that a substantial group of policyholders is willing to invest in self-insurance. These results support the ongoing debates and reforms aimed at linking flood insurance coverage with risk reduction in the European Union (Surminski et al., 2015; Hochrainer-Stigler et al., 2017) and the United States (Tullos, 2018).

Furthermore, the analysis in this thesis may justify the strengthening of purchase requirements for flood insurance as I found no support for moral hazard throughout multiple experiments and voluntary take-up rates were low. Furthermore, Chapter 2.6 showed that less cautious individuals (who do not believe that flood risk will increase, nor that they should take action) select less insurance coverage, which could lead to substantial claims for government support which may drain public resources. These could be important topics for informational campaigns aimed at improving flood preparedness, which should be focused on explaining possible cost-effective measures, rather than on increasing awareness about flood risk in general. The results further indicated that individuals who used calculations in the decision-making process were more inclined to select insurance coverage and (over-)invest in self-insurance. The fact that reporting a calculating strategy does not increase optimal investments may indicate either miscalculation or preferences for over-investment. A potential policy recommendation would be to provide calculation tools in communication about cost-effective self-insurance measures.

Chapter 3.6 demonstrated that a majority of respondents underestimates the water level of a flood. This implies that many Dutch homeowners may underestimate the cost-effectiveness of damage reduction measures. Policymakers could respond with information campaigns for homeowners in the river delta, to highlight that homes can be improved with cost-effective measures. Moreover, these campaigns could specifically target homeowners in low-lying areas as they are currently over-represented in the share of under-estimators of flood risk. This analysis also showed that worry about flooding may increase flood risk perceptions, although one should be careful to prevent over-estimations. One important policy recommendation is to use risk-related emotions in risk communication. Chapter 5.6 showed that a disaster experience can be effectively simulated in a VR lab and that this intervention may be effective in increasing investments through an increase in worry.

Another potential powerful theme to highlight in risk communication is the effectiveness of damage mitigation measures. This recommendation follows from the strong positive correlation between response efficacy and investments in damage-reduction in Chapter 2.6 and Chapter 6.4. Furthermore, risk communication campaigns could focus on consequential factors of risk, such as damage estimates and the maximum water level, since they are salient and rather easy to imagine, rather than communicating difficult to interpret probabilities or return periods.

I further found a significant negative relationship between present bias and investments in damage-reducing measures in Chapter 2.6 and 6.4. This finding is in line with previous literature about myopia in the context of preparedness for low-probability/high-impact events, such as floods (Royal and Walls, 2019; Botzen et al., 2019a). One way to overcome this bias is through offering low-interest loans that spread the investment costs over time (Meyer and Kunreuther, 2017; Kunreuther and Pauly, 2018), which could stimulate flood preparedness. However, Chapter tested whether a mitigation loan could increase self-insurance investments in the flood risk investment game, which was not supported by the data. This finding could be explained by the dislike for the mandatory 1% interest in the loan, or a general dislike of lending among the students in the sample. Alternatively, this finding suggests that the operationalization of a loan treatment in the lab lacks external validity. Further research on loans in the context of disaster risk reduction should therefore focus on field rather than lab experiments.

Finally, policy makers should pay particular attention to activating personal norms, which were found to be associated with flood risk preparedness in Chapter 6.4. These results are in line with Wenzig and Gruchmann (2018), who showed that personal norms generally have a much larger influence on pro-environmental behavior than social norms, and with Botzen et al. (2019b), who showed that personal norms matter more than social norms in a flood risk context. Schwartz (2012) argued that norms need to be activated to be able to influence intention or behavior by being aware of the consequences of actions and feeling responsible. Further work could thus examine the interactions between these antecedents of personal norms and message design to explore how personal norms can be effectively activated.

Limitations and directions for future research

Some limitations of my work should be addressed. The experiments in this thesis developed from a theory-driven lab experiment, to more applied work using virtual reality technology. Nevertheless, none of the experiments in this thesis examined real-world investment behavior in the form of natural field experiments. Instead, this thesis offers an overview of the most (and least) promising incentives and nudges to be tested in the field. The laboratory approach of this thesis could be particularly relevant in interpreting the non-significant results of the Loan treatment in Chapter , where the investment costs were spread over 12 rounds taking several minutes in the lab, rather than years in the real world. The fact that the mitigation loan has no impact in the lab may signal either that mitigation loans are not an effective incentive to increase investments in self-insurance, or simply that its operationalization in the lab failed. The fact that the results of the incentivized time-preferences task did not correlate with investments in self-insurance points in the direction of the former, but this should be re-examined by field experimental work. Alternatively, true intertemporal payoffs could be incorporated in a lab experiment (see e.g. Attema et al., 2016), although this would make the game more complex. One way to increase external validity of the laboratory game would be to add uncertainty about the future. For simplicity, our participants played a fixed number of rounds in the game. An interesting possibility would be to add a random stopping rule to the game to mimic the indefinite time horizon of real-world policyholders.

One other potential problem in the generalizability of the results is the high-stakes nature of the flood preparedness context. The flood risk investment game developed in this thesis aspired to mimic this high stakes context as closely as possible in a laboratory environment, by incorporating a payoff scheme in which one participant was paid a large sum based on their decisions in the game (up to €650). Compared to the regular payoffs of the panel company (approximately €1), the stakes are large. Nevertheless, a loss in the game corresponding to a loss in potential payout of €550 is still not even close to the size of flood damage in the real world. This limitation should be acknowledged, but it does not undermine the aim of this thesis, which was to offer an overview of the most (and least) promising incentives and nudges to be tested in the field, as well as giving a thorough understanding of the mechanisms driving the willingness to invest in flood-risk preparedness.

It should further be noted that all experiments were individual decision-making experiments, whereas homeowners rarely take large investment decisions in real life without consultation of neighbors and family members. Moreover, the limited length of the post-experimental surveys in Chapter 2.6 and Chapter 6.4 restricted the explanatory variables to simple survey questions, while it would have been interesting to take a closer look at risk attitudes, by differentiating between utility curvature, probability weighting and loss aversion as in Prospect Theory (Kahneman and Tversky, 1979). Even though probability (under)weighting of flood risks may be explained by Prospect Theory (Botzen and van den Bergh, 2012), literature suggests that probability weighting is different for precautionary decisions compared to simple monetary gambles (Kusev et al., 2009). Future research could examine the exact interplay of loss aversion, utility curvature and probability weighting in the context of flood preparedness. Lastly, this thesis examined behavior from a majority of Dutch and some Spanish participants. A more heterogeneous sample with regards to culture and public flood risk management practice could lead to different results.

It could also be interesting to examine beliefs about weather warnings and trust in flood management institutions in more detail. Furthermore, our finding that personal norms are strongly correlated with investments in self-insurance suggests that it would be interesting to test whether personal norms with regard to flood risk preparedness can be activated or strengthened. One possibility would be to develop a norm-nudge message based on personal, rather than social norms (see e.g. Bilancini et al., 2020).

Finally, quite some policy recommendations are related to risk communication campaigns. Future research could focus on the effectiveness of these informational campaigns, as well as the use of calculation tools to help to increase investments in cost-effective self-insurance measures among several types of decision-makers. This thesis examined the possibilities of high-tech virtual reality solutions to help people visualize a disaster situation. Future research could examine whether lower tech approaches, such as VR set-ups that rely on a smartphone, can be equally effective as expensive head-mounted displays. Such a lower tech approach would also allow for larger sample sizes, which is particularly useful when examining the mechanism behind the intervention effect.

References

Abrahamse, W. and Steg, L. (2013). Social influence approaches to encourage resource conservation: A meta-analysis. *Global Environmental Change*, 23(6):1773–1785.

Achenbach, J., Waltemate, T., Latoschik, M. E., and Botsch, M. (2017). Fast generation of realistic virtual humans. *Proceedings of the 23rd ACM Symposium on Virtual Reality Software and Technology - VRST '17 - retrieved from http://dl.acm.org/citation.cfm?doid=3139131.3139154*, pages 1–10.

Aerts, J. C. J. H., Botzen, W. J. W., de Moel, H., and Bowman, M. (2013). Cost estimates for flood resilience and protection strategies in New York City. *Annals of the New York Academy of Sciences*, 1294(1):1–104.

Aerts, J. C. J. H., Sprong, T., and Bannink, B. (2008). Aandacht voor Veiligheid. Technical report, Leven met Water, Klimaat voor Ruimte, DG Water.

Allcott, H. (2011). Social norms and energy conservation. *Journal of Public Economics*, 95(9-10):1082–1095.

Arad, A. and Rubinstein, A. (2018). The people's perspective on libertarian-paternalistic policies. *The Journal of Law and Economics*, 61(2):311–333.

Araujo, F. A., Erin Carbone, B., Dunietz, M. W., Ania Jaroszewicz, B., Rachel Landsman, B., Diego Lamé, B., Lise Vesterlund, B., Wang, S. W., and Alistair Wilson, B. J. (2016). The slider task: an example of restricted inference on incentive effects. *Journal of the Economic Science Association*, 2:1–12.

Arnott, R. J. and Stiglitz, J. E. (1988). The basic analytics of moral hazard. *Scandinavian Journal of Economics*, 90(3):383–413.

Arrow, K. J. (1963). Uncertainty and the welfare economics of medical care. *American Economic Review*, 53(5):141–149.

Ash, K. D., Schumann, R. L., and Bowser, G. C. (2014). Tornado warning trade-offs: Evaluating choices for visually communicating risk. *Weather, Climate, and Society*, 6(1):104–118.

Attema, A. E., Bleichrodt, H., Gao, Y., Huang, Z., and Wakker, P. P. (2016). Measuring discounting without measuring utility. *American Economic Review*, 106(6):1476–1494.

Azrieli, Y., Chambers, C. P., and Healy, P. J. (2018). Incentives in experiments: A theoretical analysis. *Journal of Political Economy*, 126(4):1472–1503.

Bailenson, J. N., Beall, A. C., Loomis, J. M., Blascovich, J., and Turk, M. (2005). Transformed social interaction, augmented gaze, and social influence in immersive virtual environments. *Human Communication Research*, 31(4):511–537.

Bailenson, J. N. and Blascovich, J. (2004). Avatars. In *Encyclopedia of Human-Computer Interaction, Berkshire Publishing Group*. Citeseer.

Bailenson, J. N., Blascovich, J., Beall, A. C., and Loomis, J. M. (2003). Interpersonal distance in immersive virtual environments. *Personality and Social Psychology Bulletin*, 29(7):819–833.

Bakker, M. H., Kerstholt, J. H., van Bommel, M., and Giebels, E. (2019). Decision-making during a crisis: the interplay of narratives and statistical information before and after crisis communication. *Journal of Risk Research*, 22(11):1409–1424.

Baldwin, M. and Mussweiler, T. (2018). The culture of social comparison. *Proceedings of the National Academy of Sciences of the United States of America*, 115(39):E9067–E9074.

Balk, S. A., Bertola, M. A., and Inman, V. W. (2013). Simulator sickness questionnaire: Twenty years later. *Proceedings of the Seventh International Driving Symposium on Human Factors in Driver Assessment, Training, and Vehicle Design*, pages 257–263.

Bamberg, S. and Möser, G. (2007). Twenty years after Hines, Hungerford, and Tomera: A new meta-analysis of psycho-social determinants of pro-environmental behaviour. *Journal of Environmental Psychology*, 27(1):14–25.

Banerjee, P., Chakravarty, S., and Mohanty, R. (2011). Equilibrium play and learning in a stochastic public goods experiment. *Journal of Quantitative Economics*, 10(2):59–77.

Barberis, N. (2013). The psychology of tail events: Progress and challenges. *The American Economic Review*, 103(3):611–616.

Bardsley, N., Cubitt, R., Loomes, G., Moffatt, P., Starmer, C., and Sugden, R. (2010). *Experimental Economics: Rethinking the Rules*. Princeton University Press.

Barnes, S. J. (2016). Understanding virtual reality in marketing: Nature, implications and potential. *SSRN Electronic Journal*, pages 1–50.

Barron, G. and Yechiam, E. (2009). The coexistence of overestimation and underweighting of rare events and the contingent recency effect. *Judgment and Decision Making*, 4(6):447–460.

Bartke, S., Friedl, A., Gelhaar, F., and Reh, L. (2017). Social comparison nudges - Guessing the norm increases charitable giving. *Economics Letters*, 152:73–75.

Bateman, I. J., Day, B. H., Jones, A. P., and Jude, S. (2009). Reducing gainloss asymmetry: A virtual reality choice experiment valuing land use change. *Journal of Environmental Economics and Management*, 58(1):106–118.

Bauernschuster, S., Falck, O., Heblich, S., Suedekum, J., and Lameli, A. (2014). Why are educated and risk-loving persons more mobile across regions? *Journal of Economic Behavior and Organization*, 98:56–69.

Bearden, W. O., Netemeyer, R. G., and Teel, J. E. (1989). Measurement of consumer susceptibility to interpersonal influence. *Journal of Consumer Research*, 15(4):473–481.

Becker, J. S., Paton, D., Johnston, D. M., Ronan, K. R., and McClure, J. (2017). The role of prior experience in informing and motivating earthquake preparedness. *International Journal of Disaster Risk Reduction*, 22(July 2016):179–193.

Bell, A. and Jones, K. (2015). Explaining fixed effects: Random effects modeling of time-series cross-sectional and panel data. *Political Science Research and Methods*, 3(01):133–153.

Benartzi, S., Beshears, J., Milkman, K. L., Sunstein, C. R., Thaler, R. H., Shankar, M., Tucker-Ray, W., Congdon, W. J., and Galing, S. (2017). Should governments invest more in nudging? *Psychological Science*, 28(8):1041–1055.

Berger, L. A. and Hershey, J. C. (1994). Moral hazard, risk seeking, and free riding. *Journal of Risk and Uncertainty*, 9(2):173–186.

Bergquist, M. and Nilsson, A. (2018). Using social norms in smart meters: the norm distance effect. *Energy Efficiency*, 11(8):2101–2109.

Bergquist, M., Nilsson, A., and Schultz, W. P. (2019). A meta-analysis of field-experiments using social norms to promote pro-environmental behaviors. *Global Environmental Change*, 59(July):101941.

Betsch, C., Böhm, R., Korn, L., and Holtmann, C. (2017). On the benefits of explaining herd immunity in vaccine advocacy. *Nature Human Behaviour*, 1(3):56.

Bicchieri, C. (2006). *The grammar of society: The nature and dynamics of social norms.* Cambridge University Press.

Bicchieri, C. (2010). Norms, preferences, and conditional behavior. *Politics, Philosophy and Economics*, 9(3):297–313.

Bicchieri, C. (2016). *Norms in the wild: How to diagnose, measure, and change social norms.* Oxford University Press.

Bicchieri, C. and Dimant, E. (2019). Nudging with care: the risks and benefits of social information. *Public Choice.*

Bicchieri, C., Lindemans, J. W., and Jiang, T. (2014). A structured approach to a diagnostic of collective practices. *Frontiers in Psychology*, 5:1418.

Biener, C., Eling, M., Landmann, A., and Pradhan, S. (2014). Ex-ante moral hazard: Experimental evidence from low-income insurance.

Biener, C., Eling, M., Landmann, A., and Pradhan, S. (2018). Can group incentives alleviate moral hazard? The role of pro-social preferences. *European Economic Review*, 101:230–249.

Bilancini, E., Boncinelli, L., Capraro, V., Celadin, T., and Paolo, R. D. (2020). The effect of norm-based messages on reading and understanding COVID-19 pandemic response governmental rules. *Journal of Behavioral Economics for Policy*, May.

Bin, O. and Landry, C. E. (2013). Changes in implicit flood risk premiums: Empirical evidence from the housing market. *Journal of Environmental Economics and Management*, 65(3):361–376.

Biocca, F. and Levy, M. R. (1995). *Communication in the age of virtual reality*. Routledge.

Bixter, M. T. and Luhmann, C. C. (2014). Shared losses reduce sensitivity to risk: A laboratory study of moral hazard. *Journal of Economic Psychology*, 42(December):63–73.

Blascovich, J., Loomis, J. M., Beall, A. C., Swinth, K. R., Hoyt, C. L., and Bailenson, J. N. (2002). Immersive virtual environment technology as a methodological tool for social psychology. *Psychological Inquiry*, 13(2):103–124.

Bolton, G., Dimant, E., and Schmidt, U. (2020). Observability and Social Image: On the Robustness and Fragility of Reciprocity. *SSRN Electronic Journal*.

Bombari, D., Schmid Mast, M., Canadas, E., and Bachmann, M. (2015). Studying social interactions through immersive virtual environment technology: virtues, pitfalls, and future challenges. *Frontiers in Psychology*, 6(June):1–11.

Botzen, W. J. W. (2013). *Managing extreme climate change risks through insurance*. Cambridge University Press.

Botzen, W. J. W., Aerts, J. C. J. H., and van den Bergh, J. C. J. M. (2009a). Dependence of flood risk perceptions on socioeconomic and objective risk factors. *Water Resources Research*, 45(10):1–15.

Botzen, W. J. W., Aerts, J. C. J. H., and van den Bergh, J. C. J. M. (2009b). Willingness of homeowners to mitigate climate risk through insurance. *Ecological Economics*, 68(8-9):2265–2277.

Botzen, W. J. W., Aerts, J. C. J. H., and van den Bergh, J. C. J. M. (2013). Individual preferences for reducing flood risk to near zero through elevation. *Mitigation and Adaptation Strategies for Global Change*, 18(2):229–244.

Botzen, W. J. W., Kunreuther, H., Czajkowski, J., and Moel, H. (2019a). Adoption of Individual Flood Damage Mitigation Measures in New York City: An Extension of Protection Motivation Theory. *Risk Analysis*, 39(10):2143–2159.

Botzen, W. J. W., Kunreuther, H., and Michel-Kerjan, E. (2019b). Protecting against disaster risks: Why insurance and prevention may be complements. *Journal of Risk and Uncertainty*, 59(2):151–169.

Botzen, W. J. W., Kunreuther, H. C., and Michel-Kerjan, E. O. (2015). Divergence between individual perceptions and objective indicators of tail risks: Evidence from floodplain residents in New York City. *Judgment and Decision Making*, 10(4):365–385.

Botzen, W. J. W. and van den Bergh, J. C. J. M. (2008). Insurance against climate change and flooding in the Netherlands: Present, future, and comparison with other countries. *Risk Analysis*, 28(2):413–426.

Botzen, W. J. W. and van den Bergh, J. C. J. M. (2012). Risk attitudes to low-probability climate change risks: WTP for flood insurance. *Journal of Economic Behavior and Organization*, 82(1):151–166.

Braun, M. and Muermann, A. (2004). The impact of regret on the demand for insurance. *The Journal of Risk and Insurance*, 71(4):737–767.

Bruns, H., Kantorowicz-Reznichenko, E., Klement, K., Luistro Jonsson, M., and Rahali, B. (2018). Can nudges be transparent and yet effective? *Journal of Economic Psychology*, 65:41–59.

Bubeck, P., Berghäuser, L., Hudson, P., and Thieken, A. H. (2020). Using Panel Data to Understand the Dynamics of Human Behavior in Response to Flooding. *Risk Analysis*.

Bubeck, P., Botzen, W. J. W., and Aerts, J. C. J. H. (2012). A review of risk perceptions and other factors that influence flood mitigation behavior. *Risk Analysis*, 32(9):1481–1495.

Bubeck, P., Botzen, W. J. W., Kreibich, H., and Aerts, J. C. J. H. (2013). Detailed insights into the influence of flood-coping appraisals on mitigation behaviour. *Global Environmental Change*, 23(5):1327–1338.

Burger, N. E. and Kolstad, C. D. (2009). Voluntary public goods provision, coalition formation, and uncertainty. *National Bureau of Economic Research Working Paper Series*, No. 15543(15543):1–35.

Buunk, A. P. and Gibbons, F. X. (2006). Social comparison orientation: a new perspective on those who do and those who don't compare with others. In Guimond, S., editor, *Social comparison and social psychology - Understanding cognition, intergroup relations, and culture*, pages 15–32. Cambridge University Press, New York.

Cai, D. A. and Fink, E. L. (2002). Conflict style differences between individualists and collectivists. *Communication Monographs*, 69(1):67–87.

Camerer, C. F. and Hogarth, R. M. (1999). The effects of financial incentives in experiments: A review and capital labor production framework. *Journal of Risk and Uncertainty*, 19:7–42.

Camerer, C. F. and Kunreuther, H. C. (1989). Decision processes for low probability events: Policy implications. *Journal of Policy Analysis and Management*, 4(8).

Capraro, V., Jagfeld, G., Klein, R., Mul, M., and de Pol, I. v. (2019). Increasing altruistic and cooperative behaviour with simple moral nudges. *Scientific Reports*, 9(1):11880.

Carlsson, F., Gravert, C. A., Kurz, V., and Johansson-Stenman, O. (2019). Nudging as an Environmental Policy Instrument. *SSRN Electronic Journal*, 756.

Carlton, S. J. and Jacobson, S. K. (2013). Climate change and coastal environmental risk perceptions in Florida. *Journal of Environmental Management*, 130:32–39.

Carson, J. M., Mccullough, K. A., and Pooser, D. M. (2013). Deciding whether to invest in mitigation measures: Evidence from Florida. *Journal of Risk and Insurance*, 80(2):309–327.

Cason, T. N., Gangadharan, L., and Maitra, P. (2012). Moral hazard and peer monitoring in a laboratory microfinance experiment. *Journal of Economic Behavior and Organization*, 82(1):192–209.

Chabé-Ferret, S., Le Coent, P., Reynaud, A., Subervie, J., and Lepercq, D. (2019). Can we nudge farmers into saving water? Evidence from a randomised experiment. *European Review of Agricultural Economics*, 46(3):393–416.

Chang, D., Chen, R., and Krupka, E. (2019). Rhetoric matters: A social norms explanation for the anomaly of framing. *Games and Economic Behavior*, 116:158–178.

Chaudhry, S. J., Hand, M., and Kunreuther, H. C. (2018). Extending the time horizon: Elevating concern for rare events by communicating losses over a longer period of time.

Chen, D. L., Schonger, M., and Wickens, C. (2016). oTreeAn open-source platform for laboratory, online, and field experiments. *Journal of Behavioral and Experimental Finance*, 9:88–97.

Chiappori, P.-A. and Salanié, B. (2000). Testing for asymmetric information in insurance markets. *Journal of Political Economy*, 108(1):56–78.

Chittaro, L., Sioni, R., Crescentini, C., and Fabbro, F. (2017). Mortality salience in virtual reality experiences and its effects on users attitudes towards risk. *International Journal of Human Computer Studies*, 101(July 2016):10–22.

Cho, J.-m., Kim, Y.-d., Jung, S. H., Shin, H., and Kim, T. (2017). 78-4 Screen door effect mitigation and its quantitative evaluation in VR display. *SID Symposium Digest of Technical Papers*, 48:1154–1156.

Cialdini, R. B. and Goldstein, N. J. (2004). Social influence: Compliance and conformity. *Annual Review of Psychology*, 55(1):591–621.

Cialdini, R. B., Reno, R. R., and Kallgren, C. A. (1990). A focus theory of normative conduct: Recycling the concept of norms to reduce littering in public places. *Journal of Personality and Social Psychology*, 58(6):1015–1026.

Cohen, A. and Siegelman, P. (2010). Testing for adverse selection in insurance markets. *Journal of Risk and Insurance*, 77(1):39–84.

Conell-Price, L. and Jamison, J. (2015). Predicting health behaviors with economic preferences & locus of control. *Journal of Behavioral and Experimental Economics*, 54:1–9.

Corcos, A., Pannequin, F., and Montmarquette, C. (2017). Leaving the market or reducing the coverage? A model-based experimental analysis of the demand for insurance. *Experimental Economics*.

Cranor, T., Goldin, J., Homonoff, T., Moore, L., Homono, T., and Moore, L. (2020). Communicating Tax Penalties to Delinquent Taxpayers: Evidence from a Field Experiment. *National Tax Journal*, 73(2):331–360.

Crosetto, P. and Filippin, A. (2013). The bomb risk elicitation task. *Journal of Risk and Uncertainty*, 47:31–65.

Cruz-Neira, C., Sandin, D. J., and DeFanti, T. A. (1993). Surround-screen projection-based virtual reality. *Proceedings of the 20th annual conference on Computer graphics and interactive techniques - SIGGRAPH '93*, pages 135–142.

Csermely, T. and Rabas, A. (2016). How to reveal peoples preferences: Comparing time consistency and predictive power of multiple price list risk elicitation methods. *Journal of Risk and Uncertainty*, 53(2-3):107–136.

Cutler, D. M., Finkelstein, A., and McGarry, K. (2008). Preference heterogeneity and insurance markets: Explaining a puzzle of insurance. *The American Economic Review*, 98(2):157–162.

Czajkowski, M., Zagórska, K., and Hanley, N. (2019). Social norm nudging and preferences for household recycling. *Resource and Energy Economics*, 58:101110.

Czura, K. (2015). Pay, peek, punish? Repayment, information acquisition and punishment in a microcredit lab-in-the-field experiment. *Journal of Development Economics*, 117(November):119–133.

Dai, Z., Galeotti, F., and Villeval, M. C. (2018). Cheating in the Lab Predicts Fraud in the Field: An Experiment in Public Transportation. *Management Science*, 64(3):1081–1100.

Dannenberg, A., Riechmann, T., Sturm, B., and Vogt, C. (2012). Inequality aversion and the house money effect. *Experimental Economics*, 15(3):460–484.

de Boer, J., Botzen, W. J. W., and Terpstra, T. (2016). Flood risk and climate change in the Rotterdam area, The Netherlands: enhancing citizen's climate risk perceptions and prevention responses despite skepticism. *Regional Environmental Change*, 16(6):1613–1622.

de Bruin, W. B., Parker, A. M., and Maurer, J. (2011). Assessing small non-zero perceptions of chance: The case of H1N1 (swine) flu risks. *Journal of Risk and Uncertainty*, 42(2):145–159.

de Meza, D. and Webb, D. C. (2001). Advantageous selection in insurance markets. *RAND Journal of Economics*, 32(2):249–262.

De Moel, H., Bouwer, L. M., and Aerts, J. C. J. H. (2014a). Uncertainty and sensitivity of flood risk calculations for a dike ring in the south of the Netherlands. *Science of the Total Environment*, 473-474:224–234.

De Moel, H., van Vliet, M., and Aerts, J. C. J. H. (2014b). Evaluating the effect of flood damage-reducing measures: A case study of the unembanked area of Rotterdam, the Netherlands. *Regional Environmental Change*, 14(3):895–908.

De Raad, B., Morales-Vives, F., Barelds, D. P., Van Oudenhoven, J. P., Renner, W., and Timmerman, M. E. (2016). Values in a cross-cultural triangle: A comparison of value taxonomies in the Netherlands, Austria, and Spain. *Journal of Cross-Cultural Psychology*, 47(8):1053–1075.

Deaton, A. and Cartwright, N. (2018). Understanding and misunderstanding randomized controlled trials. *Social Science and Medicine*, 210(October 2017):2–21.

Deck, C. A. and Reyes, J. (2008). An experimental investigation of moral hazard in costly investments. *Southern Economic Journal*, 74(January 2007):725–746.

DeHoratius, N., Gürerk, Ö., Honhon, D., and Hyndman, K. (2018). Execution failures in retail supply chains A virtual reality experiment. *SSRN Electronic Journal*, pages 1–45.

Demuth, J. L., Morss, R. E., Lazo, J. K., and Trumbo, C. (2016). The effects of past hurricane experiences on evacuation intentions through risk perception and efficacy beliefs: A mediation analysis. *Weather, Climate, and Society*, 8(4):327–344.

Den, X. L., Persson, M., Benoist, A., Hudson, P., Ruiter, M. C. d., and Ruig, L. d. (2017). Insurance of weather and climate- related disaster risk: Inventory and analysis of mechanisms to support damage prevention in the EU. Technical Report August, European Commission, Brussels.

Dhillon, A., Peeters, R., and Yuksel, A. M. (2014). Overcoming moral hazard with social networks in the workplace: An experimental approach.

Di Mauro, C. (2002). Ex ante an ex post moral hazard in compensation for income losses: results from an experiment. *Journal of Socio-Economics*, 31(3):253–271.

Di Mauro, C. and Anna, M. (1996). An experimental investigation of the impact of ambiguity on the valuation of self-insurance and self-protection. *Journal of Risk and Uncertainty*, 13(1):53–71.

Dibbets, P. and Schulte-Ostermann, M. A. (2015). Virtual reality, real emotions: a novel analogue for the assessment of risk factors of post-traumatic stress disorder. *Frontiers in Psychology*, 6(May):1–8.

Dickert, S., Västfjäll, D., Mauro, R., and Slovic, P. (2014). The Feeling of Risk: Implications for Risk Perception and Communication. In *The SAGE Handbook of Risk Communication*, pages 41–54. SAGE Publications, Inc., 2455 Teller Road, Thousand Oaks California 91320.

Dielman, T. E., Campanelli, P. C., Shope, J. T., and Butchart, A. T. (1987). Susceptibility to Peer Pressure, Self-Esteem, and Health Locus of Control as Correlates of Adolescent Substance Abuse. *Health Education & Behavior*, 14(2):207–221.

Diemer, J., Alpers, G. W., Peperkorn, H. M., Shiban, Y., and Mühlberger, A. (2015). The impact of perception and presence on emotional reactions: A review of research in virtual reality. *Frontiers in Psychology*, 6(JAN):1–9.

Dillingh, R., Kooreman, P., and Potters, J. J. M. (2016). Probability numeracy and health insurance purchase. *Economist (Netherlands)*, 164(1):19–39.

Dimant, E., van Kleef, G. A., and Shalvi, S. (2020). Requiem for a Nudge: Framing effects in nudging honesty. *Journal of Economic Behavior & Organization*, 172:247–266.

Dixit, V., Harrison, G. W., and Rutström, E. E. (2014). Estimating the subjective risks of driving simulator accidents. *Accident Analysis and Prevention*, 62:63–78.

Dohmen, T., Falk, A., Huffman, D., Sunde, U., Schupp, J., and Wagner, G. G. (2011). Individual risk attitudes: Measurement, determinants, and behavioral consequences. *Journal of the European Economic Association*, 9(3):522–550.

Doran, R., Böhm, G., Pfister, H. R., Steentjes, K., and Pidgeon, N. (2019). Consequence evaluations and moral concerns about climate change: insights from nationally representative surveys across four European countries. *Journal of Risk Research*, 22(5):610–626.

Doran, R. and Larsen, S. (2016). The Relative Importance of Social and Personal Norms in Explaining Intentions to Choose EcoFriendly Travel Options. *International Journal of Tourism Research*, 18(2):159–166.

Drichoutis, A. C. and Lusk, J. L. (2016). What can multiple price lists really tell us about risk preferences? *Journal of Risk and Uncertainty*, 53(2-3):89–106.

Du, N., Shelley, M. K., and Zhao, H. (2008). Group identification and decision disclosure effects.

Dur, R. A. J., Fleming, D., van Garderen, M., and van Lent, M. (2019). A social norm nudge to save more: A field experiment at a retail bank.

Działek, J., Biernacki, W., Konieczny, R., Fiedeń, ., Franczak, P., Grzeszna, K., and Listwan-Franczak, K. (2019). *Understanding Flood Preparedness*. SpringerBriefs in Geography. Springer International Publishing, Cham.

Eckel, C., Grossman, P. J., Johnson, C. A., de Oliveira, A. C., Rojas, C., and Wilson, R. (2011). Social norms of sharing in high school: Teen giving in the dictator game. *Journal of Economic Behavior and Organization*, 80(3):603–612.

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

Eijgenraam, C., Kind, J., Bak, C., Brekelmans, R., den Hertog, D., Duits, M., Roos, K., Vermeer, P., and Kuijken, W. (2014). Economically Efficient

Standards to Protect the Netherlands Against Flooding. *INFORMS Journal on Applied Analytics*, 44(1):7–21.

Einav, L., Finkelstein, A., Ryan, S. P., Schrimpf, P., and Cullen, M. R. (2013). Selection on moral hazard in health insurance. *The American Economic Review*, 103(1):178–219.

Ert, E. and Erev, I. (2018). It won't happen to me: The behavioral impact of extreme risks. In Bier, V., editor, *Risk in extreme environments: preparing, avoiding, mitigating, and managing*, chapter 7, pages 111–128. Routledge.

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.

European Commission, . (2013). Green Paper on the insurance of natural and man-made disasters. Technical report, European Commission, Strasbourg.

Falk, A., Becker, A., Dohmen, T., Enke, B., Huffman, D., and Sunde, U. (2018). Global Evidence on Economic Preferences*. *The Quarterly Journal of Economics*, 133(4):1645–1692.

Falk, A., Becker, A., Dohmen, T., Huffman, D., Uwe, S., and Sunde, U. (2016). The preference survey module: A validated instrument for measuring risk, time, and social preferences. *IZA Discussion Paper Series*, No. 9674(9674):1–66.

Falk, A. and Heckman, J. J. (2009). Lab experiments are a major source of knowledge in the social sciences. *Science*, 326(5952):535–538.

Fang, H., Keane, M. P., and Silverman, D. (2008). Sources of advantageous selection: Evidence from the medigap insurance market. *Journal of Political Economy*, 116(2):303–350.

Farrow, K., Grolleau, G., and Ibanez, L. (2017). Social Norms and Pro-environmental Behavior: A Review of the Evidence. *Ecological Economics*, 140:1–13.

Felicio, S. P. A. S., Silva, V. S. R., Dargains, A. R., Souza, P. R. A., Sampaio, F., Carvalho, P. V. R., Gomes, J. O., and Borges, M. R. S. (2014). Stop disasters game experiment with elementary school students in Rio de Janeiro: Building safety culture. In *ISCRAM 2014 Conference Proceedings - 11th International Conference on Information Systems for Crisis Response and Management*, number May 2014, pages 585–591.

Fellner, G., Sausgruber, R., and Traxler, C. (2013). Testing enforcement strategies in the field: Threat, moral appeal and social information. *Journal of the European Economic Association*, 11(3):634–660.

Feng, Z., González, V. A., Amor, R., Lovreglio, R., and Cabrera-Guerrero, G. (2018). Immersive virtual reality serious games for evacuation training and research: A systematic literature review. *Computers & Education*, 127(May):252–266.

Fiala, L. and Noussair, C. N. (2017). Charitable giving, emotions, and the default effect. *Economic Inquiry*, 55(4):1792–1812.

Field, A. (2009). *Discovering statistics using SPSS*. SAGE Publications, London, 3rd edition.

Finkelstein, A. and McGarry, K. (2006). Multiple dimensions of private information: Evidence from the long-term care insurance market. *The American Economic Review*, 96(4):938–958.

Fiore, S. M., Harrison, G. W., Hughes, C. E., and Rutström, E. E. (2009). Virtual experiments and environmental policy. *Journal of Environmental Economics and Management*, 57(1):65–86.

FitzSimons, D., Hendrickx, G., Lernout, T., Badur, S., Vorsters, A., and Van Damme, P. (2014). Incentives and barriers regarding immunization against influenza and hepatitis of health care workers. *Vaccine*, 32(38):4849–4854.

Floyd, D. L., Prentice-Dunn, S., and Rogers, R. W. (2000). A meta-analysis of research on Protection Motivation Theory. *Journal of Applied Social Psychology*, 30(2):407–429.

Fonseca, D. and Kraus, M. (2016). A comparison of head-mounted and hand-held displays for 360° videos with focus on attitude and behavior change. *AcademicMindtrek 2016 - Proceedings of the 20th International Academic Mindtrek Conference*, pages 287–296.

Fouarge, D., Kriechel, B., and Dohmen, T. (2014). Occupational sorting of school graduates: The role of economic preferences. *Journal of Economic Behavior and Organization*, 106:335–351.

Fox, J., Arena, D., and Bailenson, J. N. (2009). Virtual reality: A survival guide for the social scientist. *Journal of Media Psychology*, 21(3):95–113.

Fox, J. and Bailenson, J. N. (2009). Virtual self-modeling: The effects of vicarious reinforcement and identification on exercise behaviors. *Media Psychology*, 12(1):1–25.

Fox, J., McKnight, J., Sun, Y., Maung, D., and Crawfis, R. (2020). Using a serious game to communicate risk and minimize psychological distance regarding environmental pollution. *Telematics and Informatics*, 46(November 2019).

Franklin, M., Folke, T., and Ruggeri, K. (2019). Optimising nudges and boosts for financial decisions under uncertainty. *Palgrave Communications*, 5(1):1–13.

Frederick, S., Loewenstein, G., and O'Donoghue, T. (2002). Time discounting and time preference: A critical review. *Journal of Economic Literature*, 40(2):351–401.

Freeman, I. J., Salmon, J. L., and Coburn, J. Q. (2016). CAD integration in virtual reality design reviews for improved engineering model interaction. *ASME International Mechanical Engineering Congress and Exposition*, 11: System(50657):V011T15A006.

Füllbrunn, S. and Neugebauer, T. (2013). Limited liability, moral hazard, and risk taking: A safety net game experiment. *Economic Inquiry*, 51(2):1389–1403.

Gallagher, J. (2014). Learning about an infrequent event: Evidence from flood insurance take-up in the United States. *American Economic Journal: Applied Economics*, 6(3):206–233.

Gamberini, L., Chittaro, L., Spagnolli, A., and Carlesso, C. (2015). Psychological response to an emergency in virtual reality: Effects of victim ethnicity and emergency type on helping behavior and navigation. *Computers in Human Behavior*, 48:104–113.

Garcia, S. M., Tor, A., and Schiff, T. M. (2013). The psychology of competition: A social comparison perspective. *Perspectives on Psychological Science*, 8(6):634–650.

Gelade, W. and Guirkinger, C. (2018). The enforcement advantage of external monitoring: Lessons from an experiment with joint-liability groups in Burkina Faso. *Journal of Economic Behavior and Organization*, 151:307–325.

Gifford, R., Lacroix, K., and Chen, A. (2018). Understanding responses to climate change. In *Psychology and Climate Change*, chapter 7, pages 161–183. Elsevier.

Gillath, O., McCall, C., Shaver, P. R., and Blascovich, J. (2008). What can virtual reality teach us about prosocial tendencies in real and virtual environments? *Media Psychology*, 11(November):259–282.

Giraudet, L., Houde, S., and Maher, J. (2018). Moral hazard and the energy efficiency gap: Theory and evidence. *Journal fo the Association of Environmental and Resource Economists*, 5(4):755–790.

Glöckner, A., Hilbig, B. E., Henninger, F., and Fiedler, S. (2016). The reversed description-experience gap: Disentangling sources of presentation format effects in risky choice. *Journal of Experimental Psychology: General*, 145(4):486–508.

Gobierno de España, . (2020). Real estate appraisals. Technical report, Ministerio de Transportes, Movilidad y Agenda Urbana.

Goldstein, N. J., Cialdini, R. B., and Griskevicius, V. (2008). A Room with a Viewpoint: Using Social Norms to Motivate Environmental Conservation in Hotels. *Journal of Consumer Research*, 35(3):472–482.

Gong, M., Baron, J., and Kunreuther, H. C. (2009). Group cooperation under uncertainty. *Journal of Risk and Uncertainty*, 39(3):251–270.

Gong, M., Heal, G., Krantz, D. H., Kunreuther, H. C., and Weber, E. U. (2014). The role of subsidies in coordination games with interconnected risk. *Journal of Behavioral Decision Making*, 27(5):395–407.

Graff, F., Grund, C., and Harbring, C. (2018). Competing on the Holodeck: The effect of virtual peers and heterogeneity in dynamic tournaments. *IZA Discussion Paper Series*, DP(11919):1–36.

Grimmelikhuijsen, S. and Knies, E. (2017). Validating a scale for citizen trust in government organizations. *International Review of Administrative Sciences*, 83(3):583–601.

Grothmann, T. and Reusswig, F. (2006). People at risk of flooding: Why some residents take precautionary action while others do not. *Natural Hazards*, 38(1-2):101–120.

Guéguen, N. and De Gail, M. (2003). The effect of smiling on helping behavior: Smiling and good Samaritan behavior. *Communication Reports*, 16(2):133–140.

Guo, Y. and Li, Y. (2016). Getting ready for mega disasters: the role of past experience in changing disaster consciousness. *Disaster Prevention and Management: An International Journal*, 25(4):492–505.

Gürerk, Ö., Bönsch, A., Kittsteiner, T., and Staffeldt, A. (2019). Virtual humans as co-workers: A novel methodology to study peer effects. *Journal of Behavioral and Experimental Economics*, 78:17–29.

Gürerk, Ö. and Kasulke, A. (2018). Does virtual reality increase charitable giving? An experimental study. *SSRN Electronic Journal*, October(08):1–13.

Hadley, W., Houck, C. D., Barker, D. H., Garcia, A. M., Spitalnick, J. S., Curtis, V., Roye, S., and Brown, L. K. (2014). Eliciting affect via immersive virtual reality: A tool for adolescent risk reduction. *Journal of Pediatric Psychology*, 39(3):358–368.

Hafner, B. R., Elmes, D., Read, D., and White, M. P. (2019). Exploring the role of normative, financial and environmental information in promoting uptake of energy efficient technologies. *Journal of Environmental Psychology*, 63:26–35.

Hafner, R., Elmes, D., and Read, D. (2017). Exploring the Role of Messenger Effects and Feedback Frames in Promoting Uptake of Energy-Efficient Technologies. *Current Psychology*, pages 1–12.

Hale, J., Payne, M. E. M., Taylor, K. M., Paoletti, D., and Hamilton, A. F. D. C. (2017). The virtual maze: A behavioural tool for measuring trust. *The Quarterly Journal of Experimental Psychology*, 0(0):1–53.

Hallsworth, M., List, J. A., Metcalfe, R. D., and Vlaev, I. (2017). The behavioralist as tax collector: Using natural field experiments to enhance tax compliance. *Journal of Public Economics*, 148:14–31.

Hanley, N., Boyce, C., Czajkowski, M., Tucker, S., Noussair, C. N., and Townsend, M. (2017). Sad or happy? The effects of emotions on stated preferences for environmental goods. *Environmental and Resource Economics*, 68(4):821–846.

Harland, P., Staats, H., and Wilke, H. A. M. (1999). Explaining proenvironmental intention and behavior by personal norms and the Theory of Planned Behavior. *Journal of Applied Social Psychology*, 29(12):2505–2528.

Harries, T. (2012). The anticipated emotional consequences of adaptive behaviour - impacts on the take-up of household flood-protection measures. *Environment and Planning A*, 44(3):649–668.

Harrison, G. W., Haruvy, E., and Rutström, E. E. (2011). Remarks on virtual world and virtual reality experiments. *Southern Economic Journal*, 78(1):87–94.

Harrison, G. W., Johnson, J. M., and Rutström, E. E. (2015). Risk perceptions in the virtual wilderness. *Center for the Economics Analysis of Risk, Working Papers, available at diva2:1247501*, 2015(03):1–49.

Harrison, G. W. and List, J. A. (2004). Field experiments. *Journal of Economic Literature*, 42(December):1009–1055.

Harrison, G. W. and Rutström, E. E. (2008). Risk aversion in the laboratory. In *Research in Experimental Economics*, volume 12, pages 41–196. Emeral Group Publishing Limited.

Haruvy, E., Li, S. X., McCabe, K., and Twieg, P. (2017). Communication and visibility in public goods provision. *Games and Economic Behavior*, 105:276–296.

Hasson, R., Löfgren, ., and Visser, M. (2012). Treatment effects of climate change risk on mitigation and adaptation behaviour in an experimental setting. *South African Journal of Economics*, 80(3):415430.

Hauser, O. P., Gino, F., and Norton, M. I. (2018). Budging beliefs, nudging behaviour. *Mind & Society*, 17(1-2):15–26.

Hershfield, H. E., Goldstein, D. G., Sharpe, W. F., Fox, J., Yeykelis, L., Carstensen, L. L., and Bailenson, J. N. (2011). Increasing saving behavior through age-progressed renderings of the future self. *JMR, Journal of marketing research*, 48:S23–S37.

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

Hertwig, R., Hogarth, R. M., and Lejarraga, T. (2018). Experience and Description: Exploring Two Paths to Knowledge. *Current Directions in Psychological Science*, 27(2):123–128.

Hertwig, R. and Ortmann, A. (2001). Experimental practices in economics: a methodological challenge for psychologists? *Behavioral and Brain Sciences*, 24(3):383–403.

Hochrainer-Stigler, S., Linnerooth-Bayer, J., and Lorant, A. (2017). The European Union Solidarity Fund: an assessment of its recent reforms. *Mitigation and Adaptation Strategies for Global Change*, 22(4):547–563.

Hoffmann, A. O. I. and Broekhuizen, T. L. J. (2009). Susceptibility to and impact of interpersonal influence in an investment context. *Journal of the Academy of Marketing Science*, 37(4):488–503.

Hofstede, G. (2001). *Culture's consequences: Comparing values, behaviors, institutions and organizations across nations.* SAGE Publications.

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

Holzmeister, F. (2017). oTree: Ready-made apps for risk preference elicitation methods. *Journal of Behavioral and Experimental Finance*.

Holzmeister, F. and Pfurtscheller, A. (2016). oTree: The bomb risk elicitation task. *Journal of Behavioral and Experimental Finance*, 10:105–108.

Hong, Y., Kim, J. S., and Xiong, L. (2019). Media exposure and individuals emergency preparedness behaviors for coping with natural and human-made disasters. *Journal of Environmental Psychology*, 63:82–91.

Hooker, S. A., Wooldridge, J. S., Ross, K. M., and Masters, K. S. (2018). Do Monetary Incentives Increase Fitness Center Utilization? It Depends. *American Journal of Health Promotion*, 32(3):606–612.

Hopfensitz, A., Mantilla, C., and Miquel-Florensa, J. (2016). Conditional rewards for sustainable behavior: targeting lessons from an open access fishery.

Hoppe, E. I. and Kusterer, D. J. (2011). Conflicting tasks and moral hazard: Theory and experimental evidence. *European Economic Review*, 55(8):1094–1108.

Hoppe, E. I. and Schmitz, P. W. (2018). Hidden action and outcome contractibility: An experimental test of moral hazard theory. *Games and Economic Behavior*, 109:544–564.

Huber, J., Kip Viscusi, W., and Bell, J. (2020). Dynamic relationships between social norms and pro-environmental behavior: evidence from household recycling. *Behavioural Public Policy*, 4(1):1–25.

Huck, S., Lünser, G., Spitzer, F., and Tyran, J.-R. (2016). Medical insurance and free choice of physician shape patient overtreatment: A laboratory experiment. *Journal of Economic Behavior & Organization*, 131:78–105.

Hudson, P., Botzen, W. J. W., Czajkowski, J., and Kreibich, H. (2017). Moral hazard in natural disaster insurance markets: Empirical evidence from Germany and the United States. *Land Economics*, 93(2):179–208.

Hudson, P., Botzen, W. J. W., Feyen, L., and Aerts, J. C. J. H. (2016). Incentivising flood risk adaptation through risk based insurance premiums: Trade-offs between affordability and risk reduction. *Ecological Economics*, 125(November):1–13.

Hudson, P., Thieken, A. H., and Bubeck, P. (2019). The challenges of longitudinal surveys in the flood risk domain. *Journal of Risk Research*, 0(0):1–22.

Hui, C. H. and Triandis, H. C. (1986). Individualism-collectivism: A study of cross-cultural researchers. *Journal of Cross-Cultural Psychology*, 17(2):225–248.

Huizinga, J., Moel, H. d., and Szewczyk, W. (2017). Global flood depth-damage functions: Methodology and the database with guidelines. Technical report, Joint Research Centre (JRC).

Hummel, D. and Maedche, A. (2019). How effective is nudging? A quantitative review on the effect sizes and limits of empirical nudging studies. *Journal of Behavioral and Experimental Economics*, 80(September 2018):47–58.

Innocenti, A. (2017). Virtual reality experiments in economics. *Journal of Behavioral and Experimental Economics*, 69(49):71–77.

Institute Nacional de Estadística, . (2020). Living Conditions Survey. Technical report.
IPCC (2012). *Managing the risks of extreme events and disasters to advance climate change adaptation*. Cambridge University Press.
IPCC (2014). Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Geneva, Switzerland.
Jansen, P. C. P., Snijders, C. C. P., and Willemsen, M. C. (2020). Playing with fire. Understanding how experiencing a fire in an immersive virtual environment affects prevention behavior. *PLoS ONE*, 15(3):1–29.
Janssens, W. and Kramer, B. (2016). The social dilemma of microinsurance: Free-riding in a framed field experiment. *Journal of Economic Behavior and Organization*.
Jaspersen, J. G. (2016). Hypothetical surveys and experimental studies of insurance demand: A review. *Journal of Risk and Insurance*, 83(1):217–255.
Jongman, B., Hochrainer-Stigler, S., Feyen, L., Aerts, J. C. J. H., Mechler, R., Botzen, W. J. W., Bouwer, L. M., Pflug, G., Rojas, R., and Ward, P. J. (2014). Increasing stress on disaster-risk finance due to large floods. *Nature Climate Change*, 4(4):1–5.
Jonkman, S. N., Voortman, H. G., Klerk, W. J., and van Vuren, S. (2018). Developments in the management of flood defences and hydraulic infrastructure in the Netherlands. *Structure and Infrastructure Engineering*, 14(7):895–910.
Kahneman, D. (2003). A perspective on judgment and choice: mapping bounded rationality. *The American psychologist*, 58(9):697–720.
Kahneman, D. and Tversky, A. (1979). Prospect Theory: An analysis of decision under risk. *Econometrica*, 47(2):263–292.
Kallgren, C. A., Reno, R. R., and Cialdini, R. B. (2000). A focus theory of normative conduct: When norms do and do not affect behavior. *Personality and Social Psychology Bulletin*, 26(8):1002–1012.
Karlan, D. and Zinman, J. (2009). Observing unobservables: Identifying information asymmetries with a consumer credit field experiment. *Econometrica*, 77(6):1993–2008.
Kellens, W., Terpstra, T., and De Maeyer, P. (2013). Perception and communication of flood risks: A systematic review of empirical research. *Risk Analysis*, 33(1):24–49.
Keller, C., Siegrist, M., and Gutscher, H. (2006). The Role of the Affect and Availability Heuristics in Risk Communication. *Risk Analysis*, 26(3):631–639.
Kelly, M. and Kleffner, A. E. (2003). Optimal loss mitigation and contract design. *The Journal of Risk and Insurance*, 70(1):53–72.
Kennedy, R. S., Lane, N. E., Berbaum, K. S., and Lilienthal, M. G. (1993). Simulator Sickness Questionnaire: An enhanced method for quantifying simulator sickness.

Keshavarz, B. and Hecht, H. (2011). Validating an efficient method to quantify motion sickness. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 53(4):415–426.

Kinateder, M., Gromer, D., Gast, P., Buld, S., Müller, M., Jost, M., Nehfischer, M., Mühlberger, A., and Pauli, P. (2015). The effect of dangerous goods transporters on hazard perception and evacuation behavior A virtual reality experiment on tunnel emergencies. *Fire Safety Journal*, 78:24–30.

Kinateder, M., Müller, M., Jost, M., Mühlberger, A., and Pauli, P. (2014). Social influence in a virtual tunnel fire - Influence of conflicting information on evacuation behavior. *Applied Ergonomics*, 45(6):1649–1659.

King, A. J. (2014). Visual Messaging and Risk Communication. In *The SAGE Handbook of Risk Communication*, pages 193–206. SAGE Publications, Inc., 2455 Teller Road, Thousand Oaks California 91320.

Kleindorfer, P. R., Kunreuther, H. C., and Ou-Yang, C. (2012). Single-year and multi-year insurance policies in a competitive market. *Journal of Risk and Uncertainty*, 45(1):51–78.

Köbis, N. C., Troost, M., Brandt, C. O., and Soraperra, I. (2019). Social norms of corruption in the field: social nudges on posters can help to reduce bribery. *Behavioural Public Policy*, pages 1–28.

Koerth, J., Vafeidis, A. T., and Hinkel, J. (2017). Household-level coastal adaptation and its drivers: A systematic case study review. *Risk Analysis*, 37(4):629–646.

Koetse, M. J. and Brouwer, R. (2016). Reference dependence effects on WTA and WTP value functions and their disparity. *Environmental and Resource Economics*, 65(4):723–745.

Kok, M. and Doef, M. v. d. (2008). Leidraad overstromingsberekeningen voor VNK2. Technical Report september, VNK.

Krantz, D. H. and Kunreuther, H. C. (2007). Goals and plans in decision making. *Judgment and Decision Making*, 2(3):137–168.

Kreibich, H., Bubeck, P., Van Vliet, M., and De Moel, H. (2015). A review of damage-reducing measures to manage fluvial flood risks in a changing climate. *Mitigation and Adaptation Strategies for Global Change*, 20(6):967–989.

Kreibich, H., Seifert, I., Thieken, A. H., Lindquist, E., Wagner, K., and Merz, B. (2011). Recent changes in flood preparedness of private households and businesses in Germany. *Regional Environmental Change*, 11(1):59–71.

Krieger, M. and Felder, S. (2013). Can decision biases improve insurance outcomes? An experiment on status quo bias in health insurance choice. *International Journal of Environmental Research and Public Health*, 10(6):2560–2577.

Krupka, E. L. and Weber, R. A. (2009). The focusing and informational effects of norms on pro-social behavior. *Journal of Economic Psychology*, 30(3):307–320.

Kugler, T., Ye, B., Motro, D. J., and Noussair, C. N. (2018). On trust and disgust: Evidence from face reading and virtual reality. Mimeo.

Kunreuther, H. C. (1996). Mitigating disaster losses through insurance. *Journal of Risk and Uncertainty*, 12:171–187.

Kunreuther, H. C. (2015). The role of insurance in reducing losses from extreme events: The need for public-private partnerships. *Geneva Papers on Risk and Insurance: Issues and Practice*, 40(4):741–762.

Kunreuther, H. C. (2018). Improving the National Flood Insurance Program. *Behavioural Public Policy*, pages 1–15.

Kunreuther, H. C., Ginsberg, R., Miller, L., Sagi, P., Slovic, P., Borkan, B., and Katz, N. (1978). *Disaster insurance protection: Public policy lessons*. Wiley, New York.

Kunreuther, H. C. and Michel-Kerjan, E. O. (2015). Demand for fixed-price multi-year contracts: Experimental evidence from insurance decisions. *Journal of Risk and Uncertainty*, 51(2):171–194.

Kunreuther, H. C. and Pauly, M. V. (2004). Neglecting disaster: Why don't people insure against large losses? *Journal of Risk and Uncertainty*, 28(1):5–21.

Kunreuther, H. C. and Pauly, M. V. (2018). Dynamic insurance decision-making for rare events: The role of emotions. *Geneva Papers on Risk and Insurance: Issues and Practice*, 43(2):335–355.

Kusev, P., van Schaik, P., Ayton, P., Dent, J., and Chater, N. (2009). Exaggerated risk: Prospect Theory and probability weighting in risky choice. *Journal of Experimental Psychology: Learning Memory and Cognition*, 35(6):1487–1505.

Lades, L. K., Laffan, K., and Weber, T. O. (2020). Do economic preferences predict pro-environmental behaviour? *Ecological Economics*, 183(May 2020):106977.

Langhinrichsen-Rohling, J., Rohde, P., Seeley, J. R., and Rohling, M. L. (2004). Individual, family, and peer correlates of adolescent gambling. *Journal of Gambling Studies*, 20(1):23–46.

Latoschik, M. E., Roth, D., Gall, D., Achenbach, J., Waltemate, T., and Botsch, M. (2017). The effect of avatar realism in immersive social virtual realities. *Proceedings of the 23rd ACM Symposium on Virtual Reality Software and Technology - VRST '17 - retrieved from https://dl.acm.org/citation.cfm?id=3139156*, pages 1–10.

Latu, I. M., Mast, M. S., Lammers, J., and Bombari, D. (2013). Successful female leaders empower women's behavior in leadership tasks. *Journal of Experimental Social Psychology*, 49(3):444–448.

Laury, S. K., McInnes, M. M., and Todd Swarthout, J. (2009). Insurance decisions for low-probability losses. *Journal of Risk and Uncertainty*, 39(1):17–44.

Lechowska, E. (2018). What determines flood risk perception? A review of factors of flood risk perception and relations between its basic elements. *Natural Hazards*, 94(3):1341–1366.

Lee, T. J., Cameron, L. D., Wünsche, B., and Stevens, C. (2011). A randomized trial of computer-based communications using imagery and text information to alter representations of heart disease risk and motivate protective behaviour. *British Journal of Health Psychology*, 16(1):72–91.

Lendering, K. T., Jonkman, S. N., and Kok, M. (2016). Effectiveness of emergency measures for flood prevention. *Journal of Flood Risk Management*, 9(4):320–334.

Lerner, J. S., Small, D. A., and Loewenstein, G. (2004). Carryover effects of emotions on economic decisions. *Psychological Science*, 15(5):337–341.

Levitt, S. D. and List, J. A. (2007). What do laboratory experiments measuring social preferences reveal about the real world? *Journal of Economic Perspective*, 21(2):153–174.

Li, C., Liang, W., Quigley, C., Zhao, Y., and Yu, L.-F. (2017). Earthquake Safety Training through Virtual Drills. *IEEE Transactions on Visualization and Computer Graphics*, 23(4):1275–1284.

Lin, C.-H., Chuang, S.-C., Kao, D. T., and Kung, C. Y. (2006). The role of emotions in the endowment effect. *Journal of Economic Psychology*, 27(4):589–597.

Lindell, M. K. and Hwang, S. N. (2008). Households' perceived personal risk and responses in a multihazard environment. *Risk Analysis*, 28(2):539–556.

Lindner, B. L., Johnson, J., and Alsheimer, F. (2018). Increasing Risk Perception and Understanding of Hurricane Storm Tides Using an Interactive, Web-Based Visualization Approach. *Journal of Coastal Research*, 34(6):1484.

Liu, J., Thomas, J. M., and Higgs, S. (2019). The relationship between social identity, descriptive social norms and eating intentions and behaviors. *Journal of Experimental Social Psychology*, 82:217–230.

Lo, A. Y. (2013). The role of social norms in climate adaptation: Mediating risk perception and flood insurance purchase. *Global Environmental Change*, 23(5):1249–1257.

Loewenstein, G. and Prelec, D. (1992). Anomalies in intertemporal choice: Evidence and an interpretation. *The Quarterly Journal of Economics*, 107(2):573–597.

Loewenstein, G. F., Hsee, C. K., Weber, E. U., and Welch, N. (2001). Risk as feelings. *Psychological Bulletin*, 127(2):267–286.

Lovreglio, R., Gonzalez, V., Amor, R., Spearpoint, M., Thomas, J., Trotter, M., and Sacks, R. (2017). The need for enhancing earthquake evacuee safety by using virtual reality serious games. *Lean and Computing in Construction Congress - Volume 1: Proceedings of the Joint Conference on Computing in Construction*, 1(July):381–389.

Macera, R. (2018). Present or future incentives? On the optimality of fixed wages with moral hazard. *Journal of Economic Behavior and Organization*, 147(11140101):129–144.

Mackay, M., Yamazaki, S., Jennings, S., Sibly, H., van Putten, I. E., and Emery, T. J. (2020). The influence of nudges on compliance behaviour in recreational fisheries: a laboratory experiment. *ICES Journal of Marine Science*, 77(6):2319–2332.

Madary, M. and Metzinger, T. K. (2016). Real Virtuality: A code of ethical conduct. Recommendations for good scientific practice and the consumers of VR-technology. *Frontiers in Robotics and AI*, 3(February):1–23.

Maidl, E. and Buchecker, M. (2015). Raising risk preparedness by flood risk communication. *Natural Hazards and Earth System Sciences*, 15(7):1577–1595.

Markowitz, D. M., Laha, R., Perone, B. P., Pea, R. D., and Bailenson, J. N. (2018). Immersive Virtual Reality field trips facilitate learning about climate change. *Frontiers in Psychology*, 9(NOV).

Martinez, L. M. F., Zeelenberg, M., and Rijsman, J. B. (2011). Behavioural consequences of regret and disappointment in social bargaining games. *Cognition and Emotion*, 25(2):351–359.

Matthews, Y., Scarpa, R., and Marsh, D. (2017). Using virtual environments to improve the realism of choice experiments: A case study about coastal erosion management. *Journal of Environmental Economics and Management*, 81:193–208.

McCall, C. and Singer, T. (2015). Facing off with unfair others: Introducing proxemic imaging as an implicit measure of approach and avoidance during social interaction. *PLoS ONE*, 10(2):e0117532.

McClelland, G. H., Schulze, W. D., and Coursey, D. L. (1993). Insurance for low-probability hazards: A bimodal response to unlikely events. *Journal of Risk and Uncertainty*, 7:95–116.

McDowell, M., Galesic, M., and Gigerenzer, G. (2018). Natural Frequencies Do Foster Public Understanding of Medical Tests: Comment on Pighin, Gonzalez, Savadori, and Girotto (2016). *Medical Decision Making*, 38(3):390–399.

McGee, T. K., McFarlane, B. L., and Varghese, J. (2009). An examination of the influence of hazard experience on wildfire risk perceptions and adoption of mitigation measures. *Society and Natural Resources*, 22(4):308–323.

McKee, M., Berrens, R. P., Jones, M. L., Helton, R., and Talberth, J. (2004). Using experimental economics to examine wildfire insurance and averting decisions in the WildlandUrban Interface. *Society & Natural Resources*, 17(6):491–507.

McKee, M., Santore, R., and Shelton, J. (2007). Contingent fees, moral hazard, and attorney rents: A laboratory experiment. *The Journal of Legal Studies*, 36(2):253–273.

McNaughton, C. D., Cavanaugh, K. L., Kripalani, S., Rothman, R. L., and Wallston, K. A. (2015). Validation of a short, 3-item version of the subjective numeracy scale. *Medical Decision Making*, 35(8):932–936.

Meißner, M., Pfeiffer, J., Pfeiffer, T., and Oppewal, H. (2017). Combining virtual reality and mobile eye tracking to provide a naturalistic experimental environment for shopper research. *Journal of Business Research*, In Press(Corrected Proof).

Menkhoff, L. and Sakha, S. (2017). Estimating risky behavior with multiple-item risk measures. *Journal of Economic Psychology*, 59:59–86.

Merz, B., Kreibich, H., Schwarze, R., and Thieken, A. H. (2010). Assessment of economic flood damage. *Natural Hazards and Earth System Science*, 10(8):1697–1724.

Meyer, R. J., Broad, K., Orlove, B., and Petrovic, N. (2013). Dynamic simulation as an approach to understanding hurricane risk response: Insights from the stormview lab. *Risk Analysis*, 33(8):1532–1552.

Meyer, R. J. and Kunreuther, H. C. (2017). *The ostrich paradox: Why we underprepare for disasters.* Wharton School Press.

Miceli, R., Sotgiu, I., and Settanni, M. (2008). Disaster preparedness and perception of flood risk: A study in an alpine valley in Italy. *Journal of Environmental Psychology*, 28(2):164–173.

Michel-Kerjan, E. O. (2010). Catastrophe economies: The National Flood Insurance Program. *The Journal of Economic Perspectives*, 24(4):165–186.

Michel-Kerjan, E. O. and Kunreuther, H. C. (2011). Redesigning flood insurance. *Science*, 333:408–409.

Midden, C. and Ham, J. (2018). Persuasive Technology to Promote Pro-Environmental Behaviour. In *Environmental Psychology*, pages 283–294. John Wiley & Sons, Ltd, Chichester, UK.

Mills, M., Mutafoglu, K., Adams, V. M., Archibald, C., Bell, J., and Leon, J. X. (2016). Perceived and projected flood risk and adaptation in coastal Southeast Queensland, Australia. *Climatic Change*, 136(3-4):523–537.

Mol, J. M. (2019). Goggles in the lab: Economic experiments in immersive virtual environments. *Journal of Behavioral and Experimental Economics*, 79(C):155–164.

Mol, J. M., Botzen, W. J. W., and Blasch, J. E. (2020a). Risk reduction in compulsory disaster insurance: Experimental evidence on moral hazard and financial incentives. *Journal of Behavioral and Experimental Economics*, 84(February):101500.

Mol, J. M., Botzen, W. W., and Blasch, J. E. (2020b). Behavioral motivations for self-insurance under different disaster risk insurance schemes. *Journal of Economic Behavior & Organization*, 180:967–991.

Mol, J. M., van der Heijden, E. C. M., and Potters, J. J. M. (2020c). (Not) alone in the world: Cheating in the presence of a virtual observer. *Experimental Economics*, 23(4):961–978.

Munafo, J., Diedrick, M., and Stoffregen, T. A. (2017). The virtual reality head-mounted display Oculus Rift induces motion sickness and is sexist in its effects. *Experimental Brain Research*, 235(3):889–901.

Munich RE, . (2018). Hurricanes cause record losses in 2017 - The year in figures.

Murray, E. G., Neumann, D. L., Moffitt, R. L., and Thomas, P. R. (2015). The effects of the presence of others during a rowing exercise in a virtual reality environment. *Psychology of Sport and Exercise*, 22:328–336.

Mutz, D. C. and Pemantle, R. (2015). Standards for experimental research: Encouraging a better understanding of experimental methods. *Journal of Experimental Political Science*, 2(2):192–215.

Natapov, A. and Fisher-Gewirtzman, D. (2016). Visibility of urban activities and pedestrian routes: An experiment in a virtual environment. *Computers, Environment and Urban Systems*, 58:60–70.

Navarrete, C. D., Mcdonald, M. M., Mott, M. L., and Asher, B. (2012). Virtual morality: Emotion and action in a simulated three-dimensional trolley problem. *Emotion*, 12(2):364–370.

Netherlands Statistics, . (2018a). Inkomensverdeling van huishoudens; nationale rekeningen. Technical report, CBS.

Netherlands Statistics, . (2018b). Waarde onroerende zaken van woningen en niet-woningen. Technical report, CBS.

Netherlands Statistics, . (2020a). Existing own homes. Technical report.

Netherlands Statistics, . (2020b). Inkomen van huishoudens. Technical report.

Neuß, S., Peter, R., and Knoller, C. (2016). Financial support, moral hazard, and other-regarding preferences.

Nieken, P. and Schmitz, P. W. (2012). Repeated moral hazard and contracts with memory: A laboratory experiment. *Games and Economic Behavior*, 75:1000–1008.

Nowak, G. J., Evans, N. J., Wojdynski, B. W., Ahn, S. J. G., Len-Rios, M. E., Carera, K., Hale, S., and McFalls, D. (2020). Using immersive virtual reality to improve the beliefs and intentions of influenza vaccine avoidant 18-to-49-year-olds: Considerations, effects, and lessons learned. *Vaccine*, 38(5):1225–1233.

Nyborg, K., Anderies, J. M., Dannenberg, A., Lindahl, T., Schill, C., Schlüter, M., Adger, W. N., Arrow, K. J., Barrett, S., Carpenter, S., Chapin, F. S. I., Crépin, A.-S., Daily, G., Ehrlich, P., Carl Folke, W., and Zeeuw, A. d. (2016). Social norms as solutions: Policies may influence large-scale behavioral tipping. *Science*, 354(6308):42 – 43.

Oh, S. H. D. (2013). Do collectivists conform more than individualists? Cross-cultural differences in compliance and internalization. *Social Behavior and Personality*, 41(6):981–994.

Oliver, A. and Ubel, P. (2014). Nudging the obese: A UK-US consideration. *Health Economics, Policy and Law*, 9(3):329–342.

O'Neill, E., Brereton, F., Shahumyan, H., and Clinch, J. P. (2016). The impact of perceived flood exposure on flood-risk perception: The role of distance. *Risk Analysis*, 36(11):2158–2186.

Ooi, S., Tanimoto, T., and Sano, M. (2019). Virtual reality fire disaster training system for improving disaster awareness. *ACM International Conference Proceeding Series*, Part F1481:301–307.

Orth, U. R. and Kahle, L. R. (2008). Intrapersonal variation in consumer susceptibility to normative influence: Toward a better understanding of brand choice decisions. *Journal of Social Psychology*, 148(4):423–448.

Ortmann, A. and Hertwig, R. (2002). The costs of deception: Evidence from psychology. *Experimental Economics*, 5(2):111–131.

Osberghaus, D. (2015). The determinants of private flood mitigation measures in Germany - Evidence from a nationwide survey. *Ecological Economics*, 110:36–50.

Osberghaus, D. (2017). The effect of flood experience on household mitigationEvidence from longitudinal and insurance data. *Global Environmental Change*, 43:126–136.

Osman, M., Fenton, N., Pilditch, T., Lagnado, D., and Neil, M. (2018). Whom do we trust on social policy interventions? *Basic and Applied Social Psychology*, 40(5):249–268.

Ostrom, E. (2010). Beyond markets and states: polycentric governance of complex economic systems. *American Economic Review*, 100(June):641–672.

Oxford Dictionaries, . (2018). Virtual Reality. *Accessed 01-09-2018 from https://www.oxforddictionaries.com/definition/english/virtual-reality.*

Ozdemir, O. (2017). Impact of Outcome Ambiguity on Self-Insurance and Self-Protection: Experimental Evidence. *Bulletin of Economic Research*, 69(4):E112–E125.

Page, L., Savage, D. A., and Torgler, B. (2014). Variation in risk seeking behaviour following large losses: A natural experiment. *European Economic Review*, 71:121–131.

Papon, T. (2008). The effect of pre-commitment and past-experience on insurance choices: An experimental study. *The Geneva Risk and Insurance Review*, 33(1):47–73.

Parboteeah, K. P., Addae, H. M., and Cullen, J. B. (2012). Propensity to Support Sustainability Initiatives: A Cross-National Model. *Journal of Business Ethics*, 105(3):403–413.

Parsons, T. D. (2015). Virtual reality for enhanced ecological validity and experimental control in the clinical, affective and social neurosciences. *Frontiers in Human Neuroscience*, 9(December):1–19.

Patel, M. S., Asch, D. A., Troxel, A. B., Fletcher, M., Osman-Koss, R., Brady, J., Wesby, L., Hilbert, V., Zhu, J., Wang, W., and Volpp, K. G. (2016). Premium-based financial incentives did not promote workplace weight loss in a 201315 study. *Health Affairs*, 35(1):71–79.

Patterson, Z., Darbani, J. M., Rezaei, A., Zacharias, J., and Yazdizadeh, A. (2017). Comparing text-only and virtual reality discrete choice experiments of neighbourhood choice. *Landscape and Urban Planning*, 157:63–74.

Paudel, Y., Botzen, W. J. W., and Aerts, J. C. J. H. (2012). A comparative study of public-private catastrophe insurance systems: Lessons from current practices. *The Geneva Papers on Risk and Insurance Issues and Practice*, 37(2):603–603.

Peacock, W. G., Brody, S. D., and Highfield, W. (2005). Hurricane risk perceptions among Florida's single family homeowners. *Landscape and Urban Planning*, 73(2-3):120–135.

Peck, T. C., Seinfeld, S., Aglioti, S. M., and Slater, M. (2013). Putting yourself in the skin of a black avatar reduces implicit racial bias. *Consciousness and Cognition*, 22(3):779–87.

Pedroni, A. F., Frey, R., Bruhin, A., Dutilh, G., Hertwig, R., and Rieskamp, J. (2017). The risk elicitation puzzle. *Nature Human Behaviour*.

Peperkorn, H. M., Diemer, J., and Mühlberger, A. (2015). Temporal dynamics in the relation between presence and fear in virtual reality. *Computers in Human Behavior*, 48:542–547.

Pérez Fernández, R. and Alonso, V. (2015). Virtual reality in a shipbuilding environment. *Advances in Engineering Software*, 81:30–40.

Perreault, G. (2017). The power of digital games in disaster preparation and post-disaster resilience The power of digital games in disaster 2. In *2017 International Communication Association Conference*, number May, pages 1–31.

Peters, E., Slovic, P., Hibbard, J. H., and Tusler, M. (2006). Why worry? Worry, risk perceptions, and willingness to act to reduce medical errors. *Health Psychology*, 25(2):144–152.

Petrolia, D. R., Hwang, J., Landry, C. E., and Coble, K. H. (2015). Wind insurance and mitigation in the coastal zone. *Land Economics*, 91(2):272–295.

Pineda, A., Hernández-Santaolalla, V., and del Mar Rubio-Hernández, M. (2015). Individualism in Western advertising: A comparative study of Spanish and US newspaper advertisements. *European Journal of Communication*, 30(4):437–453.

Portman, M. E., Natapov, A., and Fisher-Gewirtzman, D. (2015). To go where no man has gone before: Virtual reality in architecture, landscape architecture and environmental planning. *Computers, Environment and Urban Systems*, pages 1–9.

Poussin, J. K., Botzen, W. W., and Aerts, J. C. (2014). Factors of influence on flood damage mitigation behaviour by households. *Environmental Science & Policy*, 40:69–77.

Poussin, J. K., Wouter Botzen, W. J., and Aerts, J. C. J. H. (2015). Effectiveness of flood damage mitigation measures: Empirical evidence from French flood disasters. *Global Environmental Change*, 31:74–84.

Price, M. K. (2014). Using field experiments to address environmental externalities and resource scarcity: Major lessons learned and new directions for future research. *Oxford Review of Economic Policy*, 30(4):621–638.

Prinstein, M. J., Brechwald, W. A., and Cohen, G. L. (2011). Susceptibility to peer influence: Using a performance-based measure to identify adolescent males at heightened risk for deviant peer socialization. *Developmental Psychology*, 47(4):1167–1172.

Puschmann, P., Horlitz, T., Wittstock, V., and Schütz, A. (2016). Risk analysis (assessment) using virtual reality technology - Effects of subjective experience: An experimental study. *Procedia CIRP*, 50:490–495.

Qu, C., Ling, Y., Heynderickx, I., and Brinkman, W.-P. (2015). Virtual bystanders in a language lesson: examining the effect of social evaluation, vicarious experience, cognitive consistency and praising on students' beliefs, self-efficacy and anxiety in a virtual reality environment. *PLoS ONE*, 10(4):e0125279.

Rebelo, F., Noriega, P., Duarte, E., and Soares, M. (2012). Using virtual reality to assess user experience. *Human Factors*, 54(6):964–982.

Reimer, T., Jones, C., and Skubisz, C. (2015). Numeric Communication of Risk. In *The SAGE Handbook of Risk Communication*, pages 166–179. SAGE Publications, Inc., 2455 Teller Road, Thousand Oaks California 91320.

Reynaud, A., Aubert, C., and Nguyen, M. H. (2013). Living with floods: Protective behaviours and risk perception of Vietnamese households. *Geneva Papers on Risk and Insurance: Issues and Practice*, 38(3):547–579.

Richert, C., Erdlenbruch, K., and Figuières, C. (2017). The determinants of households flood mitigation decisions in France - on the possibility of feedback effects from past investments. *Ecological Economics*, 131:342–352.

Richter, I., Thøgersen, J., and Klöckner, C. (2018). A social norms intervention going wrong: Boomerang effects from descriptive norms information. *Sustainability*, 10(8):2848.

Ripberger, J. T., Jenkins-Smith, H. C., Silva, C. L., Czajkowski, J., Kunreuther, H. C., and Simmons, K. M. (2018). Tornado damage mitigation: Homeowner support for enhanced building codes in Oklahoma. *Risk Analysis*, 38(11):2300–2317.

Ritterfeld, U., Cody, M., and Vorderer, P. (2009). *Serious Games: Mechanisms and effects*. Routledge.

Robinson, P. J. and Botzen, W. (2020). Flood insurance demand and probability weighting: The influences of regret, worry, locus of control and the threshold of concern heuristic. *Water Resources and Economics*, 30:100144.

Robinson, P. J. and Botzen, W. J. W. (2018). The impact of regret and worry on the threshold level of concern for flood insurance demand: Evidence from Dutch homeowners. *Judgment and Decision Making*, 13(2):237–245.

Rogers, R. W. (1975). A protection motivation theory of fear appeals and attitude change. *The Journal of Psychology*, 91(1):93–114.

Rosenberg, R. S., Baughman, S. L., and Bailenson, J. N. (2013). Virtual superheroes: Using superpowers in virtual reality to encourage prosocial behavior. *PLoS ONE*, 8(1):1–9.

Rowell, D. and Connelly, L. B. (2012). A history of the term "moral hazard". *Journal of Risk and Insurance*, 79(4):1051–1075.

Royal, A. and Walls, M. (2019). Flood risk perceptions and insurance choice: Do decisions in the floodplain reflect overoptimism? *Risk Analysis*, 39(5):1088–1104.

Rud, O. A., Rabanal, J. P., and Horowitz, J. (2018). Does competition aggravate moral hazard? A Multi-Principal-Agent experiment. *Journal of Financial Intermediation*, 33:115–121.

Ruin, I., Gaillard, J. C., and Lutoff, C. (2007). How to get there? Assessing motorists' flash flood risk perception on daily itineraries. *Environmental Hazards*, 7(3):235–244.

Sanchez-Vives, M. V. and Slater, M. (2005). From presence to consciousness through virtual reality. *Nature reviews. Neuroscience*, 6(4):332–339.

Sattler, D. N., Kaiser, C. F., and Hittner, J. B. (2000). Disaster preparedness: Relationships among prior experience, personal characteristics, and distress. *Journal of Applied Social Psychology*, 30(7):1396–1420.

Savage, L. J. (1954). *The Foundations of Statistics*. Wiley, New York.

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

Schram, A. (2005). Artificiality: The tension between internal and external validity in economic experiments. *Journal of Economic Methodology*, 12(2):225–237.

Schubert, T. A., Friedmann, F., and Regenbrecht, H. (2001). The experience of presence: factor analytic insights. *Presence*, 10(3):266–281.

Schutte, N. S. and Stilinović, E. J. (2017). Facilitating empathy through virtual reality. *Motivation and Emotion*, 41(6):1–5.

Schwartz, S. H. (1977). Normative influences on altruism. *Advances in Experimental Social Psychology*, 10(C):221–279.

Schwartz, S. H. (2012). An Overview of the Schwartz Theory of Basic Values. *Online Readings in Psychology and Culture*, 2(1):1–20.

Schwarze, R., Schwindt, M., Weck-Hannemann, H., Raschky, P. A., Zahn, F., and Wagner, G. G. (2011). Natural hazard insurance in Europe: Tailored responses to climate change are needed. *Environmental Policy and Governance*, 21(1):14–30.

Scussolini, P., Aerts, J. C. J. H., Jongman, B., Bouwer, L. M., Winsemius, H. C., De Moel, H., and Ward, P. J. (2016). FLOPROS: an evolving global database of flood protection standards. *Natural Hazards and Earth System Sciences*, 16(5):1049–1061.

Serge, S. R. and Fragomeni, G. (2017). Assessing the relationship between type of head movement and simulator sickness using an immersive virtual reality head mounted display: A pilot study. In *Virtual, Augmented and Mixed Reality.VAMR 2017. Lecture Notes in Computer Science*, volume 10280, pages 556–566.

Shafran, A. P. (2011). Self-protection against repeated low probability risks. *Journal of Risk and Uncertainty*, 42(3):263–285.

Shao, W., Xian, S., Lin, N., Kunreuther, H., Jackson, N., and Goidel, K. (2017). Understanding the effects of past flood events and perceived and estimated flood risks on individuals voluntary flood insurance purchase behavior. *Water Research*, 108:391–400.

Sharples, S., Cobb, S., Moody, A., and Wilson, J. R. (2007). Virtual reality induced symptoms and effects (VRISE): Comparison of head mounted display (HMD), desktop and projection display systems. *Displays*, 29:58–69.

Shogren, J. F. (1990). The impact of self-protection and self-insurance on individual response to risk. *Journal of Risk and Uncertainty*, 3(2):191–204.

Siegrist, M. (2012). The necessity for longitudinal studies in risk perception research. *Risk Analysis*, 33(1):1–2.

Siegrist, M. and Gutscher, H. (2006). Flooding risks: A comparison of lay people's perceptions and expert's assessments in Switzerland. *Risk Analysis*, 26(4):971–979.

Siegrist, M. and Gutscher, H. (2008). Natural hazards and motivation for mitigation behavior: People cannot predict the affect evoked by a severe flood. *Risk Analysis*, 28(3):771–778.

Simon, H. A. (1959). Theories of decision-making and behavioral science. *The American Economic Review*, 49(3):253–283.

Sitter, B., Yang, J., Thielen, J., Naismith, N., and Lonergan, J. (2017). 78-3: Screen door effect reduction with diffractive film for virtual reality and augmented reality displays. *SID Symposium Digest of Technical Papers*, 48:1150–1153.

Sjöberg, L. (2000). Factors in risk perception. *Risk Analysis*, 20(1):1–12.

Sjöberg, L. (2007). Emotions and risk perception. *Risk Management*, 9(4):223–237.

Slater, M. (2004). How colorful was your day? Why questionnaires cannot assess presence in virtual environments. *Presence: Teleoperators and Virtual Environments*, 13(4):484–493.

Slater, M. (2009). Place illusion and plausibility can lead to realistic behaviour in immersive virtual environments. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 364(1535):3549–3557.

Slater, M., Antley, A., Davison, A., Swapp, D., Guger, C., Barker, C., Pistrang, N., and Sanchez-Vives, M. V. (2006). A virtual reprise of the Stanley Milgram obedience experiments. *PLoS ONE*, 1(1):e39.

Slater, M., Rovira, A., Southern, R., Swapp, D., Zhang, J. J., Campbell, C., and Levine, M. (2013). Bystander responses to a violent incident in an immersive virtual environment. *PLoS ONE*, 8(1):1–13.

Slater, M. and Sanchez-Vives, M. V. (2016). Enhancing our lives with immersive virtual reality. *Frontiers in Robotics and AI*, 3(December):1–47.

Slovic, P. (2010). *The feeling of risk: New perspectives on risk perception.* Routledge.

Slovic, P., Finucane, M. L., Peters, E., and MacGregor, D. G. (2004). Risk as analysis and risk as feelings. *Risk Analysis*, 24(2):311–322.

Slovic, P., Fischhoff, B., Lichtenstein, S., Corrigan, B., and Combs, B. (1977). Preference for insuring against probable small losses: Insurance implications. *The Journal of Risk and Insurance*, 44(2):237–258.

Slovic, P., Flynn, J. H., and Layman, M. (1991). Perceived risk, trust, and the politics of nuclear waste. *Science*, 254(5038):1603–1607.

Smerecnik, C. M., Mesters, I., Kessels, L. T., Ruiter, R. A., De Vries, N. K., and De Vries, H. (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.

Spence, A., Poortinga, W., Butler, C., and Pidgeon, N. F. (2011). Perceptions of climate change and willingness to save energy related to flood experience. *Nature Climate Change*, 1(1):46–49.

Stanney, K. M., Hale, K. S., Nahmens, I., and Kennedy, R. S. (2003). What to expect from immersive virtual environment exposure: influences of gender, body mass index, and past experience. *Human Factors*, 45(3):504–520.

Stevenson, M., Harris, A., Mortimer, D., Wijnands, J. S., Tapp, A., Peppard, F., and Buckis, S. (2018). The effects of feedback and incentive-based insurance on driving behaviours: Study approach and protocols. *Injury Prevention*, 24(1):89–93.

Stiglitz, J. E. (1974). Theories of discrimination and economic policy. In Fustenberg, G. v., editor, *Patterns of racial discrimination*, pages 5–26. D.C. Heath and Company (Lexington Books).

Stöckli, S. and Hofer, D. (2020). Susceptibility to social influence predicts behavior on Facebook. *PLoS ONE*, 15(3):1–20.

Stoffel, S. T., Goodwin, M., Sieverding, M., Vlaev, I., and von Wagner, C. (2019). Testing verbal quantifiers for social norms messages in cancer screening: evidence from an online experiment. *BMC Public Health*, 19(1):658.

Sunstein, C. R. (2017). Nudges that fail. *Behavioural Public Policy*, 1(1):4–25.

Sunstein, C. R. and Reisch, L. A. (2013). Green by default. *Kyklos*, 66(3):398–402.

Surminski, S., Aerts, J. C. J. H., Botzen, W. J. W., Hudson, P., Mysiak, J., and Pérez-Blanco, C. D. (2015). Reflections on the current debate on how to link flood insurance and disaster risk reduction in the European Union. *Natural Hazards*, 79(3):1451–1479.

Tajfel, H. (1982). Social psychology of intergroup relations. *Annual Review of Psychology*, 33(1):1–39.

Talberth, J., Berrens, R. P., McKee, M., and Jones, M. L. (2006). Averting and insurance decisions in the Wildland-Urban interface: Implications of survey and experimental data for wildfire risk reduction policy. *Contemporary Economic Policy*, 24(2):203–223.

Tambor, M., Pavlova, M., Golinowska, S., Arsenijevic, J., and Groot, W. (2016). Financial incentives for a healthy life style and disease prevention among older people: a systematic literature review. *BMC Health Services Research*.

Terpstra, T. (2011). Emotions, trust, and perceived risk: Affective and cognitive routes to flood preparedness behavior. *Risk Analysis*, 31(10):1658–1675.

Thaler, R. H. and Johnson, E. J. (1990). Gambling with the house money and trying to break even: The effects of prior outcomes on risky choice. *Management Science*, 36(6):643–660.

Thaler, R. H. and Sunstein, C. R. (2008). *Nudge: Improving decisions about health, wealth, and happiness*, volume 47. Penguin Group, New York.

Thieken, A. H., Petrow, T., Kreibich, H., and Merz, B. (2006). Insurability and mitigation of flood losses in private households in Germany. *Risk Analysis*, 26(2):383–395.

Thistlethwaite, J., Henstra, D., Brown, C., and Scott, D. (2018). How flood experience and risk perception influences protective actions and behaviours among Canadian homeowners. *Environmental Management*, 61(2):197–208.

Tobin, G. A. (1995). The levee love affair: A stormy relationship? *JAWRA Journal of the American Water Resources Association*, 31(3):359–367.

Toppenberg, H. L., Bos, A. E. R., Ruiter, R. A. C., Wigboldus, D. H. J., and Pryor, J. B. (2015). HIV-related stigma in social interactions: Approach and avoidance behaviour in a virtual environment. *European Journal of Social Psychology*, 45:169–179.

Triandis, H. C. (1989). The self and social behavior in differing cultural contexts. *Psychological Review*, 96(3):506–520.

Tsai, M.-H., Chang, Y.-L., Shiau, S., and Wang, S.-M. (2019). Exploring the effects of a serious game-based learning package for disaster prevention education: The case of Battle of Flooding Protection. *International Journal of Disaster Risk Reduction*, 43(January 2019):101393.

Tullos, D. (2018). Opinion: How to achieve better flood-risk governance in the United States. *Proceedings of the National Academy of Sciences*, 115(15):3731–3734.

Turner, W. A. and Casey, L. M. (2014). Outcomes associated with virtual reality in psychological interventions: where are we now? *Clinical Psychology Review*, 34(8):634–644.

Tversky, A. and Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology*, 5(2):207–232.

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

Tyers, R. (2018). Nudging the jetset to offset: voluntary carbon offsetting and the limits to nudging. *Journal of Sustainable Tourism*, 26(10):1668–1686.

UNISDR (2015). The human costs of weather related disasters. Technical report, UNISDR.

UNISDR (2017). Public communication for disaster risk reduction. Technical report.

Van der Linden, S. (2015). The social-psychological determinants of climate change risk perceptions: Towards a comprehensive model. *Journal of Environmental Psychology*, 41:112–124.

van Gelder, J. L., Hershfield, H. E., and Nordgren, L. F. (2013). Vividness of the future self predicts delinquency. *Psychological Science*, 24(6):974–980.

van Herpen, E., van den Broek, E., van Trijp, H. C. M., and Yu, T. (2016). Can a virtual supermarket bring realism into the lab? Comparing shopping behavior using virtual and pictorial store representations to behavior in a physical store. *Appetite*, 107:196–207.

van Valkengoed, A. M. and Steg, L. (2019). Meta-analyses of factors motivating climate change adaptation behaviour. *Nature Climate Change*, 9(2):158–163.

Vergouwe, R. (2015). The national flood risk analysis for the Netherlands. Technical report, Rijkswaterstaat.

Verhallen, P., Brrggen, E., Post, T., and Odekerken-Schrrder, G. (2018). Norms in Behavioral Interventions: Peer or Anchoring Effects? *SSRN Electronic Journal*.

Verlynde, N., Voltaire, L., and Chagnon, P. (2019). Exploring the link between flood risk perception and public support for funding on flood mitigation policies. *Journal of Environmental Planning and Management*, 0(0):1–22.

Vieider, F. M., Lefebvre, M., Bouchouicha, R., Chmura, T., Hakimov, R., Krawczyk, M., and Martinsson, P. (2015). Common components of risk and uncertainty attitudes across contexts and domains: Evidence from 30 countries. *Journal of the European Economic Association*, 13(3):421–452.

Viscusi, W. K. and Zeckhauser, R. J. (2006). National survey evidence on disasters and relief: Risk beliefs, self-interest, and compassion. *Journal of Risk and Uncertainty*, 33(1-2):13–36.

Visschers, V. H. M., Meertens, R. M., Passchier, W. W. F., and De Vries, N. N. K. (2009). Probability information in risk communication: A review of the research literature. *Risk Analysis*, 29(2):267–287.

von Neumann, J. and Morgenstern, O. (1947). *Theory of Games and Economic Behavior*. Princeton University Press, 2nd edition.

Weinstein, N. D. (1989). Effects of personal experience on self-protective behavior. *Psychological Bulletin*, 105(1):31–50.

Weinstein, N. D., Kwitel, A., McCaul, K. D., Magnan, R. E., Gerrard, M., and Gibbons, F. X. (2007). Risk perceptions: Assessment and relationship to influenza vaccination. *Health Psychology*, 26(2):146–151.

Wenzig, J. and Gruchmann, T. (2018). Consumer preferences for local food: Testing an extended norm taxonomy. *Sustainability*, 10(5):1–23.

Wiederhold, B. K. and Bouchard, S. (2014). *Advances in virtual reality and anxiety disorders*. Springer.

Willems, S. J. W., Albers, C. J., and Smeets, I. (2019). Variability in the interpretation of Dutch probability phrases-a risk for miscommunication. *arXiv preprint arXiv:1901.09686.*, pages 1–22.

Winter, R. A. (2013). Optimal insurance contracts under moral hazard. In Dionne, G., editor, *Handbook of Insurance*, volume 68, chapter 9, page 205230. Springer, New York, second edi edition.

Witmer, B. G. and Singer, M. J. (1998). Measuring presence in virtual environments: A presence questionnaire. *Presence: Teleoperators and Virtual Environments*, 7(3):225–240.

Woloshin, S., Schwartz, L. M., Byram, S., Fischhoff, B., and Welch, H. G. (2000). A new scale for asessing perceptions of chance: a validation study. *Medical Decision Making*, 20(3):298–307.

Wu, H.-C., Lindell, M. K., and Prater, C. S. (2015). Process Tracing Analysis of Hurricane Information Displays. *Risk Analysis*, 35(12):2202–2220.

Wulff, D. U., Mergenthaler-Canseco, M., and Hertwig, R. (2018). A meta-analytic review of two modes of learning and the description-experience gap. *Psychological Bulletin*, 144(2):140–176.

Yazdanpanah, M. and Forouzani, M. (2015). Application of the Theory of Planned Behaviour to predict Iranian students' intention to purchase organic food. *Journal of Cleaner Production*, 107:342–352.

Yechiam, E., Barron, G., and Erev, I. (2005). The role of personal experience in contributing to different patterns of response to rare terrorist attacks. *Journal of Conflict Resolution*, 49(3):430–439.

Yee, N. and Bailenson, J. N. (2007). The Proteus Effect: The effect of transformed self-representation on behavior. *Human Communication Research*, 33(3):271–290.

Young, S. D., Adelstein, B. D., and Ellis, S. R. (2006). Demand characteristics of a questionnaire used to assess motion sickness in a virtual environment. *Proceedings - IEEE Virtual Reality*, 2006:13.

Zaalberg, R. and Midden, C. J. H. (2013). Living behind dikes: Mimicking flooding experiences. *Risk Analysis*, 33(5):866–876.

Zhang, Y., Hwang, S. N., and Lindell, M. K. (2010). Hazard proximity or risk perception? Evaluating effects of natural and technological hazards on housing values. *Environment and Behavior*, 42(5):597–624.

Zhao, M., Rosoff, H., and John, R. S. (2019). Media disaster reporting effects on public risk perception and response to escalating tornado warnings: A natural experiment. *Risk Analysis*, 39(3):535–552.

About the author

Jantsje Mol was born in Bolsward, the Netherlands on August 17th, 1990. She obtained a BSc degree in Liberal Arts and Sciences at Tilburg University and a research MSc degree Cognitive and Clinical Neuroscience at Maastricht University, with a specialization in neuroeconomics.

In October 2016, Jantsje joined the Institute for Environmental Studies at the Vrije Universiteit Amsterdam for a PhD project on individual flood-preparedness for low-probability/high-impact flood risk. The project was funded by an NWO-VIDI grant that was awarded to prof. dr. W. J. Wouter Botzen. During her PhD, Jantsje visited the Wharton Risk Management and Decision Processes Center at the University of Pennsylvania as a visiting researcher.

Over the course of her PhD, Jantsje (co-)authored over nine publications in international peer-reviewed journals. She contributed to a report on longitudinal survey data collected from Florida during and after Hurricane Dorian. She presented her work at various international scientific conferences, including the 2018 Economic Science Association (ESA) World Meeting, the 2019 Subjective Probability, Utility and Decision Making Conference (SPUDM) and the 2020 Society of Advancement of Behavioral Economics (SABE) Annual Conference. She received the Runner Up Best Poster Award for her poster presentation at SPUDM 2019. She was invited as a guest speaker at the Virtual Reality for experimental economics workshop at RWTH Aachen in 2017 and at the European conference on risk perception at Université de Cergy Pontoise in Paris 2019.

In her current position as postdoctoral researcher at CREED at the University of Amsterdam, Jantsje works on behavioral ethics and the sharing economy. Her research interests include experimental economics, behavioral economics and virtual reality experiments.

List of publications

Publications on which this thesis is based

Mol, J. M. (2019). Goggles in the lab: Economic experiments in immersive virtual environments. *Journal of Behavioral and Experimental Economics*, 79(C): 155 - 164.

Mol, J. M., Botzen, W. J. W., & Blasch, J. E. (2020). Risk reduction in compulsory disaster insurance: Experimental evidence on moral hazard and financial incentives. *Journal of Behavioral and Experimental Economics*, 84: 101500.

Mol, J. M., Botzen, W. J. W., & Blasch, J. E. (2020). Behavioral motivations for self-insurance under different disaster risk insurance schemes. *Journal of Economic Behavior & Organization*, 180: 967 – 991.

Mol, J. M., Botzen, W. J. W., Blasch, J. E., & de Moel, H. (2020). Insights into flood risk misperceptions of homeowners in the Dutch river delta. *Risk Analysis*, 40(7): 1450 - 1468.

Mol, J. M., Botzen, W. J. W., Blasch, J. E., Kranzler, E. C., & Kunreuther, H. C. (2021). All by myself? Testing descriptive social norm-nudges to increase flood preparedness among homeowners. *Behavioural Public Policy*, 1-33.

Mol, J. M., Botzen, W. J. W., & Blasch, J. E. (2021). After the virtual flood: risk perceptions and flood preparedness after virtual reality risk communication. In review with *Judgment and Decision Making.*

Other publications

Mol, J. M., van der Heijden, E. C. M., & Potters, J. J. M. (2020). (Not) alone in the world: Cheating in the presence of a virtual observer. *Experimental Economics*, 23: 961 - 978.

Bloemendaal, N., de Moel, H., **Mol, J. M.**, Bosma, P. R. M., Polen, A. N., & Collins, J. M. (2021). Adequately reflecting the severity of tropical cyclones using the new Tropical Cyclone Severity Scale. *Environmental Research Letters*, 16(1): 014048.

Morren, M., **Mol, J. M.**, Blasch, J. E., & Malek, Ž. (2021). Changing diets - Testing the impact of knowledge and information nudges on sustainable dietary choices. *Journal of Environmental Psychology*, 75: 101610.

Acknowledgements

Dear reader, thanks for opening this book on the most important section. This is where I look back at the past five years, get a little sentimental and try to recall some funny anecdotes.

I first would like to thank prof. dr. Wouter Botzen, my promotor, for giving me the opportunity to do a PhD in environmental (and experimental) economics. Wouter, thanks for being such a great mentor and for teaching me how to write interesting academic papers. I am always impressed by your fast but thoughtful emails and your optimism when interpreting the comments of any reviewer (but especially reviewer two). This thesis would not have been possible without my co-promotor dr. Julia Blasch. Julia, thanks for being such a supportive and dedicated supervisor. Your door was always open for me to discuss methodological details and I really appreciate that. Although in retrospect I guess we often overestimated the level of detail required for publication and policy advice. But that is the joy of doing science, I believe.

I want to thank the members of my reading committee for taking the time and effort to read my thesis and provide me with very valuable feedback: prof. dr. Jeroen Aerts, prof. dr. Martijn van den Assem, prof. dr. Eline van der Heijden, dr. Özgür Gürerk and dr. Claudia Schwirplies. To prof. dr. Pieter van Beukering, I really appreciate your willingness to act as chair.

Without participants, experimental economics would be impossible. Dear participants, thanks for joining my experiments, and thanks for your quirky answers to the final feedback question. Some examples: "*I am thrilled after completing such a mind refreshing task and would love to accept such challenges in the future.*" "*Utopia does not exist on earth.*" "*I once lent 25.000 EUR to a family member. Never got it back, now I know you have to be careful with all investments.*" Participant in response to VR voiceover: "*Girl, don't tell me how to handle sandbags, I used to be a construction worker!* These answers helped me realize that you are all unique human beings and that society is an interesting mix of climate change deniers, mansplainers, happy campers, supportive fellow researchers and many more.

My first academic adventures were at Tilburg University and I am very grateful for the fact that Arnoud-Jan convinced me to join LAS. Jan, thanks for supervising my Bachelor thesis with such enthousiasm and Eline, thanks for being such a critical reader of that thesis that I was glad to have you on my side for the Master thesis. Thanks both for the opportunity to do a master thesis internship on such an exciting but risky method and for the many intellectually stimulating meetings in Jan's *balzaal*. I am still looking forward to a sunny day at Esplanade where we can toast to our publication in Experimental Economics!

I have great memories from my time in Maastricht with the Neuroeconomics Research Master. Arno, Matthias and my cohort: thanks for all the intense discussions, both in the classroom and in one of Maastricht's many eetcafés. Dear Tilburg Economics Research Master cohort, thanks for your company in the first office on the right. Thijs and Lenka: I hope we will always make time for beer, ribs and desserts. Lenka, doctorsister, I really look forward to a celebration of our amazing double defense day.

I also would like to thank my co-authors at the VU: Meike Morren, Žiga Malek and Hans de Moel. It was a pleasure working on a side project with you and I learned a lot about qualitative research, land use change and hydrology on the go. I would like to thank Howard Kunreuther for the invitation to visit the Wharton Risk Center at the University of Pennsylvania in Fall 2019. Elissa, Penny, Dana, Helen, Zoë, Erin and Ebony: thanks for the warm welcome at the Center and for showing me the American way of things, in particular potlucks and madlibs. Shaul, Ivan, Margarita, Linh, Catherine, Joël and all other CREED-ers, thanks for the warm welcome and for listening to all my flood-related stories.

Roland and Rudi from IT voor Onderzoek, I have often referred to you as 'my IT heroes' when explaining your work to my colleagues. Thanks for making the oTree environment possible in a secure way and for the (totally unrealistic) trust that I would understand all your jargon. Marco and Nadine from the Network Institute, thanks for developing such a realistic flood in VR and for organizing a Covid-safe protocol to run this experiment in October (!) 2020.

Mark, Justin, Paul, Pete and Teun, we worked our asses off together to make the EE course work, but I really believe it got better with all that effort. I enjoyed teaching those tutorials and coordinating those games, but it was largely the teacher chat that made it so much fun. Cheers to an empty sea! Ralph, thanks for showing me how to effectively supervise students, and for the courage to hit that mouse with a PhD thesis that was lingering on my desk in the W&N building (sorry Sanne).

Marleen and Jens, thanks for enjoying boardgames and dinners as much as we do. I still count you as part of team Den Haag. Marleen, thanks for joining me to all university activities, with and without free lunches. It is just great to have many breaks together, even though our *koe* did not survive every week in the IVM fridge. And thanks for being so active on Slack, it really makes me feel I still have a bunch of IVM officemates hiding in my computer. Jens, thanks for being available for quick python advice 24/7 and for coming back to the office telling crazy travel stories as if they were nothing extraordinary.

I also would like to thank Ted, Paul, Pete, Toon, Sanne, Pedja and Job for being great office mates. It was a pleasure meeting many IVM PhDs and colleagues over the years. Thank you Theo, Elco, Hanna, Lisa, Lars, Johanna, Kostas, Erick, Liselotte, Shutong, Bianca, Sadhana, Gabi, Marthe, Niels, Timothy, Kate, Sem, Sem, Vita, Giorgia, Tadzio, Max, Rhian, Eric, Rosa and many others, for making IVM such a fun and inspiring workplace.

Anaïs, thanks for being my favorite colleague in the amazing Korenbeurs library of Schiedam. I hope we will be able to go there again to finish our cappuccino *strippenkaart*. Paul, thanks for teaching me how to teach and to maximize the number of cartoons in my slides. My past and future students are forever grateful. Dirk, thanks for encouraging me to set up an academic Twitter account. Job, thanks for representing team Den Haag at IVM and for all the lunch talk about volleyball and badminton. Pete, thanks for tolerating my sometimes not so subtle views on our profession and our reviewers. And thanks for a lot of native English writing advice which was always accompanied by the words 'I don't know why but it just sounds better'. I hope my modest amount of Dutch translations for you makes up for it at least a little bit. Nynke, Yue, Katharina, Anne, Anna, Marije and Verena, thanks for the great company at the writing group. I got to finish this thesis and we planted trees on the go!

Marjolijn, without you IVM would not be the same. Thanks for teaching me the magic of powerpoint in exchange for a little bit of InDesign. That 10 second Christmas jingle still haunts me sometimes, but now in a positive way.

Over the years, my Tall Ship family at Eendracht, Morgenster and SPANT has been a huge support. Ships and crew, thanks for the great adventures, the deep conversations on night watch, surviving extreme weather and for being my back-up career plan in case of a real disaster. Anna, thanks for showing that tall ship life and university life can always be combined. Nienke, I'm glad you are as exited about my defense as about EU exchange sail training programs. Both life-changing experiences, I suppose.

Leonie, thanks for tolerating me as the most critical client of your design studio. I am really in love with the final result! Roos, thanks for teaching me the value of good company and dice games already at high school.
Team Eindhoven: Ymke, Martijn, Maartje, Hylke, Wilma, Peter, Daisy, Wouter, Charlotte, Wim, Nikolai and Erik, thanks for keeping in touch after we left Hajraa and the city of light. It is a pleasure to eat, drink, borrel, walk, klus, ski, laser-game, play volleyball, beach volleyball or boardgames with you.

Team Den Haag: Jonne, Wouter, Anna, Niek, Tom, Freek, Carolien, Martine, Roely, Florien, Rianne, Ilona, Rosalie, Nynke, Joël and D1 t/m D4, thanks for making Den Haag a place to call home. Merel, thanks for being my *maat* in sand and snow and indoor and beyond, and for calling everyone *kanjer* and yourself a *race tiger*, which is truly what you are.

A great thanks to my amazing paranimphs. Nadia, *Miss (but almost Dr.) Hurricane*, thanks for the endless stream of GIFjes on Slack, for the best first-hand IVM gossip, for taking my social science perspective serious, for defeating me with Azul, for fine-tuning graphs and figures until eternity and for inviting me to co-author on your paper. I look forward to many more collaborations, because clearly tropical cyclones need more behavioral economics and vice versa. Mehmet, *pannenkoek*, from the first time I walked into my neighbor office I felt welcome. Thank you for the countless times you patiently answered my questions about econometrics or comparative statics, for teaching me the unwritten rules of the academic game and for being impressed by my oTree code regardless of its quality. Yay also for coffee and for organizing an IVM coffee tasting experiment with proper R-code and the irony of standing by environmental standards, which meant that there was no choice whatsoever...

I would also like to thank my family, the Mollen and Zeemannen and also the Verdenii for your support over the years. My love of games, water and research would not have been the same without you all. Marta, thanks for being my biggest fan (sorry Stefan, sorry mam). Oma Maatje, thanks for being the most awesome grandmother on this planet. I cannot thank you enough for all emails, *slibtongetjes*, card games, and trips through the *zak van Zuid-Beveland*, let alone the detailed stories about the *Ramp* which helped to shape the virtual reality experience in Chapter 6.

Pap & mam, I am very grateful for your unconditional support. In high school I was pretty convinced that I never wanted to follow in your footsteps by becoming a teacher or a scientist. Thanks for letting me seek my own path. I think it is only fitting that the acorn doesn't fall far from the tree, after all. And surely all credits for the clear Dutch summary go to team Bolsward.

Stefan, these last words are for you. I really appreciate your courage to join my adventures, such as leaving *Brabant* for the big city and selecting holiday locations based on conference and research visits. Thanks for always believing in me and my work, even though my social sciences are way too noisy for the engineer you are. Sometimes I think that steel bridge constructions and people's behavior can be equally *vermoeiend*. Thank you for helping me to focus and to be distracted, for surviving more than a year of home-office with me and for countless other things I cannot express in words.