NBER WORKING PAPER SERIES

PREDICTING INSURANCE DEMAND FROM RISK ATTITUDES

Johannes G. Jaspersen
Marc A. Ragin
Justin R. Sydnor

Working Paper 26508
http://www.nber.org/papers/w26508

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
November 2019

We thank the Alfred P. Sloan Foundation for financial support under grant number G-2016-7312. For helpful feedback and comments, we thank Levon Barseghyan, Mark Browne, Sebastian Ebert, Julian Reif, Charles Sprenger and seminar participants at the University of Hannover, the MRIC Winter Workshop, and the American Risk and Insurance Association annual meeting. We also thank Josh Frederick and Joëlle Näger for research assistance. All remaining errors are ours. This study was pre-registered with the American Economic Association's randomized controlled trial registry under ID AEARCTR-0002783 The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Predicting Insurance Demand from Risk Attitudes
Johannes G. Jaspersen, Marc A. Ragin, and Justin R. Sydnor
NBER Working Paper No. 26508
November 2019
JEL No. D01,D81,G22

ABSTRACT

Can measured risk attitudes and associated structural models predict insurance demand? In an experiment (n = 1,730), we elicit measures of utility curvature, probability weighting, loss aversion, and preference for certainty and use them to parameterize seventeen common structural models (e.g., expected utility, cumulative prospect theory). Subjects also make twelve insurance choices over different loss probabilities and prices. The insurance choices show coherence and some correlation with various risk-attitude measures. Yet all the structural models predict insurance poorly, often less accurately than random predictions. Simpler prediction heuristics show more promise for predicting insurance choices across different conditions.

Johannes G. Jaspersen
LMU Munich
jaspersen@bwl.lmu.de

Justin R. Sydnor
University of Wisconsin at Madison
jsydnor@bus.wisc.edu

Marc A. Ragin
University of Georgia
mragin@uga.edu

A randomized controlled trials registry entry is available at
https://www.socialscienceregistry.org/trials/2783/history/

1 Introduction

In this paper, we consider the relationship between risk attitudes and insurance demand from a new angle. Much of the research in this area has focused on inferring risk attitudes from observed insurance choices. We evaluate this relationship from the other direction and instead measure risk attitudes and then examine the extent to which these measures predict insurance decisions. We ask two interrelated questions.[1] First, do measured risk attitudes correlate with insurance demand? Second, do structural models incorporating different combinations of risk attitudes (such as expected utility or cumulative prospect theory) predict insurance choices well?

We are motivated by an underlying tension in the literature analyzing insurance demand in market settings. A number of studies use observed insurance choices to estimate structural models of risk preferences, and then use those structural models to inform policy by predicting demand in alternative market conditions (e.g., Cohen and Einav, 2007; Decarolis et al., 2019; Handel et al., 2015). The value of these counterfactual policy evaluations hinges on whether the model predictions are accurate. Papers in this literature tend to adjust for observable heterogeneity and factors such as inertia, and these adjustments have important effects on model predictions. Yet there are rarely opportunities to validate the underlying risk preference structure assumed in the models using out-of-sample tests. Having the right structural model of risk attitudes is important both for the model's out-of-sample predictions and for the validity of any welfare conclusions drawn about policy.

There is evidence that a range of behavioral factors may affect insurance demand, leading to substantial misspecification for models incorporating only classic risk aversion (e.g., Abaluck and Gruber, 2011; Abito and Salant, 2018; Barseghyan et al., 2013; Bhargava et al., 2017; Handel, 2013; Handel and Kolstad, 2015; Sydnor, 2010). This motivates an interest in models incorporating alternative sources of risk attitudes, such as loss aversion and probability weighting (Barseghyan et al., 2018). A leading example is Barseghyan et al. (2013), who estimate a more flexible structural model and conclude that probability distortions explain choices better than classic risk aversion. It is not yet clear, though, whether this model or other suggested behavioral models predict insurance choices out-of-sample more accurately. It is also fundamentally challenging to separately identify different possible behavioral models and motivations from insurance data alone. A given willingness to pay for insurance can typically be rationalized by many different preference motives, and in many studies there is not sufficient underlying variation in choice sets and loss probabilities to differentiate these motives (Barseghyan et al., 2013; Sydnor, 2010).

[1] We pre-specified these questions in our randomized controlled trial registration with the American Economic Association (AEARCTR-0002783).

Our focus is on measuring risk attitudes and developing structural models from one set of data and studying how well those measures and models predict insurance decisions out-of-sample. We use an incentivized experiment to measure risk attitudes using lottery choices and to elicit demand for insurance in different scenarios. We conduct our experiment with students at the University of Wisconsin–Madison and in an online experiment via Amazon's Mechanical Turk, recruiting a total of 1,730 subjects across the two platforms. While the ultimate goal is to inform our understanding of real insurance markets, the advantage of using a lab experiment to study this question is that insurance demand in market settings is partly influenced by factors other than risk attitudes, including awareness, institutional details (e.g., sales channels), and variation in subjective beliefs about loss probabilities (Collier and Ragin, 2019). A laboratory experiment offers a controlled environment to isolate the links between underlying risk attitudes and insurance decisions.[2] In addition, the laboratory allows us to exogenously vary the price of insurance; truly exogenous variation in prices is rare in market data.

We measure risk attitudes using decisions in a lottery task. The lotteries are adapted from procedures used previously in the literature (e.g., Tanaka et al., 2010) and are designed to allow us to jointly measure different components of risk attitudes: utility curvature, probability weighting, loss aversion, and a preference for certainty. We evaluate utility curvature and probability weighting in both the gain domain and the loss domain, leading to six primitive preference motives. For each of these motives, we calculate a nonparametric measure (an ordinal preference measure, with no assumed functional form) and a parametric measure (with an assumed functional form). Our parametric measures are generally in line with estimates from other studies. For instance, we find that the majority of subjects display inverse-S probability weighting in both the gain domain and loss domain. We also observe a median coefficient of loss aversion of 2.7, which is broadly consistent with estimates in prior studies.

In the insurance task, we expose participants to the possibility of a modest loss of their earnings from a real-effort task conducted at the beginning of the experiment. Participants choose the proportion of the potential loss they want to insure (from 0% to 100%).[3] They make decisions in twelve scenarios, which differ in the loss probability and the price of insurance (i.e., loading). While there are no equivalent studies to benchmark these choices against, the choice patterns vary significantly across different combinations of loads and probability, indicating that subjects

[2] The laboratory experiment also allows us to study attitudes toward risky lotteries and insurance choices over the same-size financial stakes.

[3] Throughout the paper, we use the terms "coverage level," "amount insured," and "proportion insured" interchangeably to describe this coverage choice. When we refer to "insurance demand," we are referring to the coverage level chosen at a given price.

actively consider the parameters of the choice decision. Subjects also obey the law of demand on average, in that they purchase less insurance when prices are higher.

The answer to our first question—Do measured risk attitudes correlate with insurance demand?—is yes, at least modestly. Within an insurance-choice scenario, we observe modest correlations between the level of insurance chosen and our measures of utility curvature, loss aversion, and probability weighting. For each of these three measures, we find that a one-standard deviation increase in the measured preference motive is associated with around a 4% increase in insurance demand. Moreover, the correlations between probability weighting and insurance choices show the expected interaction with the loss probability—stronger probability weighting correlates more positively with insurance demand when the loss probability is low. We find no correlation between preference for certainty and insurance demand. These findings suggest that insurance choices are related to a number of different underlying preference motives. In particular, our results support the idea that better understanding motives such as loss aversion and probability weighting can have value for understanding insurance demand. None of the three main preference motives, however, is clearly more important than the others. Moreover, when evaluated jointly, they explain only a small part of the variation in insurance choices.

To answer to our second question—Do structural models incorporating different combinations of risk attitudes predict insurance choices well?—we use the lottery choices to parameterize seventeen utility models that have been used previously in the literature, including classic expected utility theory, cumulative prospect theory, rank-dependent expected utility theory, and a number of other models.[4] We then predict insurance choices by determining the level of insurance which maximizes the target function of each parameterized model for each person in each insurance scenario.

If we look at the correlation of these model predictions and the amount of insurance purchased *within* a particular insurance scenario, the results are consistent with the modest correlations we observe when looking at the primitive preference motives alone. The highest correlations are for models which combine probability weighting and utility curvature (i.e., rank-dependent expected utility and prospect theory). The parameterized utility models have some ability to predict the *relative* amount of insurance that different subjects select within a scenario. However, in all cases the correlations are quite modest—in the utility model with the strongest within-scenario correlations, a 10 percentage-point increase in the predicted level of insurance purchased is associated with only a 0.8 percentage-point increase in actual insurance purchased.

[4] For prospect theory, we consider a number of different reference points that might be employed when considering insurance decisions. All of the models we consider are "consequentialist" in the sense that (decision) utility is a function of the distribution of potential final wealth outcomes experienced with the choice.

A key question for structural models applied to insurance demand, though, is whether they are valuable for predicting the *level* of demand for insurance (as opposed to the *relative* demand) and whether they can do so well *across* different economic scenarios. Parameterized utility models are useful primarily if they allow one to predict behavior in a range of counterfactual environments. We find, however, that all of the utility models, including the ones incorporating loss aversion and probability weighting, have low overall correlation with insurance demand. In fact, each of the parameterized utility models performs worse than random choice in predicting the level of insurance purchased.

We attribute the poor predictive validity of these models to two primary inconsistencies between the model predictions and our subjects' choices. First, the models predict strong sensitivity to the price of insurance. Subjects' actual insurance choices, however, respond to price in the expected direction, but much more modestly. For example, a simple expected utility model calibrated from lottery choices predicts subjects to be more than six times as sensitive to price than they actually are. Second, all the models predict that a (weakly) higher share of the loss will be insured at low loss probabilities. Subjects' actual decisions show the reverse pattern—they purchase more insurance when the probability of loss is higher.

Our results add to growing evidence on the limited value and consistency of risk preference estimates in the lab (e.g., Chapman et al., 2018; Friedman et al., 2019; Friedman and Sunder, 2011). Our controlled experimental setting involves similar stakes between the lottery choices used to elicit risk attitudes and the potential loss scenarios used to model insurance demand. We find that a broad range of structural models parameterized with these risk attitudes have limited external validity for predicting insurance choices.

An important question, though, is whether these results primarily highlight the limits of trying to model risk attitudes across context (e.g., between abstract financial lotteries and insurance), or if they point to deeper issues in using structural models to predict insurance choices. We explore this issue by looking within our insurance-choice data. For this exercise we adopt the framework from Barseghyan et al. (2013) that allows for both classic risk aversion via a concave utility function and generalized probability distortions. We estimate the model for each individual using a subset of their insurance choices and then use the structural model to predict their decisions at other price levels or other loss probabilities. The models parameterized using only risk attitudes elicited from insurance choices show reasonably strong correlations between observed and predicted demand. However, the structural model predictions perform worse than simple alternatives that are not grounded in risk preferences. Both using an individual's average observed level of demand as a fixed prediction for that person and creating individual-level linear models extrapolating demand as a function of price or probability outperform the structural model. Similar to our main findings, the structural model estimated using insurance choices predicts too much sensitivity to price and struggles to capture how people select insurance at different levels of probability.

Our findings offer new insights to the question of whether insurance demand reflects a stable underlying set of risk attitudes. Prior literature has documented that there are meaningful rank-order correlations between the choices individuals make in different insurance contexts, though the correlations can be low across decisions with different financial stakes (Abito and Salant, 2018; Einav et al., 2012). Despite this underlying consistency in choices, though, the limited attempts to compare the fit of structural models across contexts or stakes have concluded that choices cannot be rationalized by a single underlying model (Barseghyan et al., 2011; Collier et al., 2018). Our results are consistent with these existing findings and extend them by highlighting that the limitations in predicting insurance demand are shared by a wide range of structural models. Our results from a controlled experimental setting also imply that the limitations of structural models are not driven only by institutional factors that may affect decisions across real-world insurance settings. There appears to be a more fundamental failure of established models of risk attitudes to capture how people make decisions about insurance.

There are a few implications of these results for both applied policy work and directions for future research. For example, the Congressional Budget Office uses models to estimate health insurance coverage under different policy environments (CBO, 2019). Historically, their models relied primarily on an "elasticity approach" that modeled the responsiveness to prices and other variables, but they have been moving toward an approach using structural models that incorporate risk preferences. Our results suggest a reason for caution in moving away from the elasticity approach. Second, we believe our results should motivate further research into approaches that incorporate features beyond consequentialist utility frameworks, such as decision frictions, limited consideration sets, and decision heuristics (e.g., Abaluck and Adams, 2017; Barseghyan et al., 2019; Handel et al., 2015).

The remainder of the paper proceeds as follows. We lay out the underlying theoretical structure for our exercise in Section 2. In Section 3, we discuss the experimental design and the choice patterns we observe for lottery choices and insurance choices separately. We provide our main results on the predictions of insurance choices from the risk attitudes and structural models estimated from lottery choices in Section 4. In Section 5, we present the results of estimating the structural model within the insurance choices. We conclude and discuss in Section 6.

2 Theoretical Basis

The paper is concerned with eliciting preference motives and using them to predict insurance demand. In our analysis of the data, we use two approaches to evaluate the ability of risk attitudes and structural models to predict insurance demand. First, we consider the relationship between insurance demand and preference motives in their raw form. In this approach, we simply correlate the elicited preferences with insurance demand to gauge their predictive power. In the second approach, we assemble the preference motives into seventeen different structural models and test the insurance demand predictions of these models against the empirically-observed demand.

2.1 Preference Motives

To elicit primitive preference motives, we first construct a decision model such that these motives have both a definition and implications for decision-making behavior.[5] We begin with a standard cumulative prospect theory (CPT, Tversky and Kahneman, 1992) preference functional. We also consider an extended version of this model that adds a parameter which can regulate a preference for (or against) certainty. This general model allows for the most common preference motives in risky decision-making, including those hypothesized to drive insurance demand (see, e.g., Barseghyan et al., 2013; Callen et al., 2014; Sydnor, 2010). The model is also useful because, for the decisions faced by subjects in our experiment, it nests several other models (see Section 3.2).

We represent risk as lotteries with n outcomes. Each outcome $x_i \in \mathbb{R}$ has a probability p_i and outcomes are indexed such that $x_1 \geq ... \geq x_n$. We describe a lottery with the notation $(x_1, p_1; ...; x_n, p_n)$. Preferences over lotteries are represented by the function

$$V(\tilde{x}) = \sum_{i=1}^{n} \pi_i U(x_i). \tag{1}$$

Here, outcomes are evaluated by the function

$$U(x) = \begin{cases} u^+(x-r) & \text{for } x \geq r \\ \lambda u^-(x-r) & \text{for } x < r \end{cases} \tag{2}$$

in which r describes the reference point, u^+ is the utility function over gains, u^- is the utility function over losses and λ is a loss aversion parameter. u^+ and u^- are assumed to be strictly

[5] Parts of the following section and Section 3.2 resemble the theoretical considerations of the companion paper to this, Jaspersen, Ragin, and Sydnor (2019). The companion paper considers a question distinct from the analysis here but is based on the same experimental data and thus utilizes the same approach to preference elicitation.

increasing and $u^+(0) = 0$. The probability weights π_i are formed according to

$$\pi_i = \begin{cases} w^+(\sum_{j=1}^{i} p_j) - w^+(\sum_{j=1}^{i-1} p_j) & \text{if } x_i \geq r \\ w^-(\sum_{j=i}^{n} p_j) - w^-(\sum_{j=i+1}^{n} p_j) & \text{if } x_i < r. \end{cases} \tag{3}$$

If one assumes the reference point to be defined by current wealth, the standard CPT preference functional includes five different behavioral motives that influence behavior in decisions under risk and thus insurance demand. We begin by outlining our "nonparametric" measures of these motives (denoted by capital letters), which do not assume a specific underlying functional form. If the utility function is not linear, it shapes an attitude towards risk. We denote this concept of utility curvature in the gain and loss domains as UC^+ and UC^-, respectively. The shape of the probability weighting functions in the gain and loss domain similarly influence risk attitudes. We assume that deviations from the identity line are in the form of either S shaped or inverse-S shaped probability weighting functions, with an inflection point around the middle of the unit interval. Our measures PW^+ and PW^- denote the extent to which the gain domain, respectively loss domain, probability weighting function is inverse-S shaped. Lower values of these measures imply an S shape of the function, while higher values imply an inverse-S shape. Lastly, a slope difference between utility in the gain domain and utility in the loss domain denotes loss aversion (LA) of the individual. Higher values of LA indicate a steeper slope in the loss domain than in the gain domain.

In addition to these five standard CPT preference motives, we add a potential sixth motive, a preference for certainty. We model this motive after Schmidt (1998), who axiomatizes a decision model and develops a measure for certainty preference.[6] While payouts of risky prospects are evaluated with utility functions u^+ or u^-, payouts in prospects with a certain outcome are evaluated by a function v. Because we are within a CPT framework, $u^+(0) = v(0) = 0$ must hold. We thus model the value function as

$$v(x) = \kappa u^+(x). \tag{4}$$

Here, $\kappa > 1$ implies a preference for certainty, while $\kappa < 1$ implies a preference for risk. We model and elicit the certainty preference CP in the gain domain. CP is elicited with the help of a separate choice table. As such, all other preference motives are elicited as if a standard CPT preference functional was assumed (i.e., they are measured independently of CP).

The motives introduced above are not parametric in nature because we have not yet assumed a fully parametric structure for Equations (2) and (3). Our elicitation procedure allows us to as-

[6] Note that this motive has distinct implications from the preference for certainty already inherent in probability weighting (Andreoni and Sprenger, 2012).

sess these motives ordinally using statements in which preference motives change *ceteris paribus* in subjects' answers to our preference elicitation task. The specific assumptions allowing us to do so are given in Section 3.2. Formal derivations of the elicitation procedures are provided in Appendix B.

In addition to the nonparametric motives, we also aim to calibrate a parametric structure of the preference functional. The parameters shaping the structure are our "parametric" measures of the six preference motives, and we refer to them with Greek letters. Our parametric structure involves assumptions similar to Tanaka et al. (2010), after which our preference elicitation procedure is modeled. Specifically, we assume a utility function shaped by γ

$$U(x) = \begin{cases} \frac{(x+1)^{1-\gamma^+}}{1-\gamma^+} - \frac{1}{1-\gamma^+} & \text{for } x \geq 0 \\ -\lambda \left(\frac{(1-x)^{1+\gamma^-}}{1+\gamma^-} - \frac{1}{1+\gamma^-} \right) & \text{for } x < 0. \end{cases} \tag{5}$$

and a probability weighting function shaped by β

$$w^{+/-}(p) = exp\left(-(-ln(p))^{2-\beta^{+/-}}\right). \tag{6}$$

The two forms are similar but not equal to standard assumptions in the literature. The utility function closely resembles the iso-elastic form, which has been the workhorse of preference elicitation (e.g., Andersen et al., 2008; Tanaka et al., 2010).[7] Its general functional form $u(x) = \frac{x^{1-\gamma}}{1-\gamma}$ is not appropriate for prospect theory, because the model requires $u(0) = 0$ and the iso-elastic form is not defined at $x = 0$. We thus opt for the expression in (5), which is very close to iso-elastic but does not match the structure completely. For the sake of easy interpretation, the γ coefficients can, however, be seen as approximate coefficients of relative risk aversion.[8] We code γ^- such that a higher value implies more concave curvature so that the interpretation of the nonparametric UC^- and the parametric γ^- are congruent. Equation (6) is a slight modification of the one-parameter Prelec (1998) probability weighting function, with $\beta^{+/-} = 1$ giving the identity (no probability weighting), $\beta^{+/-} > 1$ implying the typical inverse-S shape, and $\beta^{+/-} < 1$ implying S shaped probability weighting.[9]

[7] Many papers in the literature use the power formulation $u(x) = x^\gamma$, which is a special case of the iso-elastic function. The power function, however, is limited to modeling relative risk aversion less than or equal to one, which is insufficient to grasp the utility curvature of many individuals over laboratory stakes.

[8] The exact coefficient of relative risk aversion is $\gamma^+ \frac{x}{x+1}$.

[9] In Prelec's original formulation, inverse-S probability weighting is decreasing in β. Our modification makes inverse-S probability weighting increasing in β, so that all primitive preference motives are increasing in their measures. The adjustment does not affect the properties of the weighting function.

Measuring loss aversion is not trivial and several measures have been proposed in the literature (Abdellaoui et al., 2007). The parameter λ has an intuitive interpretation only when using power utility. In the utility function used here, γ^+ and γ^- also shape the absolute values of the utility function in the gain and in the loss domain such that λ must be interpreted in combination with these parameters (Balcombe et al., 2019). We thus create an index $\hat{\lambda}$, equal to the average absolute ratio of utility values over a range $[-x_{max}, x_{max}]$ such that

$$\hat{\lambda} = -\int_0^{x_{max}} \frac{\lambda u^-(-x)}{u^+(x)} dx. \tag{7}$$

Similarly, the preference for certainty is not fully captured by κ in the certainty value function $v(x)$. We use the index $\hat{\kappa}$ developed by Schmidt (1998), where

$$\hat{\kappa} = \frac{v(x) - u^+(x)}{u'^+}. \tag{8}$$

This index measures the difference between the two utility functions in units of marginal utility over lottery prospect outcomes.

2.2 Structural Models

In addition to analyzing how the primitive motives explain insurance demand, we assemble the preference motives in various structural decision models to predict insurance demand. The models involve individuals with wealth y who face a loss L with probability p. They can purchase insurance with a coverage level $\alpha \in [0, 1]$. Once purchased, the policy provides coverage for a share α of the loss if it occurs. The insurance premium for coverage level α is calculated as αqpL, where q is a relative loading factor. At $q = 1$, the premium is actuarially fair. If q exceeds one, the premium includes a risk premium. A structural model predicts insurance demand by determining the level of α at which the model's objective function is maximized. In the experiment, subjects choose α in twelve different insurance scenarios with varying levels of p and q. The models are thus used to make twelve predictions of α^* for each subject.

We test seventeen different structural models which have been proposed in the literature. The models, their generalized objective functions, and their possible variants are summarized in Table 1. This is a comprehensive list of the possible structural models for which our preference elicitation procedure provides us information. All of these models are context-independent—the evaluation of one action does not depend on the other available actions.[10] For each model, we

[10] Because they involve psychologically distinct concepts, context-dependent models, the most prominent of which are regret theory (Loomes and Sugden, 1982) and salience theory (Bordalo et al., 2012), cannot easily be parameterized with our elicitation procedure. The preference elicitation tables in our experiment, which are summarized in Tables 2

Table 1: Summary of structural models

Model	Objective Function	Variants
Expected Value	$p(y-(1-\alpha)L-\alpha qpL)+(1-p)(y-\alpha qpL)$	EV
Expected Utility	$pu(y-(1-\alpha)L-\alpha qpL)+(1-p)u(y-\alpha qpL)$	EU^+, EU^-
Dual Theory	$w(p)(y-(1-\alpha)L-\alpha qpL)+(1-w(p))(y-\alpha qpL)$	DT^+, DT^-
Reference Dependent Preferences	$y-pL-(q-1)\alpha pL+p(1-p)(1-\lambda)(1-\alpha)L$ Axiom A3' of Kőszegi and Rabin (2007). Model: CPE Stochastic reference point $\{p, y-\alpha qpL-(1-\alpha)L; 1-p, y-\alpha qpL\}$	KR
Rank Dependent Expected Utility	$w(p)u(y-(1-\alpha)L-\alpha qpL)+(1-w(p))u(y-\alpha qpL)$	RDEU^+, RDEU^-
CPT loss domain	$w^-(p)u^-(-(1-\alpha)L-\alpha qpL)+(1-w^-(p))u^-(-\alpha qpL)$ Reference point $r=y$	CPT^-
CPT no loss in buying	$w^+(1-p)u^+((1-\alpha)qpL)+\lambda w^-(p)u^-((\alpha-1)L+(1-\alpha)qpL)$ Reference point $r=y-qpL$	CPT^{NLIB}
Certainty Preference	Above models with $\kappa(y-\alpha qpL)$ for $\alpha=1$	EV_{CP}, DT^+_{CP}, DT^-_{CP}
	Above models with $\kappa u(y-\alpha qpL)$ for $\alpha=1$	EU^+_{CP}, EU^-_{CP}, $\text{RDEU}^+_{\text{CP}}$, $\text{RDEU}^-_{\text{CP}}$

Note: The table lists the seventeen structural models tested in this study. p denotes loss probability. y denotes wealth. L denotes loss amount. α is the proportion of the loss insured. q is the relative premium loading factor. Superindices $+$ and $-$ on the model abbreviations indicate whether preference parameters are elicited in the gain domain or in the loss domain. The subindex CP indicates that the otherwise unchanged model evaluates certain outcomes with utility function v instead of u. The two CPT models are listed separately because the difference in the reference point changes the objective function.

assume that individuals are narrowly focusing on the decision at hand in our experiment and do not integrate the outcomes of this experiment with background wealth.

We begin with a simple expected value calculation (EV) and then consider increasingly complex models that increase the number of parameters involved. The first three models, expected utility theory (EU), dual theory (DT) and reference-dependent preferences under property A3' of Kőszegi and Rabin (KR, 2007) each feature one primitive preference motive. These are utility curvature, probability weighting, and loss aversion, respectively.[11] Rank-dependent expected

and 3, render six indifference relationships between prospects which would enable us to calibrate both a regret theory and a salience theory preference functional within some margin of error. However, the tables were not designed with this goal in mind, which might lead to framing and boundary effects. We thus refrain from the exercise.

[11] For modeling probability weighting in DT and RDEU, we use the lower Choquet integral such that the probability of the loss event is weighted first. This seems appropriate as insurance decisions regard so called "bad news events" (Sarin and Wakker, 1998).

utility theory (RDEU) and cumulative prospect theory in the loss domain (CPT^{-}) combine the two motives of utility curvature and probability weighting. Other combinations of two motives have seldom been proposed in the literature. More general reference-dependent preferences of Kőszegi and Rabin (2007, property A3) can be seen as a model combining utility curvature and loss aversion. The model, however, requires further assumptions to calibrate using our data (Argyris et al., 2019). All three preference motives are combined in the CPT^{NLIB} model in which purchasing full insurance is taken as a reference point (see Novemsky and Kahneman, 2005, for a psychological rationale for this reference point).

For some of the models, we include different variants. We elicit utility curvature and probability weighting in the gain domain and in the loss domain, and consider EU, DT and RDEU using parameters in both domains. For most models, we also add a variant with a preference for certainty. In the reference dependence model, adding a preference for certainty seems infeasible and there are no prior implementations of the concept.[12]

It is worth noting that the parameters of utility function, probability weighting function, and loss aversion are not the same for the different structural models. While CPT^{NLIB} and CPT^{-} can simply utilize the parameters of the model in equation (1), all other structural models restrict some elements of (1). In expected utility theory, for example, the probability weighting function is restricted to the identity line and the preference elicitation procedure is adjusted accordingly. Similarly, some theories only draw from a select number of preference elicitation questions, such as considering only choices among lottery prospects in the gain domain.

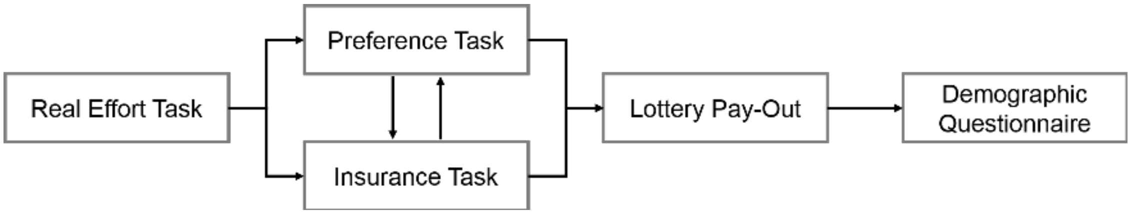
3 Experimental Design and Resulting Choice Patterns

3.1 General Structure and Subject Recruitment

After an initial screen with a consent form (and a bot check for online participants), our experiment consists of the five stages outlined in Figure 1. All subjects begin with the typing task, where they can earn $5.00 in virtual currency by correctly typing two passages of text from an image into a text box. Subjects are told they must correctly transcribe each passage to earn their payment and continue to the next stage (we allowed transcriptions with errors to continue after three minutes).

[12] If such a preference motive were to be added, it seems in the spirit of the model to add it in the gain-loss utility function. However, we use the choice-acclimated personal equilibrium (CPE) to predict insurance demand in this model, as suggested by Kőszegi and Rabin (2007, p. 1058). In this equilibrium, gain-loss utility disappears if full insurance is bought, and so a preference for certainty would not affect the purchasing decision. In CPT^{NLIB}, purchasing full insurance is the reference point and thus implies a value of 0 by definition, also implying no role for certainty preferences.

Figure 1: Structural outline of the experimental design



Subjects then complete the "preference task" and the "insurance task" as described in following subsections.[13]

After both tasks are complete, the computer randomly selects one question to play out, collecting the insurance premium (if applicable) and drawing a virtual ball to determine whether any additional money is gained or lost. After the outcomes are determined, subjects complete a questionnaire including the "general risk question" of Dohmen et al. (2011) and demographic information (namely, age, sex, location, income, and education). Finally, we provide a validation code to enter in mTurk or provide to the experimenter.

Experimental sessions with a total 1,730 subjects were conducted in March of 2018. 1,352 subjects were recruited on the online recruitment platform Amazon Mechanical Turk (mTurk) and conducted the experiment online.[14] The other 378 subjects were recruited without any restrictions from the subject pool of the BRITE laboratory at the University of Wisconsin-Madison. The sample sizes were based on an *ex ante* power calculation which rendered target sizes of 1,350 and 350 subjects for mTurk and in-person experiments, respectively. These targets were registered in the American Economic Association Randomized Control Trial Registry. Subjects were on average 33 years old and 53% of subjects were male. Summary statistics of demographic characteristics and other experimental data are provided in Appendix C. On average, experimental sessions were about 25 minutes long and subjects earned \$6.27.[15]

3.2 Preference Task

Our preference measurement tasks are based on the lottery tables of Tanaka et al. (2010). The Tanaka et al. study involved three lottery tables designed to elicit probability weighting and util-

[13] We randomly vary whether the preference task or the insurance task is presented first. Each task begins with a set of instructions, a practice task, and several questions to ensure subjects understand the instructions.

[14] To participate in our mTurk human intelligence task (HIT), mTurk workers must have completed at least 100 prior HITs with an approval rate of 95% or greater. There were no other filters, though some participants failed a browser compatibility check at the beginning of our experiment. While our experiment was live, we monitored the popular mTurk worker forums (e.g., TurkerHub, mTurkCrowd) and found no threads discussing our experiment to any other degree than that it was a good opportunity to earn money.

[15] Subjects in the in-person experiments were additionally paid a flat fee of \$6 to comply with laboratory standards. To avoid income effects, this additional payment was not mentioned until the end of the experiment.

ity curvature in the gain domain, as well as loss aversion. We make two major adjustments to elicit additional preference motives and make fewer assumptions. First, we change the stakes and probabilities to permit a greater range of elicitable utility curvature and probability weighting. Second, we add three additional lottery tables. The first two allow us to elicit utility curvature and probability weighting in the loss domain, separately from those measures in the gain domain. The third table allows us to elicit a preference for (or against) certainty in the sense of Schmidt (1998). Our preference elicitation task thus includes six lottery tables.

We outline the stakes and probabilities for each set of lotteries in Tables 2 and 3. These lotteries involve potential gains between \$0.50 and \$60.00 and potential losses between \$0.10 and \$4.00 (taken from the \$5.00 typing task earnings). In the experiment, each of the six tables is provided on one screen, with probabilities illustrated with urns containing colored balls.[16] Each row indicates the question number for possible selection at the end of the experiment, and subjects click radio buttons between the lotteries to indicate their choice. To encourage subjects to think carefully about their choices, the "Next" button is hidden for twenty seconds. We randomize the order of the six tables, the left-right display of the lotteries, and whether the stakes are ascending or descending.

There are a total of 101 choices across the six tables. To lower the cognitive load from making so many decisions, we automate monotonic choices—if a subject switches from Lottery A to Lottery B, we automatically select Lottery B for all subsequent choices on the screen. If the subject switches a Lottery B choice back to Lottery A, we automatically switch all higher choices back to Lottery A as well. We also simplify the choice sets by specifying the same probabilities for Lottery A and Lottery B (except for the CP table, where one lottery is a certain outcome).

Each lottery table, with the exception of the certainty preference table, includes a choice with a first-order stochastically dominated (FOSD) option (highlighted in italics in Tables 2 and 3). Choosing these options may indicate a lack of attention by participants. We construct an indicator for whether a subject violated FOSD in any of these five choices and use this as a proxy for inattentive participants. In Table 4, we report the Spearman rank correlations of the lottery choices. We designed the tables so that they measure different aspects of the preference functional, so the correlations are relatively low. One might be concerned that low correlations are the result of random choice. The high correlation between the choices in the GD1 and CP lottery tables, however, indicates deliberate behavior by subjects, as the lotteries in these two tables are similar.

We use the subjects' answers in the preference task to construct the "nonparametric" and "parametric" preference measures briefly described in Section 2.1. Below, we describe the construction of both measures from the choice tables and the resulting choice patterns in detail. In

[16] Screenshots of all decision screens are provided in Appendix G.

Table 2: Preference motive elicitation tables for utility curvature and probability weighting in the gain and loss domain

	Gain Domain Table 1 (GD1)					**Gain Domain Table 2 (GD2)**			
	Lottery A		Lottery B			Lottery A		Lottery B	
p	0.2	0.8	0.2	0.8	*p*	0.9	0.1	0.9	0.1
$	*2.50*	*2.00*	*2.50*	*1.00*	$	*2.00*	*1.50*	*2.00*	*0.50*
	2.50	2.00	4.50	1.00		2.00	1.50	2.05	0.50
	2.50	2.00	4.75	1.00		2.00	1.50	2.10	0.50
	2.50	2.00	5.00	1.00		2.00	1.50	2.15	0.50
	2.50	2.00	5.50	1.00		2.00	1.50	2.20	0.50
	2.50	2.00	6.00	1.00		2.00	1.50	2.25	0.50
	2.50	2.00	6.50	1.00		2.00	1.50	2.30	0.50
	2.50	2.00	7.00	1.00		2.00	1.50	2.35	0.50
	2.50	2.00	8.00	1.00		2.00	1.50	2.45	0.50
	2.50	2.00	9.00	1.00		2.00	1.50	2.55	0.50
	2.50	2.00	10.00	1.00		2.00	1.50	2.65	0.50
	2.50	2.00	12.00	1.00		2.00	1.50	2.80	0.50
	2.50	2.00	15.00	1.00		2.00	1.50	3.00	0.50
	2.50	2.00	20.00	1.00		2.00	1.50	3.25	0.50
	2.50	2.00	30.00	1.00		2.00	1.50	3.50	0.50
	2.50	2.00	60.00	1.00		2.00	1.50	3.75	0.50

	Loss Domain Table 1 (LD1)					**Loss Domain Table 2 (LD2)**			
	Lottery A		Lottery B			Lottery A		Lottery B	
p	0.1	0.9	0.1	0.9	*p*	0.8	0.2	0.8	0.2
$	*-0.75*	*-0.25*	*-0.75*	*-0.50*	$	*-1.75*	*-0.10*	*-1.75*	*-0.50*
	-1.20	-0.25	-0.75	-0.50		-1.95	-0.10	-1.75	-0.50
	-1.25	-0.25	-0.75	-0.50		-2.00	-0.10	-1.75	-0.50
	-1.30	-0.25	-0.75	-0.50		-2.05	-0.10	-1.75	-0.50
	-1.40	-0.25	-0.75	-0.50		-2.10	-0.10	-1.75	-0.50
	-1.50	-0.25	-0.75	-0.50		-2.15	-0.10	-1.75	-0.50
	-1.60	-0.25	-0.75	-0.50		-2.20	-0.10	-1.75	-0.50
	-1.70	-0.25	-0.75	-0.50		-2.30	-0.10	-1.75	-0.50
	-1.85	-0.25	-0.75	-0.50		-2.40	-0.10	-1.75	-0.50
	-2.00	-0.25	-0.75	-0.50		-2.50	-0.10	-1.75	-0.50
	-2.15	-0.25	-0.75	-0.50		-2.60	-0.10	-1.75	-0.50
	-2.35	-0.25	-0.75	-0.50		-2.75	-0.10	-1.75	-0.50
	-2.65	-0.25	-0.75	-0.50		-2.90	-0.10	-1.75	-0.50
	-3.00	-0.25	-0.75	-0.50		-3.05	-0.10	-1.75	-0.50
	-3.40	-0.25	-0.75	-0.50		-3.25	-0.10	-1.75	-0.50
	-4.00	-0.25	-0.75	-0.50		-3.50	-0.10	-1.75	-0.50

Note: A row in each of the four tables above represents a choice set presented to the subject. Row values are the possible dollar outcomes from the displayed lotteries. Column headings are the probability of obtaining the given outcome. Losses and gains are relative to the $5.00 earned in the real effort task to begin the experiment. Italicized rows indicate a stochastically-dominated choice.

Table 3: Preference motive elicitation tables for certainty preference and loss aversion

	Certainty Preference Table (CP)				**Loss Aversion Table (LA)**			
	Lottery A	Lottery B			Lottery A		Lottery B	
p	1.00	0.20	0.80	$	+0.50	−0.20	+5.00	−2.00
$	2.00	2.50	1.00	*p*	*0.00*	*1.00*	*0.00*	*1.00*
	2.00	4.50	1.00		0.05	0.95	0.05	0.95
	2.00	4.75	1.00		0.10	0.90	0.10	0.90
	2.00	5.00	1.00		0.15	0.85	0.15	0.85
	2.00	5.50	1.00		0.20	0.80	0.20	0.80
	2.00	6.00	1.00		0.25	0.75	0.25	0.75
	2.00	6.50	1.00		0.30	0.70	0.30	0.70
	2.00	7.00	1.00		0.35	0.65	0.35	0.65
	2.00	8.00	1.00		0.40	0.60	0.40	0.60
	2.00	9.00	1.00		0.45	0.55	0.45	0.55
	2.00	10.00	1.00		0.50	0.50	0.50	0.50
	2.00	12.00	1.00		0.55	0.45	0.55	0.45
	2.00	15.00	1.00		0.60	0.40	0.60	0.40
	2.00	20.00	1.00		0.65	0.35	0.65	0.35
	2.00	30.00	1.00		0.70	0.30	0.70	0.30
	2.00	60.00	1.00		0.75	0.25	0.75	0.25
					0.80	0.20	0.80	0.20
					0.85	0.15	0.85	0.15
					0.90	0.10	0.90	0.10
					0.95	0.05	0.95	0.05
					1.00	*0.00*	*1.00*	*0.00*

Note: A row in each of the two tables above represents a choice set presented to the subject. In the certainty preference table (left), row values are the possible dollar outcomes and column headings are the probabilities of obtaining the given outcome. In the loss aversion table (right), column headings are the dollar outcomes and row values are the probabilities of obtaining each outcome. Losses and gains are relative to the $5.00 earned in the real effort task to begin the experiment. Italicized rows indicate a stochastically-dominated choice.

Table 4: Spearman rank correlations of lottery choices

	GD1	GD2	LD1	LD2	LA
GD1	1.000	.	.	.	.
GD2	0.046	1.000	.	.	.
LD1	-0.266	0.110	1.000	.	.
LD2	-0.010	-0.161	0.148	1.000	.
LA	0.197	0.202	0.011	0.037	1.000
CP	0.605	0.015	-0.231	0.028	0.182

Note: For our sample size of 1,730 subjects, a correlation is statistically significant at the 10% level if it is larger than 0.040, at the 5% level if it is larger than 0.047, and at the 1% level if it is larger than 0.062 (in absolute values).

Appendix B, we offer formal derivations of each preference measure under our elicitation procedure.

3.2.1 Nonparametric Preference Measures

In the nonparametric approach, we refrain from making further structural assumptions on the elements of Equations (1) through (4) and use a *ceteris paribus* approach for the elicitation. That is, we assume for the sake of measuring each preference motive that choices in the corresponding tables are guided *only* by that preference motive. For example, we assume that a subject has a more concave utility function if he chooses the less risky Lottery A in tables GD1 and GD2 more often. The nonparametric measure of gain-domain utility curvature (UC^+) is constructed as $UC^+ = GD1_A + GD2_A$. (For reference, the captions in Figure 2 summarize the construction and interpretation of each measure.) The only necessary structural assumption on the gain-domain utility function u^+ is that there is a single parameter regulating its global risk aversion. Under the corresponding assumptions, more concave loss-domain utility is implied by more Lottery B choices in LD1 and LD2 and thus $UC^- = LD1_B + LD2_B$.

For the nonparametric measure of probability weighting, we assume that the probability weighting function has an inflection point towards the middle of the unit interval and is either S shaped or inverse-S shaped. For the gain domain, the shape is dictated by a single parameter such that the degree of inverse-S shape of the function can be measured by adding those choices in the gain domain tables which have a better outcome with a low probability (i.e., which would be overweighted if the probability weighting function was inverse-S shaped). The measure is calculated as $PW^+ = GD1_B + GD2_A$. The corresponding loss domain measure is determined according to the same logic, so $PW^- = LD1_B + LD2_A$.

To measure a preference for or against certainty, we adopt a procedure similar to that of Callen et al. (2014), who recently highlighted this preference motive. Subjects face table CP which is equal to table GD1 except that the less risky Lottery A in table CP is a certain payment of the lower outcome of Lottery A in table GD1. If a subject makes more Lottery A choices in CP than in GD1, we can conclude that an outcome paid with certainty has a higher value than that outcome has in a lottery setting. The nonparametric measure for certainty is calculated as $CP = CP_A - GD1_A$.

Our loss aversion table LA differs from the other preference tables—rather than varying outcomes along its rows, the table varies the probabilities of fixed outcomes. This allows us to elicit loss aversion from a greater possible range. Lottery A involves a small gain and a small loss, while both gain and loss are multiplied by 10 for Lottery B. The probability of the gain increases down the rows, increasing the difference in expected value between the lotteries. Under our *ce-*

teris paribus approach, the more loss averse a subject is, the larger will be the required difference in expected value to switch from Lottery A to Lottery B, so $LA = LA_A$.

Figure 2 shows the distributions of our nonparametric preference measures based on the subjects' choices in the experiment. The distributions show few extreme or corner values, and most seem to be single-peaked. Recall that the gain (loss) utility curvature measure is a simple sum of the number of safe choices in tables GD1 and GD2 (LD1 and LD2). If we had chosen a too-narrow spectrum of the riskiness of the lotteries in those tables, we would observe high frequencies of subjects with UC^+ and UC^- at either 0 or 32. While such values occur, they are infrequent compared to other choice patterns and alternative measurement scales in the literature. The measures of PW^+, PW^-, and CP subtract the choices in two tables and thus could lead to a bunching around 0 if subjects chose the same number of safe choices in multiple tables. While such an effect seems to occur in 10–20% of the probability weighting measures, more than 30% of subjects have a CP value of 0. Such an answer pattern in CP, however, is not a sign of inattentive choices by the subjects. In the CP measure, a value of 0 implies that subjects have no distinct preference for or against certainty, which makes a concentration of answers on this point natural. The nonparametric measure of loss aversion results from a single lottery table, which may raise concerns about corner solutions and a tendency of subjects to switch in the middle of the table. Both tendencies can be observed to a small degree but neither of them appear to dominate the choices of most subjects.

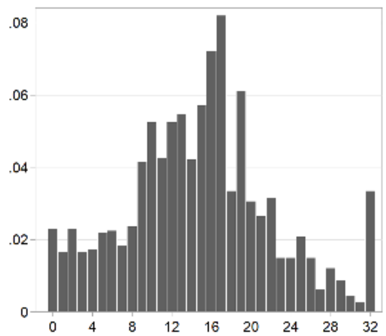
3.2.2 Parametric Preference Measures

For the parametric preference measures, we assume the functional forms detailed in Equations (5) and (6). With the help of these functional forms, a subject's choices in the preference measurement tables imply six indifference statements with six open parameters (γ^+, γ^-, β^+, β^-, λ and κ).[17] The tables are designed such that any combination of choices in the six preference measurement tables will lead to a solvable system of equations as long as the subject did not make a choice that violated FOSD.

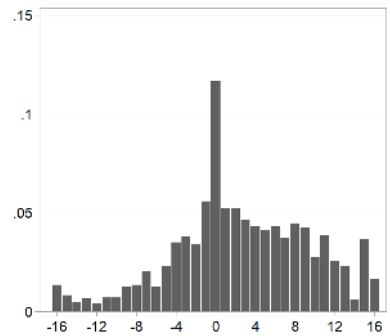
We solve the system by first determining (γ^+, β^+) and (γ^-, β^-) from the indifference relation implied by the choices in the gain domain tables and the loss domain tables, respectively. The implied values from the possible combinations of answers both in the gain domain and in the loss domain are displayed in Figure 3. Here, the upward-sloping lines indicate choices in the table

[17] We make the assumption that switching from Lottery A in one line to Lottery B in the next line implies indifference for a hypothetical lottery statement between those two lines. Consider, for example, a subject who chooses Lottery A in the first four rows of GD1 and Lottery B in all subsequent rows. This implies the preference relations $(2.50, 0.8, 2.00) \succ (5.00, 0.2, 1.00)$ and $(2.50, 0.8, 2.00) \prec (5.50, 0.2, 1.00)$. By our assumption, it also implies $(2.50, 0.8, 2.00) \sim (5.25, 0.2, 1.00)$.

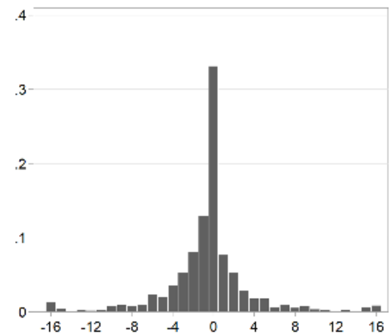
Figure 2: Distributions of nonparametric preference measures



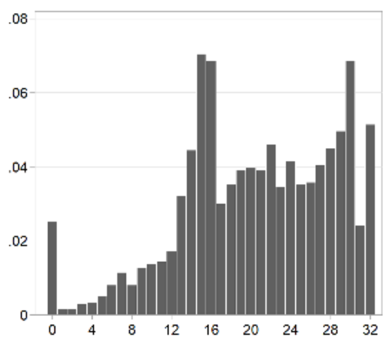
(a) Utility curvature gains
$UC^{+} = GD1_A + GD2_A$
Higher → more concave



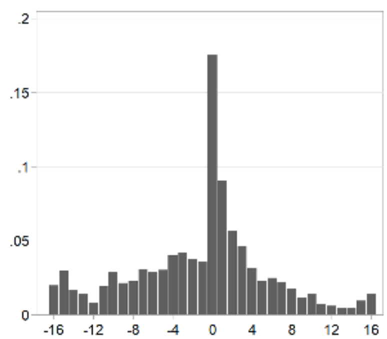
(b) Probability weighting gains
$PW^{+} = GD1_B + GD2_A$
Higher → more inverse-S



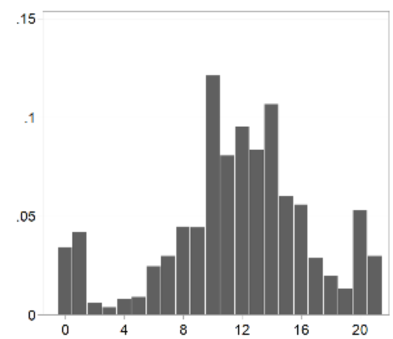
(c) Certainty preference
$CP = CP_A - GD1_A$
Higher → stronger preference



(d) Utility curvature losses
$UC^{-} = LD1_B + LD2_B$
Higher → more concave



(e) Probability weighting losses
$PW^{-} = LD1_B + LD2_A$
Higher → more inverse-S

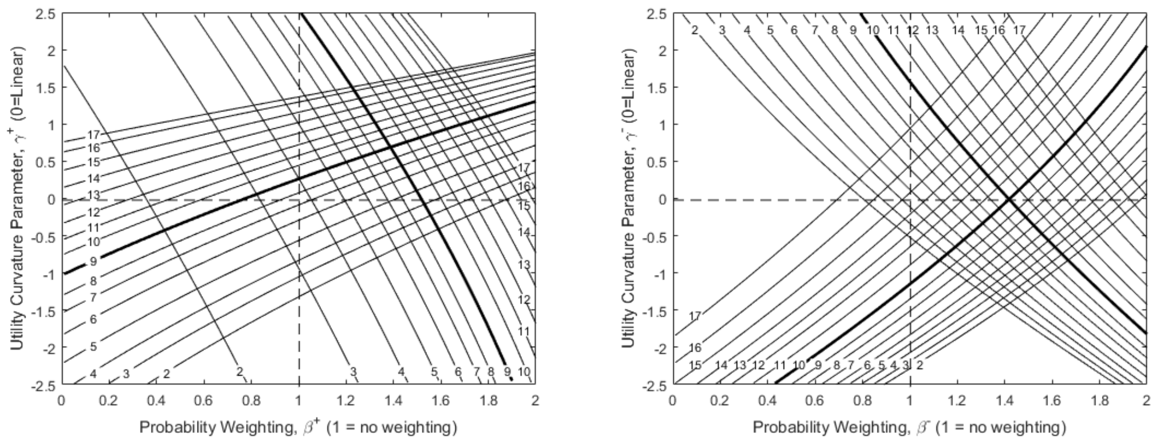


(f) Loss aversion
$LA = LA_A$
Higher → more loss averse

Note: Histograms include observations for all 1,730 subjects (both online and in-person) who completed the experiment. In each histogram, the y-axis is the fraction of subjects. Note that the scales of the y-axes differ for the different preference motives.

GD1 (LD2) and downward-sloping lines indicate choices in GD2 (LD1). We designed the tables such that switching in the middle is consistent with common elicited values of the respective parameters, and we indicate these choices with thick lines. The lines for the tables differ in the sign of their slope because in tables GD1 and LD2 the safer lottery has a worse outcome in the low-probability event than the riskier lottery while in tables GD2 and LD1, the safer lottery has a better outcome in the low-probability event than the risky lottery. As such, more safe choices in GD1 and LD2 can either be caused by more concave utility or by less inverse-S probability weighting, while more safe choices in GD2 and LD1 can either be caused by more concave utility or by more inverse-S probability weighting. The difference in slopes along with the chosen parameters for

Figure 3: Range of elicitable parameters for utility curvature and probability weighting



(a) Possible gain domain preferences implied by the number of safe choices in tables GD1 (upward-sloping lines) and GD2 (downward-sloping lines).

(b) Possible loss domain preferences implied by the number of safe choices in tables LD1 (downward-sloping lines) and LD2 (upward-sloping lines).

Note: The figure displays the parameters for utility curvature and probability weighting implied by the choices in tables GD1 and GD2 (panel(a)) and LD1 and LD2 (panel (b)). The number on each line denotes the number of safe lottery choices in the respective table. The intersection of two lines sets the parameter pair.

the tables ensure that each choice combination has exactly one intersection in Figure 3 and thus implies a unique utility curvature/probability weighting pair.

As can be seen in Figure 3, the preference measurement tables are designed with two goals in mind. First, we want to identify a sufficiently broad range of parameter values such that utility could be both concave and convex and that probability weighting could be both inverse-S and S shaped. Second, we want to minimize the required number of decisions while also limiting the number of corner choices made by subjects. We thus calibrate our tables such that the implied parameters are centered around previous measures from the literature (e.g., Harbaugh et al., 2010; Tanaka et al., 2010; Tversky and Kahneman, 1992). Choices in the middle of the tables thus imply concave utility and inverse-S shaped probability weighting in the gain domain (panel (a)) and linear utility and inverse-S shaped probability weighting in the loss domain (panel (b)). Once γ^+, γ^-, β^+ and β^- are calibrated from the choices in GD1, GD2, LD1 and LD2, we use the indifference relationships implied by choices in CP and LA to calculate κ and λ as well as the corresponding indices $\hat{\kappa}$ and $\hat{\lambda}$. Because these two parameters are dependent on the prior values of utility curvature and probability weighting, there is no clear range for them preset by the choice tables.

Distributions of the parametric preference measures are displayed in Figure 4. As with the nonparametric measures, they show relatively few corner solutions and subjects do not seem to

Table 5: Summary statistics for parametric preference measures

	UW Only $n = 326$		mTurk Only $n = 950$		Full Sample $n = 1,276$	
	Mean	Median	Mean	Median	Mean	Median
γ^+	0.38	0.45	0.37	0.39	0.38	0.39
β^+	1.36	1.46	1.47	1.54	1.44	1.53
$\hat{\kappa}$	0.06	0.01	0.23	-0.01	0.19	-0.00
γ^-	0.64	0.47	0.72	0.68	0.70	0.59
β^-	1.29	1.29	1.38	1.42	1.36	1.41
$\hat{\lambda}$	3.22	2.47	4.25	2.79	3.98	2.68

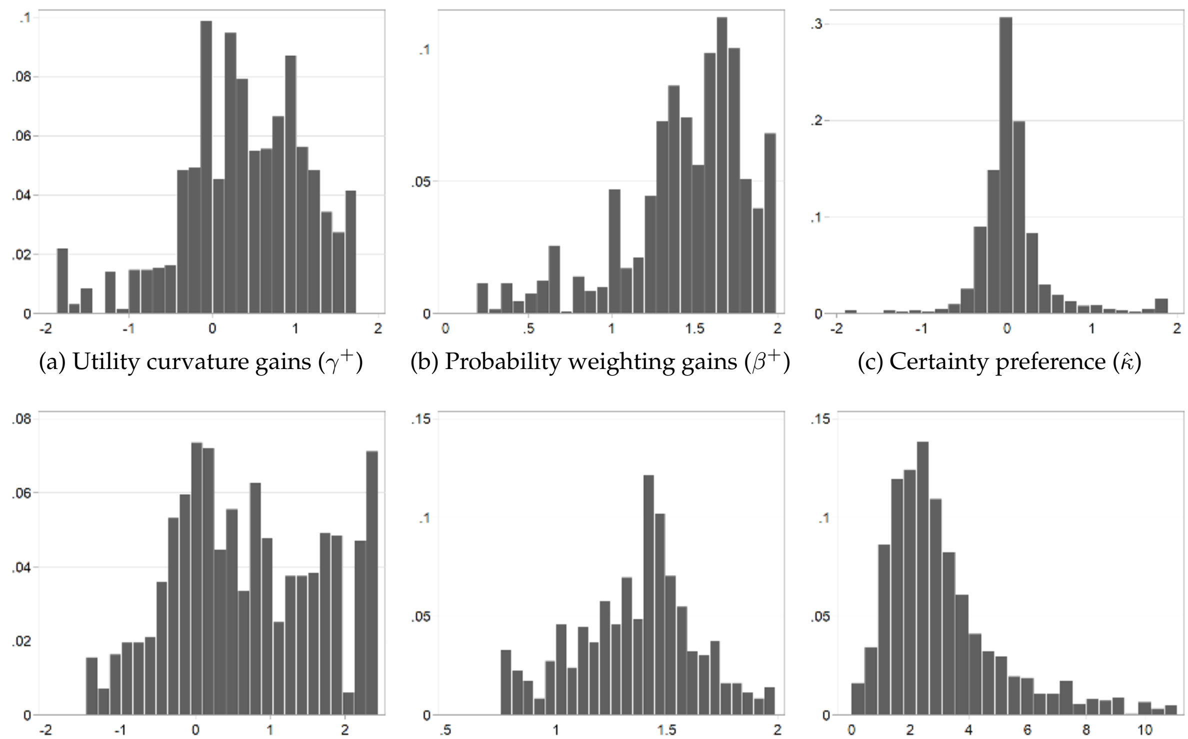
Note: The parameters $\gamma, \beta, \hat{\kappa}$, and $\hat{\lambda}$ denote our measures of utility curvature, probability weighting, certainty preference, and loss aversion, respectively. Superscripts $^+$ ($^-$) indicate that the parameter was elicited in the gain (loss) domain. The number of subjects summarized above is less than the total 1,730 experiment subjects because parametric measures are not calculable for subjects who make FOSD-violating choices. Due to the elicitation procedure and its associated calculations, some extreme values could appear for both $\hat{\kappa}$ and $\hat{\lambda}$. To limit the influence of such outliers on our results, we winsorize both measures at the 99$^{\text{th}}$ percentile and $\hat{\kappa}$ values additionally at the 1$^{\text{st}}$ percentile. We report summary statistics for the nonparametric preference measures in Appendix C.

be biased strongly towards choosing in the middle of the table. The distributions further appear single-peaked. In the gain domain, we observe concave utility functions ($u^+ > 0$) and inverse-S shaped probability weighting functions ($\beta^+ > 1$) for most subjects. In the loss domain, probability weighting is commonly inverse-S shaped ($\beta^- > 1$), but fewer subjects show concave utility ($u^- > 0$). Nevertheless, the majority of subjects still display risk aversion due to utility curvature in the loss domain, which makes our results at odds with the common assumptions in models such as cumulative prospect theory. As expected from the nonparametric CP measure, the index of marginal preference for certainty is near zero for most subjects, indicating that this preference motive plays a minor role.[18] We report means and medians of the parametric preference measures in Table 5.[19] The means and medians for the mTurk sample and the UW sample appear very similar. Overall, the median subject has concave utility and inverse-S shaped probability weighting in both domains, is loss averse with $\hat{\lambda}$ being about 2.5, and does not have a distinct preference for or against certainty. These parameters are broadly consistent with other estimates in the literature.

[18] Other studies have found significant preference for certainty among their subjects (e.g., Callen et al., 2014; Stango and Zinman, 2019). However, Vieider (2018) suggests that this could stem from randomness in subject responses. We are susceptible to this criticism as well—if choices were made randomly, roughly 65% of our subjects would be classified as having a preference for certainty. This is evident in Table 5, which shows that the mTurk sample (which is susceptible to more noisy decision making) demonstrates a slightly higher preference for certainty on average than the in-person sample. In addition, our elicitation procedure varies payouts while the elicitation procedure of Callen et al. (2014) varies probabilities, so differences in the elicitation tasks might also be the cause of the discrepancy in results.

[19] Summary statistics of the nonparametric measures are not particularly informative because the motives are ordinally scaled. We provide those summary statistics in Table C.2 of Appendix C. We also report the pairwise Pearson correlations of the nonparametric preference measures in Table C.3.

Figure 4: Distributions of parametric preference measures



(a) Utility curvature gains (γ^+) (b) Probability weighting gains (β^+) (c) Certainty preference ($\hat{\kappa}$)

(d) Utility curvature losses (γ^-) (e) Probability weighting losses (β^-) (f) Loss aversion ($\hat{\lambda}$)

Note: Histograms include observations for the 1,276 subjects (both online and in-person) who did not violate FOSD in their preference measurement choices. In plots (c) and (f), a small number of outliers make the central part of the distribution difficult to see. For display purposes, we exclude outliers from these plots. Certainty preference is plotted from -2 (2$^{\text{nd}}$ percentile) to 2 (96$^{\text{th}}$ percentile). We also omit loss aversion parameters above 11 (95$^{\text{th}}$ percentile). In each histogram, the y-axis is the fraction of subjects. Note that the scales of the y-axes differ for the different preference measures.

We report the pairwise Pearson correlations of the parametric preference measures in Table 6. For these and all other inductive statistics using the preference measures, we standardize the measures such that they are measured in units of standard deviations (indicated by the subscript std). As would be expected from a behavioral perspective, the probability weighting parameters are positively and significantly correlated between the two domains ($\rho = 0.338$). For utility curvature, we observe a similarly strong, but negative correlation between domains ($\rho = -0.230$). This is interesting as it provides evidence that there is an underlying motive of diminishing marginal sensitivity for some subjects (Tversky and Kahneman, 1992). However, this motive is not strong enough to lead to convex utility in the loss domain for most subjects. We see that a more concave utility function is typically accompanied by more inverse S probability weighting ($\rho^+ = 0.288, \rho^- = 0.249$), while loss aversion is negatively correlated with probability weighting and has a mixed association with utility curvature.

Table 6: Correlation table for standardized parametric preference measures

		γ^+_{std}	β^+_{std}	$\hat{\kappa}_{std}$	γ^-_{std}	β^-_{std}	$\hat{\lambda}_{std}$
Utility curvature gains	γ^+_{std}	1.000	.	.	.	.	.
Probability weighting gains	β^+_{std}	0.288	1.000	.	.	.	.
Certainty preference	$\hat{\kappa}_{std}$	-0.043	0.039	1.000	.	.	.
Utility curvature losses	γ^-_{std}	-0.230	0.008	-0.045	1.000	.	.
Probability weighting losses	β^-_{std}	-0.002	0.338	-0.007	0.249	1.000	.
Loss aversion index	$\hat{\lambda}_{std}$	-0.292	-0.258	-0.004	0.195	-0.101	1.000

Note: Preference motives are standardized by subtracting the mean and dividing by standard deviation. For our sample size of 1,276 subjects with estimable parametric preferences, a correlation is statistically significant at the 10% level if it is larger than 0.047, at the 5% level if it is larger than 0.055, and at the 1% level if it is larger than 0.073 (in absolute values). We report correlations for the nonparametric preference measures in Appendix C.

The calibrated parameters are also used in the structural models introduced in Section 2.2. As stated above, some of the models (namely RDEU and CPT) can utilize the elicited preference parameters directly, while others restrict some elements of the preference functional (such as setting $\beta = 1$ in the case of EU). For these restrictive models, we derive model-specific parameters taking the restriction(s) into account. Applying such restrictions can, however, lead to situations in which two choices by a subject imply different values for the same preference parameter. For example, in EU, a choice in GD1 can imply one value of γ^+_{EU}, while a choice in GD2 can imply a different value of γ^+_{EU}. A similar inconsistency can occur in DT models with conflicting estimates of β^+_{DT}. If such inconsistencies arise, we calculate the parameter separately for each relevant lottery table and use the average parameter value in the structural model.

Eliciting preference parameters from choice lists allows for a quick elicitation procedure based on relatively few choices of the subjects. One potential drawback is that using few choices to elicit risk preferences can lead to inference from noise in the decision process. Alternative procedures use a larger set of choices and base preference inference on maximum likelihood estimation (e.g., l'Haridon and Vieider, 2019). Such procedures, however, have the risk of leading to fatigue by experimental subjects which is particularly a concern for experiments carried out online. We thus follow much of the recent literature (Callen et al., 2014; Sprenger, 2015; Tanaka et al., 2010) and use fewer choices with the potential drawback of over-interpreting noise. We discuss the robustness of our main findings in light of the potential for noisy decisions by subjects at the end of Section 4.2.

3.3 Insurance Task

In the insurance task, subjects make twelve insurance decisions. For each decision, they face the potential to lose \$3 of their \$5 typing task earnings under various loss probabilities. Loss proba-

Table 7: Insurance scenarios faced by the experimental subjects.

		Loading Factor				
		0.80	1.00	1.25	1.50	2.50
Probability of loss	5%				×	×
	10%		×	×	×	×
	20%			×	×	
	40%				×	
	70%	×	×	×		

Note: Subjects faced twelve insurance scenarios with a potential loss of $3 and the probability of loss and loading factor displayed above.

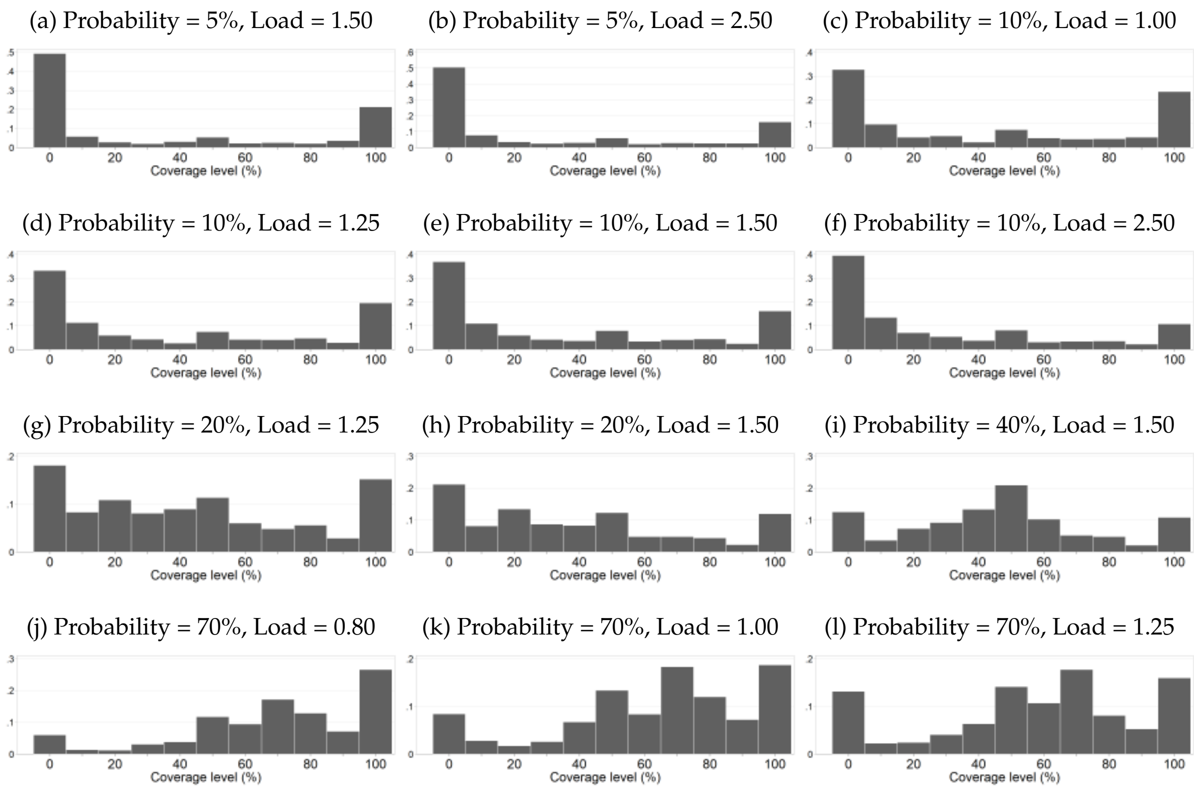
bilities are stated in percentage terms and visualized with an urn of 20 white and red balls, where drawing a red ball results in a loss. Subjects can insure against the potential loss by purchasing a level of insurance ranging from 0% to 100%, in one-percent increments. In the event of a loss, the insurance covers the selected percentage of the $3 loss.[20] Subjects change their insurance choice by moving a slider on the screen (see screenshots of the task starting on page 29 of Appendix G). Every time the slider is set to a value, the screen displays the chosen level of insurance coverage, the corresponding insurance premium, and the covered and uncovered loss amounts in case of a loss.

We present the twelve insurance scenarios in random order. These scenarios differ in the probability of loss and the relative loading charged for the insurance policy. The combinations of loss probability and loading in our experiment are outlined in Table 7. In each insurance scenario, we hide and disable the "Next" button for ten seconds to encourage subjects to explore various coverage levels.

We plot the distributions of coverage levels chosen for each scenario in Figure 5. It is apparent that insurance for low-probability losses tends to result in subjects choosing corner solutions of full or no insurance. For the scenarios with a 5% (10%) loss probability, 36.8% (27.1%) of choices are for 0% coverage and 18.6% (17.5%) are for 100% coverage. This is consistent with prior studies showing a bimodal distribution of insurance demand at low probabilities (McClelland et al., 1993). As loss probabilities (and thus absolute premium amounts) increase, subjects select corner solutions less often—at 70% loss probability, 71.7% of choices were for partial coverage. Insur-

[20] One could be concerned that this response format, while being consistent with theoretical analyses of optimal insurance demand (see Schlesinger, 2013), is different from those insurance decisions commonly observed in real insurance markets where individuals typically choose between a small set of available policies. We also conduct the main analyses of this paper treating insurance demand more discretely by (i) rounding subjects' replies into bins of 25% coverage and (ii) rounding them into either no or full insurance coverage. For both discrete formats, we make a new set of structural model predictions. Results are consistent with the main text across all analyses. A more detailed description of these analyses and their results is given in Appendix F.

Figure 5: Distributions of observed coverage levels chosen by scenario



Note: Histograms include observations for all 1,730 subjects (both online and in-person) who completed the experiment. In each histogram, the y-axis is the fraction of subjects. Note that the scales of the y-axes differ for the different scenarios.

ance demand also appears to decrease as loading increases—the mean coverage level is 71.3% at a load of 0.80 and decreases monotonically to a mean of 31.5% at a load of 2.50. Interestingly, when loading was actuarially fair or better, subjects chose partial (no) insurance 65.1% (12.1%) of the time.

In Table 8, we report the Spearman rank correlations of the coverage level selected for each insurance scenario. All coverage level choices are positively correlated with each other. Choices over scenarios with a common probability generally have the highest correlation. Scenarios with common loading factors do not appear to be substantially more correlated than scenarios with different loading factors. For example, among the scenarios with a loading factor of 1.50, the scenario with loss probability of 5% has a high correlation with the 10% loss probability scenario. However, as the loss probability is increased to 20% or 40%, the correlation with the 5% scenario steadily decreases. This observation illustrates an important point to note from the correlation table—correlations generally decrease as the difference in loss probabilities increases.

Table 8: Spearman rank correlations of coverage levels

Loss prob	5%		10%				20%		40%	70%	
Loading	1.50	2.50	1.00	1.25	1.50	2.50	1.25	1.50	1.50	0.80	1.00
	P5L150	P5L250	P10L100	P10L125	P10L150	P10L250	P20L125	P20L150	P40L150	P70L080	P70L100
P5L150	1.000	.	.	.	.	.	.	.	.	.	.
P5L250	0.729	1.000	.	.	.	.	.	.	.	.	.
P10L100	0.676	0.631	1.000	.	.	.	.	.	.	.	.
P10L125	0.699	0.670	0.706	1.000	.	.	.	.	.	.	.
P10L150	0.709	0.705	0.689	0.739	1.000	.	.	.	.	.	.
P10L250	0.660	0.706	0.641	0.687	0.719	1.000	.	.	.	.	.
P20L125	0.583	0.566	0.628	0.677	0.631	0.623	1.000	.	.	.	.
P20L150	0.576	0.564	0.632	0.686	0.654	0.646	0.732	1.000	.	.	.
P40L150	0.347	0.329	0.386	0.432	0.414	0.433	0.557	0.591	1.000	.	.
P70L080	0.102	0.068	0.211	0.189	0.139	0.149	0.279	0.296	0.443	1.000	.
P70L100	0.076	0.048	0.147	0.131	0.120	0.154	0.233	0.289	0.495	0.644	1.000
P70L125	0.035	0.053	0.066	0.107	0.106	0.113	0.204	0.255	0.485	0.519	0.642

Note: Insurance scenarios are labeled in the first column and the third row as P[probability]L[loading * 100]. For our sample size of 1,730 subjects, a correlation is statistically significant at the 10% level if it is larger than 0.040, at the 5% level if it is larger than 0.047, and at the 1% level if it is larger than 0.062 (in absolute values).

To illustrate the demand response to price more explicitly, we plot average insurance demand as a function of the loading in Figure 6. Average insurance demand is downward sloping in price at all loss probabilities, showing general adherence to the law of demand by our subjects. Further, insurance demand increases with loss probability at all loads except for a load of 2.5, where average insurance demand for loss probabilities of 5% and 10% is virtually identical. These graphical observations are corroborated by statistical analysis. A simple pooled OLS model estimated on all insurance choices (Table 9) shows statistically significant effects of both loading factor (negative) and loss probability (positive) on insurance demand.

Figure 6: Insurance demand reactions to loading and loss probabilities

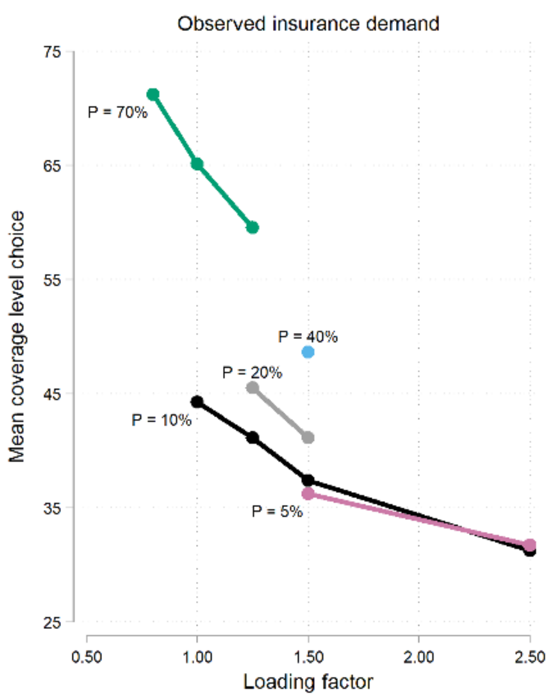


Table 9: Regression of coverage level on loss probability and insurance loading

	(1)
Probability of loss	0.37***
	(0.02)
Loading factor	−8.18***
	(0.40)
Constant	47.75***
	(1.24)
R^2	0.11
N choices	20,760
N subjects	1,730

Note: Dependent variable is the coverage level selected by the subject in the given insurance scenario. Coverage level ranges from 0 to 100 with a mean of 46.2. Probability of loss takes values 5, 10, 20, 40, or 70. Loading factor takes values 0.80, 1.00, 1.25, 1.50, or 2.50. Standard errors clustered by subject are in parentheses. Stars *, **, and *** denote statistical significance at the 0.10, 0.05, and 0.01 levels, respectively.

4 Results

4.1 Predicting Insurance Demand from Preference Motives

We now consider whether the preference motives are correlated with insurance demand. For this purpose, we run a series of linear regressions with coverage level as the dependent variable (integers between 0 and 100).[21] In the first analysis, we regress the coverage level on each preference motive separately, clustering standard errors by subject and correcting p-values for multiple hypothesis tests (Table 10).[22] We include a fixed effect for each of the twelve insurance scenarios, so the coefficient estimates indicate whether a given preference motive can explain variation of the subjects' insurance demand within each scenario. The results indicate that utility curvature in both domains, loss domain probability weighting, and loss aversion have statistically significant correlations with insurance demand, whether they are measured nonparametrically or parametrically. For probability weighting in the gain domain, we also observe a significant positive effect when it is measured parametrically. A preference for certainty does not seem to have any significant effect on insurance demand. These results are in accordance with theory for most behavioral models which have been proposed in the literature, as long as these models do not feature a preference for certainty.

These regressions, however, do not account for the correlation between different preference motives. Particularly for the parametric measures, which are calibrated dependent on the values of the other parameters, this can pose a significant bias for inference. To address this, our next analysis specifies regressions in which all preference motives are included together. As with the single-preference regressions, we include insurance scenario fixed effects and cluster standard errors on the subject level.

We report the results of these estimations in Table 11, with the nonparametric measures in columns (1) and (2) and the parametric measures in columns (3) and (4). In column (1), loss domain probability weighting and loss aversion significantly correlate with insurance demand. The linear term for probability weighting, however, does not correctly capture the theoretical prediction that those with S-shaped probability weighting should overweight lower probabilities but underweight higher probabilities. To allow for this dynamic, in column (2) we introduce an interaction between probability weighting measures and an indicator for whether the probability of loss was greater or equal to 40%. Consistent with theoretical predictions for inverse-S probability weighting, we find that probability weighting has a positive and significant correlation with insurance demand for scenarios with a low probability of loss. The effect is offset once the probability of loss is greater than or equal to 40%. This construction also leads to a negative and significant ef-

[21] It could be argued both theoretically and from the empirical distributions that the dependent variable insurance demand is left- and right-truncated. When this is taken into account via a Tobit estimator, the results remain virtually unchanged in sign and significance as can be seen in Appendix D.

[22] In all our regressions involving multiple hypotheses, we correct p-values using the Šidák (1967) method. Details of each adjustment are in the respective table note.

Table 10: Regression of coverage level on each standardized preference motive separately

Panel (a): Nonparam. measures	(1)	(2)	(3)	(4)	(5)	(6)
UC^{+}_{std}	1.36*					
	(0.62)					
PW^{+}_{std}		0.61				
		(0.61)				
CP_{std}			−0.63			
			(0.58)			
UC^{-}_{std}				1.49*		
				(0.63)		
PW^{-}_{std}					2.30***	
					(0.58)	
LA_{std}						2.09***
						(0.67)
FOSD violators	Yes	Yes	Yes	Yes	Yes	Yes
Fixed effects	Scenario	Scenario	Scenario	Scenario	Scenario	Scenario
Clustered SEs	Subject	Subject	Subject	Subject	Subject	Subject
R^2	0.11	0.11	0.11	0.11	0.11	0.11
N choices	20,760	20,760	20,760	20,760	20,760	20,760
N subjects	1,730	1,730	1,730	1,730	1,730	1,730

Panel (b): Param. measures	(7)	(8)	(9)	(10)	(11)	(12)
γ^{+}_{std}	1.70**					
	(0.68)					
β^{+}_{std}		1.92**				
		(0.69)				
$\hat{\kappa}_{std}$			−1.26			
			(0.78)			
γ^{-}_{std}				1.96**		
				(0.72)		
β^{-}_{std}					1.90**	
					(0.66)	
$\hat{\lambda}_{std}$						1.72*
						(0.80)
FOSD violators	No	No	No	No	No	No
Fixed effects	Scenario	Scenario	Scenario	Scenario	Scenario	Scenario
Clustered SEs	Subject	Subject	Subject	Subject	Subject	Subject
R^2	0.13	0.13	0.12	0.13	0.13	0.13
N choices	15,312	15,312	15,312	15,312	15,312	15,312
N subjects	1,276	1,276	1,276	1,276	1,276	1,276

Note: Dependent variable is the coverage level selected by the subject in the given insurance scenario. Preference motives are standardized by subtracting the mean and dividing by standard deviation. Parametric preference measures cannot be calculated for subjects who make FOSD-violating choices, so those subjects are excluded from regressions involving the parametric preference measures. For comparison, we also exclude FOSD violators from regressions using nonparametric preference measures, with results in Appendix E. All models contain fixed effects for each of the 12 insurance scenarios. In each panel, p-values have been adjusted using the Šidák (1967) method for the six hypotheses tested. Standard errors clustered by subject are in parentheses. Stars *, **, and *** denote statistical significance at the 0.10, 0.05, and 0.01 levels, respectively.

fect of gain domain probability weighting on insurance demand in high loss probability scenarios. There is, however, no significant effect in low probability scenarios.[23]

The results for the nonparametric preference motives are corroborated and extended by the results for the parametric preference motives, reported in columns (3) and (4) of Table 11. In these regressions, we only consider those subjects who do not violate FOSD because we cannot calibrate a preference functional for those who violate FOSD. For most preference motives, the coefficients increase in their value indicating a higher predictive power for the (standardized) parametric preference motives. The exceptions are the preference for certainty and probability weighting in the loss domain. However, when considering the estimation including the interaction structure in column (4), it is clear that probability weighting in both domains has the predicted effect on insurance demand. The biggest difference between the nonparametric and the parametric preference motives is the significant effect of gain domain utility curvature on insurance demand in the latter. The nonlinear transformation due to the parametric structure and the sample including only non-FOSD violating subjects increase the predictive power of this element.[24]

While the statistical significance of the coefficients in Tables 10 and 11 seems generally encouraging for the predictive power of behavioral decision theories on insurance demand, the estimated coefficients are modest in scale. Using coefficient estimates from Tables 9 and 11 allows for some back-of-the-envelope calculations regarding effect size. Increasing the loss aversion of a subject by one standard deviation has the same positive effect on insurance demand as decreasing the loading q by about 32 percentage points ($2.60/8.18$).

4.2 Predicting Insurance Demand from Structural Models

We next assemble the preference motives into the seventeen structural models outlined in Table 1 and see how well these models predict insurance demand. For each model, we calculate the optimal coverage level for each of the twelve insurance scenarios. Summary statistics for the observed insurance demand and for each model's prediction are given in Table 12. Cursory observation shows that while some models predict similar average demand levels (such as EU, DT, RDEU, and CPT^{-}), other models predict too much (KR, CPT$^{\text{NLIB}}$, and almost all models with a preference for certainty) or too little (EV) demand.

[23] One interpretation of this joint result is that there is a single underlying probability weighting motive which does not differ between the gain and loss domains. Our measures exhibit similar distributions between the two domains and correlate strongly with each other ($\rho = 0.338$, see Table 6), providing evidence of this relationship. Under this interpretation, our joint result implies a positive and significant effect of probability weighting on insurance demand at low probabilities and a negative and significant effect at high probabilities.

[24] For comparison, we conduct the nonparametric regressions from Tables 10 and 11 and restrict the sample to non-FOSD violating subjects. We report our estimations in Table E.1 of Appendix E. There is no substantial difference in results.

Table 11: Regressions of coverage level on all standardized preference motives jointly

Nonparam./Param. measures	Nonparametric (1)	(2)	Parametric (3)	(4)
$UC^+_{std}/\gamma^+_{std}$	0.81	0.81	2.34***	2.34**
	(0.69)	(0.69)	(0.74)	(0.74)
PW^+_{std}/β^+_{std}	−0.07	1.00	1.53	2.68**
	(0.68)	(0.85)	(0.76)	(0.95)
$CP_{std}/\hat{\kappa}_{std}$	−0.39	−0.39	−1.13	−1.13
	(0.65)	(0.65)	(0.79)	(0.79)
$UC^-_{std}/\gamma^-_{std}$	1.12	1.12	1.62	1.62
	(0.65)	(0.65)	(0.75)	(0.75)
PW^-_{std}/β^-_{std}	2.15***	3.28***	1.24	2.30*
	(0.62)	(0.79)	(0.72)	(0.93)
$LA_{std}/\hat{\lambda}_{std}$	1.91**	1.91**	2.60**	2.60**
	(0.69)	(0.69)	(0.87)	(0.87)
Prob ≥ 40		34.99***		37.05***
		(1.12)		(1.28)
Prob $\geq 40 \times PW^+_{std}/\beta^+_{std}$		−3.20***		−3.47***
		(0.83)		(0.96)
Prob $\geq 40 \times PW^-_{std}/\beta^-_{std}$		−3.41***		−3.18**
		(0.84)		(0.99)
FOSD violators	Yes	Yes	No	No
Fixed effects	Scenario	Scenario	Scenario	Scenario
Clustered SEs	Subject	Subject	Subject	Subject
R^2	0.12	0.12	0.14	0.14
N choices	20,760	20,760	15,312	15,312
N subjects	1,730	1,730	1,276	1,276

Note: Dependent variable is the coverage level selected by the subject in the given insurance scenario. Preference motives are standardized by subtracting the mean and dividing by standard deviation. Parametric preference measures cannot be calculated for subjects who make FOSD-violating choices, so those subjects are excluded from regressions involving the parametric preference measures. For comparison, we also exclude FOSD violators from regressions using nonparametric preference measures, with results in Appendix E. All models contain fixed effects for each of the 12 insurance scenarios. Column 1 contains only the nonparametric preference measures. Column 2 adds an interaction between the PW preferences and a dummy for "high probability" (Prob $\geq 40 = 1$ if the probability of loss is 40% or 70%). Columns 3 and 4 replicate columns 1 and 2 using the parametric preference measures. In each column, *p*-values have been adjusted using the Šidák (1967) method. Standard errors clustered by subject are in parentheses. Stars *, **, and *** denote statistical significance at the 0.10, 0.05, and 0.01 levels, respectively.

To assess each model's predictive validity more thoroughly, we first examine the correlation between the prediction and the empirically-observed insurance demand. The Pearson correlation coefficients over all 1,276 non-FOSD violating subjects are provided in column (1) of Table 13. The results show a positive and significant correlation between model predictions and observed demand for all decision models which do not feature probability weighting (EV, EU, and KR).

Table 12: Summary statistics of observed and predicted coverage level choices

Model	Mean	Median	SD	Percent 0% cov.	Percent 100% cov.
Observed	46.51	50.00	37.35	21.03	17.22
EV	16.67	0.00	31.18	75.00	8.33
EU^{+}	43.94	44.00	39.70	29.64	25.00
EU^{-}	38.47	22.00	40.93	40.22	23.92
DT^{+}	40.85	0.00	49.16	59.15	40.85
DT^{-}	44.87	0.00	49.74	55.13	44.87
KR	88.83	100.00	31.50	11.17	88.83
$RDEU^{+}$	55.61	93.00	47.57	38.63	48.75
$RDEU^{-}$	54.85	78.00	46.40	36.91	42.86
CPT^{-}	57.98	84.00	45.27	32.84	40.59
CPT^{NLIB}	73.61	100.00	38.31	16.02	54.37
EV_{CP}	48.21	12.00	48.76	46.86	39.17
DT^{+}_{CP}	70.43	100.00	45.44	29.38	52.32
DT^{-}_{CP}	70.77	100.00	45.28	29.05	52.43
EU^{+}_{CP}	81.58	100.00	37.59	14.45	74.42
EU^{-}_{CP}	83.20	100.00	34.95	11.47	74.07
$RDEU^{+}_{CP}$	69.42	99.00	44.65	27.23	42.05
$RDEU^{-}_{CP}$	71.16	99.00	43.06	23.97	43.16

Note: Table displays summary statistics of the observed insurance demand and the insurance demand predictions of the 17 analyzed structural models for the 1,276 non-FOSD violating subjects. Model predictions are made by choosing the coverage level which maximizes the respective target function given in Table 1.

Models which feature probability weighting do not significantly correlate with observed demand or even have a negative correlation. Adding the preference for certainty to any of the models decreases the correlation with observed demand. This corroborates the results in Section 4.1, which show CP to have little explanatory value. Lastly, the correlation coefficients show that using preference parameters elicited in the loss domain often leads to a slight increase in the explanatory value of a given structural model.

Positive and significant correlations show that some of the models can explain part of the variation in observed insurance demand. However, positive correlations do not imply predictive accuracy—for example, they ignore the absolute level of the prediction. An alternative procedure is to compare the models' predictions with randomly generated ones in a "horse race" (Gathergood et al., 2019). In our horse race (columns (2) to (4) of Table 13), we compare the predictive ability of three non-behavioral benchmarks to each of the structural models. In each column, we report the proportion of choices where the structural model denoted in the row outperforms a particular benchmark (i.e., is closer to the observed coverage level). The structural model is supe-

Table 13: Goodness of fit analysis for structural models

Model	(1) Correlation with observed	Horse Race (2) Random choice	(3) Overall average	(4) Scenario average
EV	0.261	0.461	0.469	0.398
EU^+	0.189	0.479	0.426	0.384
EU^-	0.205	0.470	0.429	0.385
DT^+	0.015	0.345	0.343	0.327
DT^-	−0.002	0.336	0.331	0.320
KR	0.053	0.319	0.320	0.291
$RDEU^+$	−0.064	0.331	0.314	0.304
$RDEU^-$	0.010	0.372	0.352	0.332
CPT^-	0.017	0.382	0.354	0.335
CPT^{NLIB}	−0.008	0.370	0.340	0.314
EV_{CP}	0.098	0.385	0.381	0.345
DT^+_{CP}	−0.025	0.312	0.310	0.293
DT^-_{CP}	−0.023	0.313	0.310	0.293
EU^+_{CP}	0.050	0.336	0.331	0.300
EU^-_{CP}	0.042	0.340	0.328	0.297
$RDEU^+_{CP}$	−0.069	0.311	0.299	0.285
$RDEU^-_{CP}$	−0.019	0.334	0.324	0.303

Note: For our sample size of 1,276 subjects for whom we can make parametric predictions, the correlation in column (1) is statistically significant at the 10% level if it is larger than 0.047, at the 5% level if it is larger than 0.055, and at the 1% level if it is larger than 0.073 (in absolute values). Table cells in the "Horse Race" columns are the frequency with which the specified model in each row predicts coverage level choice at least as well as the prediction rule in the column header. "Random choice" specifies a random uniform distribution of predicted coverage from 0 to 100, which we repeated 1,000 times for each choice (the average random prediction across repetitions ranged from 46.3% to 53.6%). "Overall average" specifies the average coverage level of 46.5% for every prediction. "Scenario average" specifies the average coverage level in each of the 12 scenarios as the prediction for all choices in that scenario. This ranges from 30.8% in the scenario with 10% loss probability and a loading of 2.5 to 73.5% in the scenario with 70% loss probability and a loading of 0.8.

rior if the proportion exceeds 0.5. Our benchmark in column (2) is choosing a random coverage level from a uniform distribution, which we perform 1,000 times for each choice. None of the seventeen models we analyze can predict insurance demand better than a random number generator the majority of the time. Those models with a higher correlation coefficient generally perform better (with KR being the exception), but even EV and EU^-—the models with the highest correlation coefficients—do not outperform predicting coverage levels randomly. Columns (3) and (4) of Table 13 report similar horse race analyses, where the benchmark predictions are the overall average insurance demand (46.5%) or the average insurance demand within a scenario (ranging from 30.8% to 73.5%, depending on the scenario), respectively. Both of these benchmark predictions perform better than the random choice benchmark and thus none of the structural models outperform them either.

The horse race results in Table 13 seemingly imply two contradictions. First, even models that show a correlation above 20% with observed insurance demand are not better at predicting demand than random choice. Second, the primitive probability weighting motive shows some explanatory power when correlated with insurance demand (Tables 10 and 11), but significantly decreases predictive power once it is integrated into a structural model. The reasons for both of these observations in part stem from the way the structural models respond to changes in loading (q) and probability (p). We examine these responses in Figure 7, where we compare observed insurance demand to predicted demand over values of q and p (holding the other constant).

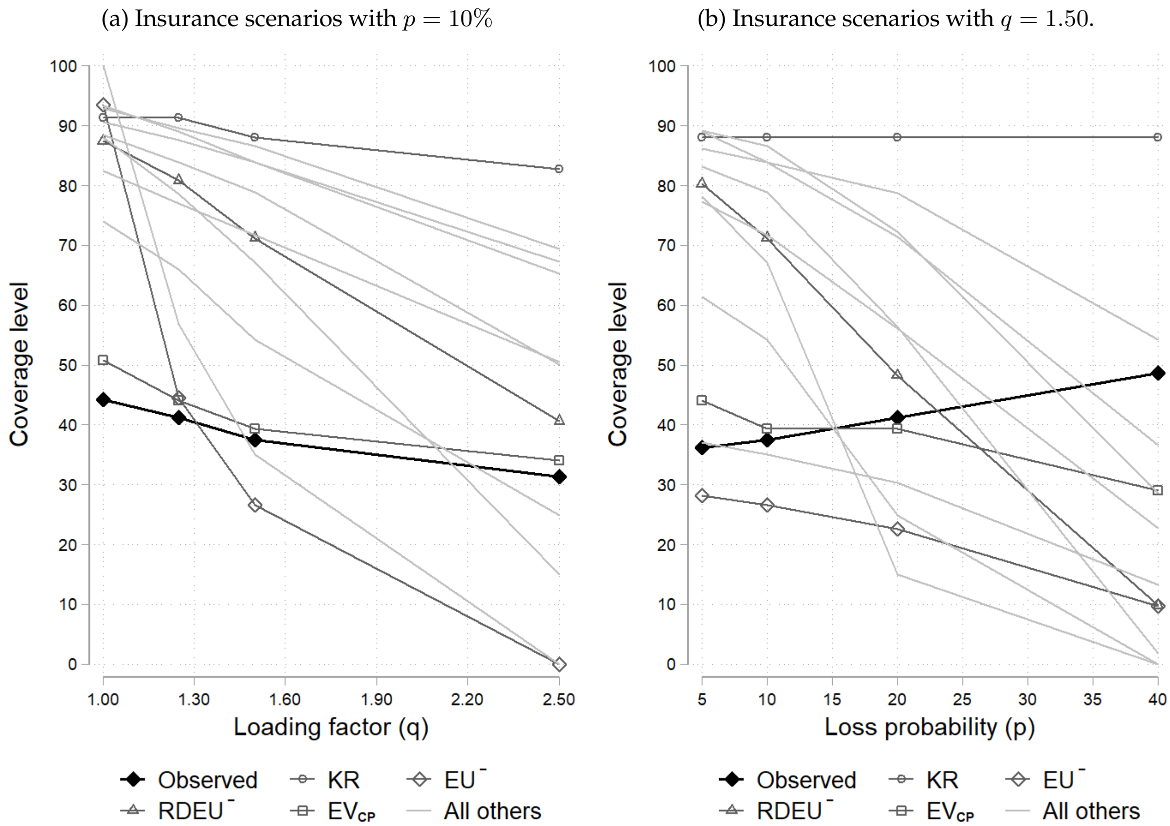
The first contradiction—a disconnect between correlations and horse race performance—is the result of models predicting a more extreme price response than what our subjects exhibit. Panel (a) of Figure 7 shows the change in insurance demand due to an increase in loading if the loss probability is fixed at 10%. Both observed insurance demand and predicted insurance demand are downward-sloping as loading increases. This leads to a positive correlation between observed demand and model predictions. The magnitude of the price response, however, differs substantially—the observed demand curve is relatively flat, while many of the structural models (such as the highlighted EU^-) predict a much stronger reaction.[25] The predicted reaction will thus be more extreme than the observed reaction. The two rightmost columns of Table 12 provide evidence for this effect, as many of the models which perform badly in the horse race analysis also predict a higher proportion of corner solutions than are observed empirically.

The extreme demand response is compounded by structural models which predict only corner solutions. Of the models analyzed here, only those featuring utility curvature (EU, RDEU and CPT) can lead to an interior solution for optimal insurance demand. All other models will always predict either 0% or 100% coverage, which strongly disadvantages them in a horse race analysis.[26] This is also the reason why the average prediction of the EV_{CP} model looks very close to average observed demand (Figure 7), but individual predictions of the EV_{CP} model perform badly in the horse race analysis. While the average prediction might be close to the average observed demand, all individual predictions will be corner solutions. Even in models which allow for interior pre-

[25] To compare these price responses numerically, we regress predicted and observed insurance demand on Probability and Loading (as in Table 9). In the "Predicted" regression, the dependent variable is the predicted demand according to EU^- and the estimated coefficient on Loading is -55.69. In the "Observed" regression, the dependent variable is the observed insurance demand (excluding FOSD violating subjects) and the estimated coefficient on Loading is -9.16. Thus, the predicted reaction to price is 6.08 times what is observed.

[26] To see this, observe in Table 1 that models without utility curvature have linear target functions in α and will thus always lead to corner solutions. Utility curvature makes the target functions non-linear and potentially concave in α, allowing for interior solutions.

Figure 7: Observed insurance demand and model predictions over values of q and p



Note: Figure displays observed and predicted insurance demand as a function of premium loading q when the loss probability is fixed at $p = 10\%$ (panel (a)) and as a function of loss probability p when the premium loading is fixed at $q = 1.50$ (panel (b)). Observed demand (thick black line) is reported for all 1,276 non FOSD-violating subjects. The other lines in the figure show the average predictions of the seventeen structural models listed in Table 1 using the individual parameters for each of the 1,276 subjects.

dictions, corner solutions can still appear frequently. If insurance is simply too expensive (and the coverage level is predicted to be 0%) or when probability weighting or loss aversion feature strongly (and the coverage level is predicted to be 100%), the structural models will overshoot observed demand reactions and perform badly in the horse race. Another contributing factor is that correlation coefficients are not affected by the absolute levels of the predicted insurance demand. In the horse race analysis, however, absolute levels directly affect the relative performance of each model. As such, any model which on average predicts too much or too little insurance demand has a positive correlation with observed demand, but performs poorly in the horse race.

The second contradiction—probability weighting performing well until it is integrated into a structural model—stems from the difference in explaining demand variation between subjects and explaining demand variation between scenarios. In panel (b) of Figure 7, we display the observed and predicted demand response to changes in the loss probability at a fixed premium loading of 1.50. As demonstrated in Figure 6, observed insurance demand increases with the loss probability. However, in models with probability weighting (such as the highlighted RDEU^{-}) predicted insurance demand decreases with increasing loss probability as long as the probability weighting function is inverse-S shaped.[27] This is the most common shape of the weighting function elicited from our preference elicitation tasks, so all structural models featuring probability weighing predict people will buy more insurance when the probability of loss is low. This discrepancy is the reason for low or even negative correlations between the models' predictions and observed demand. The structural models without probability weighting do not predict a negative influence of loss probability on insurance demand (such as KR) or do so to a smaller extent (such as EU). Their predictions thus correlate positively with observed demand across scenarios.

The analyses of Table 10 and columns (1) and (3) of Table 11 show a positive correlation between probability weighting and insurance demand. However, because probability weighting in those analyses is a fixed parameter for each subject and does not further interact with the decision situation, those analyses only examine whether the preference motive can explain variation in insurance demand between subjects within each scenario. Even when an interaction structure with the loss probability is specified in columns (2) and (4) of Table 11, the analysis features insurance scenario fixed effects such that again only between-subject variation is studied. For between-subject variation, probability weighting has explanatory power. When considering the results of the two analyses in combination, we can thus conclude that probability weighting does influence insurance demand, but that an additional mechanism connected to loss probabilities is likely also at work. The nature of this additional mechanism is a compelling question for future research.

To further investigate how loading and probability affect the predictive validity of the tested structural models, we use regression to control for these factors. Specifically, we regress observed insurance demand on the predicted insurance demand for each model. In Table 14, we report the results of our estimations for the ten structural models that do not include a certainty preference.[28] We begin with a set of univariate regressions in panel (a). We then incorporate fixed effects for the loss probability (panel (b)), the premium loading (panel (c)), or the twelve scenarios (panel (d)). The regression results corroborate our interpretation above. Results in panel (a) are by construction very close to the correlation coefficients displayed in column (1) of Table 13. Panel (b) shows

[27] Technically, the required condition is not an inverse-S shape, but decreasing relative overweighting of the probability weighting function. See Jaspersen, Peter, and Ragin (2019) for a discussion of this property and a formal derivation in the RDEU framework.

[28] Adding a preference for certainty decreases the predictive validity across all models and leads to the same qualitative conclusions otherwise. See Appendix A for results on the seven models involving a certainty preference.

that once probabilities are controlled for, the predictive validity of models including probability weighting increases and becomes statistically significant. Controlling instead for the premium loading in panel (c) (and thus reducing variation between the scenarios solely to variation in loss probability), the predictive validity of all models decreases to the point of negative correlations for all models except KR. Panel (d) displays the regression analyses with scenario fixed effects, such that both probability and loading are controlled for. With scenario fixed effects, the decision models explain only differences between subjects, but not between scenarios. In this case, only models which feature probability weighting (DT, RDEU, and CPT) have positive and significant correlations between their predictions and observed insurance demand. For this last analysis, models which combine probability weighting and utility curvature (i.e., the RDEU and the CPT^{-} models) show the best predictive validity.

Table 14: Regressions of observed coverage level on predicted coverage level

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Panel (a): No fixed effects										
$PredictedCovLvl$	0.31***	0.18***	0.19***	0.01	−0.00	0.06**	−0.05***	0.01	−0.01	0.01
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)
Model	EV	EU^{+}	EU^{-}	DT^{+}	DT^{-}	KR	$RDEU^{+}$	$RDEU^{-}$	CPT^{NLIB}	CPT^{-}
Fixed effect	None	None	None	None	None	None	None	None	None	None
R^2	0.07	0.04	0.04	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Panel (b): Probability fixed effects										
$PredictedCovLvl$	0.15***	0.10***	0.11***	0.07***	0.09***	0.04**	0.10***	0.11***	0.09***	0.10***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)
Model	EV	EU^{+}	EU^{-}	DT^{+}	DT^{-}	KR	$RDEU^{+}$	$RDEU^{-}$	CPT^{NLIB}	CPT^{-}
Fixed effect	Prob	Prob	Prob	Prob	Prob	Prob	Prob	Prob	Prob	Prob
R^2	0.12	0.12	0.12	0.12	0.12	0.11	0.12	0.12	0.12	0.12
Panel (c): Loading fixed effects										
$PredictedCovLvl$	–	−0.18***	−0.07***	−0.03**	−0.04***	0.03	−0.04***	−0.02**	−0.03**	−0.01
		(0.02)	(0.02)	(0.01)	(0.01)	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)
Model	EV	EU^{+}	EU^{-}	DT^{+}	DT^{-}	KR	$RDEU^{+}$	$RDEU^{-}$	CPT^{NLIB}	CPT^{-}
Fixed effect	Load	Load	Load	Load	Load	Load	Load	Load	Load	Load
R^2	0.09	0.10	0.09	0.09	0.10	0.09	0.10	0.09	0.09	0.09
Panel (d): Scenario fixed effects										
$PredictedCovLvl$	–	−0.02	0.04	0.04***	0.06***	0.03	0.07***	0.08***	0.05***	0.07***
		(0.04)	(0.03)	(0.01)	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Model	EV	EU^{+}	EU^{-}	DT^{+}	DT^{-}	KR	$RDEU^{+}$	$RDEU^{-}$	CPT^{NLIB}	CPT^{-}
Fixed effect	Scenario	Scenario	Scenario	Scenario	Scenario	Scenario	Scenario	Scenario	Scenario	Scenario
R^2	0.12	0.12	0.12	0.13	0.13	0.12	0.13	0.13	0.13	0.13

Note: Dependent variable is the coverage level selected by the subject. Explanatory variable of interest is the predicted level of coverage using the specified model. The only other explanatory variables are fixed effects as noted. The sample in each regression is the 1,276 subjects with estimated preference parameters (who did not violate FOSD). In each panel, p-values have been adjusted using the Šidák (1967) method for the ten hypotheses tested. Standard errors clustered by subject are in parentheses. Stars *, **, and *** denote statistical significance at the 0.10, 0.05, and 0.01 levels, respectively.

Some further conclusions can be drawn from the results in Table 14. The relatively high predictive ability of the EV and EU models in panel (a) is due to their ability to explain demand response to increases in the loading, as evidenced by the change in coefficients in panel (c). Panel (d) shows that little explanatory power in these models is due to their ability to explain between-subject variation in insurance demand. Further, the preferences measured in the loss domain appear to have higher predictive validity than those in the gain domain. Finally, loss aversion does not fare well when included in the structural models to explain insurance demand. There are three models which each feature only one behavioral motive—EU includes only utility curvature, DT uses only probability weighting, and KR features only loss aversion. Of those three, the KR model with loss aversion has one of the lowest reported coefficients in panel (d). Of the more complex models, the one with the lowest coefficient, CPT^{NLIB}, is the only one that features loss aversion. However, loss aversion has one of the largest coefficients in Table 11. This suggests that loss aversion has a significant relationship to insurance demand, but that there is no well-established structural model that leverages this predictive power well.

One possible explanation for the low observed correlations could be randomness in the insurance decisions. To investigate this, we divide the subjects in our sample roughly into quartiles based on the number of times the subject violated the law of demand (increasing insurance demand at higher loads). We consider subjects who never violated the law of demand as exhibiting the most consistent choices, and subjects who violated three or more times as making the least consistent choices (the other quartiles being subjects who violated only once and subjects who violated twice). For each subsample, we calculate the correlation between observed demand and each structural model's predicted demand, as in column (1) of Table 13. In addition, we regress observed demand on predicted demand with scenario fixed effects, as in Panel (d) of Table 14. Our findings are in Appendix E, with the correlations in Table E.2 and the regressions in Table E.3. Even for the most consistent group, the relationships between observed and predicted insurance demand are moderate and, in line with our analysis above, no model has a higher correlation than expected value maximization. The only difference we find is that insurance demand for the most consistent subjects now has small positive correlations with models incorporating probability weighting. The least consistent subjects (i.e., those who violate the law of demand three or more times) tend to make decisions which have much less correlation with structural model predictions.

5 Robustness: Eliciting Preferences Using Insurance Choices

Our results show that structural models estimated from lottery choices have limited ability to predict insurance choices. This suggests that we cannot yet identify a portable model of risk attitudes that can be applied across these contexts.

An open question, though, is whether this is because there is something about decision making in insurance environments that is not well captured by the prevailing structural models or if the problem is primarily in the transfer of preferences across domains. Our setting eliminates some obvious concerns with using models across domains by having the choices made in the same setting, at nearly the same time and for similar financial stakes. Nonetheless, it is possible that other issues limit the portability across context, yet within a particular financial context (e.g., lottery choices or insurance) structural models may have strong predictive power.

In this section, we test this possibility by analyzing the predictive power of structural models when we work only with the insurance data. This exercise was not part of our original pre-specified research plan, but strikes us as valuable for exploring the implications of our main results. However, because this analysis was not pre-specified, we focus our attention on estimating a more limited set of structural models than in our main analysis.

Our goal is to identify a subset of the insurance choices made by subjects that can be used to fit a structural model and then use the resulting fitted model to predict the choices that subject made under other insurance conditions. To simplify, we focus on two cases in our data that allow us to test the ability of the models to predict either across different prices (i.e., loads) or across different probabilities.

For the first case, we analyze the four choices subjects made under different loads when the probability of loss was 10%. We use the choices the subjects made at the middle two loads of 1.25 and 1.50 as our observed data for the modeling exercise, and then use the model to predict the choices under the other two loads of 1.00 and 2.50.

We adopt the structural model of Barseghyan et al. (2013), which incorporates both a standard utility function and generalized probability distortions without assuming a form on the probability weighting function. The model of the (decision) utility for an insurance choice in this framework is

$$\Omega(p)u\left(w-(1-\alpha)L-\alpha qpL\right)+\left(1-\Omega(p)\right)u(w-\alpha qpL), \tag{9}$$

where $\Omega(p)$ is the generalized probability distortion. We maintain our assumptions from the original analysis, and assume that subjects' choices are non-stochastic. This lets us use the result of Jaspersen et al. (2019), who show that two coverage level decisions at two different values of $q > 1$

are sufficient to elicit $\Omega(p)$ and u, provided that u can be described by a parametric form with a single parameter that is monotonically related to the function's curvature. The condition for this result is that subjects purchase more insurance at lower levels of q. To further simplify our elicitation procedure, we limit the sample of subjects to those who purchase less than full insurance at both loadings and we assume that subjects who purchase no insurance at $q = 1.5$ maximize their utility at exactly $\alpha = 0$. Our sample restrictions lead to a sample size of 481 subjects for this analysis.

To assess the quality of our predictions, we report the same two metrics which we already considered in Table 13: the correlation between observed and predicted demand and a horse race of the prediction against a benchmark. For the latter, we only use random choice as a benchmark for the sake of brevity. The results of the analysis are shown for each predicted insurance scenario in the top panel of Table 15. The first row shows the results for the utility function used in our main analysis. The second and third rows use a CRRA and a CARA utility function, respectively, as a robustness check. Since the results are consistent between the different functional forms, we focus our discussion on our initial parametric assumption.

On both reported measures, the model predictions perform markedly better than the predictions from the models calibrated using lottery choices. For example, there is a correlation coefficient of 0.42 between the predicted and actual choices when subjects faced high loads of 2.50 and the predictions of the structural model out-perform random predictions 64% of the time.

Although this result looks more favorable for structural modeling, another way to assess the value of the structural model is to compare it to simpler predictions that do not incorporate modeled risk attitudes. We test two such alternatives and show the results in the bottom panel of Table 15. The first model simply calculates the average insurance demand for each subject across the $q = 1.25$ and $q = 1.50$ scenario and uses this personalized average to predict insurance demand for the individual in the other two scenarios. This model performs better than the structural model predictions in both scenarios, considering both the correlation and the horse race. The difference in performance is large and significant ($p < 0.001$).[29] This suggests that using a structural model calibrated on the information provided by two insurance scenarios to predict a different one performs worse than simply using the information in its raw form. In addition, the difference in performance between the structural model and the average demand is larger at $q = 2.50$ than it is at $q = 1.00$ such that predicting insurance demand from a structural model becomes less informative the further the predicted scenario is removed from the elicitation scenarios.

[29] Significance in the differences was calculated using the web utility from Lee and Preacher (2013). The correlations between observed and the Eq. (5) predictions are 0.7945 and 0.6786 for $q = 1.00$ and $q = 2.50$, respectively.

Table 15: Goodness of fit analysis at 10% loss probability for structural models calibrated with insurance choices

	Loading Factor $q = 1.00$		Loading Factor $q = 2.50$	
Prediction Method	Correlation	Horse Race	Correlation	Horse Race
Utility Models				
As in main analysis	0.55	0.63	0.42	0.64
CRRA	0.54	0.63	0.43	0.64
CARA	0.55	0.64	0.39	0.62
Non-Utility Models				
Average demand	0.64	0.74	0.59	0.76
Linear demand function	0.90	0.77	0.59	0.60

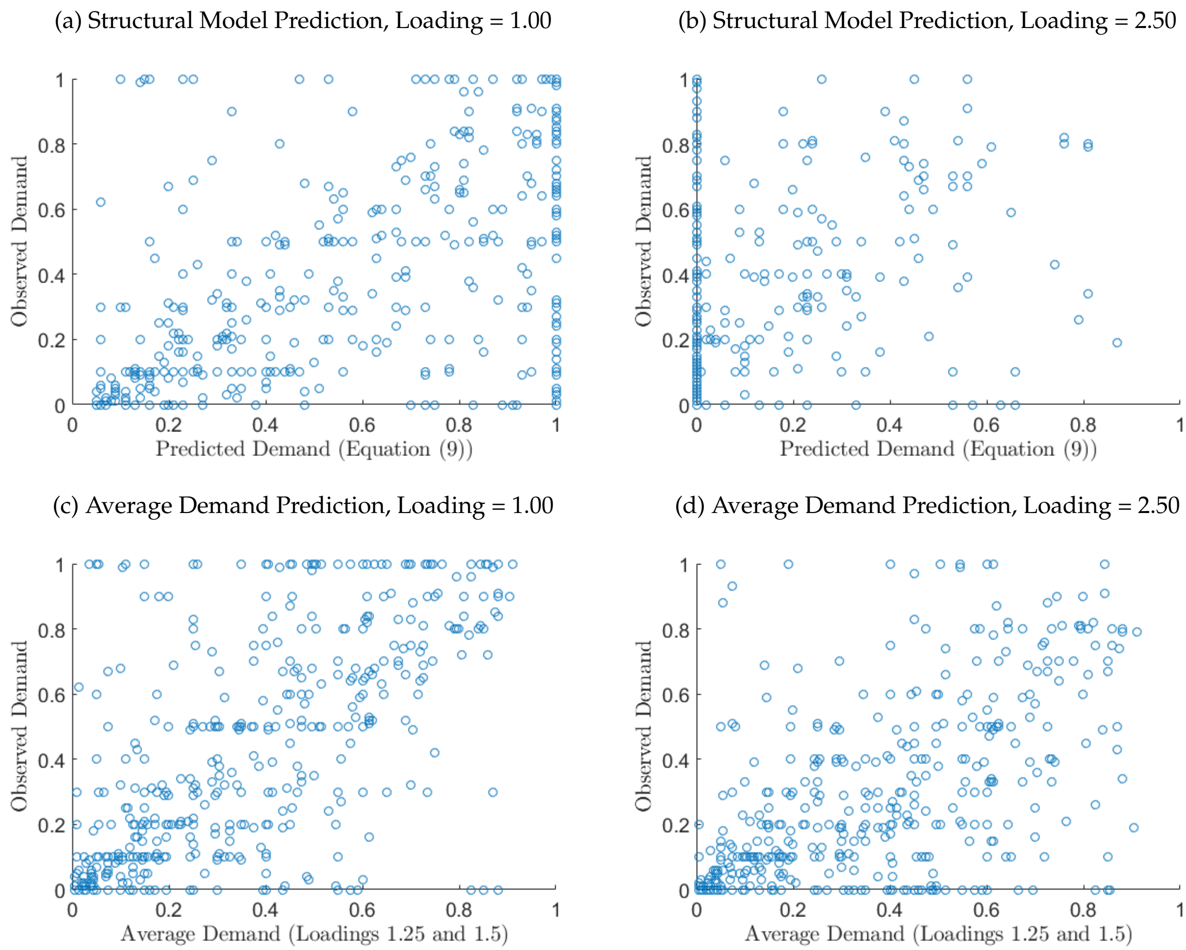
Note: The table reports the results of our calibration analysis in which we use the coverage level decisions in the scenarios with loss probability 10% and loading factors 1.25 and 1.50 to calibrate the decision model in Equation (5) and use it to predict insurance demand in the scenarios with loss probability 10% and loading factors 1.00 and 2.50. Results are reported for the 481 subjects with $\alpha_{(q=1.5)} < \alpha_{(q=1.25)} < 1$ at $p = 10\%$. Rows in the top panel indicate the utility function used in the analysis. In the bottom panel, we offer two non-utility models for comparison. The first non-utility model uses the average demand for loading factors 1.25 and 1.50 to predict demand in the other two scenarios. The second non-utility model calibrates an individual linear demand function $\alpha = bq + c$ for predictions. Correlations are Pearson coefficients. Horse race results are reported against a uniformly distributed random number.

The second simple alternative model is a linear demand function of insurance coverage in the loading. Specifically, we use the demand function $\alpha = bq + c$. We calibrate b and c at the individual level based on the subject's choices for $q = 1.25$ and $q = 1.50$. We then use the individual's estimated line to predict insurance demand for $q = 1.00$ and $q = 2.50$, bounding the predicted coverage choices at 0 and 1. This model performs best among all tested procedures for predicting insurance demand with $q = 1.00$ and shows the highest correlation for $q = 2.50$. In this high loading scenario, it does however have the worst performance in the horse race analysis due to predicting too many corner solutions.

To understand the results of our analysis better, we plot the predicted and observed insurance demand for the structural model prediction and the Average Demand prediction for both scenarios in Figure 8. We can see that the poorer predictive quality of the structural model is again rooted in the over-responsiveness to price—it predicts too many cases of full insurance at low loadings (panel (a)) and too many cases of no insurance at high loadings (panel (b)). The average demand shows no such overreactions and is thus the better predictor.

Our insurance demand data also feature four scenarios at different loss probabilities (5%, 10%, 20% and 40%) but the same loading of $q = 1.50$. We again use the two middle choices (10% and 20%) to calibrate the insurance choices for 5% and 40% loss probability. For this analysis, we cannot use the probability distortion model of Barseghyan et al. (2013), because we require a model

Figure 8: Scatter plots of predicted and observed insurance demand at 10% loss probability



Note: Predictions are for the loss scenario with a 10% loss probability. The structural model prediction is the coverage level which maximizes Equation (9), with parameters estimated based on scenarios with loading of 1.25 and 1.50. The average demand prediction is the mean coverage level for scenarios with loading of 1.25 and 1.50.

of the subject's reaction to a change in probability. Instead, we incorporate probability weighting using the common rank dependent expected utility format with a one-parameter Prelec (1998) weighting function. Though not strictly necessary for the calibration, we limit the analysis to the same 481 subjects used above to ensure comparability. The results of the analysis are shown in Table 16. We can see that the results for predicting the 5% loss probability scenario look similar to the analysis with varying loading factors. As mentioned in Section 3.3, the insurance choices for low loss probabilities are highly correlated. However, when considering the predictions for insurance demand with a 40% loss probability, the model has a strongly decreased performance in terms of both correlation and the horse race. As in the main analysis, the model structure has a hard time predicting insurance behavior when loss probabilities change strongly. The most

Table 16: Goodness of fit analysis at loading factor 1.50 for structural models calibrated with insurance choices

	Loss Probability $p = 5\%$		Loss Probability $p = 40\%$	
Prediction Method	Correlation	Horse Race	Correlation	Horse Race
Utility Models				
As in main analysis	0.52	0.60	0.21	0.32
CRRA	0.52	0.60	0.21	0.32
CARA	0.53	0.61	0.20	0.30
Non-Utility Models				
Average demand	0.59	0.67	0.45	0.60
Linear demand function	0.62	0.72	0.40	0.49

Note: The table reports the results of our calibration analysis in which we use the coverage level decisions in the scenarios with loading factor 1.50 and loss probabilities 10% and 20% to calibrate the decision model in Equation (5) and use it to predict insurance demand in the scenarios with loading factor 1.50 and loss probabilities 5% and 40%. Results are reported for the 481 subjects with $\alpha_{(q=1.5)} < \alpha_{(q=1.25)} < 1$ at $p = 10\%$. Rows in the top panel indicate the utility function used in the analysis. In the bottom panel, we offer two non-utility models for comparison. The first non-utility model uses the average demand for loss probabilities 10% and 20% to predict demand in the other two scenarios. The second non-utility model calibrates an individual linear demand function $\alpha = bp + c$ for predictions. Correlations are Pearson coefficients. Horse race results are reported against a uniformly distributed random number.

important conclusion to be taken from the results in Table 16 is, however, that the model structure is again outperformed by the simple models using average insurance demand and linear demand over probabilities.

In summary, using insurance choices to calibrate the parameters of a decision model seems to lead to better predictions than using lottery choices. However, aggregating the information gained from insurance choices in a decision model is less informative for other insurance decisions than simply aggregating the information in an average. This seems discouraging for model-based policy recommendations, which we elaborate on in the next section.

6 Discussion and Conclusion

The results of our experiment call into question the ability of common consequentialist structural models of risk preferences to predict insurance demand. We find moderately-sized and statistically significant predictive power for the three most commonly discussed primitive preference motives of insurance demand: utility curvature, probability weighting, and loss aversion. Despite these correlations, however, assembling these motives into structural decision models leads to poor predictive validity in our setting, particularly across different insurance scenarios.

Our work adds to mounting evidence that there may simply be no single underlying model of risk preferences that can adequately capture attitudes toward financial risk in different domains, and toward insurance in particular. Our results are consistent both with a broader literature that has shown inconsistency in measures of risk aversion across different elicitation strategies in the lab (e.g., Chapman et al., 2018; Friedman et al., 2019; Friedman and Sunder, 2011) and with Barseghyan et al. (2011) who find that choices across insurance context in field data cannot be rationalized by an expected utility model with stable underlying preferences. A distinguishing feature of our approach is that we are considering simultaneously a very wide range of structural models that have been posited in the literature.

This suggests that there may be significant value in research that aims to better illuminate the alternative decision processes that shape decisions about insurance and other financial risks. For example, research has suggested that decision editing phases (Kahneman and Tversky, 1979; McClelland et al., 1993), information frictions (Handel and Kolstad, 2015), and limited consideration sets (Abaluck and Adams, 2017; Barseghyan et al., 2019) may play an important role in how people make decisions about insurance. These approaches could potentially help explain why many people in our experiment purchased no insurance at low loss probabilities—these events might have been ignored. There is likely also value in better understanding the types of systematic heuristics people use when evaluating insurance, similar in spirit to recent work examining debt-repayment heuristics (e.g., Gathergood et al., 2019). The evidence from these studies does not yet provide a conclusive picture of how people make insurance choices and there may be significant heterogeneity across people, but research in this area is starting to develop techniques for better identifying relevant decision processes.

Until we have a better understanding of the decision processes at play in insurance choices, economists may benefit from being more cautious about using structural models. Yet this does not imply that there is no scope for counterfactual predictions in insurance markets. Prior research has shown that insurance choices correlate to some degree across contexts. Our work adds to that point and helps to highlight that it may be possible to use data on insurance choices to predict choices in somewhat different economic conditions. The key implication, though, is that these predictions may be better if they are based on the direct patterns of observed choices rather than structural models of risk attitudes.

References

Abaluck, J. and A. Adams (2017). What do consumers consider before they choose? Identification from asymmetric demand responses. *NBER Working Paper No. 23566*.

Abaluck, J. and J. Gruber (2011). Choice inconsistencies among the elderly: Evidence from plan choice in the Medicare Part D program. *American Economic Review 101*(4), 1180–1210.

Abdellaoui, M., H. Bleichrodt, and C. Paraschiv (2007). Loss aversion under prospect theory: A parameter-free measurement. *Management Science 53*(10), 1659–1674.

Abito, J. M. and Y. Salant (2018). The effect of product misperception on economic outcomes: Evidence from the extended warranty market. *Review of Economic Studies Forthcoming*.

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

Andreoni, J. and C. Sprenger (2012). Risk preferences are not time preferences. *American Economic Review 102*(7), 3357–3376.

Argyris, N., J. Jaspersen, and A. Richter (2019). Risk aversion in additive multivariate utility functions. Technical report, Munich Risk and Insurance Center.

Balcombe, K., N. Bardsley, S. Dadzie, and I. Fraser (2019). Estimating parametric loss aversion with prospect theory: Recognising and dealing with size dependence. *Journal of Economic Behavior & Organization 162*, 106–119.

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

Barseghyan, L., F. Molinari, T. O'Donoghue, and J. C. Teitelbaum (2018). Estimating risk preferences in the field. *Journal of Economic Literature 56*(2), 501–64.

Barseghyan, L., F. Molinari, and M. Thirkettle (2019). Discrete choice under risk with limited consideration. *Working Paper*.

Barseghyan, L., J. Prince, and J. C. Teitelbaum (2011). Are risk preferences stable across contexts? Evidence from insurance data. *American Economic Review 101*(2), 591–631.

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

Bordalo, P., N. Gennaioli, and A. Shleifer (2012). Salience theory of choice under risk. *The Quarterly Journal of Economics 127*(3), 1243–1285.

Callen, M., M. Isaqzadeh, J. D. Long, and C. Sprenger (2014). Violence and risk preference: Experimental evidence from Afghanistan. *American Economic Review 104*(1), 123–148.

Chapman, J., M. Dean, P. Ortoleva, E. Snowberg, and C. Camerer (2018). Willingness to pay and willingness to accept areprobably less correlated than you think. *NBER Working Paper No. 23954*.

Chiu, W. H. (2010). Skewness preference, risk taking and expected utility maximisation. *The Geneva Risk and Insurance Review 35*(2), 108–129.

Cohen, A. and L. Einav (2007). Estimating risk preferences from deductible choice. *American Economic Review 97*(3), 745–788.

Collier, B., D. Schwartz, H. Kunreuther, and E. Michel-Kerjan (2018). Risk preferences in small and large stakes: Evidence from insurance contract decisions. *NBER Working Paper No. 23579*.

Collier, B. L. and M. A. Ragin (2019). The influence of sellers on contract choice: Evidence from flood insurance. *The Journal of Risk and Insurance Early View*, 1–35.

Decarolis, F., M. Polyakova, and S. Ryan (2019). Subsidy design in privately-provided social insurance: Lessons from Medicare Part D. *NBER Working Paper No. 21298*.

Dohmen, T., A. Falk, D. Huffman, U. Sunde, J. Schupp, and G. G. Wagner (2011). Individual risk attitudes: Measurement, determinants, and behavioral consequences. *Journal of the European Economic Association 9*(3), 522–550.

Ebert, S. (2015). On skewed risks in economic models and experiments. *Journal of Economic Behavior & Organization 112*(4), 85–97.

Einav, L., A. Finkelstein, I. Pascu, and M. R. Cullen (2012). How general are risk preferences? Choices under uncertainty in different domains. *American Economic Review 102*(6), 2606–38.

Friedman, D., S. Habib, D. James, and B. Williams (2019). Varieties of risk elicitation. *Working Paper*.

Friedman, D. and S. Sunder (2011). Risky curves: From unobservable utility to observable opportunity sets. *Working Paper*.

Gathergood, J., N. Mahoney, N. Stewart, and J. Weber (2019). How do individuals repay their debt? The balance-matching heuristic. *American Economic Review 109*(3), 844–875.

Handel, B., I. Hendel, and M. D. Whinston (2015). Equilibria in health exchanges: Adverse selection versus reclassification risk. *Econometrica 83*(4), 1261–1313.

Handel, B. R. (2013). Adverse selection and inertia in health insurance markets: When nudging hurts. *American Economic Review 103*(7), 2643–2682.

Handel, B. R. and J. Kolstad (2015). Health insurance for "humans": Information frictions, plan choice, and consumer welfare. *American Economic Review 105*(8), 2449–2500.

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

Jaspersen, J., R. Peter, and M. Ragin (2019). The role of probability weighting in optimal insurance demand: A double-edged sword. Technical report, Munich Risk and Insurance Center.

Jaspersen, J., M. Ragin, and J. R. Sydnor (2019). Linking subjective and incentivized risk attitudes: The importance of the loss domain. Technical report, Munich Risk and Insurance Center.

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

Kim, G. and S. Lyons (2019, June). HSIM2—CBO's new health insurance simulation model. Washington, D.C. American Society of Health Economists Conference.

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

Lee, I. A. and K. J. Preacher (2013). Calculation for the test of the difference between two dependent correlations with one variable in common [computer software]. Available from http://quantpsy.org/corrtest/corrtest2.htm.

l'Haridon, O. and F. M. Vieider (2019). All over the map: A worldwide comparison of risk preferences. *Quantitative Economics 10*(1), 185–215.

Loomes, G. and R. Sugden (1982). Regret theory: An alternative theory of rational choice under uncertainty. *The Economic Journal 92*(368), 805–824.

McClelland, G. H., W. D. Schulze, and D. L. Coursey (1993). Insurance for low-probability hazards: A bimodal response to unlikely events. *Journal of Risk and Uncertainty 7*(1), 95–116.

Novemsky, N. and D. Kahneman (2005). The boundaries of loss aversion. *Journal of Marketing Research 42*(2), 119–128.

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

Sarin, R. and P. Wakker (1998). Revealed likelihood and Knightian uncertainty. *Journal of Risk and Uncertainty 16*(3), 223–250.

Schlesinger, H. (2013). The Theory of Insurance Demand. In *Handbook of Insurance*, pp. 167–184. Springer.

Schmidt, U. (1998). A measurement of the certainty effect. *Journal of Mathematical Psychology 42*(1), 32–47.

Šidák, Z. (1967). Rectangular confidence regions for the means of multivariate normal distributions. *Journal of the American Statistical Association 62*(318), 626–633.

Sprenger, C. (2015). An endowment effect for risk: Experimental tests of stochastic reference points. *Journal of Political Economy 123*(6), 1456–1499.

Stango, V. and J. Zinman (2019). We are all behavioral, more or less: Measuring and using consumer-level behavioral sufficient statistics. *NBER Working Paper No. 25540*.

Sydnor, J. (2010). (Over)insuring modest risks. *American Economic Journal: Applied Economics 2*(4), 177–199.

Tanaka, T., C. Camerer, and Q. Nguyen (2010). Risk and time preferences: Experimental and household survey data from Vietnam. *American Economic Review 100*(1), 557–571.

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

Vieider, F. M. (2018). Certainty preference, random choice, and loss aversion: A comment on "Violence and Risk Preference: Experimental Evidence from Afghanistan". *American Economic Review 108*(8), 2366–2382.

Appendices

[For Online Publication]

Appendix A Predictive models incorporating certainty preference

Table A.1: Regressions of observed coverage level on predicted coverage level, models with certainty preference

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel (a): No fixed effects							
$PredictedCovLvl$	0.08***	0.05***	0.04**	−0.02	−0.02	−0.06***	−0.02
	(0.01)	(0.02)	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)
Model	EV_{CE}	EU^{+}_{CE}	EU^{-}_{CE}	DT^{+}_{CE}	DT^{-}_{CE}	$RDEU^{+}_{CE}$	$RDEU^{-}_{CE}$
Fixed effect	None	None	None	None	None	None	None
R^2	0.01	0.00	0.00	0.00	0.00	0.00	0.00
Panel (b): Probability fixed effects							
$PredictedCovLvl$	0.01	0.02	0.02	0.06***	0.07***	0.06***	0.07***
	(0.01)	(0.02)	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)
Model	EV_{CE}	EU^{+}_{CE}	EU^{-}_{CE}	DT^{+}_{CE}	DT^{-}_{CE}	$RDEU^{+}_{CE}$	$RDEU^{-}_{CE}$
Fixed effect	Prob	Prob	Prob	Prob	Prob	Prob	Prob
R^2	0.11	0.11	0.11	0.11	0.12	0.11	0.12
Panel (c): Loading fixed effects							
$PredictedCovLvl$	−0.01	−0.02	−0.03*	−0.03**	−0.04***	−0.05***	−0.04***
	(0.01)	(0.02)	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)
Model	EV_{CE}	EU^{+}_{CE}	EU^{-}_{CE}	DT^{+}_{CE}	DT^{-}_{CE}	$RDEU^{+}_{CE}$	$RDEU^{-}_{CE}$
Fixed effect	Load	Load	Load	Load	Load	Load	Load
R^2	0.09	0.09	0.09	0.09	0.09	0.10	0.09
Panel (d): Scenario fixed effects							
$PredictedCovLvl$	−0.01	−0.01	−0.01	0.04***	0.04***	0.03**	0.04***
	(0.02)	(0.02)	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)
Model	EV_{CE}	EU^{+}_{CE}	EU^{-}_{CE}	DT^{+}_{CE}	DT^{-}_{CE}	$RDEU^{+}_{CE}$	$RDEU^{-}_{CE}$
Fixed effect	Scenario	Scenario	Scenario	Scenario	Scenario	Scenario	Scenario
R^2	0.12	0.12	0.12	0.12	0.12	0.12	0.12

Note: Dependent variable is the coverage level selected by the subject. Explanatory variable of interest is the predicted level of coverage using the specified model. The only other explanatory variables are fixed effects as noted. The sample in each regression is the 1,276 subjects with estimated preference parameters (who did not violate FOSD). In each panel, p-values have been adjusted using the Šidák (1967) method for the seven hypotheses tested. Standard errors clustered by subject are in parentheses. Stars *, **, and *** denote statistical significance at the 0.10, 0.05, and 0.01 levels, respectively.

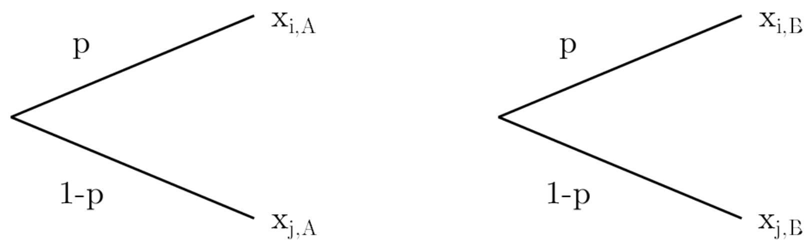
Appendix B Theory on Preference Elicitation

This paper uses the same experimental data and preference measures as the study in Jaspersen, Ragin, and Sydnor (2019). Since the analysis there does not consider the insurance demand tasks, both studies analyze distinct research questions. However, because they both use the same preference measures, the contents of this appendix are to a large part identical to a corresponding appendix in Jaspersen, Ragin, and Sydnor (2019).

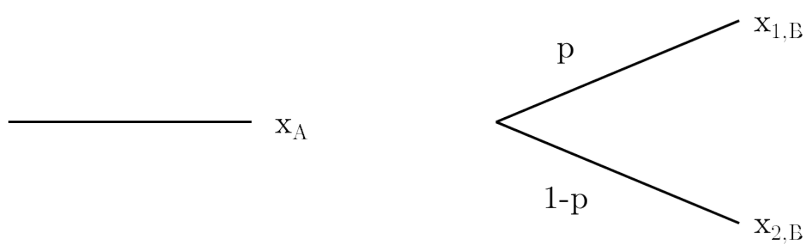
B.1 Nonparametric Preference Measures

There are six tables in the elicitation procedure: GD1, GD2, LD1, LD2, LA and CP. The first five are repeated choices between two binary lotteries, the last one is the choice between a binary lottery and a certain outcome. The choices are depicted in Figure B.1.

Figure B.1: Graphical illustrations of the different choice tables



(a) Schematic representation of choices in the GD1, GD2, LD1, LD2 and LA tables.



(b) Schematic representation of choices in the CP table

In accordance with Section 2.1, $i = 1$ and $j = 2$ in tables GD1, GD2, and LA and $i = 2$ and $j = 1$ in tables LD1 and LD2.

B.1.1 Utility Curvature and Probability Weighting in the Gain Domain

Assuming a full CPT preference functional, we evaluate the lotteries in the GD1/ GD2 tables as $EV = w^+(p)u^+(large\ gain) + (1 - w^+(p))u^+(small\ gain)$. In these tables, only the large gain of Lottery B changes. Adopting the notation that the large gain is x_2 and the small gain is x_1, the indifference large gain on Lottery B fulfills

$$w^+(p)u^+(x_{2,A}) + (1 - w^+(p))u^+(x_{1,A}) = w^+(p)u^+(x_{2,B}) + (1 - w^+(p))u^+(x_{1,B}). \tag{B.1}$$

We first consider the *ceteris paribus* elicitation of utility curvature. When changing the indifference value for $x_{2,B}$ to $\hat{x}_{2,B}$ and keeping everything but the utility function fixed, we signify the change in the utility function by the transformation $g(\cdot)$ of which we know that it must be increasing. The new indifference relationship implies

$$w^+(p)g(u^+(x_{2,A})) + (1 - w^+(p))g(u^+(x_{1,A})) = w^+(p)g(u^+(\hat{x}_{2,B})) + (1 - w^+(p))g(u^+(x_{1,B})). \tag{B.2}$$

From $g'(\cdot) > 0$, we know that $g(u^+(\hat{x}_{2,B})) - g(u^+(x_{2,B})) > 0$. We can thus see from (B.2) that

$$w^+(p)g(u^+(x_{2,A})) + (1 - w^+(p))g(u^+(x_{1,A})) > w^+(p)g(u^+(x_{2,B})) + (1 - w^+(p))g(u^+(x_{1,B})). \tag{B.3}$$

We proceed by using w^+ to define ω as a probability measure of the random variables $\tilde{x}$ and, respectively, $u^+(\tilde{x})$. This can be done because $w^+(p) \in [0, 1]$ and all payoffs are in the gain domain, so the decision weights add to one. Under this probability measure, we know from (B.1) that the expectation of $u^+(\tilde{x})$ is equal between lotteries A and B. From p being equal for both lotteries, we further know the skewness of $u^+(\tilde{x})$ under the probability measure ω to be equal between both lotteries (Ebert, 2015). From Chiu (2010) we thus know that Lottery B is a mean-preserving spread in $u^+(\tilde{x})$ under the probability measure ω. As such $\mathbb{E}_\omega[g(u^+(\tilde{x}_A))] > \mathbb{E}_\omega[g(u^+(\tilde{x}_B))]$ is true if and only if $g'' < 0$. Thus, a higher value of $x_{2,B}$ in GD1 and GD2 implies more concave utility curvature.

For the *ceteris paribus* elicitation of the probability weighting function, we again consider the indifference relationship at two points $x_{2,B}$ and $\hat{x}_{2,B}$ with $x_{2,B} < \hat{x}_{2,B}$. The respective probability weighting functions for the two indifference relationships are w^+ and $\hat{w}^+$. We have the two equalities:

$$w^+(p)u^+(x_{2,A}) + (1 - w^+(p))u^+(x_{1,A}) = w^+(p)u^+(x_{2,B}) + (1 - w^+(p))u^+(x_{1,B}). \tag{B.4}$$

$$\hat{w}^+(p)u^+(x_{2,A}) + (1 - \hat{w}^+(p))u^+(x_{1,A}) = \hat{w}^+(p)u^+(\hat{x}_{2,B}) + (1 - \hat{w}^+(p))u^+(x_{1,B}). \tag{B.5}$$

We subtract (B.4) from (B.5) and add $\hat{w}^+(p)[u(x_{2,B}) - u(x_{2,B})]$ on the right hand side to obtain

$$\begin{aligned}[\hat{w}^+(p) - w^+(p)][u^+(x_{2,A}) - u^+(x_{1,A}))] = \quad & [\hat{w}^+(p) - w^+(p)][u^+(x_{2,B}) - u^+(x_{1,B}))] \\ & + \hat{w}^+(p)[u^+(\hat{x}_{2,B}) - u^+(x_{2,B}))]. \end{aligned} \tag{B.6}$$

Since $\hat{w}^+(p)[u^+(\hat{x}_{2,B}) - u^+(x_{2,B})) > 0$, (B.6) becomes

$$[\hat{w}^+(p) - w^+(p)][u^+(x_{2,A}) - u^+(x_{1,A}))] > [\hat{w}^+(p) - w^+(p)][u^+(x_{2,B}) - u^+(x_{1,B}))]. \tag{B.7}$$

From the construction of the table, we know $u^+(x_{2,A}) - u^+(x_{1,A})) < u^+(x_{2,B}) - u^+(x_{1,B}))$. (B.7) thus implies $\hat{w}^+(p) < w^+(p)$. A switching row lower in the GD1/GD2 tables thus leads to a smaller decision weight on the large outcome of each lottery.

We have structured the tables such that $p^{GD1} < p^{GD2}$ and assume the inflection point of the probability weighting function to be between these two probabilities. Under this construction, if the decision weight on p^{GD1} becomes smaller, we have less inverse S-shape or more S-shape in the weighting function. If the decision weight on p^{GD2} becomes smaller, we have more inverse S-shape or less S-shape in the weighting function.

B.1.2 Utility Curvature and Probability Weighting in the Loss Domain

Assuming a full CPT preference functional, we evaluate the lotteries in the LD1/LD2 tables as $EV = w^-(p)u^-(large\ loss) + (1 - w^-(p))u^-(small\ loss)$. In these tables only the large loss in Lottery A changes. The indifference gain fulfills

$$w^-(p)u^-(x_{2,A}) + (1 - w^-(p))u^-(x_{1,A}) = w^-(p)u^-(x_{2,B}) + (1 - w^-(p))u^-(x_{1,B}). \tag{B.8}$$

We change the indifference value to $\hat{x}_{2,A} < x_{2,A}$ and the utility function gets transformed by function $g(\cdot)$. The new indifference value now fulfills

$$w^-(p)g(u^-(\hat{x}_{2,A})) + (1 - w^-(p))g(u^-(x_{1,A})) = w^-(p)g(u^-(x_{2,B}) + (1 - w^-(p))g(u^-(x_{1,B})). \tag{B.9}$$

We again use the positive first derivatives of $u(\cdot)$ and $g(\cdot)$ and transform (B.9) such that

$$w^-(p)g(u^-(x_{2,A})) + (1 - w^-(p))g(u^-(x_{1,A})) > w^-(p)g(u^-(x_{2,B}) + (1 - w^-(p))g(u^-(x_{1,B})). \tag{B.10}$$

We proceed as in the gain domain. Under the probability measure ω defined by w^-, the utility lottery of Lottery A is a mean-preserving spread of that of Lottery B. As such, $\mathbb{E}_\omega[g(u^-(\tilde{x}_A))] > \mathbb{E}_\omega[g(u^-(\tilde{x}_B))]$ if and only if $g'' < 0$. Thus, a more negative value of $x_{2,A}$ (and thus a later switching row) is, ceteris paribus, associated with more convex utility curvature/ less risk aversion.

For probability weighting, we again consider two different indifference relationships. One at $x_{A,2}$ and one at $\hat{x}_{A,2}$ with $\hat{x}_{A,2} < x_{A,2}$ and different probability weighting functions w^- and $\hat{w}^-$. We have the two equalities:

$$w^-(p)u^-(x_{2,A}) + (1 - w^-(p))u^-(x_{1,A}) = w^-(p)u^-(x_{2,B}) + (1 - w^-(p))u^-(x_{1,B}). \tag{B.11}$$

$$\hat{w}^-(p)u^-(\hat{x}_{2,A}) + (1 - \hat{w}^-(p))u^-(x_{1,A}) = \hat{w}^-(p)u^-(x_{2,B}) + (1 - \hat{w}^-(p))u^-(x_{1,B}). \tag{B.12}$$

We subtract (B.11) from (B.12) and add $\hat{w}(p)[u(x_{2,A}) - u(x_{2,A})]$ on the left hand side to obtain

$$\begin{aligned} [\hat{w}^-(p) - w^-(p)][u^-(x_{2,A}) - u^-(x_{1,A}))] = \quad & [\hat{w}^-(p) - w^-(p)][u^-(x_{2,B}) - u^-(x_{1,B}))]. \\ +\hat{w}^-(p)[u^-(\hat{x}_{2,A}) - u^-(x_{2,A}))] \quad & \end{aligned} \tag{B.13}$$

Since $\hat{w}^-(p)[u^-(\hat{x}_{2,A}) - u^-(x_{2,A})) < 0$, (B.13) becomes

$$[\hat{w}^-(p) - w^-(p)][u^-(x_{2,A}) - u^-(x_{1,A}))] > [\hat{w}^-(p) - w^-(p)][u^-(x_{2,B}) - u^-(x_{1,B}))]. \tag{B.14}$$

From the construction of the table, we know $u^-(x_{2,A}) - u^-(x_{1,A})) < u^-(x_{2,B}) - u^-(x_{1,B})) < 0$. (B.7) thus implies $\hat{w}^-(p) < w^-(p)$. A switching row lower in the LD1/LD2 tables thus leads to a smaller decision weight on the more negative outcome of each lottery.

Since in CPT, the probability weighting function uses the upper Choquet integral in the gain domain and the lower Choquet integral in the loss domain, the implications of this theoretical result can be interpreted as for the gain domain described above. We have structured the tables such that $p^{LD1} < p^{LD2}$ and assume the inflection point of the probability weighting function to be between these two probabilities. Under this construction, if the decision weight on p^{LD1} becomes smaller, we have less inverse S-shape or more S-shape in the weighting function. If the decision weight on p^{LD2} becomes smaller, we have more inverse S-shape or less S-shape in the weighting function.

B.1.3 Loss Aversion

Assuming a full CPT preference functional, we evaluate the lotteries in the LA table as $EV = w^+(p)u^+(gain) + \lambda w^-(1-p)u^-(loss)$. As such, the probability of the gain leading to indifference fulfills

$$w^+(p)u^+(x_{2,A}) + \lambda w^-(1-p)u^-(x_{1,A}) = w^+(p)u^+(x_{2,B}) + \lambda w^-(1-p)u^-(x_{1,B}) \tag{B.15}$$

or equivalently

$$\lambda = \frac{w^+(p)(u^+(x_{2,B}) - u^+(x_{2,A}))}{w^-(1-p)(u^-(x_{1,A}) - u^-(x_{1,B}))}. \tag{B.16}$$

Given functions w^+, w^-, u^+, u^-, we can see that a higher indifference probability always leads to a higher λ. Observe

$$\frac{\partial \lambda}{\partial p} = \frac{\frac{\partial w^+(p)}{\partial p}(u^+(x_{2,B}) - u^+(x_{2,A}))}{w^-(1-p)(u^-(x_{1,A}) - u^-(x_{1,B}))} - \frac{\partial w^-(1-p)}{\partial p} \frac{w^+(p)(u^+(x_{2,B}) - u^+(x_{2,A}))}{[w^-(1-p)(u^-(x_{1,A}) - u^-(x_{1,B}))]^2} > 0. \tag{B.17}$$

B.1.4 Certainty Preference

To accommodate a preference for certainty, we utilize the model of Schmidt (1998). We integrate this model into CPT similarly as he has integrated it into EU. All evaluations of uncertain payments are made by a utility function $u(x)$ and evaluations of certain payments are made by a value function $v(x)$.[30] For a preference regarding certainty to appear, the utility and value functions must differ, whereas the difference is usually assumed to be monotonic for all values of x (Vieider, 2018). As such, when we can observe $v(x) > u(x)$ ($v(x) < u(x)$) for some value of x, we assume the preference for (against) certainty to persist everywhere in the preferences of the individual except at zero. Zero is excluded such that the prospect-theoretic convention that the utility/value function crosses at the origin is upheld.

The only table in which a certain payment appears is the CP table. It is structured such that $x_{2,B} > x_A > x_{1,B} > 0$. It is obvious that when we assume that $u(x)$ does not change, a higher value of $x_{2,B}$ implies a stronger certainty effect. Consider the two equalities implied by indifference values $x_{2,B}$ and $\hat{x}_{2,B}$ with $\hat{x}_{2,B} > x_{2,B}$ and the corresponding value functions $v(x)$ and $\hat{v}(x)$.

$$v(x_A) = w^+(p)u^+(x_{2,B}) + (1 - w^+(p))u^+(x_{1,B}). \tag{B.18}$$

$$\hat{v}(x_A) = w^+(p)u^+(\hat{x}_{2,B}) + (1 - w^+(p))u^+(x_{1,B}). \tag{B.19}$$

[30] See Schmidt (1998) for a discussion of the terminology "utility function" and "value function."

Subtracting (B.18) from (B.19), we gain

$$\hat{v}(x_A) - v(x_A) = w^+(p)[u^+(\hat{x}_{2,B}) - u^+(x_{2,B})]. \tag{B.20}$$

By assumption, $\hat{x}_{2,B} > x_{2,B}$ renders $\hat{v}(x_A) > v(x_A)$ and thus with a fixed utility function of $u(x)$ an increase in the certainty preference.

B.2 Analysis Using Parametric Preference Measures

For the parametric measures of the preference parameters, we assume that a switch from Lottery A to Lottery B at row h implies that the subject would be indifferent between the average lotteries A and B between rows $h-1$ and h. In every table, only one value varies with each row. As such, the average lotteries are the simple mid-point between the two different values. We signify this indifference value with a bar. When the switch occurred at a corner solution (which did not violate FOSD), then we assume the indifference value to be the same distance below the value indicated in the row as it would be above it if the switch occurred one row before the corner solution.

For the utility function and the probability weighting function, we assume the parametric forms summarized in Section 2.1. Using these assumptions, we identify the values of γ^+, γ^-, β^+, β^-, λ and κ which solve the six equations

$$\begin{array}{lcl}
w^+(p^{GD1})u^+(x^{GD1}_{1,A}) + (1 - w^+(p^{GD1}))u^+(x^{GD1}_{2,A}) & = & w^+(p^{GD1})u^+(\bar{x}^{GD1}_{1,B}) + (1 - w^+(p^{GD1}))u^+(x^{GD1}_{2,B}) \\
w^+(p^{GD2})u^+(x^{GD2}_{1,A}) + (1 - w^+(p^{GD2}))u^+(x^{GD2}_{2,A}) & = & w^+(p^{GD2})u^+(\bar{x}^{GD2}_{1,B}) + (1 - w^+(p^{GD2}))u^+(x^{GD2}_{2,B}) \\
w^-(p^{LD1})u^-(\bar{x}^{LD1}_{1,A}) + (1 - w^-(p^{LD1}))u^-(x^{LD1}_{2,A}) & = & w^-(p^{LD1})u^-(x^{LD1}_{1,B}) + (1 - w^-(p^{LD1}))u^-(x^{LD1}_{2,B}) \\
w^-(p^{LD2})u^-(\bar{x}^{LD2}_{1,A}) + (1 - w^-(p^{LD2}))u^-(x^{LD2}_{2,A}) & = & w^-(p^{LD2})u^-(x^{LD2}_{1,B}) + (1 - w^-(p^{LD2}))u^-(x^{LD2}_{2,B}) \\
w^+(\bar{p}^{LA})u^+(x^{LA}_{1,A}) + \lambda w^-(1 - \bar{p}^{LA})u^-(x^{LA}_{2,A}) & = & w^+(\bar{p}^{LA})u^+(x^{LA}_{1,B}) + \lambda w^-(1 - \bar{p}^{LA})u^-(x^{LA}_{2,B}) \\
v(x^{CP}_A) & = & w^+(p^{CP})u^+(\bar{x}^{CP}_{1,B}) + (1 - w^+(p^{CP}))u^+(x^{CP}_{2,B}).
\end{array} \tag{B.21}$$

Subjects who chose a lottery which was first-order dominated by another one were excluded from the analysis. The first two equations jointly determine γ^+ and β^+. Every pair of choices in both tables lead to a unique vector of these two variables. Similarly, the third and fourth equation jointly determine γ^- and β^-. Based on the inferred values of γ^+, β^+, γ^- and β^- and the known parametric forms of u^+, w^+, u^- and w^-, λ and κ can be inferred from the equations

$$\lambda = \frac{w^+(\bar{p}^{LA})(u^+(x^{LA}_{2,B}) - u^+(x^{LA}_{2,A}))}{w^-(1 - \bar{p}^{LA})(u^-(x^{LA}_{1,A}) - u^-(x^{LA}_{1,B}))} \tag{B.22}$$

and

$$\kappa^{1-\gamma} = \frac{w^+(p^{CP/GD1})u^+(\bar{x}^{CP}_{2,B}) + (1 - w^+(p^{CP/GD1}))u^+(x^{CP/GD1}_{1,B})}{w^+(p^{CP/GD1})u^+(\bar{x}^{GD1}_{2,B}) + (1 - w^+(p^{CP/GD1}))u^+(x^{CP/GD1}_{1,B})}. \tag{B.23}$$

In Equation (B.23), the index $CP/GD1$ indicates that the values are equal in both tables.

Directly eliciting values of λ and κ results in bad measures of loss aversion and certainty preference (Abdellaoui et al., 2007; Schmidt, 1998). For loss aversion, the shapes of u^+ and u^- also influence how utility is traded off between the gain domain and the loss domain. We thus use an index for loss aversion that takes this into account. We use the negative average ratio of utility in the loss domain and the gain domain over the spectrum of relevant values for the LA table.

Formally, our loss aversion index is described by

$$\hat{\lambda} = -\int_0^5 \frac{\lambda u^-(-x)}{u^+(x)dx} dx. \tag{B.24}$$

For certainty preferences, we use the marginal certainty preference index suggested by Schmidt (1998):

$$\hat{\kappa} = \frac{v(x) - u(x)}{u'(x)}. \tag{B.25}$$

Appendix C Summary statistics

Table C.1: Summary statistics for demographic and experiment variables

	UW		mTurk		Total	
	Mean	SD	Mean	SD	Mean	SD
Age	21.48	2.34	36.14	10.81	32.94	11.37
Dummy for female	0.72	0.45	0.40	0.49	0.47	0.50
Dummy for US	1.00	0.00	0.85	0.36	0.88	0.32
Dummy for white	0.69	0.46	0.69	0.46	0.69	0.46
Dummy for black	0.04	0.20	0.06	0.24	0.06	0.23
Dummy for asian	0.29	0.45	0.21	0.41	0.23	0.42
Dummy for latino	0.02	0.15	0.06	0.25	0.06	0.23
GRQ	5.28	2.13	5.23	2.76	5.24	2.63
Final payment to subject	6.41	4.65	6.23	5.05	6.27	4.97
Dummy for made $\geq$ 1 FOSD choice	0.14	0.34	0.30	0.46	0.26	0.44
Num times violated FOSD	1.44	0.73	1.75	0.98	1.72	0.96
Dummy for pref task instr wrong	0.12	0.32	0.14	0.35	0.13	0.34
Dummy for ins task instr wrong	0.17	0.37	0.18	0.38	0.18	0.38
Understanding score	1.94	0.80	1.76	0.82	1.80	0.82
Education:						
Less than high school	0.00		0.01		0.00	
High school graduate	0.24		0.10		0.13	
Some college, no degree	0.57		0.20		0.28	
Associate's college degree	0.02		0.11		0.09	
Bachelor's college degree	0.13		0.44		0.37	
Master's degree	0.03		0.12		0.10	
Doctoral degree	0.00		0.01		0.01	
Professional degree (JD, MD)	0.00		0.02		0.01	
Income ($ USD):						
Less than 5,000	0.00		0.06		0.04	
5,000–9,999	0.00		0.06		0.04	
10,000–24,999	0.00		0.16		0.13	
25,000–49,999	0.00		0.28		0.22	
50,000–74,999	0.00		0.20		0.16	
75,000–99,999	0.00		0.13		0.10	
100,000–149,999	0.00		0.08		0.06	
150,000 or greater	0.00		0.04		0.03	

Note: Summary statistics are presented for the 378 in-person subjects at UW-Madison, the 1,352 online subjects recruited through Amazon mTurk, and all 1,730 subjects together. GRQ is the self-reported risk aversion measure of Dohmen et al. (2011). Final payment to the subject does not include the $6 "show-up fee" paid to UW participants. Num times violated FOSD is conditional on having violated FOSD at least once. Understanding score is the subject's rating of how easy or difficult the study was to understand, with 1 indicating very easy, 5 indicating very difficult, and 3 indicating a neutral response.

Table C.2: Summary statistics for nonparametric preference measures

	UW Only $n = 378$		mTurk Only $n = 1,352$		Full Sample $n = 1,730$	
	Mean	Median	Mean	Median	Mean	Median
UC^+	14.83	15.00	14.77	15.00	14.78	15.00
PW^+	1.04	1.00	2.91	2.00	2.50	2.00
CP	-0.51	0.00	-0.51	0.00	-0.51	0.00
UC^-	20.06	20.00	20.59	21.00	20.47	21.00
PW^-	-3.00	-3.00	-0.67	0.00	-1.18	0.00
LA	11.30	11.00	11.83	12.00	11.71	12.00

Note: The labels $UC, PW, CP,$ and LA denote our measures of utility curvature, probability weighting, certainty preference, and loss aversion, respectively. Superscripts $^+$ ($^-$) indicate that the measure was elicited in the gain (loss) domain.

Table C.3: Correlation table for nonparametric preference measures

	UC^+_{std}	PW^+_{std}	CP_{std}	UC^-_{std}	PW^-_{std}	LA_{std}
UC^+_{std}	1.000	.	.	.	.	.
PW^+_{std}	0.044	1.000	.	.	.	.
CP_{std}	-0.304	0.268	1.000	.	.	.
UC^-_{std}	-0.120	0.120	0.034	1.000	.	.
PW^-_{std}	0.008	0.294	0.005	0.214	1.000	.
LA_{std}	0.289	-0.010	-0.015	0.016	-0.038	1.000

Note: For our sample size of 1,730 subjects, a correlation is statistically significant at the 10% level if it is larger than 0.040, at the 5% level if it is larger than 0.047, and at the 1% level if it is larger than 0.062 (in absolute values).

Appendix D Tobit regressions

Table D.1: Tobit regressions of coverage level on each preference separately

Panel (a): Nonparametric	(1)	(2)	(3)	(4)	(5)	(6)
UC^{+}_{std}	2.31*					
	(0.98)					
PW^{+}_{std}		0.69				
		(0.99)				
CP_{std}			−0.93			
			(0.90)			
UC^{-}_{std}				2.45*		
				(1.00)		
PW^{-}_{std}					3.86***	
					(0.92)	
LA_{std}						3.17**
						(1.04)
Pseudo R^2	0.01	0.01	0.01	0.01	0.01	0.01
N subjects	1,730	1,730	1,730	1,730	1,730	1,730
N choices	20,760	20,760	20,760	20,760	20,760	20,760
N left-censored	4,321	4,321	4,321	4,321	4,321	4,321
N uncensored	12,871	12,871	12,871	12,871	12,871	12,871
N right-censored	3,568	3,568	3,568	3,568	3,568	3,568
Panel (b): Parametric	**(7)**	**(8)**	**(9)**	**(10)**	**(11)**	**(12)**
γ^{+}_{std}	2.90**					
	(1.11)					
β^{+}_{std}		2.85**				
		(1.14)				
$\hat{k}_{std}$			−1.92			
			(1.22)			
γ^{-}_{std}				3.04**		
				(1.14)		
β^{-}_{std}					3.44***	
					(1.07)	
$\hat{\lambda}_{std}$						2.79*
						(1.31)
Pseudo R^2	0.01	0.01	0.01	0.01	0.01	0.01
N subjects	1,276	1,276	1,276	1,276	1,276	1,276
N choices	15,312	15,312	15,312	15,312	15,312	15,312
N left-censored	3,220	3,220	3,220	3,220	3,220	3,220
N uncensored	9,456	9,456	9,456	9,456	9,456	9,456
N right-censored	2,636	2,636	2,636	2,636	2,636	2,636

Note: Tobit regressions are specified with a left-censored limit of 0 and a right-censored limit of 100. Dependent variable is the coverage level selected by subject i. All models contain fixed effects for each of the 12 insurance scenarios. Regressions of parametric preference motives are limited to the 1,276 non FOSD-violating subjects. In each panel, p-values have been adjusted using the Šidák (1967) method, adjusting for the six hypotheses tested in each panel. Standard errors clustered by subject are in parentheses. Stars *, **, and *** denote statistical significance at the 0.10, 0.05, and 0.01 levels, respectively.

Table D.2: Tobit regressions of coverage level on joint preferences

Nonparam./Param.	Nonparametric		Parametric	
	(1)	(2)	(3)	(4)
$UC^{+}_{std}/\gamma^{+}_{std}$	1.62	1.63	4.01***	4.03***
	(1.07)	(1.07)	(1.20)	(1.20)
$PW^{+}_{std}/\beta^{+}_{std}$	−0.57	0.96	1.92	3.86*
	(1.10)	(1.37)	(1.25)	(1.59)
$CP_{std}/\hat{k}_{std}$	−0.33	−0.34	−1.67	−1.67
	(1.02)	(1.02)	(1.23)	(1.23)
$UC^{-}_{std}/\gamma^{-}_{std}$	1.91	1.92	2.40	2.41
	(1.02)	(1.02)	(1.20)	(1.20)
$PW^{-}_{std}/\beta^{-}_{std}$	3.73***	5.76***	2.62*	4.52**
	(0.99)	(1.27)	(1.15)	(1.51)
$LA_{std}/\hat{\lambda}_{std}$	2.82**	2.81*	4.21**	4.20**
	(1.06)	(1.05)	(1.41)	(1.42)
Prob ≥ 40		49.95***		53.24***
		(1.88)		(2.17)
Prob $\geq 40 \times PW^{+}_{std}/\beta^{+}_{std}$		−4.48***		−5.63***
		(1.32)		(1.55)
Prob $\geq 40 \times PW^{-}_{std}/\beta^{-}_{std}$		−5.92***		−5.51***
		(1.33)		(1.59)
FOSD violators	Yes	Yes	No	No
Fixed effects	Scenario	Scenario	Scenario	Scenario
Clustered SEs	Subject	Subject	Subject	Subject
Pseudo R^2	0.01	0.01	0.02	0.02
N subjects	1,730	1,730	1,276	1,276
N choices	20,760	20,760	15,312	15,312
N left-censored	4,321	4,321	3,220	3,220
N uncensored	12,871	12,871	9,456	9,456
N right-censored	3,568	3,568	2,636	2,636

Note: Tobit regressions are specified with a left-censored limit of 0 and a right-censored limit of 100. Dependent variable is the coverage level selected by subject i. All models contain fixed effects for each of the 12 insurance scenarios. Column 1 contains only the nonparametric preference scales. Column 2 adds an interaction between the PW preferences and a dummy for "high probability" (Prob $\geq 40 = 1$ if the probability of loss is 40% or 70%). Columns 3 and 4 replicate columns 1 and 2 using the parametric preferences. In each column, p-values have been adjusted using the Šidák (1967) method. Standard errors clustered by subject are in parentheses. Stars *, **, and *** denote statistical significance at the 0.10, 0.05, and 0.01 levels, respectively.

Appendix E Analysis with alternative samples

Table E.1: Regression of coverage level on nonparametric preferences, excluding subjects who made FOSD-violating choices

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
UC^+_{std}	1.99**						1.32	1.32
	(0.78)						(0.86)	(0.86)
PW^+_{std}		1.10					0.65	1.83
		(0.77)					(0.82)	(1.05)
CP_{std}			−0.76				−0.72	−0.72
			(0.81)				(0.90)	(0.90)
UC^-_{std}				2.25**			1.99*	1.99
				(0.79)			(0.82)	(0.82)
PW^-_{std}					2.00**		1.13	2.25
					(0.70)		(0.78)	(1.00)
LA_{std}						4.57***	4.20***	4.20***
						(0.91)	(0.92)	(0.92)
Prob ≥ 40								37.00***
								(1.28)
Prob $\geq 40 \times PW^+_{std}$								−3.54***
								(1.06)
Prob $\geq 40 \times PW^-_{std}$								−3.38**
								(1.06)
FOSD violators	No	No	No	No	No	No	No	No
Fixed effects	Scenario	Scenario	Scenario	Scenario	Scenario	Scenario	Scenario	Scenario
Clustered SEs	Subject	Subject	Subject	Subject	Subject	Subject	Subject	Subject
R^2	0.13	0.12	0.12	0.13	0.13	0.13	0.14	0.14
N choices	15,312	15,312	15,312	15,312	15,312	15,312	15,312	15,312
N subjects	1,276	1,276	1,276	1,276	1,276	1,276	1,276	1,276

Note: Sample is restricted to subjects who did not violate FOSD. Dependent variable is the coverage level selected by subject i. All models contain fixed effects for each of the 12 insurance scenarios. Columns 1-6 replicate columns 1-6 in Panel A of Table 10, with p-values adjusted for six hypotheses using the Šidák (1967) method. Columns 7 and 8 replicate columns 1 and 2 in Table 11, with Šidák-adjusted p-values for the hypotheses in each column. Standard errors clustered by subject are in parentheses. Stars *, **, and *** denote statistical significance at the 0.10, 0.05, and 0.01 levels, respectively.

Table E.2: Correlation between actual and predicted coverage levels, by number of law of demand violations

Model	0 violations (24.4%)	1 violation (26.4%)	2 violations (22.5%)	3+ violations (26.6%)
EV	0.249	0.288	0.261	0.252
EU^{+}	0.185	0.217	0.230	0.131
EU^{-}	0.220	0.245	0.204	0.148
DT^{+}	0.049	0.029	−0.001	−0.028
DT^{-}	0.055	0.028	0.021	−0.127
KR	0.063	0.024	0.108	0.029
$RDEU^{+}$	0.026	−0.070	−0.087	−0.146
$RDEU^{-}$	0.094	0.029	0.022	−0.126
CPT^{-}	0.105	0.033	0.017	−0.108
CPT^{NLIB}	0.069	−0.019	0.019	−0.110
EV_{CP}	0.098	0.114	0.089	0.096
DT^{+}_{CP}	−0.003	−0.028	−0.042	−0.034
DT^{-}_{CP}	−0.007	−0.007	−0.010	−0.079
EU^{+}_{CP}	0.006	0.066	0.101	0.041
EU^{-}_{CP}	0.003	0.067	0.090	0.021
$RDEU^{+}_{CP}$	−0.012	−0.074	−0.093	−0.114
$RDEU^{-}_{CP}$	0.020	−0.008	−0.011	−0.085

Table E.3: Regressions of observed coverage level on predicted coverage level with scenario fixed effects, by number of law of demand violations

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Panel (a): 0 law of demand violations ($N = 5,208$)										
$PredictedCovLvl$	–	−0.19	0.08	0.05*	0.05	0.05	0.10***	0.12***	0.10**	0.13***
		(0.12)	(0.07)	(0.03)	(0.04)	(0.06)	(0.04)	(0.04)	(0.04)	(0.04)
Model	EV	EU^+	EU^-	DT^+	DT^-	KR	$RDEU^+$	$RDEU^-$	CPT^{NLIB}	CPT^-
Fixed effect	Scenario	Scenario	Scenario	Scenario	Scenario	Scenario	Scenario	Scenario	Scenario	Scenario
R^2	0.08	0.09	0.09	0.09	0.09	0.09	0.09	0.10	0.09	0.10
Panel (b): 1 law of demand violation ($N = 5,184$)										
$PredictedCovLvl$	–	−0.04	0.09	0.07***	0.07***	−0.01	0.07**	0.09***	0.04	0.08***
		(0.08)	(0.05)	(0.02)	(0.02)	(0.04)	(0.03)	(0.03)	(0.03)	(0.03)
Model	EV	EU^+	EU^-	DT^+	DT^-	KR	$RDEU^+$	$RDEU^-$	CPT^{NLIB}	CPT^-
Fixed effect	Scenario	Scenario	Scenario	Scenario	Scenario	Scenario	Scenario	Scenario	Scenario	Scenario
R^2	0.15	0.15	0.15	0.16	0.16	0.15	0.16	0.16	0.15	0.16
Panel (c): 2 law of demand violations ($N = 4,680$)										
$PredictedCovLvl$	–	0.20***	0.01	0.03	0.07***	0.10**	0.05*	0.08***	0.09***	0.07**
		(0.07)	(0.05)	(0.03)	(0.03)	(0.05)	(0.03)	(0.03)	(0.03)	(0.03)
Model	EV	EU^+	EU^-	DT^+	DT^-	KR	$RDEU^+$	$RDEU^-$	CPT^{NLIB}	CPT^-
Fixed effect	Scenario	Scenario	Scenario	Scenario	Scenario	Scenario	Scenario	Scenario	Scenario	Scenario
R^2	0.13	0.13	0.13	0.13	0.13	0.13	0.13	0.14	0.13	0.13
Panel (d): 3+ law of demand violations ($N = 5,688$)										
$PredictedCovLvl$	–	−0.04	−0.00	0.01	0.03	0.01	0.05**	0.03	−0.00	0.03
		(0.06)	(0.04)	(0.02)	(0.02)	(0.03)	(0.02)	(0.02)	(0.03)	(0.02)
Model	EV	EU^+	EU^-	DT^+	DT^-	KR	$RDEU^+$	$RDEU^-$	CPT^{NLIB}	CPT^-
Fixed effect	Scenario	Scenario	Scenario	Scenario	Scenario	Scenario	Scenario	Scenario	Scenario	Scenario
R^2	0.19	0.19	0.19	0.19	0.19	0.19	0.19	0.19	0.19	0.19

Note: Dependent variable is the coverage level selected by the subject. Explanatory variable of interest is the predicted level of coverage using the specified model. The only other explanatory variables are scenario fixed effects. The sample in each regression is the subjects who violated the law of demand the number of times noted in each panel heading. In each panel, p-values have been adjusted using the Šidák (1967) method for the ten hypotheses tested. Standard errors clustered by subject are in parentheses. Stars *, **, and *** denote statistical significance at the 0.10, 0.05, and 0.01 levels, respectively.

Appendix F Alternatively scaled insurance choices

The analysis in the main text treats insurance demand as a continuous and cardinal scaled measure. This appendix demonstrates that our results are robust to alternative assumptions about the scale of the variable. The first two columns of Table F.1 compare the correlations between observed demand and model predictions when treating insurance demand as cardinal (column 1, equal to analysis in main text) and treating it as ordinal (column 2).

The other analyses reported in this appendix treat insurance demand as a discrete measure. In the first analysis, we round answers of subjects in the experiment to the nearest increment of 25%. We adjust our predictions such that they indicate which of five possible discrete alternatives (0%, 25%, 50%, 75%, 100%) maximizes the objective function of the subject. Analyses are then carried out using estimators that take the ordinal nature of the reported scales into account. The correlations reported in the rightmost two columns of Table F.1 are calculated as Kendall's Tau. The regressions in Tables F.2 and F.3 use an ordered logit estimator. In the analyses using predicted insurance demand as the independent variable, we use dummy variables to indicate the predicted value. The second analysis of insurance demand as a discrete measure restricts observations and predictions even more and only considers no insurance (0%) or full insurance (100%) as possible options. The results in Tables F.4 and F.5 use logit estimators and a dummy variable for 100% coverage to reflect this binary nature of insurance demand.

Table F.1: Correlation between actual and predicted coverage levels

	Continuous		25% bins	0% or 100%
Model	Pearson	Kendall's Tau	Kendall's Tau	Kendall's Tau
EV	0.261	0.201	0.212	0.247
EU^+	0.189	0.123	0.144	0.123
EU^-	0.205	0.142	0.157	0.150
DT^+	0.015	0.011	0.009	0.004
DT^-	−0.002	0.001	−0.007	−0.016
KR	0.053	0.043	0.047	0.050
$RDEU^+$	−0.064	−0.050	−0.059	−0.088
$RDEU^-$	0.010	0.008	0.004	−0.015
CPT^-	0.017	0.013	0.008	−0.008
CPT^{NLIB}	−0.008	−0.014	−0.019	−0.037
EV_{CP}	0.098	0.051	0.029	0.021
DT^+_{CP}	−0.025	−0.004	−0.009	−0.008
DT^-_{CP}	−0.023	−0.006	−0.013	−0.020
EU^+_{CP}	0.050	−0.003	−0.001	0.003
EU^-_{CP}	0.042	−0.005	−0.003	0.009
$RDEU^+_{CP}$	−0.069	−0.036	−0.043	−0.036
$RDEU^-_{CP}$	−0.019	−0.015	−0.019	−0.025

Note: For our sample size of 1,276 subjects for whom we can make parametric predictions, the correlation is statistically significant at the 10% level if it is larger than 0.047, at the 5% level if it is larger than 0.055, and at the 1% level if it is larger than 0.073 (in absolute values).

Table F.2: Ordered logit regressions of coverage level (rounded to nearest 25%) on joint preferences

Nonparam./Param.	Nonparametric		Parametric	
	(1)	(2)	(3)	(4)
$UC^{+}_{std}/\gamma^{+}_{std}$	0.05	0.04	0.12***	0.12**
	(0.03)	(0.03)	(0.04)	(0.04)
$PW^{+}_{std}/\beta^{+}_{std}$	−0.00	0.05	0.08	0.15**
	(0.03)	(0.05)	(0.04)	(0.05)
$CP_{std}/\hat{\kappa}_{std}$	−0.02	−0.02	−0.06	−0.06
	(0.03)	(0.03)	(0.04)	(0.04)
$UC^{-}_{std}/\gamma^{-}_{std}$	0.05	0.05	0.07	0.08
	(0.03)	(0.03)	(0.04)	(0.04)
$PW^{-}_{std}/\beta^{-}_{std}$	0.10***	0.18***	0.06	0.13*
	(0.03)	(0.04)	(0.04)	(0.05)
$LA_{std}/\hat{\lambda}_{std}$	0.10**	0.10**	0.13**	0.13*
	(0.03)	(0.03)	(0.05)	(0.05)
Prob ≥ 40		1.78***		1.90***
		(0.07)		(0.09)
Prob $\geq 40 \times PW^{+}_{std}/\beta^{+}_{std}$		−0.14**		−0.18***
		(0.04)		(0.05)
Prob $\geq 40 \times PW^{-}_{std}/\beta^{-}_{std}$		−0.19***		−0.17***
		(0.04)		(0.05)
FOSD violators	Yes	Yes	No	No
Fixed effects	Scenario	Scenario	Scenario	Scenario
Clustered SEs	Subject	Subject	Subject	Subject
Pseudo R^2	0.039	0.041	0.045	0.047
N choices	20,760	20,760	15,312	15,312
N subjects	1,730	1,730	1,276	1,276

Note: Dependent variable is the coverage level selected by subject i, rounded to the nearest 25%. All models contain fixed effects for each of the 12 insurance scenarios. Column 1 contains only the nonparametric preference scales. Column 2 adds an interaction between the PW preferences and a dummy for "high probability" (Prob $\geq 40 = 1$ if the probability of loss is 40% or 70%). Columns 3 and 4 replicate columns 1 and 2 using the parametric preferences. In each column, p-values have been adjusted using the Šidák (1967) method. Standard errors clustered by subject are in parentheses. Stars *, **, and *** denote statistical significance at the 0.10, 0.05, and 0.01 levels, respectively.

Table F.3: Regressions of observed coverage level on predicted coverage level (25% bins)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Panel (a): No fixed effects										
$PredictedCovLvl(25\%)$	–	0.099	0.222***	–	–	–	0.203*	0.146	0.269***	0.462***
		(0.051)	(0.056)				(0.079)	(0.073)	(0.073)	(0.067)
$PredictedCovLvl(50\%)$	0.627***	−0.221***	−0.102*	–	–	–	−0.008	0.269***	0.275***	0.284***
	(0.032)	(0.055)	(0.061)				(0.081)	(0.067)	(0.069)	(0.064)
$PredictedCovLvl(75\%)$	–	0.016	0.037	–	–	–	0.118	0.173***	0.198**	0.277***
		(0.061)	(0.069)				(0.076)	(0.065)	(0.078)	(0.065)
$PredictedCovLvl(100\%)$	1.354***	0.860***	0.902***	0.060	0.006	0.157***	−0.157***	0.042	0.025	0.062
	(0.044)	(0.040)	(0.039)	(0.056)	(0.051)	(0.060)	(0.056)	(0.057)	(0.059)	(0.060)
Model	EV	EU^+	EU^-	DT^+	DT^-	KR	$RDEU^+$	$RDEU^-$	CPT^{NLIB}	CPT^-
Fixed effect	None	None	None	None	None	None	None	None	None	None
Pseudo R^2	0.014	0.013	0.012	0.000	0.000	0.001	0.001	0.000	0.001	0.001
Panel (b): Probability fixed effects										
$PredictedCovLvl(25\%)$	–	0.040	0.232***	–	–	–	0.073	0.114	0.069	0.024
		(0.058)	(0.061)				(0.077)	(0.073)	(0.071)	(0.064)
$PredictedCovLvl(50\%)$	0.330***	0.065	0.175***	–	–	–	0.028	0.145**	0.188***	0.116*
	(0.037)	(0.061)	(0.067)				(0.077)	(0.065)	(0.069)	(0.063)
$PredictedCovLvl(75\%)$	–	0.209***	0.234***	–	–	–	0.149*	0.210***	0.242***	0.122*
		(0.071)	(0.077)				(0.077)	(0.069)	(0.079)	(0.065)
$PredictedCovLvl(100\%)$	0.603***	0.430***	0.440***	0.300***	0.364***	0.124**	0.341***	0.406***	0.236***	0.379***
	(0.043)	(0.040)	(0.039)	(0.058)	(0.054)	(0.063)	(0.064)	(0.062)	(0.063)	(0.065)
Model	EV	EU^+	EU^-	DT^+	DT^-	KR	$RDEU^+$	$RDEU^-$	CPT^{NLIB}	CPT^-
Fixed effect	Prob	Prob	Prob	Prob	Prob	Prob	Prob	Prob	Prob	Prob
Pseudo R^2	0.035	0.035	0.035	0.034	0.035	0.033	0.035	0.036	0.034	0.035

TABLE CONTINUES ON NEXT PAGE

[Cont'd] Regressions of observed coverage level on predicted coverage level (25% bins)

CONTINUED FROM PREVIOUS PAGE	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Panel (c): Loading fixed effects										
$PredictedCovLvl(25\%)$	–	0.108*	0.238***	–	–	–	0.167**	0.246***	0.271***	0.454***
		(0.062)	(0.062)				(0.079)	(0.075)	(0.075)	(0.067)
$PredictedCovLvl(50\%)$	0.234***	−0.278***	−0.130*	–	–	–	−0.056	0.241***	0.281***	0.343***
	(0.077)	(0.065)	(0.067)				(0.077)	(0.068)	(0.073)	(0.065)
$PredictedCovLvl(75\%)$	–	−0.413***	−0.351***	–	–	–	−0.021	0.223***	0.167**	0.235***
		(0.073)	(0.080)				(0.076)	(0.068)	(0.081)	(0.068)
$PredictedCovLvl(100\%)$	0.376***	0.286***	0.275***	−0.100*	−0.164***	0.110*	−0.144**	−0.077	−0.031	−0.015
	(0.079)	(0.069)	(0.065)	(0.058)	(0.054)	(0.063)	(0.056)	(0.059)	(0.062)	(0.062)
Model	EV	EU^+	EU^-	DT^+	DT^-	KR	$RDEU^+$	$RDEU^-$	CPT^{NLIB}	CPT^-
Fixed effect	Load	Load	Load	Load	Load	Load	Load	Load	Load	Load
Pseudo R^2	0.026	0.028	0.027	0.026	0.026	0.026	0.026	0.026	0.026	0.027
Panel (d): Scenario fixed effects										
$PredictedCovLvl(25\%)$	–	0.006	0.206***	–	–	–	0.065	0.120	0.083	0.120*
		(0.064)	(0.065)				(0.079)	(0.075)	(0.073)	(0.063)
$PredictedCovLvl(50\%)$	0.223***	−0.011	0.140*	–	–	–	0.015	0.142**	0.187**	0.125*
	(0.077)	(0.077)	(0.077)				(0.079)	(0.068)	(0.073)	(0.064)
$PredictedCovLvl(75\%)$	–	0.003	0.055	–	–	–	0.108	0.238***	0.217***	0.128*
		(0.098)	(0.097)				(0.081)	(0.072)	(0.083)	(0.068)
$PredictedCovLvl(100\%)$	0.381***	0.279***	0.267***	0.191***	0.239***	0.111*	0.248***	0.296***	0.159**	0.289***
	(0.080)	(0.070)	(0.066)	(0.064)	(0.065)	(0.063)	(0.070)	(0.069)	(0.068)	(0.070)
Model	EV	EU^+	EU^-	DT^+	DT^-	KR	$RDEU^+$	$RDEU^-$	CPT^{NLIB}	CPT^-
Fixed effect	Scenario	Scenario	Scenario	Scenario	Scenario	Scenario	Scenario	Scenario	Scenario	Scenario
Pseudo R^2	0.036	0.036	0.037	0.037	0.037	0.036	0.037	0.038	0.037	0.037

Note: Dependent variable is the coverage level selected by the subject, rounded to the nearest 25%. Explanatory variables of interest are a set of dummy variables indicating the specified model predicted a coverage level of 25%, 50%, 75%, or 100% (a prediction of 0% coverage being the excluded category). The only other explanatory variables are fixed effects as noted. The sample in each regression is the 1,276 subjects with estimated preference parameters (who did not violate FOSD). In each panel, *p*-values have been adjusted using the Šidák (1967) method for the ten hypotheses tested. Standard errors clustered by subject are in parentheses. Stars *, **, and *** denote statistical significance at the 0.10, 0.05, and 0.01 levels, respectively.

Table F.4: Ordered logit regressions of coverage level (rounded to 0% or 100%) on joint preferences

Nonparam./Param.	Nonparametric		Parametric	
	(1)	(2)	(3)	(4)
$UC^{+}_{std}/\gamma^{+}_{std}$	0.04	0.04	0.11**	0.12*
	(0.04)	(0.04)	(0.04)	(0.04)
$PW^{+}_{std}/\beta^{+}_{std}$	0.01	0.06	0.11**	0.15**
	(0.04)	(0.04)	(0.04)	(0.05)
$CP_{std}/\hat{\kappa}_{std}$	−0.02	−0.02	−0.05	−0.05
	(0.04)	(0.04)	(0.04)	(0.04)
$UC^{-}_{std}/\gamma^{-}_{std}$	0.05	0.05	0.09*	0.09
	(0.04)	(0.04)	(0.04)	(0.04)
$PW^{-}_{std}/\beta^{-}_{std}$	0.12***	0.17***	0.06	0.11
	(0.04)	(0.04)	(0.04)	(0.05)
$LA_{std}/\hat{\lambda}_{std}$	0.11**	0.11**	0.13**	0.14**
	(0.04)	(0.04)	(0.05)	(0.05)
Prob ≥ 40		2.26***		2.47***
		(0.08)		(0.10)
Prob $\geq 40 \times PW^{+}_{std}/\beta^{+}_{std}$		−0.18***		−0.14
		(0.05)		(0.06)
Prob $\geq 40 \times PW^{-}_{std}/\beta^{-}_{std}$		−0.16**		−0.20**
		(0.05)		(0.06)
FOSD violators	Yes	Yes	No	No
Fixed effects	Scenario	Scenario	Scenario	Scenario
Clustered SEs	Subject	Subject	Subject	Subject
Pseudo R^2	0.097	0.100	0.113	0.115
N choices	20,760	20,760	15,312	15,312
N subjects	1,730	1,730	1,276	1,276

Note: Dependent variable is coverage level selected by subject i, rounded to the closer of 0% or 100%. All models contain fixed effects for each of the 12 insurance scenarios. Column 1 contains only the nonparametric preference scales. Column 2 adds an interaction between the PW preferences and a dummy for "high probability" (Prob $\geq 40 = 1$ if the probability of loss is 40% or 70%). Columns 3 and 4 replicate columns 1 and 2 using the parametric preferences. In each column, p-values have been adjusted using the Šidák (1967) method. Standard errors clustered by subject are in parentheses. Stars *, **, and *** denote statistical significance at the 0.10, 0.05, and 0.01 levels, respectively.

Table F.5: Regressions of observed coverage level on predicted coverage level (0% or 100%)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Panel (a): No fixed effects										
$PredictedCovLvl(100\%)$	2.009***	0.420***	0.544***	0.036	−0.031	0.147	−0.240***	−0.021	−0.060	0.003
	(0.084)	(0.039)	(0.041)	(0.055)	(0.050)	(0.063)	(0.053)	(0.052)	(0.055)	(0.054)
Model	EV	EU^+	EU^-	DT^+	DT^-	KR	$RDEU^+$	$RDEU^-$	CPT^{NLIB}	CPT^-
Fixed effect	None	None	None	None	None	None	None	None	None	None
Pseudo R^2	0.029	0.007	0.011	0.000	0.000	0.001	0.003	0.000	0.000	0.000
Panel (b): Probability fixed effects										
$PredictedCovLvl(100\%)$	0.815***	0.319***	0.350***	0.328***	0.372***	0.124*	0.283***	0.355***	0.201***	0.316***
	(0.080)	(0.046)	(0.049)	(0.062)	(0.058)	(0.071)	(0.066)	(0.062)	(0.063)	(0.063)
Model	EV	EU^+	EU^-	DT^+	DT^-	KR	$RDEU^+$	$RDEU^-$	CPT^{NLIB}	CPT^-
Fixed effect	Prob	Prob	Prob	Prob	Prob	Prob	Prob	Prob	Prob	Prob
Pseudo R^2	0.088	0.088	0.088	0.088	0.089	0.085	0.087	0.089	0.086	0.088
Panel (c): Loading fixed effects										
$PredictedCovLvl(100\%)$	0.670***	−0.236***	−0.149**	−0.134**	−0.221***	0.095	−0.241***	−0.140**	−0.126**	−0.068
	(0.141)	(0.058)	(0.060)	(0.059)	(0.056)	(0.068)	(0.057)	(0.057)	(0.060)	(0.059)
Model	EV	EU^+	EU^-	DT^+	DT^-	KR	$RDEU^+$	$RDEU^-$	CPT^{NLIB}	CPT^-
Fixed effect	Load	Load	Load	Load	Load	Load	Load	Load	Load	Load
Pseudo R^2	0.062	0.063	0.062	0.062	0.063	0.062	0.064	0.062	0.062	0.062
Panel (d): Scenario fixed effects										
$PredictedCovLvl(100\%)$	0.670***	0.114	0.148**	0.217***	0.236***	0.107	0.194***	0.255***	0.126*	0.239***
	(0.141)	(0.070)	(0.068)	(0.067)	(0.069)	(0.071)	(0.070)	(0.067)	(0.067)	(0.067)
Model	EV	EU^+	EU^-	DT^+	DT^-	KR	$RDEU^+$	$RDEU^-$	CPT^{NLIB}	CPT^-
Fixed effect	Scenario	Scenario	Scenario	Scenario	Scenario	Scenario	Scenario	Scenario	Scenario	Scenario
Pseudo R^2	0.092	0.092	0.092	0.093	0.093	0.092	0.093	0.094	0.092	0.094

Note: Dependent variable is the coverage level selected by the subject, rounded to the closer of 0% or 100%. Explanatory variable of interest is a dummy variable equal to 1 if the specified model predicts 100% coverage and 0 otherwise. The only other explanatory variables are fixed effects as noted. The sample in each regression is the 1,276 subjects with estimated preference parameters (who did not violate FOSD). In each panel, p-values have been adjusted using the Šidák (1967) method for the ten hypotheses tested. Standard errors clustered by subject are in parentheses. Stars *, **, and *** denote statistical significance at the 0.10, 0.05, and 0.01 levels, respectively.

Appendix G Collection of experimental data

Here, we provide screenshots of the full experiment given to the mTurk subjects. In-person subjects completed an identical experiment, without some mTurk-specific questions (e.g., entering the mTurk worker ID). Each box represents a different screen.

The experiment begins with a browser compatibility test and confirmation of the mTurk worker ID. This also prevents bots from continuing the experiment.

Before we begin, we need to make sure your browser is compatible with our study.

Set the slider to 75.

| 1417500 |

0 10 20 30 40 50 60 70 80 90 100

Slider 75

What is the red number that appears between the lines above the slider? If you don't see a number, enter 100.

Please verify that the ID in the field below is your correct Amazon Mechanical Turk ID.

- If it is your ID, please click "Next."
- If it is not your ID, or if no ID is displayed, please enter your ID and click "Next."

Then, subjects are presented with an overview of the experiment.

Overview

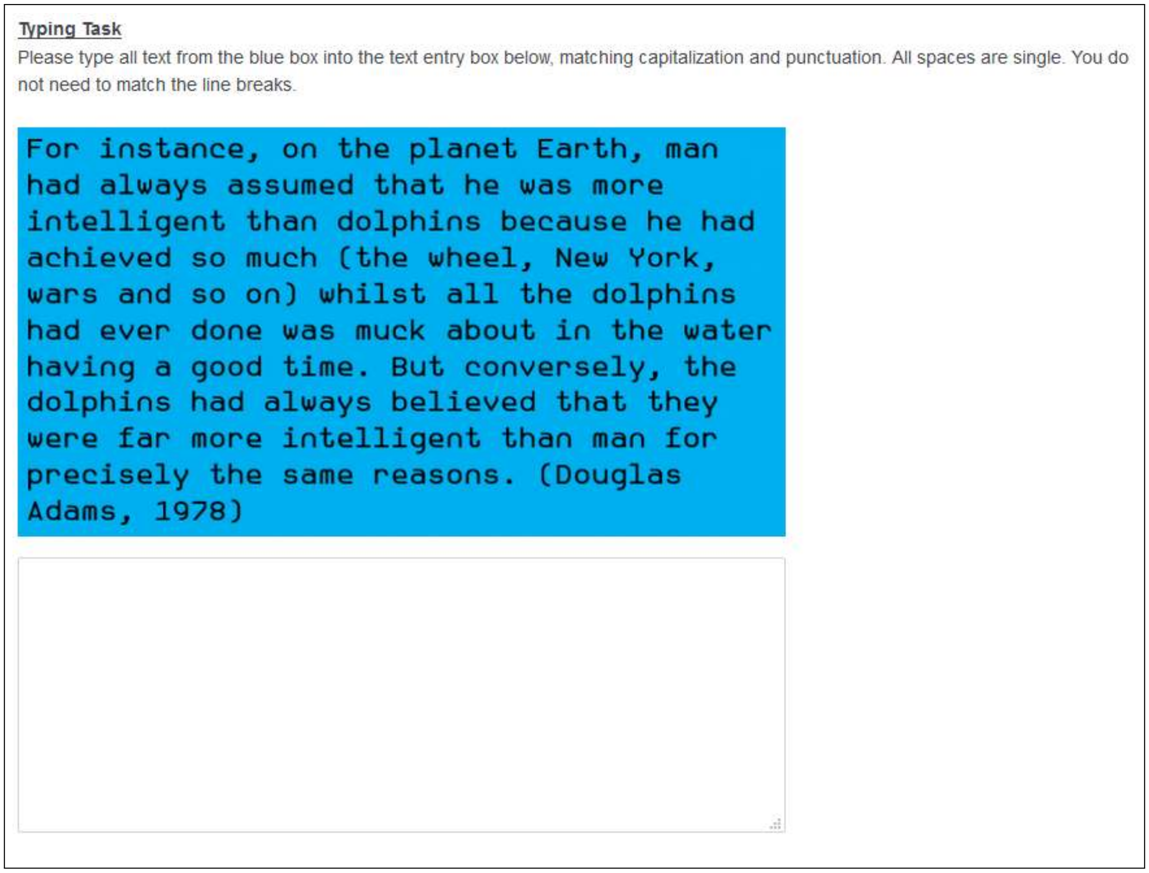
Thank you for participating in this study. You will begin this experiment by completing a typing task to earn $5.00. Your final payment at the end of this experiment may be more or less than $5.00, depending partly on your later choices and partly on chance.

After you complete the typing task, we will ask you a series of economic questions. In these questions, you may gain additional money or lose some of your $5.00. Whether you have a gain or a loss is based on the draw of colored balls from a bucket. In different questions you will be asked to make decisions related to these buckets that can affect how much you earn. Each question is designated with a number (Q1, Q2, Q3, etc…). At the end of the experiment, the computer will randomly select a question number. We will apply your choice in that question and the computer will play it out for real money.

Even though only one of your choices will count, you will not know in advance which question will be used to determine your ultimate earnings. Therefore, you should think about each of them carefully before submitting your choice.

There are no correct answers - we are simply interested in your preferences for risky versus safe outcomes.

Following the overall instructions, subjects complete the real-effort task. This involves manually typing a passage correctly into the provided text box. Every character must be correct to pass.[31] The font used in this passage is resistant to optical character recognition. There were five possible passages, and each subject was randomly assigned to complete two of them.

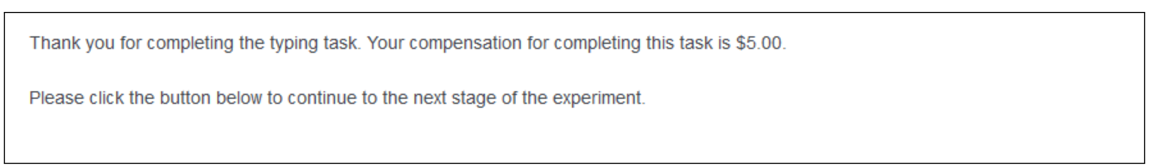


Typing Task

Please type all text from the blue box into the text entry box below, matching capitalization and punctuation. All spaces are single. You do not need to match the line breaks.

For instance, on the planet Earth, man had always assumed that he was more intelligent than dolphins because he had achieved so much (the wheel, New York, wars and so on) whilst all the dolphins had ever done was muck about in the water having a good time. But conversely, the dolphins had always believed that they were far more intelligent than man for precisely the same reasons. (Douglas Adams, 1978)

Subjects are then notified of their payment for the task.



Thank you for completing the typing task. Your compensation for completing this task is $5.00.

Please click the button below to continue to the next stage of the experiment.

31 If the subject took longer than three minutes for a particular passage, we allowed them to continue to the next passage. This was not disclosed to subjects. In every case, the subject had entered the full text but had made a typo.

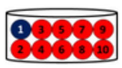
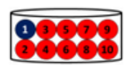
The lottery task begins with a set of instructions and three questions to ensure subjects understood them.

Part 1 Instructions

In this part, you must decide whether you would prefer to draw a ball from Bucket A or Bucket B. Each bucket contains a mixture of colored balls, and the color of the ball you draw will determine how much you gain or lose. We will tell you the mixture of the balls and the dollar value of each colored ball, and you should consider this information in choosing a bucket.

Here is an example of the buckets you will see in this part. Use the buttons in the middle to indicate whether you would prefer to draw from Bucket A or Bucket B.

You currently have $5.00.

	Bucket A			Bucket B	
	● Gain $0.50 ● Lose $0.20 Ball mixture: 10% ●, 90% ●			● Gain $5.00 ● Lose $2.00 Ball mixture: 10% ●, 90% ●	
Example		● Gain $0.50 ● Lose $0.20	○ ○	● Gain $5.00 ● Lose $2.00	

Here are a few questions to make sure you understand the instructions.

What would happen if you drew a red ball from Bucket B?

- ○ Gain $0.50
- ○ Lose $0.20
- ○ Gain $5.00
- ◉ Lose $2.00

What is the chance of drawing a blue ball from Bucket A?

- ◉ 10%
- ○ 20%
- ○ 50%
- ○ 90%

If you lost $2.00, what would your total payment be at the end of this experiment?

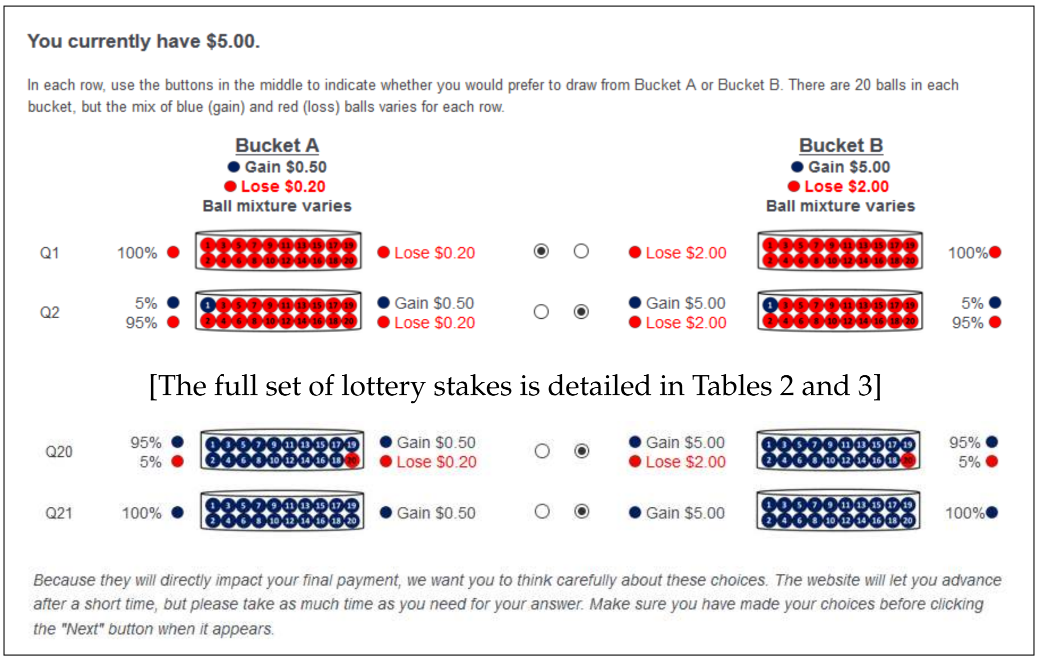
- ○ $0.00
- ○ $2.00
- ◉ $3.00
- ○ $5.00

Subjects who answer the questions correctly may continue to the lottery task. Subjects who answer incorrectly may not continue until they have answered the questions correctly.

> Yes, that's right. Now we will ask you to make similar choices between two buckets.

Lotteries are presented in random order, and question numbers update automatically. In this example, the LA questions are presented first.

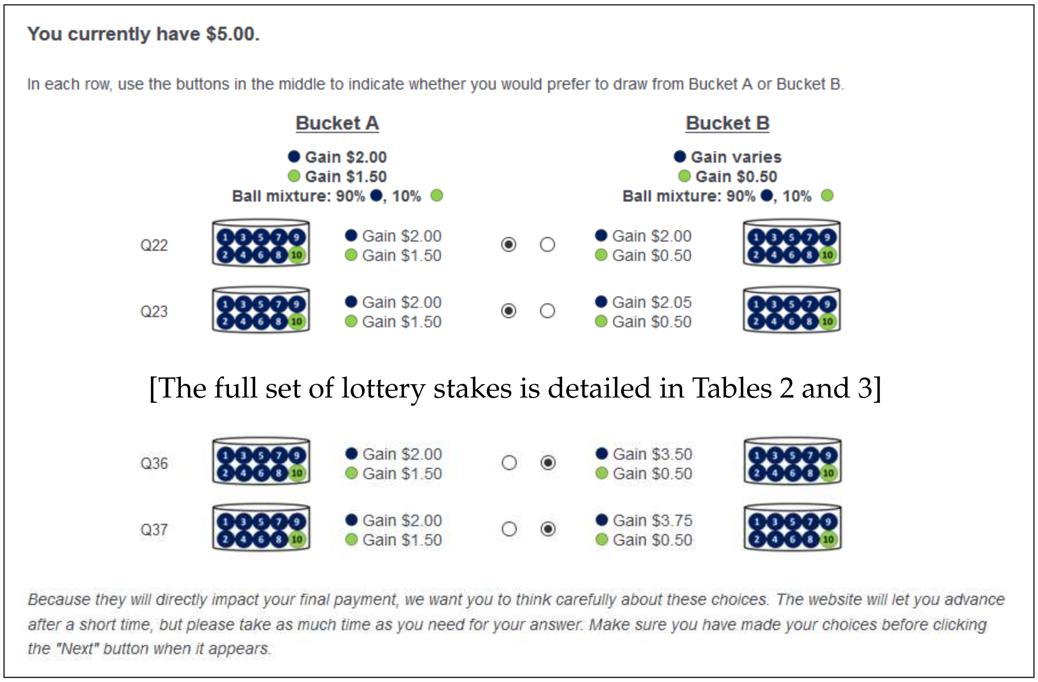
LA



To encourage subjects to make careful choices, the "Next" button is hidden for 20 seconds on each lottery page.

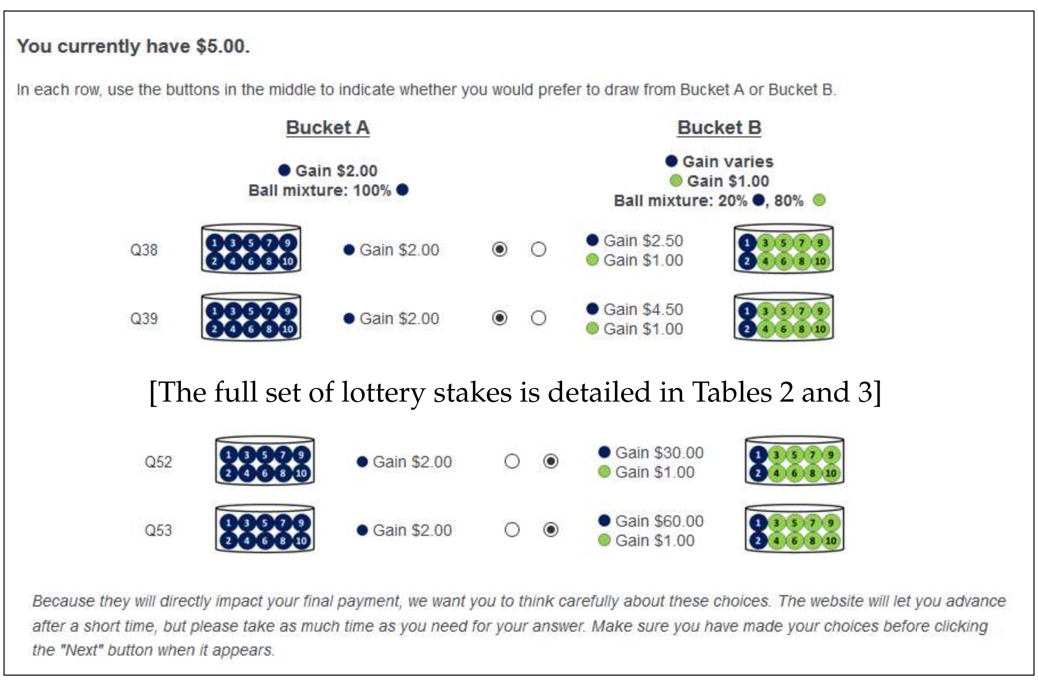
In this example, the GD2 lottery is randomly-selected to appear second.

GD2



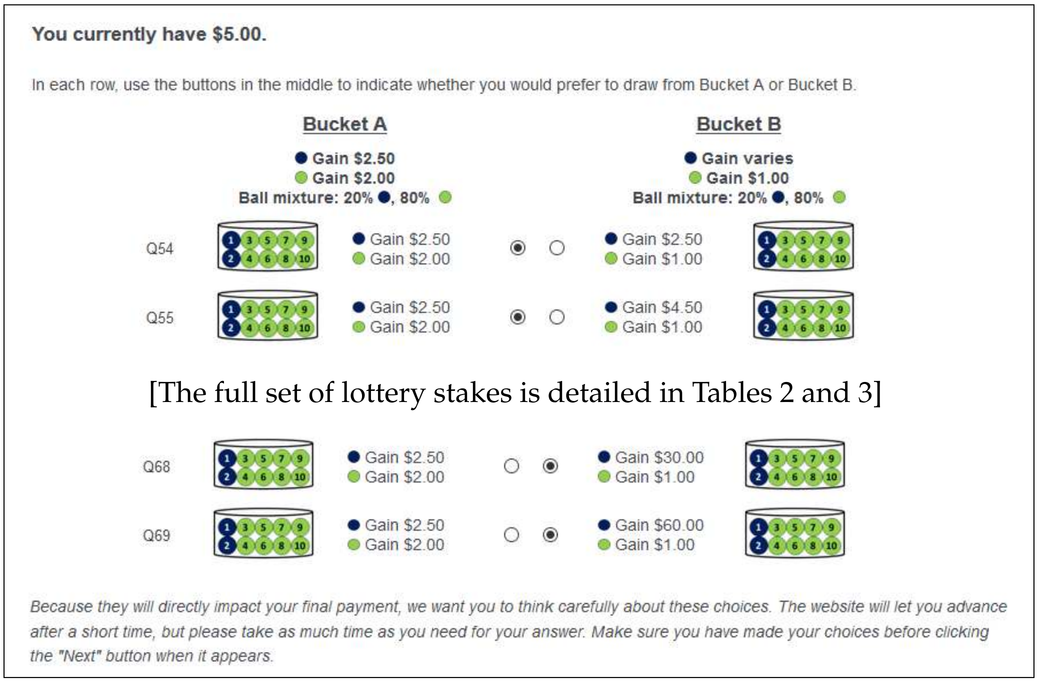
In this example, the CP table is randomly-selected to appear third.

CP



In this example, the GD1 table is randomly-selected to appear fourth.

GD1



You currently have $5.00.

In each row, use the buttons in the middle to indicate whether you would prefer to draw from Bucket A or Bucket B.

	Bucket A			Bucket B	
		● Gain $2.50 ● Gain $2.00 Ball mixture: 20% ●, 80% ●		● Gain varies ● Gain $1.00 Ball mixture: 20% ●, 80% ●	
Q54		● Gain $2.50 ● Gain $2.00	◉ ○	● Gain $2.50 ● Gain $1.00	
Q55		● Gain $2.50 ● Gain $2.00	◉ ○	● Gain $4.50 ● Gain $1.00	

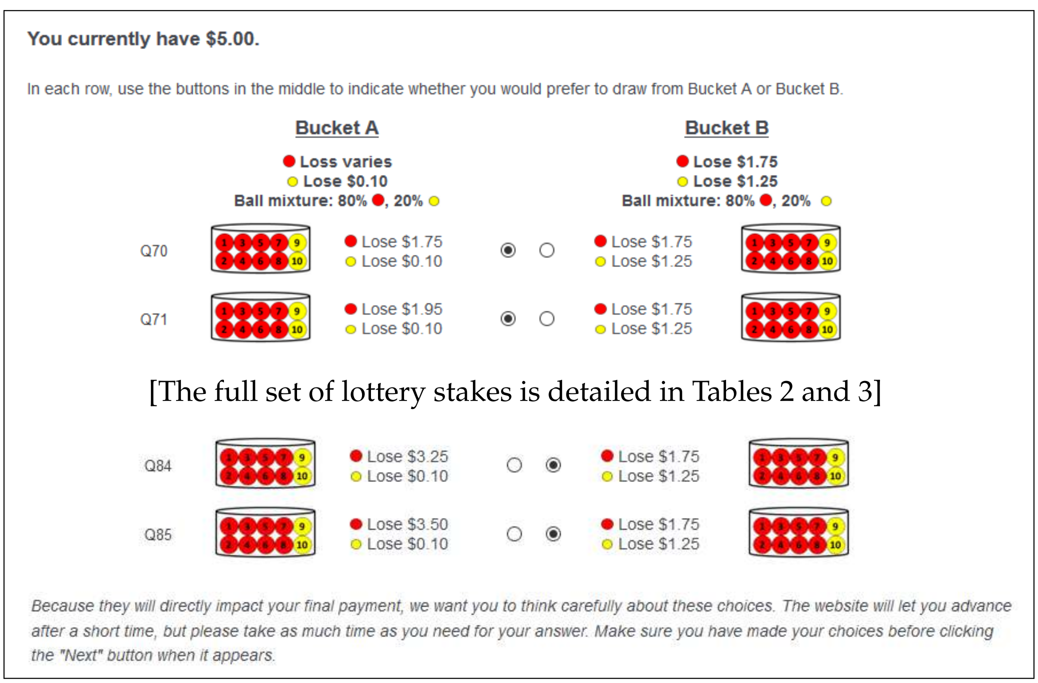
[The full set of lottery stakes is detailed in Tables 2 and 3]

Q68		● Gain $2.50 ● Gain $2.00	○ ◉	● Gain $30.00 ● Gain $1.00	
Q69		● Gain $2.50 ● Gain $2.00	○ ◉	● Gain $60.00 ● Gain $1.00	

Because they will directly impact your final payment, we want you to think carefully about these choices. The website will let you advance after a short time, but please take as much time as you need for your answer. Make sure you have made your choices before clicking the "Next" button when it appears.

In this example, the LD2 table is randomly-selected to appear fifth.

LD2



You currently have $5.00.

In each row, use the buttons in the middle to indicate whether you would prefer to draw from Bucket A or Bucket B.

	Bucket A			Bucket B	
		● Loss varies ○ Lose $0.10 Ball mixture: 80% ●, 20% ○		● Lose $1.75 ○ Lose $1.25 Ball mixture: 80% ●, 20% ○	
Q70		● Lose $1.75 ○ Lose $0.10	◉ ○	● Lose $1.75 ○ Lose $1.25	
Q71		● Lose $1.95 ○ Lose $0.10	◉ ○	● Lose $1.75 ○ Lose $1.25	

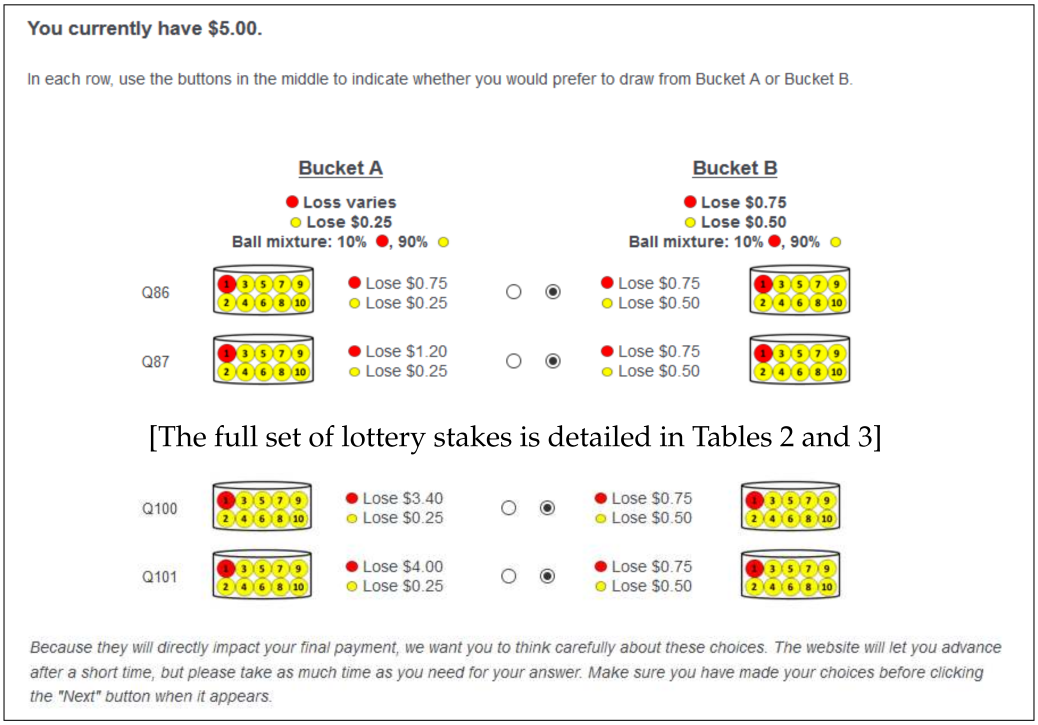
[The full set of lottery stakes is detailed in Tables 2 and 3]

Q84		● Lose $3.25 ○ Lose $0.10	○ ◉	● Lose $1.75 ○ Lose $1.25	
Q85		● Lose $3.50 ○ Lose $0.10	○ ◉	● Lose $1.75 ○ Lose $1.25	

Because they will directly impact your final payment, we want you to think carefully about these choices. The website will let you advance after a short time, but please take as much time as you need for your answer. Make sure you have made your choices before clicking the "Next" button when it appears.

Finally, the LD1 table is presented sixth.

LD1



You currently have $5.00.

In each row, use the buttons in the middle to indicate whether you would prefer to draw from Bucket A or Bucket B.

	Bucket A			Bucket B	
		Loss varies Lose $0.25 Ball mixture: 10% red, 90% yellow		Lose $0.75 Lose $0.50 Ball mixture: 10% red, 90% yellow	
Q86		Lose $0.75 Lose $0.25	○ ◉	Lose $0.75 Lose $0.50	
Q87		Lose $1.20 Lose $0.25	○ ◉	Lose $0.75 Lose $0.50	

[The full set of lottery stakes is detailed in Tables 2 and 3]

Q100		Lose $3.40 Lose $0.25	○ ◉	Lose $0.75 Lose $0.50	
Q101		Lose $4.00 Lose $0.25	○ ◉	Lose $0.75 Lose $0.50	

Because they will directly impact your final payment, we want you to think carefully about these choices. The website will let you advance after a short time, but please take as much time as you need for your answer. Make sure you have made your choices before clicking the "Next" button when it appears.

The insurance task begins with a set of instructions and two questions to ensure subjects understood them.

Part 2 Instructions

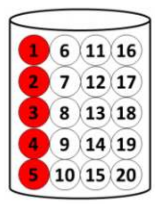
In this part, you will face a number of scenarios in which you may lose $3.00 of your money. Whether you have a loss depends on the draw of one ball from a bucket filled with 20 balls. Drawing a red ball results in a loss of $3.00, while drawing a white ball results in no loss.

You have the option to buy insurance to protect some or all of your money from loss. The price of insurance will be different in every scenario. Scenarios may also have a different chance of loss.

If you want to buy insurance, you must pay for it out of your $5.00 before you draw a ball at the end of the experiment. You may choose to buy varying levels of insurance, from 0% of the potential loss (no insurance) to 100% of the potential loss (full insurance). You may click the slider as many times as you like to see the different options. Your final earnings will depend on the price of the insurance you choose and any uninsured loss you may experience.

Here is an example question so you can see how your choices influence your final payment. Click the slider to practice selecting an insurance amount.

You currently have $5.00.



You have a 25% chance of drawing a red ball and losing $3.00.

For every 10% of the loss you want to cover with insurance, you would have to pay about 9 cents. Please select your desired amount of insurance (if any) using the slider. The exact price of your choice will display below.

- **Insurance coverage: 50%**
- **Before drawing a ball, pay $0.47 for insurance.**

If you draw ●:	If you draw ○:
• $3.00 loss ∘ $1.50 covered by insurance ∘ $1.50 loss to you	• No loss

0% 10% 20% 30% 40% 50% 60% 70% 80% 90% 100%

Set Insurance Level 50

Here are a few questions to make sure you understand the instructions:

Select 50% insurance using the slider above. How much would you need to pay for this insurance before drawing a ball?

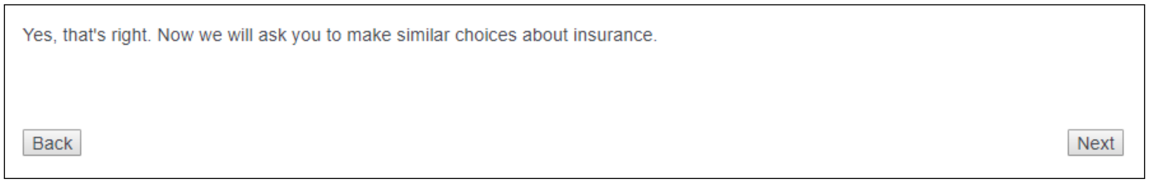
- ○ $0.09
- ○ $0.30
- ◉ $0.47
- ○ $0.94
- ○ $1.50

Select 50% insurance using the slider above. If you had this level of insurance and drew a red ball, how much of the $3.00 loss is covered by insurance?

- ○ $0.25
- ○ $0.47
- ○ $0.94
- ◉ $1.50
- ○ $3.00

Next

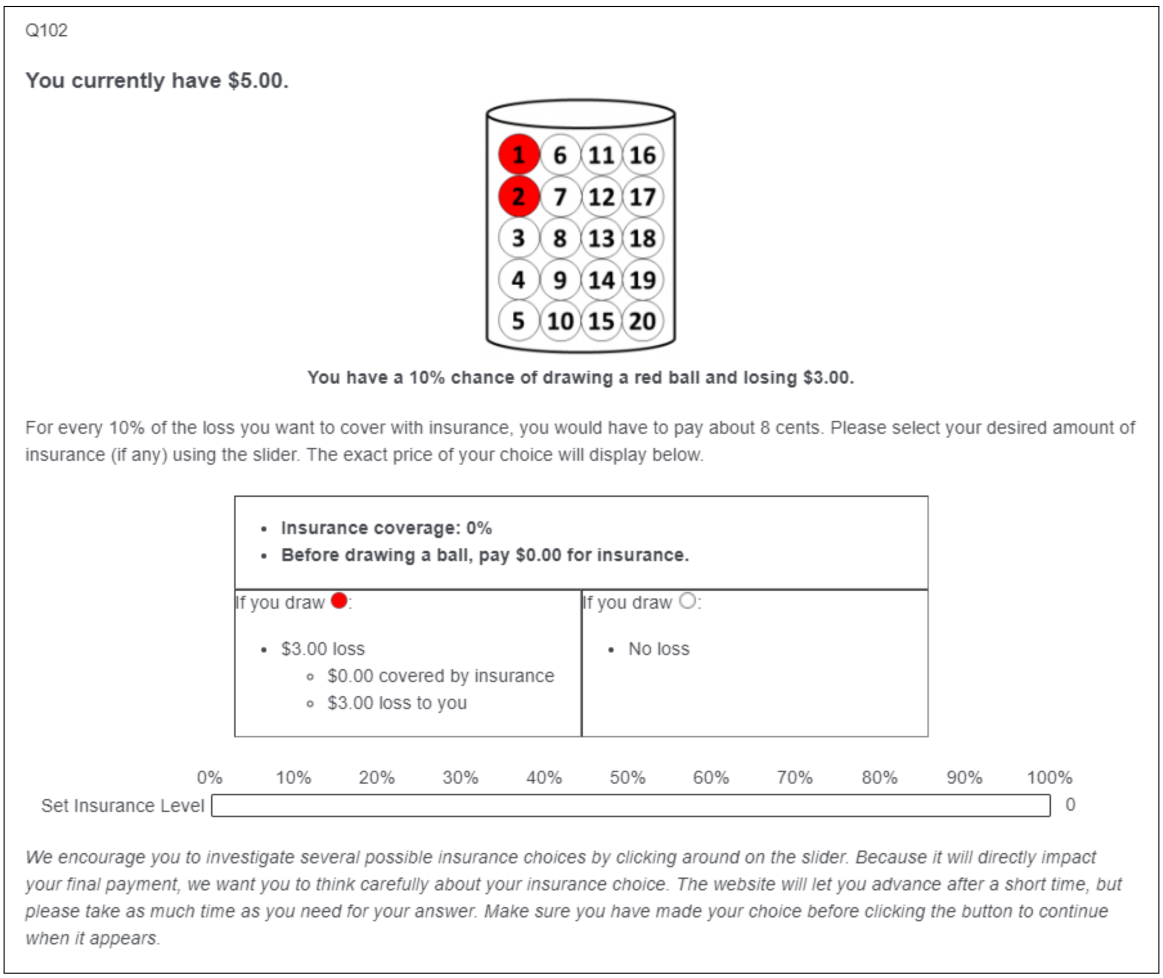
Subjects who answer the questions correctly may continue to the insurance task. Subjects who answer incorrectly may not continue until they have answered the questions correctly.



Yes, that's right. Now we will ask you to make similar choices about insurance.

Back | Next

Insurance scenarios are presented in random order, and question numbers update automatically. In this example, an insurance scenario with a 10% probability of loss is presented first. We have selected 0% coverage here.



Q102

You currently have $5.00.

You have a 10% chance of drawing a red ball and losing $3.00.

For every 10% of the loss you want to cover with insurance, you would have to pay about 8 cents. Please select your desired amount of insurance (if any) using the slider. The exact price of your choice will display below.

- **Insurance coverage: 0%**
- **Before drawing a ball, pay $0.00 for insurance.**

If you draw (red ball):
- $3.00 loss
  - $0.00 covered by insurance
  - $3.00 loss to you

If you draw (white ball):
- No loss

Set Insurance Level: 0% 10% 20% 30% 40% 50% 60% 70% 80% 90% 100% — 0

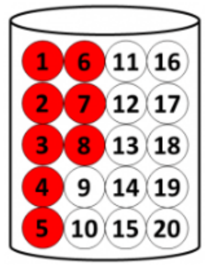
We encourage you to investigate several possible insurance choices by clicking around on the slider. Because it will directly impact your final payment, we want you to think carefully about your insurance choice. The website will let you advance after a short time, but please take as much time as you need for your answer. Make sure you have made your choice before clicking the button to continue when it appears.

To encourage subjects to make careful choices, the "Next" button is hidden for 10 seconds on each insurance page.

The next randomly-selected insurance scenario involves a 40% chance of loss. We have selected 50% coverage here.

Q103

You currently have $5.00.



You have a 40% chance of drawing a red ball and losing $3.00.

For every 10% of the loss you want to cover with insurance, you would have to pay about 18 cents. Please select your desired amount of insurance (if any) using the slider. The exact price of your choice will display below.

- **Insurance coverage: 50%**
- **Before drawing a ball, pay $0.90 for insurance.**

If you draw ●:	If you draw ○:
• $3.00 loss ∘ $1.50 covered by insurance ∘ $1.50 loss to you	• No loss

0% 10% 20% 30% 40% 50% 60% 70% 80% 90% 100%

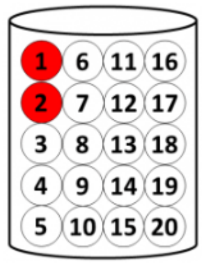
Set Insurance Level 50

We encourage you to investigate several possible insurance choices by clicking around on the slider. Because it will directly impact your final payment, we want you to think carefully about your insurance choice. The website will let you advance after a short time, but please take as much time as you need for your answer. Make sure you have made your choice before clicking the button to continue when it appears.

The next randomly-selected insurance scenario again involves a 10% chance of loss. We have selected 100% coverage here.

Q104

You currently have $5.00.



You have a 10% chance of drawing a red ball and losing $3.00.

For every 10% of the loss you want to cover with insurance, you would have to pay about 5 cents. Please select your desired amount of insurance (if any) using the slider. The exact price of your choice will display below.

- **Insurance coverage: 100%**
- **Before drawing a ball, pay $0.45 for insurance.**

If you draw (red ball):	If you draw (white ball):
• $3.00 loss ∘ $3.00 covered by insurance ∘ $0.00 loss to you	• No loss

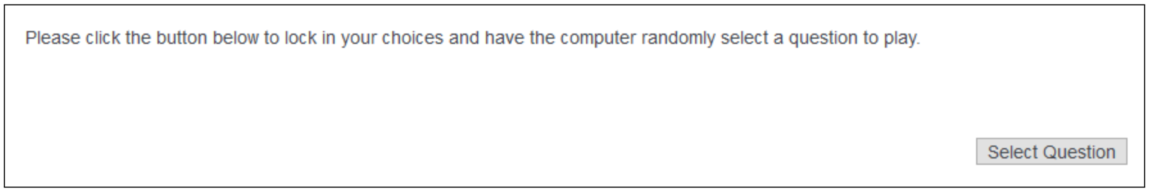
0% 10% 20% 30% 40% 50% 60% 70% 80% 90% 100%

Set Insurance Level 100

We encourage you to investigate several possible insurance choices by clicking around on the slider. Because it will directly impact your final payment, we want you to think carefully about your insurance choice. The website will let you advance after a short time, but please take as much time as you need for your answer. Make sure you have made your choice before clicking the button to continue when it appears.

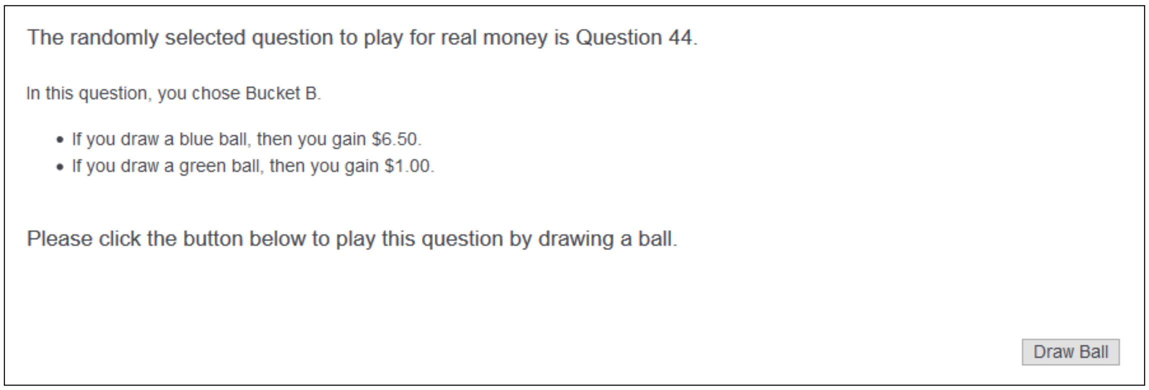
The nine remaining insurance scenarios have identical layouts, so we omit them from this appendix. The last question is numbered Q113.

After all choices are made in the experiment, the computer randomly selects one decision to play out.


Please click the button below to lock in your choices and have the computer randomly select a question to play.

Select Question

In this example, Question 44 is the randomly-selected decision. We confirm to the subject which selection he/she originally made and the possible outcomes. Then, they click the button to draw a "ball."


The randomly selected question to play for real money is Question 44.

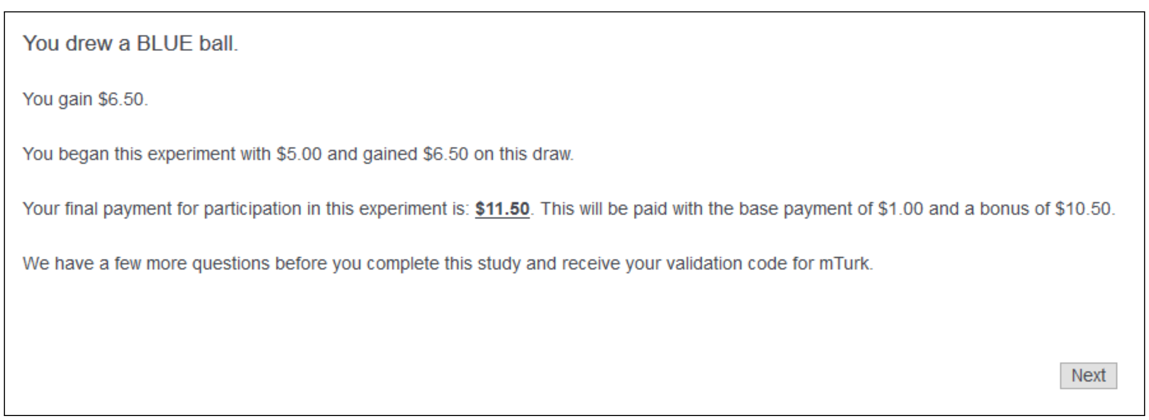
In this question, you chose Bucket B.

- If you draw a blue ball, then you gain $6.50.
- If you draw a green ball, then you gain $1.00.

Please click the button below to play this question by drawing a ball.

Draw Ball

The computer then displays the outcome of the draw and outlines the subject's payment. Amazon mTurk requires variable payments to be paid as a "base" plus a "bonus." We were clear in the mTurk posting that our task involved variable payments paid as bonuses. This appears a common way to compensate subjects, and no subjects expressed confusion about the payments. During our pilot studies, we had built a positive reputation in a number of third-party websites for paying bonuses quickly, so nonpayment risk was a minimal concern. All pilot participants were excluded from our main experiment.


You drew a BLUE ball.

You gain $6.50.

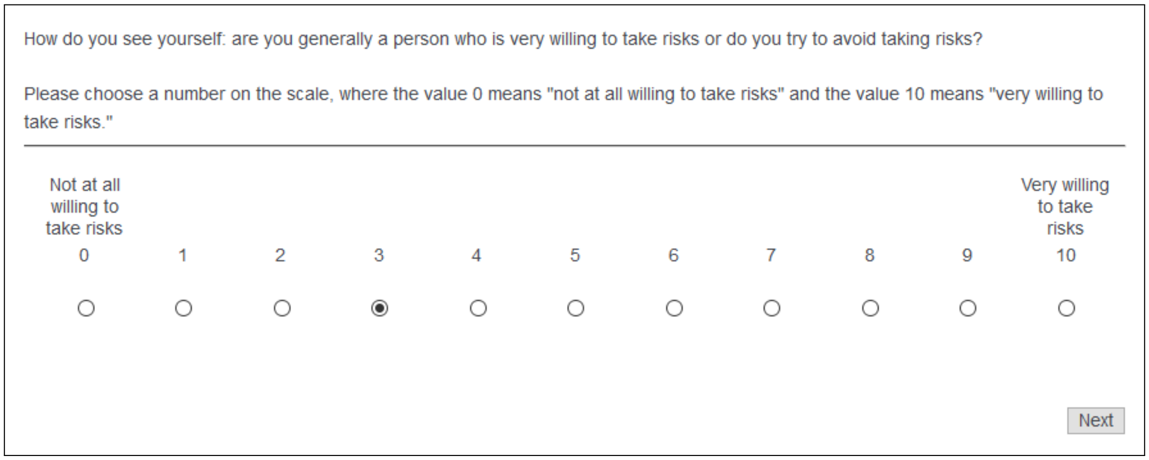
You began this experiment with $5.00 and gained $6.50 on this draw.

Your final payment for participation in this experiment is: **$11.50**. This will be paid with the base payment of $1.00 and a bonus of $10.50.

We have a few more questions before you complete this study and receive your validation code for mTurk.

Next

Immediately after the draw, we ask the GRQ. The direction of the scale (0-10 or 10-0) is randomly assigned.



Finally, we ask the following demographics questions and follow-up questions about the experiment.

What is your year of birth?

1983

What is your gender?

- (●) Male
- ( ) Female
- ( ) Transgender male
- ( ) Transgender female
- ( ) Gender variant/non-conforming
- ( ) Other
- ( ) Prefer not to answer

In which country do you currently reside?

In which U.S. state do you currently reside?

I do not reside in the United States

What is your race? You may choose all that apply.

- [ ] White or Caucasian
- [ ] Black or African American
- [ ] Hispanic or Latino
- [ ] Asian
- [ ] American Indian or Alaska Native
- [ ] Native Hawaiian or Pacific Islander
- [ ] Other

What was your approximate household income (in U.S. dollars, before taxes) in 2017? If you need to convert your income from another currency, you may calculate it here (opens in a new window).

- ( ) Less than $5,000
- ( ) $5,000 to $9,999
- ( ) $10,000 to $24,999
- ( ) $25,000 to $49,999
- ( ) $50,000 to $74,999
- ( ) $75,000 to $99,999
- ( ) $100,000 to $149,999
- ( ) $150,000 or greater

What is the highest level of school you have completed or the highest degree you have received?

- ○ Less than high school degree
- ○ High school graduate (high school diploma or equivalent including GED)
- ○ Some college but no degree
- ○ Associate degree in college (2-year)
- ○ Bachelor's degree in college (4-year)
- ○ Master's degree
- ○ Professional degree (JD, MD)
- ○ Doctoral degree

We have some final questions about your experience with this study.

Please rate your feelings on how easy or difficult this study was to UNDERSTAND:

Very easy to understand	Easy to understand	Neither easy nor difficult to understand	Difficult to understand	Very difficult to understand
○	○	○	○	○

If you had any technical difficulties completing this study, please explain below:

What do you think we were trying to find out in this study?

After all questions are answered, we provide a final "Submit" screen to remind subjects about the upcoming validation code.

Thank you for completing this study. Please click the button below to submit your responses. The next page will include a validation code. You must enter this validation code in mTurk so we can connect your final outcome in this survey to your mTurk account.

Submit Responses

For mTurk subjects, the study concludes with a validation code to connect their experiment outcomes to their mTurk account. In-person subjects were provided with a validation code to write down and bring to the experimenter for payment.

Thank you for your responses in this survey! Your validation code is:

7472154

Write this number down. Once you close this window, the validation code will disappear and cannot be recovered.

Please return to Amazon Mechanical Turk to enter the above validation code.