

**Sure-Thing vs. Probabilistic Charitable Giving: Experimental
Evidence on the Role of Individual Differences in Risky and
Ambiguous Charitable Decision-Making**

Miguel Costa-Gomes¹ & Philipp Schoenegger*²

¹ School of Economics and Finance, University of St Andrews

² School of Economics and Finance & School of Philosophical, Anthropological and Film Studies, University of St Andrews

*Corresponding author,

E-mail: ps234@st-andrews.ac.uk

Abstract

Charities differ, among other things, alongside the likelihood that their interventions succeed and produce the desired outcomes and alongside the extent that such likelihood can even be articulated numerically. In this paper, we investigate what best explains charitable giving behaviour regarding charities that have interventions that will succeed with a quantifiable and high probability (sure-thing charities) and charities that have interventions that only have a small and hard to quantify probability of bringing about the desired end (probabilistic charities). We study individual differences in risk/ambiguity attitudes, empathy, numeracy, optimism, and donor type (warm glow vs. pure altruistic donor type) as potential predictors of this choice. We conduct a money incentivised, pre-registered experiment on Prolific on a representative UK sample (n=1,506) to investigate participant choices (i) between these two types of charities and (ii) about one randomly selected charity. Overall, we find little to no evidence that individual differences predict choices regarding decisions about sure-thing and probabilistic charities, with the exception that a purely altruistic donor type predicts donations to probabilistic charities in (ii). Conducting exploratory equivalence tests, we find that the data provide robust evidence in favour of the absence of an effect (or a negligibly small effect) where we fail to reject the null. This is corroborated by exploratory Bayesian analyses. We take this paper to be contributing to the literature on charitable giving via this comprehensive null-result in pursuit of contributing to a cumulative science.

Keywords: Charitable Giving, Decision-Making under Uncertainty, Risk Aversion, Ambiguity Aversion, Individual Differences, Psychology of Economics

Introduction

Charitable giving has been the subject of much research across the social sciences. After all, substantial resources are spent every year on charitable ventures. For example, in 2019, US Americans donated just over 2% of GDP (“Giving USA 2020”, 2020) to a total of up to 1.5 million registered non-profit organisations (“Nonprofit Sector in Brief 2019”, 2019). The vast majority of these decisions are made under conditions of risk (with known outcome distributions) as well as ambiguity (with unknown outcome distributions). Plausibly, some charitable interventions have a high probability of providing the outcome promised; think of soup kitchens – where there is little uncertainty as to the intervention’s ability to produce the outcome: If one donates a certain amount to a soup kitchen, a few additional litres of soup will probably be dispensed quite soon. However, other charitable interventions are more probabilistic in nature, i.e., they only have a (sometimes very) small chance of making an impact and some aspects of this calculation are difficult to articulate numerically; think of interventions aimed at reducing the risk of nuclear war. Here, the donor is confronted with substantial additional uncertainty over the likelihood of an event (like nuclear war inciting incidents) arising and ambiguity over whether the charity’s interventions will have an impact on them if they should arise: For all the donor knows, there might not be an incident risking nuclear war this year and even if there was, the charity might not be able to meaningfully influence the outcome. The former level of uncertainty is frequently statable numerically, while the latter often cannot.

In this paper, we study actual donor decisions about these two kinds of charities: The first we call “sure-thing charities”, where donations are technically still made under conditions of some risk, though their risk is quantifiable and comparatively minor. The second we call “probabilistic charities”, where donations are also made under risk (e.g., relating to the likelihood of their event of concern arising), though other aspects of this choice are made under ambiguity (e.g. relating to the chance that an interventions will help address the issue that is itself extremely rare). When making choices between highly reliable and more uncertain options, variation in individual differences may drive heterogenous behaviour. In this paper, we specifically want to investigate what best explains donor decisions when faced with these types of charities, and whether individual differences in risk and/or ambiguity attitudes, empathy, numeracy, or donor type (warm glow vs. pure altruistic type) predict donation behaviour. Understanding why donors make the choices they do about these two types of charities has implications both for the academic literature and for fundraising efforts more directly.

There has been some previous work on the role of risk and ambiguity attitudes on pro-social behaviour in more abstract contexts. For example, Vives & FeldmanHall (2018) find that ambiguity attitudes robustly predict pro-social behaviour, though that risk attitudes do not. Similarly, Chen & Zhong (2021) find that uncertainty increases sharing behaviour in dictator games and reduces lying in dice games. In the context of other-regarding behaviour in binary dictator games, Haisley & Weber (2010) find that dictators are less likely to choose an unfair distribution when the outcome allocation is dependent upon an ambiguous lottery rather than a risky one. For further similar research see Cettolin, Riedly, Tran (2017) and Olschewski, Dietsch, & Ludvig (2019).^{1,2} Additionally, Mesa-Vazquez, Rodriguez-Lara, & Urbano (2021) find that when there is uncertainty about whether a dictator's choices are implemented, their actions are more generous.

Our research builds on the literature on risk that has so far mostly employed directly controllable levels of risk in the lab. For example, in abstract game scenarios, risk can be controlled and stated precisely, for example by imposing a 50% chance of one's donation not going through, or by introducing a 5% chance that one's donation is matched. Crucially though, our research is substantially different from the discussed literature primarily because we move the level of risk from directly calculable interventions in the lab (as outlined above) to the actual charities themselves. While this introduces several design challenges, we argue that this step leads to an increased level of ecological validity of any potential finding. However, note that there is already a large literature on charitable giving generally that has a similar or higher level of external validity (Bekkers & Wiepking, 2011; Karlan & List, 2007; Vesterlund & Sonevi, 2006). However, our paper's main contribution is the moving of our focus on risk and ambiguity to actual organisations and their interventions and away from aspects that can be

¹ A related but distinct literature looks at the effect of risk over donations themselves. Exley (2016) find that in situations of risky donations where donors trade off personal payoffs and donations, donors give less (cf. also Garcia, Massoni, & Villeval, 2020). However, there is also some evidence for the opposite effect, namely that some types of uncertainty actually increase pro-social behaviour (cf. Kappes et al., 2018).

² Further work investigates another type of uncertainty investigated in pro-social behaviour, namely uncertainty over whether any given recipient needs one's help. Engel & Goerg (2018) find that uncertainty of this type "does not deter generosity" (Engel & Goerg, 2018, 51). When excluding selfish dictators, they find that in conditions of uncertainty donations are higher than under certainty. Further, Niehaus (2014) suggests that because outcomes of charitable giving are usually not observed by the donors, decisions are primarily made on the perception of the outcome rather than the actual outcome. This would also be consistent with findings that show how many factors explain donor's reluctance to give to the most effective charities (Caviola, Schubert, & Nemirow, 2020).

controlled in the lab. Having risk and ambiguity at the level of actual charities is the level at which risk and ambiguity typically enter people's decision-making processes; rarely are we uncertain as to whether our donation will randomly increase when we donate (as in some experimental lab studies), but we are almost always acting under uncertainty about the charity's interventions that we consider donating to.

Additionally, there has also been a recently growing literature that adds individual differences in personality traits to standardly used predictors like risk preferences in a variety of economic contexts. For example, Knapp, Wuepper, & Finger (2021) analyse the interrelations of risk preferences and personality as well as their predictive power in the context of farmer behaviour, finding that personality measures are among the best predictors of some of their economic behaviour. See also Müller & Schwieren (2012) on the role of the Big Five and Hilbig & Zettler (2009) on the HEXACO model (cf. also Antoncic et al., 2015) on economic behaviour. The role of personality has also been more directly studied in the specific context of charitable giving. Some find that personality traits can play a central role in charitable decision-making (Kasser, 2002) while others (Sargeant & Woodliffe, 2007) see the effect as markedly smaller. Specifically, some have also found associations between empathetic character traits and donations as well as volunteering behaviour (Davis et al., 1999) while others have established a link between the Big Five personality trait of agreeableness and prosocial behaviour (Ashton et al., 1998). For further recent work on the relationship between personality traits and charitable giving, see O'Reilly et al. (2012) and Kang & Lakshmanan (2018). Our paper adds to this literature that aims to integrate personality traits into economics related studies and aims to advance our understanding of the role of individual difference measures to standard predictors of economic behaviour like risk and ambiguity attitudes concretely in the context of charitable giving.

In attempting the construction of an externally valid donation choice that allows us to capture the difference between sure-thing charities and probabilistic charities in actual charities that participants can donate money to, our main outcome variables of interest are (i) the choice between two real charities, one sure-thing charity and one probabilistic charity and (ii) the choice about one of those charities that has been randomly selected. Both of these charitable decision-making scenarios have strengths and weaknesses from an experimental design perspective, but we hope that they jointly allow us to better understand the role of individual differences in charitable decision-making scenarios like these.

In our experiment, each participant is first presented with a randomly selected pair of charities consisting of one charity of each type to control for accidental confounds relating the charity's context as each are presented with substantial additional accurate information to increase the naturalness of the choice. Participant choices with respect to this randomly selected charity pair then allows us to isolate and capture the element of probability between the two charity types. In the second part of this experiment, we study participant behaviour when they are shown only one randomly selected charity (either sure-thing or probabilistic), which more narrowly captures the predictive value of individual differences on donation choices to charities of specific types. Overall, we find little to no evidence that individual differences in risk/ambiguity attitudes, numeracy, optimism, and donor type predict charitable giving behaviour. However, we do find that a purely altruistic donor type predicts donations to probabilistic charities. As such, we take this paper to be primarily reporting a null-result.

Hypotheses

In this paper, we are centrally interested in what best explains donations choices regarding sure-thing and probabilistic charities as we offer an account of moving the level of risk and ambiguity from risk over donations or outcomes in abstract lab environments to risk and ambiguity over actual charities' interventions, thus increasing the level of ecological validity. To provide data on these questions, we pre-registered five concrete null hypotheses.

First, we investigate whether donor choices can be explained by individual differences in risk and ambiguity attitudes. Previous work in domains such as stock market participation (Bianchi & Tallon, 2018) and health-related field behaviours (Sutter et al., 2013) has found that attitudes to risk and ambiguity can play significant roles. In context of pro-social behaviour in game environments the results show that ambiguity aversion may play a role (while risk aversion sometimes does not) (Vives & FeldmanHall, 2018), though risk aversion has also been found to be "predictive for giving" (Cettolin, Riedl, & Tran, 2017, 95). As such, we argue that given risk and ambiguity aversion have been shown to impact behaviour in many contexts including charitable giving, this makes it an a priori interesting relation to test. This hypothesis is also theoretically grounded, in that it might be the case that one's preference not to give to charities that have a low chance of making an impact might be driven by an individual's general risk aversion profile, or it might be that given the ambiguous nature of charitable interventions that it

is only ambiguity aversion that impacts this choice. As such it is plausible that either (or both) of them may play a role in this choice. Our directional pre-registered prediction is that ambiguity and risk aversion predict donations towards sure-thing charities because those averse to risk and ambiguity may prefer charities that have clearly stated, and low-risk interventions as opposed to more ambiguous and risky ones. This is our first null hypothesis.

Null Hypothesis #1: Ambiguity and risk attitudes do not predict choices between sure-thing and probabilistic charities.

Second, we investigate whether there are a few potential further individual difference measures that may help explain behaviour in the context studied here. Specifically, we will focus on numeracy, optimism, donor type (warm glow vs. pure altruistic type), and empathy. First, it may be the case that basic individual differences in numeracy explain a potential preference for probabilistic charities, primarily because decision-making in situations of small probabilities (and big potential payoffs) might be particularly difficult to comprehend for those less numerically versed. For example, previous research has found that those lower in numeracy were more insensitive to proportions of donation recipients (Kleber, Dickert, Peters, & Florack, 2013) and that they showed higher susceptibility to changes in numeric presentation (Dickert, Kleber, Peters, & Slovic, 2011). It may as such be the case that one's level of numeracy also meaningfully impacts behaviour in the context studied here as the probabilistic charities include interventions that have a small chance of making a large impact. Understanding these proportions plausibly requires a certain level of numeracy. Further, one may also think that a general proclivity to optimism may bias individuals towards overestimating the success of probabilistic charitable interventions, or conversely that higher pessimism may explain a preference for sure-thing charities as those promise to have a reliable impact even in the worst-case scenario. This is corroborated by previous research that draws on the German socioeconomic panel and finds that optimism predicts charitable giving in some of their models (Boenigk & Mayr, 2016). Additionally, we investigate whether estimates of donor type, i.e. warm-glow vs. pure altruistic donor type, may predict behaviour too: For those on the warm-glow part of the spectrum, giving to a charity that has an immediate and reliable impact may have a higher chance of producing such a warm glow than giving to a charity that most likely will have no direct impact (or one that will not be observable

for quite a while). Conversely, pure altruists would presumably primarily care about the perceived (expected) impact on social welfare and may as such be more likely to choose probabilistic charities on expected value grounds. Lastly, empathy has previously been shown to predict charitable giving frequency in a variety of contexts (Verhaert & Van den Poel, 2011; Kim & Kou, 2014). It may thus be that empathy also explains choices in this study, for example if participants are preferring sure-thing charities due to their easier-to-relate-to interventions.

Our directional pre-registered prediction is that high numeracy, purely altruistic motives, and tendency towards optimism predict donations towards probabilistic charities. Conversely, our directional pre-registered prediction is that empathy and warm glow motives predicts donations towards sure-thing charities. These predictions are encapsulated in our second null hypothesis.

Null Hypothesis #2: Individual differences in numeracy, optimism, empathy, and donor type do not predict choices between sure-thing and probabilistic charities.

One worry with the present design that might be raised is that despite the randomisation of charity pairs, any potential effect may still be driven by some level of contextual confounding present in the charity descriptions. To address this concern, we also have a condition in which participants make a choice between two ‘context-free’ charities, in which only fundamental information as to the underlying probabilistic aspects are preserved and the remainder of context is removed, though participants are informed about all aspects of the charity after their decision. This is our third null hypothesis.

Null Hypothesis #3: The factors predicting behaviour in NH1/NH2 do not predict behaviour in the blinded choice condition.

Fourth, it may be the case that if much of the potential preference for sure-thing charities is explainable by individual difference measures related to numeracy or risk attitudes, this might open up the possibility of shifting behaviour via informational interventions aimed at this. Specifically, we investigate how exposure to expected value reasoning affects donation behaviour. It might be the case that learning about or being made aware of expected value reasoning may meaningfully impact behaviour. Our directional pre-registered prediction is that those shown the expected value reasoning treatment text will

be more likely to donate to probabilistic charities than those shown a made-up theory of decision-making (midontic decision-theory), which introduces participants to a made-up theory of decision-making with the same directional recommendations, aimed at controlling for experimenter demand.

Null Hypothesis #4: The expected-value treatment does not impact behaviours in the sure-thing vs. probabilistic charity choice more than the control condition.

Lastly, one may also be interested in donation behaviour not between these two types of charities, but rather just in the context where potential donors are presented with one such charity. This may reduce the chance of additional confounds (like worrying that the design that presents two charities is artificial in its dichotomous presentation; after all, most naturalistic decisions are not decisions between two distinct choices). Because the central goal of this paper was to provide data on donor behaviour that has reasonably high external validity to better understand donor behaviour regarding their choices with respect to sure-thing and probabilistic charities, we also investigate this in the context where they are presented only with a single, randomly selected charity (equiprobable that it is a sure-thing or a probabilistic charity). This allows us to control for the type of charity and any potential confounds that this dichotomisation and variance in background knowledge might bring with it. As above, we study both the frequency and amount of giving with the same predictors as outlined in the previous sections. Our fifth null hypothesis below states this in detail and basically pools null hypotheses #1 and #2 as one single hypothesis.

Null Hypothesis #5: Ambiguity and risk attitudes (as well as individual differences in numeracy, optimism, empathy, and donor type) do not predict frequency and size of donations when presented with either a sure-thing or a probabilistic charity.

Methods

Participants and Procedure

We pre-registered this project on the Open Science Framework.³ For this study, we recruited a total of 1506 participants via Prolific (47.8% male) at a mean age of $M=45.09$, $SD=15.49$ from the United Kingdom. This sample is representative of the UK's population according to census data from the UK Office of National Statistics via Prolific's representative samples tool. This study received ethics approval from the University of St Andrews.⁴ Overall, participants receive £2 for participation and can earn up to £1.83 in addition (over three sets of tasks), depending on their choices in the experiment. We arrived at sample size of 1506 by conducting a pre-registered a priori power analysis in which we calculate that in order to obtain .95 power to detect an effect of the size $f^2=.02$ (a standard measure of a small effect size which we picked as our smallest effect size of interest) at 5% alpha error probability in a linear multiple regression model, we would need to recruit a total of 652 participants. In order to account for those not wanting to donate (estimated at 33% as previous work has documented this rate to be between roughly 20% (Butera & Horn, 2017) and 40% (Metzger & Günther, 2019)) and those failing comprehension questions, we aimed to recruit 1050 participants for the main condition and a further 450 for the remaining secondary conditions (for full specifications of all additional power analyses for further conditions and analyses, please see our pre-registration). In all results reported, we excluded a total of 261 participants who got more than one of the comprehension questions or attention checks incorrect.⁵ As such, the final sample size used for all analyses is 1,245 participants.

In the first part of the experiment, participants provide fully informed consent, and then complete several individual difference measures relating to numeracy skills

³ This project was pre-registered at the Open Science Framework: https://osf.io/dsfyk/?view_only=c1e7aa3b52c5429da810f126a05f4446.

⁴ This project has received ethics approval from the University of St Andrews (approval code SA15429).

⁵ Our final exclusion criteria were: We excluded all participants who got more than comprehension/attention check incorrect. For all analyses in Final Choice, we also, in addition, excluded all those who indicated a donation that exceeded their earned endowment. This choice deviates from our pre-registered data exclusion plan. First, we did not pre-register that we would exclude data in Final Choice from participants who were stating a final donation higher than their earned endowment. However we deem this necessary to provide high quality data in Final Choice. Second, we stated that we would exclude anyone who had failed at least one check. We did not anticipate that the combination of the comprehension check regarding final donation and earned endowment, the comprehension check for the risk attitude elicitation task, and the comprehension check for the ambiguity attitude elicitation task as well attention check would result in such high attrition rates. Specifically, excluding anybody who had gotten at least one wrong would result in an exclusion of 42.7% of participants, whereas only excluding those who had gotten more than one wrong would result in an exclusion of 17.4%. While we report all analyses with this changed exclusion criteria, we also reproduce the main regression model in Appendix B with the initially pre-registered exclusion criteria and find no differences in results.

(Likpus et al. 2001), tendency for optimism (Scheier, Carver, & Bridges 1994), and empathy (Carre et al. 2013). For the completion of each these surveys, they are paid a fixed endowment of £0.25 per survey (£0.75 in total) which is at their disposal for the remainder of the study. All participants then move onto the next part, where they are randomly selected into either the Main Choice condition (70% chance), the Expected-Value Treatment condition (20%), or the Context-Free Choice condition (10%). There, they decide to allocate some or all of their endowment to one of the two randomly chosen charities (one sure-thing and one probabilistic charity), or alternatively allocate none of it to either charity. After this choice, participants are asked to state their views on the impact of the charities that they had just seen on a 100-point scale. Further, they are also asked to estimate the average views of other participants on these charities. This task was incentivised by awarding participants an additional £0.10 if their estimates are within 5 percentage points of the actual average of judgements.⁶ Following this task, they enter the second main section of the experiment, in which they complete two sets of choices that measure their risk and ambiguity attitudes, specifically a bomb risk elicitation task (Crosetto & Filippin, 2013; Nielsen, 2019) and an Ellsberg urn task. Across these two tasks, they can earn up to £0.98 in a second endowment, depending on their decisions and luck. All participants are then presented with one additional donation choice at Final Choice where they can decide to allocate all, some, or none of their earned endowment to one randomly selected charity (either a sure-thing or a probabilistic charity). For a complete overview of the experimental outline, see Figure 1. For an in-detail description of all steps and measures, see the section below.

⁶ This additional payment was paid out after the experiment ended and was not part of any donation decision in the experiment, because determination of estimation accuracy could not be done within the experiment itself.

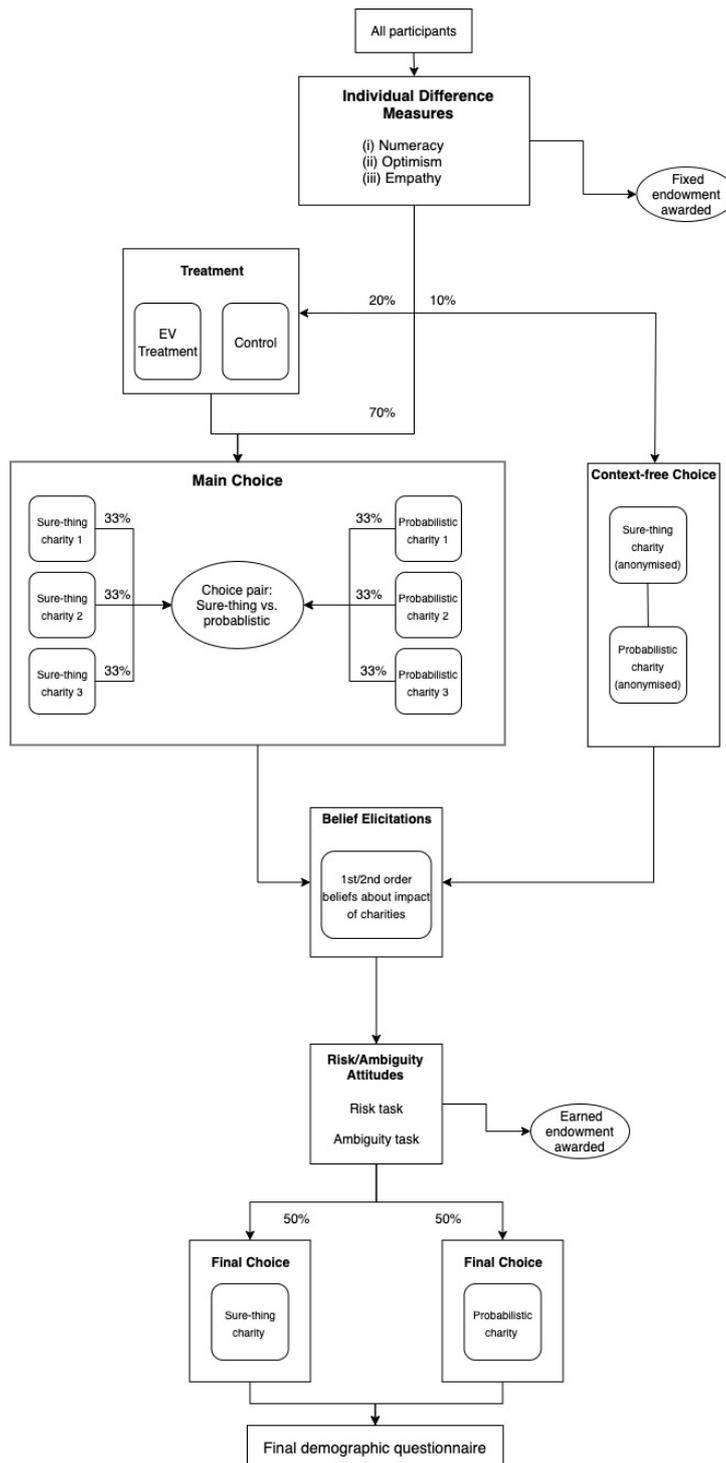


Figure 1. Experimental Outline

Measures

Numeracy, Empathy, Optimism, and Donor Type Measures

In order to acquire measures of empathy, numeracy, and optimism, we use three validated, standardly used scales. Our numeracy measure is the 11-point numeracy scale by Likpus et al. (2001) which includes three items on general numeracy and 7 items on risk

numeracy specifically. We collect tendency towards optimism with the Life Orientation Test (Scheier, Carver, & Bridges 1994) which is a standard measure for optimism and pessimism. Further, we use the Basic Empathy Scale in Adults (BES-A) as our measure of empathy (Carre et al. 2013). Participants are incentivised to complete these surveys by being rewarded with £0.25 for the full completion of each of the three surveys. We also collect self-reported donor type classification. To determine donor type (warm glow vs. pure altruist) we draw on Carpenter's (2018) methodology that relies on a self-reported recollection of previous charitable behaviour, classifying donors as either warm glow, pure altruist, or other motivations.

Risk and Ambiguity Attitudes

We estimate risk and ambiguity attitudes via two distinct, fully incentivised tasks. In order to ensure participant comprehension, participants complete a trial run of the risk attitude measure, are shown the mechanism for the ambiguity mechanism in detail via an example, and answer comprehension questions on both measures. Specifically, we measure risk attitudes via the 'Bomb risk elicitation task' (Crosetto & Filippin, 2013; Nielsen, 2019), adopting the mechanism for Qualtrics directly from Nielsen (2019). In this task, participants are presented with a 5x5 matrix of boxes. They are informed that one box contains a bomb, and that all other boxes yield an additional income of 2 pence. Participants can select any numbers of boxes, but if the box containing a bomb is selected and they confirm that they have selected as many boxes as they wish, their income is zero. As Crosetto & Filippin (2013) point out, this measure of risk attitudes has several upsides over other alternatives in that it "requires minimal numeracy skills [and] avoids truncation of the data" (Crosetto & Filippin, 2013, 31). Specifically, it "measures individual-level risk attitudes by a single parameter $k \in \{0, 1, \dots, n\}$, the number of boxes collected" (Holzmeister & Stefan, 2021). Risk aversion is indicated by selecting fewer than 13 boxes and risk-seeking by selecting more than 14.⁷

We measure ambiguity attitudes by recourse to standard Ellsberg urns: Participants are presented with two urns. They can bet on both the outcome of a risky urn

⁷ In effect, participants make choices between 26 lotteries ranging from choosing 0 boxes with a certain outcome of £0 to choosing all 25 boxes with a certain outcome of £0, while any number of boxes in between has a different chance to lead to additional income. The expected value of these lotteries is bow-shaped: playing lotteries with 13 and 14 boxes has the highest expected value, with choosing more or fewer boxes reducing expected value. As such, risk-neutral participants would choose 13 or 14 boxes, risk averse participants fewer than 13, and risk-seeking participants more than 14.

containing 5 red and 5 black balls and on the outcome of an ambiguous urn containing an unknown quantity of both red and black balls. We then collect reservation prices for each urn by using a Becker-DeGroot-Marschak mechanism with random integers being uniformly drawn from the interval $\{1,2,\dots,25\}$. This design is adapted from Halevy (2007). The ambiguity premium is calculated by taking the difference between the ambiguous urn's reservation price and the risky urn's reservation price (Halevy 2007, 522-523) with negative outcomes indicating ambiguity aversion (and positive outcomes indicating ambiguity proneness).

Demographic Factors

We also collect a number of additional demographic factors that are chosen as additional control variables. These are partially drawn from Bekkers & Wiepking (2011) and Wiepking & Bekkers (2012) who survey the literature and present a number of factors that predict charitable giving in a variety of contexts. Specifically, we collect data on age, gender, education, employment, religious affiliation and practice, marital status, children, and subjective financial well-being.

Main Choice and Final Choice Outcomes

We collect two main outcome variables for this study. These rely on participant choices when presented with the option to donate all, some, or none of their earned endowment to charity. Specifically, our two main outcomes are Main Choice (where participants make choices between a sure-thing and a probabilistic charity) and Final Choice (where participants make choices about a randomly selected charity). In both cases, we collect data on whether they donate at all, if they do, which charity they choose in Main Choice (participants can choose up to one charity), and how much they donate.

In Main Choice, each participant is shown exactly two charities, one sure-thing charity and one probabilistic charity. Each of these charities is drawn randomly from a pool of three charities (for a total of six individual charities and nine total charity pairs). This level of randomisation was chosen to avoid confounding this choice by the specifics of any given charity or any possible comparison effects between two specific charities. By randomising the pairs, we aim to isolate the factor of probability between the choices (sure-thing vs. probabilistic interventions) without having any potential result be influenced by the specific characteristics of any of these charities all the while keeping

this choice as naturalistic as possible. Their choice here relates to null hypotheses #1 and #2.

In Final Choice, participants can again make a choice about their new endowment that they earned from the risk and ambiguity attitude tasks (ranging from £0 to £0.98, depending on their choices and luck). They are presented with one randomly selected charity (either a sure-thing charity or a probabilistic charity with equal likelihood) from the same pool as before and can again decide how much of their endowment to donate on this. We also ask participants before the uncertainty over their endowment in this section is resolved to state how much they would donate to this charity for different levels of endowments. This condition controls for a number of potential confounds in Main Choice which allows for it to answer the paper's central question more directly, though it being the last task of the experiment, we cannot rule out potential order effects. This choice relates to null hypothesis #5.

All six charities that are used in both Main Choice and Final Choice are recommended by GivingWhatWeCan in their 2021 charity recommendation titled "What are the best charities to donate to in 2021?" (2021) to ensure that all of these charities are relatively similar in quality and reputation, and participants are informed about this fact prior to their choice. The three sure-thing charities are the SCI Foundation (deworming), GiveDirectly (direct cash transfers), and the Against Malaria Foundation (insecticide-treated bed nets). The three probabilistic charities are the Center for Health Security (epidemic/global health disaster prevention), Nuclear Threat Initiative (nuclear war prevention), and the Machine Intelligence Research Institute (artificial intelligence risk).

Our theoretical justification for this grouping is that all three sure-thing charities describe interventions that have a substantially high likelihood of success. All three interventions (deworming drugs, cash transfers, and bed nets) are administered by highly effective charities and have interventions that have been shown to have a high probability of being deployed successfully (i.e. that the cash transfer arrives or the drug is administered). On the other side, all three probabilistic interventions (epidemic prevention, nuclear war prevention, artificial intelligence risk mitigation) are made under substantial ambiguity as to the probability distributions both of the intervention itself and the corresponding event occurring. These interventions are highly probabilistic in a way that cannot be resolved ex ante, making them qualitatively different than the sure-thing charities on a theoretical level, which is also recognised by the charities themselves and is as such part of their mission (after all, some of those risks are hard to quantify and even

harder to address, but if such an intervention would be successful, it would bring with it substantial outcomes). For full treatment texts, see Appendix F.

However, we also provide an empirical justification for this grouping that supports the theoretical reasoning outlined above. We empirically tested our pre-selection into sure-thing and probabilistic charities by running a small auxiliary study in which we had a new set of participants ($n=101$) rate these six charities on a number of dimensions including riskiness and ambiguity. We find that, in aggregate, participants rate sure-thing charities' interventions as significantly more likely to succeed (less risky) and as significantly more quantifiable (less ambiguous) than they do probabilistic ones. Both effects are significant at $p<.001$ and the effect size in Cohen's d for the difference in riskiness is $d=.98$ and $d=1.01$ for ambiguity. For a full specification of this study and its results, see Appendix A.⁸

Belief Elicitations

We elicit 1st and 2nd order beliefs regarding the charity's impact to investigate if participants who donate to one set of charities simply think that they are more impactful overall. Specifically, we ask participants to state their views on how impactful both charities they were presented with were on a percentage scale from 0 (Not impactful at all) to 100 (Extremely impactful). We also ask participants to estimate the average judgements regarding the impact of each charity as ascertained by the responses of all other participants. This is incentivised by paying those participants whose estimation is within 5 percentage points of the actual average an additional £0.10.

Additional Treatments

We also had two additional treatments, an Expected Value Informational Treatment and a Context-Free Treatment. These were chosen to provide data that can speak to null hypothesis #4 and #3 respectively.

For our Expected Value Informational Treatment, we present participants with an introductory text of expected value reasoning. This is to test whether this simple information treatment can meaningfully shift donor behaviour. In order to control for experimenter demand effects, half of the participants in this condition are presented with a control text that introduces them to a made-up nonsense mathematical decision theory.

⁸ This study was conducted ex-post at the request of an anonymous reviewer.

The control paragraphs introduce participants to this theory of decision-making via mathematically formulated but nonsensical arguments that are illustrated with the same example as the treatment text. It recommends the same type of action, namely that taking risks can be worth it if the potential outcomes are good enough. This allows us to identify whether a given effect is due to experimenter demand or not by comparing a shift in frequency of donation to sure-thing/probabilistic charities between the expected-value treatment with the Main Choice and the made-up theory treatment (midontic decision-theory) and the Main Choice. See Appendix F for full texts. The choices here relate to null hypothesis #4.

In the Context-free condition, participants are given the option to choose between two charities that had all context (area of intervention, geographic focus, name, etc) removed. The only characteristics left are of the probabilistic nature of their interventions as well as some generic filler text. The participants are told that the full information would be made available to them right after they made their choice. See Appendix F for full texts. This condition is used to provide data about to null hypothesis #3.

However, because the number of people who made donations was unexpectedly small, both of these conditions did not have the power that we calculated prior to running this study to detect a meaningful effect. This means that results of these conditions are inconclusive. We still report the full pre-registered analyses in the appendix, see Appendix D, but do not discuss them in the main results and discussion sections.

Results

In Table 1, we summarise the general demographic variables and their frequencies with our sample. Generally, the sample shows high numeracy skills, with only 5.5% of participants not scoring at least 8 out of 10, and we observe a mean empathy score of $M=76.54$, $SD=9.75$ and mean optimism scores of $M=33.08$, $SD=7.20$. Further, according to Carpenter’s (2018) classification of donor types, we can classify 24.3% of participants as warm-glow donors, 11.9% as pure altruist donors, and the remaining 63.9% as being motivated by some other reason.

TABLE 1—DEMOGRAPHICS

	n	%		n	%
<i>Age</i>			<i>Religious Participation</i>		
18-28	247	.198	Yes	112	.090
29-38	216	.174	No	1133	.910

39-48	228	.183			
49-58	228	.183	<i>Marriage Status</i>		
59 and above	326	.262	Married	586	.471
			Not married	659	.529
<i>Gender</i>			<i>Children</i>		
Male	617	.496	Has children	692	.556
Female	618	.496	Does not have children	553	.444
Other	10	.008			
<i>Education</i>			<i>Financial Wellbeing</i>		
High school	498	.400	Finding it very difficult	58	.047
Undergraduate	446	.358	Finding it quite difficult	149	.120
Graduate/Professional	301	.242	Just about getting by	408	.328
			Doing alright	479	.385
			Living comfortably	151	.121
<i>Religious Affiliation</i>			<i>Employment</i>		
No affiliation	893	.717	Unemployed	174	.140
Protestantism	162	.130	Out of the workforce	223	.179
Catholicism	120	.096	Part-time employment	279	.224
Islam	34	.027	Full-time employment	569	.457
Hinduism	15	.012			
Judaism	10	.008			
Buddhism	8	.006			
Sikhism	3	.002			

Notes: Demographics for the full sample after exclusion of 261 participants who failed more than one attention or comprehension check.

Our results for risk and ambiguity attitudes show a strong tendency towards risk aversion and a moderately strong tendency towards ambiguity aversion. Following (Holzmeister & Stefan, 2021), we measure individual-level risk attitudes by the number of boxes collected, k . Based on the cumulative distribution of participant choices, we find that 83.7% of participants exhibit risk aversion ($k < 13$), 8.9% of participants exhibit risk neutrality ($13 \ll k \ll 14$), and 7.4% exhibit risk-seeking behaviour ($k > 14$). Our ambiguity attitude measure is calculated by subtracting the ambiguous urn's reservation price from the risky urn's reservation price with negative outcomes indicating ambiguity aversion (and positive outcomes indicating ambiguity proneness). We find that 34.4% of participants exhibit ambiguity aversion, 46.6% of participants exhibit ambiguity neutrality, and the remaining 19.0% participants exhibit ambiguity-seeking attitudes. Additionally, we find no significant correlation between these two measures in the full sample with Pearson's $r = -.05$, $p = .079$.

Concerning the charitable giving outcomes in Main Choice, we find that across all conditions, 35.8% of participants make a donation. Of those making a donation, 84.7% donate to a sure-thing charity, and 15.3% donate to a probabilistic charity. The average donation of those who donate is 50.60 pence ($SD = 23.98$). For those donating to a sure-thing charity, the average donation is 51.50 pence ($SD = 23.83$), while for those donating to a probabilistic charity, it is 45.60 pence ($SD = 24.43$). This difference is not statistically significant, $t(305) = 1.557$, $p = .121$. We further find a strong relationship between choosing

a type of charity and one's belief in its impact as well as its estimation of the general consensus of its impact by those making a donation. Specifically, we find that choosing to donate to a probabilistic charity stands in a strong positive point-biserial correlation to judging the probabilistic charity as impactful, $r_{pb}=.424$, $p<.001$, as well as estimating its average impact judgement, $r_{pb}=.346$, $p<.001$. Conversely, the same choice also stands in a strong negative correlation with judging the sure-thing charity as impactful, $r_{pb}=-.314$, $p<.001$, which is again mirrored in its estimated average impact judgement, $r_{pb}=-.267$, $p<.001$. Looking at donation behaviour in Final Choice, we find that 29.1% of participants made a donation. 36.3% of those being presented with a sure-thing charity donated, compared to only 22.8% of those being shown a probabilistic charity. This difference is statistically significant, $\chi^2(1, N=1177)=25.66$, $p<.001$. Of those making a donation, those being presented with a sure-thing charity donated an average of 26.58 pence (SD=17.92) while those who were shown a probabilistic charity donated an average of 20.55 pence (SD=16.55), a difference that is again statistically significant, $t(340)=3.45$, $p<.001$.

First, we investigate general donation behaviour relating to Main Choice. The results presented in Table 2 speak to the central null hypotheses #1, #2, and #3. Model (1) reports the results for the Main Condition. The outcome variable is the type of charity conditional on a donation being made, with 0 being coded as the sure-thing charity and 1 as the probabilistic charity. The gender variable is coded 1 for female, all other categorical variables are coded 1 for the affirmative. As outlined above, the risk attitude measure is a discrete variable of the number of boxes opened, the ambiguity aversion is the result of the subtraction of the reservation prices. All other scales are the sum of the (re-reversed) individual items. As specified in our pre-registration, we report main regression results for binary outcomes using an OLS model and also for the corresponding logit model as a robustness check in Appendix B (Appendix Table 2) to check for sensitivity to functional form choice, where we find no impact of model choice.⁹ We also report a robustness check in Appendix B (Appendix Table 3) where we report a regression with random effects at the stimulus level, finding that our null result is also robust to this model choice.

⁹ We made this choice primarily to enable clear and straightforward a priori power calculations, present more easily interpretable results, and because results are generally insensitive to model choice in most situations because the main reason against the use of linear models is that predicted probabilities may be greater than 1 or smaller than 0 which mainly occurs when the probabilities of the outcomes are extreme, not when they are in a roughly 80/20 distribution as is present in this paper (Cleary & Angel, 1984; Dey, 1993; Angrist & Pischke, 2008; Hellevik, 2009; Gomila, 2020).

TABLE 2—REGRESSION RESULTS FOR CHARITABLE GIVING BEHAVIOUR IN MAIN CHOICE
PREDICTING CHOICE BETWEEN SURE-THING AND PROBABILISTIC CHARITIES

	(1)
Risk Attitude	.001 (.005)
Ambiguity Aversion	-.002 (.005)
Numeracy	.003 (.024)
Empathy	-.001 (.002)
Optimism	-.005 (.003)
Donor Type	
Warm-Glow	-.020 (.051)
Pure Altruism	-.056 (.064)
Age	.003 (.002)
Gender	.014 (.050)
Education	
Undergraduate degree	.039 (.051)
Postgraduate/Professional degree	-.019 (.056)
Religion	
Protestantism	-.077 (.067)
Catholicism	.011 (.074)
Islam	-.114 (.122)
Judaism	-.131 (.379)
Buddhism	-.133 (.287)
Hinduism	.345 (.276)
Religious Participation	.041 (.089)
Marriage Status	.043 (.050)
Children	-.004 (.054)
Financial Wellbeing	.005 (.024)
Employment	
Out of the workforce	-.159 (.097)
Part-time employment	.001 (.085)
Full-time employment	-.060 (.079)
R ²	.055
Sample size	307

Notes: OLS regression reporting unstandardised coefficients and standard errors. Outcome variable is charity choice (0 = sure-thing charity, 1 = probabilistic charity). *p<.1, **p<.05, ***p<.01, ****p<.001

We find that none of the main variables nor the demographic control variables predict charity choice in the Main Condition. We can straightforwardly conclude from Model (1) that we do not have evidence to reject the null hypotheses #1 and #2 as none of the independent variables meaningfully predict donor behaviour.

Second, we investigate general donation behaviour in Final Choice where participants were either presented with a sure-thing or a probabilistic charity. Here, we report two further regression models relevant to null hypothesis #5. These are not the pre-registered ones but instead include interaction terms that we did not pre-register. For specifications of the regression models as pre-registered, see Appendix C (Appendix Table 6). Model (2) explains frequency of donation, and Model (3) predicts size of donation. For Model (2), not making a donation is coded as 0 and making a donation is coded as 1. The central variables are interaction terms, where we interact the main individual difference measures with the condition. Specifically, they are coded with 0 = sure-thing charity and 1 = probabilistic charity for all our main explanatory variables. In these analyses, participants from all conditions are included¹⁰ and are split only by which type of charity they were presented with at Final Choice; recall that each participant here was only presented with one randomly selected charity.

TABLE 3—REGRESSION RESULTS FOR CHARITABLE GIVING BEHAVIOUR IN FINAL CHOICE
PREDICTING FREQUENCY OF GIVING AND SIZE OF DONATION

	(2)	(3)
Risk Attitude	.002 (.005)	.124 (.120)
Ambiguity Aversion	-.002 (.004)	-.127 (.109)
Numeracy	.015 (.017)	-.193 (.453)
Empathy	.007**** (.002)	.212**** (.048)
Optimism	.003 (.003)	.062 (.071)
Donor Type		
Warm-Glow	.061* (.047)	2.170* (1.237)
Pure Altruism	.045 (.058)	-.144 (1.525)
Condition X Risk Attitude	.003 (.006)	.019 (.161)
Condition X Ambiguity Aversion	.002 (.005)	.085 (.143)
Condition X Numeracy	.006 (.018)	.275 (.482)
Condition X Empathy	-.002* (.002)	-.086 (.056)
Condition X Optimism	-.004 (.003)	-.033 (.088)
Donor Type		
Condition X Warm-Glow	-.034 (.063)	-.993 (1.649)
Condition X Pure Altruism	.177** (.083)	2.955 (2.176)
Age	.001 (.001)	.047 (.030)

¹⁰ For all analyses of Final Choice, we both apply the general exclusion criteria (excluding anyone who has answered more than one comprehension/attention task incorrectly) as well as the additional exclusion criterion where we exclude participants who indicated a donation that was higher than their earned endowment for Final Choice.

Gender	-.050* (.030)	-.747 (.788)
Education		
Undergraduate degree	-.050 (.031)	-1.129 (.800)
Postgraduate/Professional degree	.030 (.034)	-.095 (.897)
Religion		
Protestantism	.071* (.043)	.153 (1.129)
Catholicism	.045 (.047)	1.468 (1.239)
Islam	.173* (.142)	6.606*** (2.317)
Judaism	.134 (.142)	2.636 (3.730)
Buddhism	.159 (.160)	4.994 (4.207)
Hinduism	.149 (.126)	7.049** (3.301)
Sikhism	.046 (.260)	-2.343 (6.819)
Religious Participation	.015 (.054)	-.265 (1.406)
Marriage Status	-.025 (.031)	-.469 (.819)
Children	.032 (.033)	.697 (.877)
Financial Wellbeing	.023 (.015)	.398 (.384)
Employment		
Out of the workforce	.017 (.055)	-.820 (1.442)
Part-time employment	-.019 (.047)	-2.211* (1.243)
Full-time employment	-.029 (.042)	-1.436 (1.107)
R ²	.076	.085
Sample size	1177	1177

Notes: OLS regressions reporting unstandardised coefficients and standard errors. Model (2) predicts frequency of donation and Model (3) predicts size of donation. Interaction terms interact the condition (0 = sure-thing charity, 1 = probabilistic charity) with the explanatory variables *p<.1, **p<.05, ***p<.01, ****p<.001

Here we find that while empathy predicts donation frequency and size overall, it is only the interaction term with pure altruism that statistically significantly predicts frequency of donation to probabilistic charities. This is in line with theoretical predictions that hold that pure altruists would be more likely to give to probabilistic charities as they primarily care about the impact on social welfare. Further, in our robustness check that reproduces Model (2) as a logit model (Appendix C, Appendix Table 5), we find the same result. All other pre-registered variables (interacting with condition) do not predict donor behaviour.

Because we have found mostly null results in both Main Choice and Final Choice, we conduct the following exploratory analyses to further investigate the failure to reject the null hypotheses. Regarding the results from Main Choice, we conduct a series of equivalence tests for the multiple linear regression coefficients from Model (1) in Table 2 following Alter & Counsell (2021). These allow us to make direct inferences about the absence of an effect or the presence of a negligibly small effect (i.e. an estimate of a precise zero) by providing two one-sided t-tests against an upper and a lower equivalence bound. In order for there to be sufficient evidence in favour of a negligible effect (i.e.,

evidence in favour of the absence of an effect or the presence of a negligibly small effect), both hypotheses have to be rejected at the .05 level. Because we did not pre-register the equivalence bounds, we report three plausible levels of the bounds in unstandardised coefficients at .01, .025, and .05 to show sensitivity of results to this choice. We report this for all our primary variables of interest from Main Choice that did not show significant effects.

TABLE 5—TOST FOR MAIN CHOICE COEFFICIENTS

	-01	.01	-.025	.025	-.05	.05
Risk Attitudes	2.2**	1.8**	5.2****	4.8****	10.2****	9.8****
Ambiguity Aversion	1.4*	2.6***	4.4****	5.6****	9.4****	10.6****
Numeracy	.54	.29	1.17	.92	2.21**	1.96**
Empathy	4.5****	5.5****	12****	13****	24.5****	25.5****
Optimism	1.67**	5****	6.67****	10****	15****	18.33***
Warm Glow	-.20	.59	.10	.88	.59	1.37*
Pure Altruism	-.72	1.03	-.48	1.27	-.09	1.66**

Notes: All t-test results for TOST procedures on a variety of lower and upper equivalence bounds (in unstandardized coefficients) from Model (1). *p<.1, **p<.05, ***p<.01, ****p<.001

These results suggest that we can be confident that the majority of our variables do not predict donor behaviour between sure-thing and probabilistic charities. Risk attitudes, empathy, and optimism show evidence in favour of a negligibly small effect at very tight equivalence bounds of (-).01, while ambiguity aversion shows this effect at (-).025, both relatively small levels. For numeracy, we can only conclude negligible effects at the (-).05 level. We remain unable to state clearly whether donor type has an effect on this choice as its results are insignificant both in the pre-registered regression analyses as well as the exploratory tests of equivalence. As such, we do not reach a conclusion as to the effect of donor type in Main Choice, but we can confidently state that the other explanatory variables of individual differences do not predict choices between sure-thing and probabilistic charities.

For the Final Choice analyses, we also conduct equivalence tests as above, see Table 6, with a focus on the interaction term coefficients from Model (2). We find that at the relatively tight (-).025 level, risk attitudes, ambiguity aversion, empathy, and optimism show a negligibly small effect in the frequency of donation model. Numeracy again shows a negligible effect at the (-).05 level and results regarding donor type (specifically warm glow) again do not allow for the conclusion of a negligible effect. We also report equivalence tests for the donation size in standardised coefficients in Appendix

E (Appendix Table 10), finding a similar pattern of results that shows numeracy and donor type as non-negligible effects while the remainder of variables show negligibly small effects.

TABLE 6—TOST FOR FINAL CHOICE COEFFICIENTS

	-.01	.01	-.025	.025	-.05	.05
Condition X Risk Attitudes	2.17**	1.17	4.67****	43.67****	8.83****	7.83****
Condition X Ambiguity Aversion	2.4***	1.6*	5.4****	4.6****	10.4****	9.6****
Condition X Numeracy	.89	.22	1.72**	1.06*	3.11****	2.44***
Condition X Empathy	4****	6****	11.5****	13.5****	24****	26****
Condition X Optimism	2**	4.67**	7****	9.67****	15.33***	18****
Condition X Warm Glow	-.38	.70	-.14	.94	.25	1.33*

Notes: All t-test results for TOST procedures on a variety of lower and upper equivalence bounds (in unstandardized coefficients) from Model (2). *p<.1, **p<.05, ***p<.01, ****p<.001

These results provide us with a broad evidence base in favour of there not being an effect across many of our central explanatory variables. Also note that the results in Table 5 and Table 6 are relatively robust to adjusting for multiple hypothesis testing. For example, if we adjusted our significance level to .007 following the Bonferroni method, we would still find that all significant effects remain significant at least at the (-).05 level.

Following Rouder & Morey (2012), we compute exploratory Bayes factors to evaluate the evidence for the null. We report Bayes factors for linear models based on Liang et al. (2008), using a Jeffrey-Zeelner-Siow mixture of g-priors. Applying the default r scale of .353 for the results regarding Main Choice in Model (1), we calculate a JZS Bayes Factor of 2.49×10^7 , suggesting very strong evidence in favour of the null hypothesis. The same analysis regarding Final Choice in Model (2) also provides evidence in favour of the null, though at a significantly smaller magnitude, at a JZS Bayes Factor of 22.15. This corroborates the results of the equivalence tests and suggests that our data provide strong evidence in favour of a robust null result.

Discussion

Overall, we find little or no evidence in favour of rejecting our null hypotheses, and as such this paper primarily reports a negative result: First, we do not find evidence that risk and ambiguity attitudes robustly predict charitable decision-making between sure-thing and probabilistic charities (null hypothesis #1), and we also do not find evidence that individual differences in numeracy, optimism, donor type, and empathy predict behaviour

in this choice as well (null hypothesis #2). Further, we only find weak evidence predicting behaviour when participants are presented with either a sure-thing or a probabilistic charity, where we find that pure altruistic donor types significantly predict donation frequency to probabilistic charities (null hypothesis #5). Further, due to a lower than planned sample size, we are unable to conclusively evaluate the data about our no-context condition (null hypothesis #3) and with regard to the effect of the informational treatment introduced participants to expected-value reasoning (null hypothesis #4).

The results thus indicate that our hypothesised variables for explaining behaviour in the context of donating to sure-thing or probabilistic charities fail to meaningfully predict actual donor behaviour. However, these results do not by themselves allow for a concrete specification of why we fail to detect an effect. It may simply be that a different factor is at play when donors make these types of decisions and that individual differences in risk/ambiguity aversion do not impact decision-making at this level or that standard laboratory measures of risk/ambiguity attitudes does not directly translate into charitable decision-making in the context studied. It may also be that while this design was meant to provide a naturally occurring decision as is possible within the constraints of an online experiment, that a contraposition of two distinct charity types impacts decision-making such that participant choices are biased in some way. For example, this could be because of distinction bias, i.e. the bias describing how our preferences and choices may be substantially distinct in conditions where we evaluate options in separation or jointly (Hsee, Loewenstein, Blount, & Bazerman, 1999; Hsee & Zhang, 2004; Anvari, Olsen, Hung, & Feldman, 2021). We have some evidence in favour of this worry as at least some of the pre-registered factors predict directionally as expected in Final Choice, where only one charity was presented to participants. As such, it may be justified to put more interpretative emphasis on the results from Final Choice compared to Main Choice.

Further, it may be that participants did not pick up on the distinction between the two charities and that our outcome variable was faulty in some way. If this was the case, this would threaten the validity of all results presented here. There are several reasons to believe that this is not the case. First, the qualitative comments at the end of this study often explicitly raise concerns of probability, risk, and chance with regard to the participants' charity choice, suggesting that they did pick up on the central difference. Second, we find that participants who favour probabilistic charities also judge them as more impactful and also estimate that the average judgements of other participants as higher than those who favour sure-thing charities. Lastly, we conducted an auxiliary study

in which a new set of participants rated the charities among other things on the dimension of risk. There, we find that sure-thing charities are rated as significantly less risky and less ambiguous than probabilistic charities ($p < .001$). This gives us some prima facie evidence in favour of the validity of our outcome variable, though they are not (and indeed cannot be) conclusive.

It is also worth pointing out that while most of the previous literature discussed above uses manipulations of risk and ambiguity that are more tightly controlled (i.e. by introducing risk over donations controlled by the experimenter), this study attempted to provide a more naturalistic and thus externally valid picture of actual risky and ambiguous charitable giving, and failing to provide evidence in favour of this may also be partly explained by the general difficulty that highly abstract concepts in perfectly controlled laboratory games have when they are moved towards more externally valid contexts (cf. e.g. Galizzi & Navarro-Martinez, 2019). As such, failing to provide further evidence in favour of the impact of individual differences in, say, risk attitudes, on charitable decision making may be best understood as being, at least in part, explainable by a failure of experimental measures to generalise above and beyond their original context (as the outcome variables in this study were actual charities and not, like similar literature before, allocation decisions in dictator games where risk can be varied between conditions quite precisely). This in itself is not a bad thing, and we take our finding to contribute to a cumulative science that aims to establish which measures generalise in what contexts and situations and where such generalisations fail (as is the case in this paper).

Importantly though, in conducting our exploratory analyses of equivalence tests and Bayesian analyses both in the context of Main Choice and Final Choice, we find that there is strong evidence in favour of negligibly small effects (or in favour of null effects) across a variety of plausibly set equivalence bounds and variable. Specifically, we find that for risk attitudes, empathy, and optimism, irrespective of the equivalence bound tested, their effects are negligibly small even at a delta level of an unstandardised coefficient at .01. For ambiguity aversion, the effect is negligible at .025, and for numeracy at .05. We do not have evidence in favour of a negligible effect for donor types. These patterns hold for both Main Choice and Final Choice with the caveat that in Final Choice, these tests were conducted with the interaction terms. We also conduct Bayesian analyses and find corroborating patterns of results. This provides evidence in favour of a null effect, something that the standard null-hypothesis testing technically cannot provide. In other words, these results suggest that risk attitudes, ambiguity attitudes, and individual

differences in numeracy, empathy, and optimism have a negligible or no effect on charitable giving behaviour with regard to sure-thing and probabilistic charities (in both Main Choice and Final Choice designs). This is in itself an important finding that deserves to be recognised in the academic literature in an effort to combat file-drawer concerns and to expand our understanding of donation behaviour with regard to real charities.

Our results in Final Choice also provide us with some statistically significant effects. First, we find that empathy predicts both frequency and size of donation overall. This is in line with the previous literature on empathy and charitable giving (Davis et al., 1999; Verhaert & Van den Poel, 2011; Neumayr & Handy, 2019). Second, with regard to the central research question at hand, we find that in predicting charitable behaviour when participants are only presented with either a sure-thing or a probabilistic charity, we find that purely altruistic donor type predicts frequency of donation to probabilistic charities. This effect is directionally as predicted and is in line with theory in that pure altruists are more likely to be primarily concerned with an intervention's impact on social welfare, and given the potentially high (expected) impact that probabilistic charities might have, the results provide evidence for this claim. This suggests that, at least compared to our analyses of the Main Choice, the data in Final Choice provide at least weak evidence in favour of our alternative hypotheses.

The results in Final Choice may also help reinforce the above claim in favour of the validity of this outcome measure and suggest that one reason for a failure to detect an effect in Main Choice may have been due to distinction bias related concerns, the bias that preferences and choices may differ between joint evaluation and separate evaluation modes (Hsee, Loewenstein, Blount, & Bazerman, 1999; Hsee & Zhang, 2004; Anvari, Olsen, Hung, & Feldman, 2021). On their own terms, these findings suggest that pure altruists are more likely to donate to probabilistic charities compared to other donor types (like warm glow donors).

Overall, given that that we find little or no evidence in favour of rejecting our null hypotheses as well as provide evidence in favour of a null effect across both of our main hypotheses, we take our paper to primarily report a negative result. Given our relatively high level of power that was calculated a priori, and the fact that all analysis steps were pre-registered (and where we deviated from the pre-registered protocol, we documented this and provided the pre-registered analyses in the appendix), and because our equivalence tests provided strong evidence in favour of negligibly small results, we can be relatively confident in this result. Because it is important for any science aiming to be

cumulative and reproducible, it is imperative that null results like these are communicated openly and clearly, and we take our paper to be doing precisely that.

Limitations

One potential limitation of this study is that our final sample sizes used for analyses of Main Choice were lower than those specified in the a priori power analyses. While we recruited the exact number that we preregistered, this reduction in sample size was primarily because a lower percentage of participants decided to donate to any charity than expected, i.e., in our data set 35.8% of participants choose to make a donation, while we expected about two-thirds of participants to make a donation (cf. Butera & Horn, 2017; Metzger & Günther, 2019). However, because our power analyses were conducted with the highly set goal of having .95 power to detect the smallest effect size of interest at $f^2=.02$, we were only slightly below the sample size required for the more conventional level of power at .8 (which would have required 395 participants), which should alleviate these concerns at least in part. Additionally, these sample size concerns do not apply to any analyses conducted with regard to our Final Choice data due to the difference in design. As such, while these worries ought to be brought to the attention of readers, we do think that they are ultimately manageable in scope and do not take away from the findings of this paper.

Conclusion

In this paper, we have attempted to investigate what best explains individual charitable giving behaviour in situations where donors are presented with highly reliable charities (sure-thing charities) and significantly riskier charities (probabilistic charities). We have failed to reject most of our null hypotheses in the context of donor behaviour when choosing between the types of charities and can only offer mixed results with regard to donor choices when they are presented with either a sure-thing or a probabilistic charity. There, we find that donor type meaningfully predicts frequency of donation to probabilistic charities. Exploratory results of equivalence tests, however, provide positive evidence in favour of the absence of an effect (or the presence of a negligibly small effect) across most of our main explanatory variables. Overall, we take this study to produce a robustly negative result in that our hypothesised variables failed to predict donor behaviour, while also producing positive evidence in favour of the claim that individual

differences in risk/ambiguity attitudes and individual differences in numeracy, optimism, and empathy do not predict choices with regard to these two types of charities.

Acknowledgements

We thank for helpful comments, suggestions, and support Kirby Nielsen, Julian Jamison, Andreas Mogensen, Maximilian Maier, Joshua Lewis, Ben Grodeck, Theron Pummer, Ruru Hoong, and Thomas Rowe. We further thank the Forethought Foundation and the Centre for Effective Altruism for generous funding.

References

- Aflaki, A., & Pedraza-Martinez, A. J. (2016). Humanitarian funding in a multi-donor market with donation uncertainty. *Production and Operations Management*, 25(7), 1274-1291.
- Alter, U., & Counsell, A. (2021). Equivalence Testing for Multiple Regression. <https://psyarxiv.com/ugc9e/>
- Andreoni, J., & Serra-Garcia, M. (2021). The pledging puzzle: How can revocable promises increase charitable giving?. *Management Science*, 67(10), 6198-6210.
- Angrist, J. D., & Pischke, J. S. (2008). *Mostly harmless econometrics*. Princeton university press.
- Antoncic, B., Bratkovic Kregar, T., Singh, G., & DeNoble, A. F. (2015). The big five personality–entrepreneurship relationship: Evidence from Slovenia. *Journal of small business management*, 53(3), 819-841.
- Anvari, F., Olsen, J., Hung, W. Y., & Feldman, G. (2021). Misprediction of affective outcomes due to different evaluation modes: Replication and extension of two distinction bias experiments by Hsee and Zhang (2004). *Journal of Experimental Social Psychology*, 92, 104052.
- Ashton, M. C., Paunonen, S. V., Helmes, E., & Jackson, D. N. (1998). Kin altruism, reciprocal altruism, and the Big Five personality factors. *Evolution and Human Behavior*, 19(4), 243-255.
- Bekkers, R., & Wiepking, P. (2011). Who gives? A literature review of predictors of charitable giving part one: Religion, education, age and socialisation. *Voluntary Sector Review*, 2(3), 337-365.
- Bianchi, M., & Tallon, J. M. (2019). Ambiguity preferences and portfolio choices: Evidence from the field. *Management Science*, 65(4), 1486-1501.

- Boenigk, S., & Mayr, M. L. (2016). The happiness of giving: Evidence from the German socioeconomic panel that happier people are more generous. *Journal of Happiness Studies*, *17*(5), 1825-1846.
- Brock, J. M., Lange, A., & Ozbay, E. Y. (2013). Dictating the risk: Experimental evidence on giving in risky environments. *American Economic Review*, *103*(1), 415-37.
- Butera, L., & Horn, J. R. (2017). Good News, Bad News, and Social Image: The Market for Charitable Giving. *George Mason University Interdisciplinary Center for Economic Science (ICES) Working Paper*.
- Carpenter, J. P., 2018. The Shape of Warm Glow: Field Experimental Evidence from a Fundraiser. *IZA Discussion Paper No. 11760*.
- Carré, A., Stefaniak, N., D'ambrosio, F., Bensalah, L., & Besche-Richard, C. (2013). The Basic Empathy Scale in Adults (BES-A): Factor structure of a revised form. *Psychological assessment*, *25*(3), 679-691.
- Caviola, L., Schubert, S., & Nemirow, J. (2020). The many obstacles to effective giving. *Judgment and Decision Making*, *15*(2), 159-172.
- Cettolin, E., Riedl, A., & Tran, G. (2017). Giving in the face of risk. *Journal of risk and uncertainty*, *55*(2), 95-118.
- Chen, Y., & Zhong, S. (2021). Uncertainty Motivates Morality: Evidence and Theory. *Available at SSRN 3737959*.
- Cleary, P. D., & Angel, R. (1984). The analysis of relationships involving dichotomous dependent variables. *Journal of Health and Social Behavior*, 334-348.
- Crosetto, P., & Filippin, A. (2013). The “bomb” risk elicitation task. *Journal of Risk and Uncertainty*, *47*(1), 31-65.
- Davis, M. H., Mitchell, K. V., Hall, J. A., Lothert, J., Snapp, T., & Meyer, M. (1999). Empathy, expectations, and situational preferences: Personality influences on the decision to participate in volunteer helping behaviors. *Journal of personality*, *67*(3), 469-503.
- Dey, E. L., & Astin, A. W. (1993). Statistical alternatives for studying college student retention: A comparative analysis of logit, probit, and linear regression. *Research in higher education*, *34*(5), 569-581.
- Dickert, S., Kleber, J., Peters, E., & Slovic, P. (2011). Numeracy as a precursor to pro-social behavior: The impact of numeracy and presentation format on the cognitive mechanisms underlying donation decisions.

- Engel, C., & Goerg, S. J. (2018). If the worst comes to the worst: Dictator giving when recipient's endowments are risky. *European Economic Review*, *105*, 51-70.
- Exley, C. L. (2016). Excusing selfishness in charitable giving: The role of risk. *The Review of Economic Studies*, *83*(2), 587-628.
- Galizzi, M. M., & Navarro-Martinez, D. (2019). On the external validity of social preference games: a systematic lab-field study. *Management Science*, *65*(3), 976-1002.
- Garcia, T., Massoni, S., & Villeval, M. C. (2020). Ambiguity and excuse-driven behavior in charitable giving. *European Economic Review*, *124*, 103412.
- Giving USA 2020 (2020, June). Giving USA. <https://givingusa.org/giving-usa-2020-charitable-giving-showed-solid-growth-climbing-to-449-64-billion-in-2019-one-of-the-highest-years-for-giving-on-record/>
- Gneezy, U., Imas, A., & List, J. (2015). *Estimating individual ambiguity aversion: A simple approach* (No. w20982). National Bureau of Economic Research.
- Gomila, R. (2020). Logistic or linear? Estimating causal effects of experimental treatments on binary outcomes using regression analysis. *Journal of Experimental Psychology: General*.
- Haisley, E. C., & Weber, R. A. (2010). Self-serving interpretations of ambiguity in other-regarding behavior. *Games and economic behavior*, *68*(2), 614-625.
- Haisley, E. C., & Weber, R. A. (2010). Self-serving interpretations of ambiguity in other-regarding behavior. *Games and economic behavior*, *68*(2), 614-625.
- Halevy, Y. (2007). Ellsberg revisited: An experimental study. *Econometrica*, *75*(2), 503-536.
- Hellevik, O. (2009). Linear versus logistic regression when the dependent variable is a dichotomy. *Quality & Quantity*, *43*(1), 59-74.
- Hilbig, B. E., & Zettler, I. (2009). Pillars of cooperation: Honesty–Humility, social value orientations, and economic behavior. *Journal of Research in Personality*, *43*(3), 516-519.
- Holzmeister, F., & Stefan, M. (2021). The risk elicitation puzzle revisited: Across-methods (in) consistency? *Experimental Economics*, *24*(2), 593-616.
- Hsee, C. K., & Zhang, J. (2004). Distinction bias: misprediction and mischoice due to joint evaluation. *Journal of personality and social psychology*, *86*(5), 680-695.
- Hsee, C. K., Loewenstein, G. F., Blount, S., & Bazerman, M. H. (1999). Preference reversals between joint and separate evaluations of options: a review and theoretical analysis. *Psychological bulletin*, *125*(5), 576-590.

- Kang, E., & Lakshmanan, A. (2018). Narcissism and self-versus recipient-oriented imagery in charitable giving. *Personality and Social Psychology Bulletin*, 44(8), 1214-1227.
- Kappes, A., Nussberger, A. M., Faber, N. S., Kahane, G., Savulescu, J., & Crockett, M. J. (2018). Uncertainty about the impact of social decisions increases prosocial behaviour. *Nature human behaviour*, 2(8), 573-580.
- Karlan, D., & List, J. A. (2007). Does price matter in charitable giving? Evidence from a large-scale natural field experiment. *American Economic Review*, 97(5), 1774-1793.
- Kim, S. J., & Kou, X. (2014). Not all empathy is equal: How dispositional empathy affects charitable giving. *Journal of Nonprofit & Public Sector Marketing*, 26(4), 312-334.
- Kleber, J., Dickert, S., Peters, E., & Florack, A. (2013). Same numbers, different meanings: How numeracy influences the importance of numbers for pro-social behavior. *Journal of Experimental Social Psychology*, 49(4), 699-705.
- Knapp, L., Wuepper, D., & Finger, R. (2021). Preferences, Personality, Aspirations, and Farmer Behavior. *Agricultural Economics*, XXX, 1-13.
- Liang, F., Paulo, R., Molina, G., Clyde, M. A., & Berger, J. O. (2008). Mixtures of g priors for Bayesian variable selection. *Journal of the American Statistical Association*, 103(481), 410-423.
- Lipkus, I. M., Samsa, G., & Rimer, B. K. (2001). General performance on a numeracy scale among highly educated samples. *Medical decision making*, 21(1), 37-44.
- Lipkus, I. M., Samsa, G., & Rimer, B. K. (2001). General performance on a numeracy scale among highly educated samples. *Medical decision making*, 21(1), 37-44.
- Mesa-Vázquez, E., Rodríguez-Lara, I., & Urbano, A. (2021). Standard vs random dictator games: On the effects of role uncertainty and framing on generosity. *Economics Letters*, 109981.
- Metzger, L., & Günther, I. (2019). Making an impact? The relevance of information on aid effectiveness for charitable giving. A laboratory experiment. *Journal of Development Economics*, 136, 18-33.
- Müller, J., & Schwioren, C. (2012). What can the Big Five personality factors contribute to explain small-scale economic behavior?
- Neumayr, M., & Handy, F. (2019). Charitable giving: What influences donors' choice among different causes? *Voluntas: International Journal of Voluntary and Nonprofit Organizations*, 30(4), 783-799.

- Niehaus, P. (2014). A theory of good intentions. *San Diego, CA: University of California and Cambridge, MA: NBER, 111.*
- Nielsen, K. (2019). Dynamic risk preferences under realized and paper outcomes. *Journal of Economic Behavior & Organization, 161*, 68-78.
- Nonprofit Sector in Brief 2019 (2020, June). National Center for Charitable Statistics. <https://nccs.urban.org/publication/nonprofit-sector-brief-2019>
- Null, C. (2011). Warm glow, information, and inefficient charitable giving. *Journal of Public Economics, 95*(5-6), 455-465.
- O'Reilly, N., Ayer, S., Pegoraro, A., Leonard, B., & Rundle-Thiele, S. (2012). Toward an understanding of donor loyalty: Demographics, personality, persuasion, and revenue. *Journal of Nonprofit & Public Sector Marketing, 24*(1), 65-81.
- Olschewski, S., Dietsch, M., & Ludvig, E. A. (2019). Anti-social motives explain increased risk aversion for others in decisions from experience. *Judgment and Decision making, 14*(1), 58-71.
- Rouder, J. N., & Morey, R. D. (2012). Default Bayes factors for model selection in regression. *Multivariate Behavioral Research, 47*(6), 877-903.
- Scheier, M. F., Carver, C. S., & Bridges, M. W. (1994). Distinguishing optimism from neuroticism (and trait anxiety, self-mastery, and self-esteem): a reevaluation of the Life Orientation Test. *Journal of personality and social psychology, 67*(6), 1063.
- Sutter, M., Kocher, M. G., Glätzle-Rützler, D., & Trautmann, S. T. (2013). Impatience and uncertainty: Experimental decisions predict adolescents' field behavior. *American Economic Review, 103*(1), 510-31.
- Verhaert, G. A., & Van den Poel, D. (2011). Empathy as added value in predicting donation behavior. *Journal of Business Research, 64*(12), 1288-1295.
- Vesterlund, L. (2006). Why Do people give? In W. Powell & R. Steinberg (Eds.), *The Nonprofit Sector: A Research Handbook* (pp. 568-590). Yale University Press.
- Vives, M. L., & FeldmanHall, O. (2018). Tolerance to ambiguous uncertainty predicts prosocial behavior. *Nature communications, 9*(1), 1-9.
- What are the best charities to donate to in 2021? (2021). Giving What We Can. <https://www.givingwhatwecan.org/best-charities-to-donate-to-2021/>
- Wiepking, P., & Bekkers, R. (2012). Who gives? A literature review of predictors of charitable giving. Part Two: Gender, family composition and income. *Voluntary Sector Review, 3*(2), 217-245.

Wu, S. Y. (2004). Tax effects on charitable giving in the presence of uncertainty. *Public Finance Review*, 32(5), 459-482.

Appendix A – Auxiliary Study

We conducted a post-hoc auxiliary study to empirically confirm our categorisation into sure-thing and probabilistic charities. This is to ensure that this distinction is not only theoretically grounded but also perceived as intended by the general public. We recruited a total of 101 participants on Prolific that had not participated in the main study, none of which failed the attention check. Participants were paid £0.75 for their participation. They were presented with all six charities and were asked to rate them on a scale from 0-10 on the likelihood that the charity’s intervention succeeds (relating to uncertainty over its interventions) and on the quantifiability of the charity’s intervention (relating to ambiguity). We also asked participants to rate the charities on their moral deservingness to keep the objective of this study opaque.

We find strong support for the distinction between sure-thing charities and probabilistic charities on the basis of both uncertainty and ambiguity. See Appendix Table 1 for means, standard deviations, and medians of the uncertainty and ambiguity ratings for all six charities, with 0 indicating low probability that the charity’s intervention will succeed and a low level of quantifiability of its interventions, and 10 indicating a high probability and quantifiability. In other words, the higher the scores, the less risky and the less ambiguous the charity’s respective intervention is.

APPENDIX TABLE 1—PROBABILITY-RATINGS FOR ALL SIX CHARITIES

Uncertainty	Mean (SD)	Median
Sure-Thing Charity 1 (SCI Foundation)	7.19 (2.08)	8
Sure-Thing Charity 2 (GiveDirectly)	5.92 (2.09)	6
Sure-Thing Charity 3 (Against Malaria Foundation)	7.36 (1.83)	8
Probabilistic Charity 1 (Machine Intelligence Research Institute)	4.15 (2.46)	4
Probabilistic Charity 2 (Nuclear Threat Initiative)	4.09 (2.64)	4
Probabilistic Charity 3 (The Center for Health Security)	4.30 (2.50)	4
Ambiguity		

	Mean (SD)	Median
Sure-Thing Charity 1 (SCI Foundation)	7.28 (1.97)	8
Sure-Thing Charity 2 (GiveDirectly)	5.42 (2.49)	5
Sure-Thing Charity 3 (Against Malaria Foundation)	7.24 (2.10)	8
Probabilistic Charity 1 (Machine Intelligence Research Institute)	3.63 (2.72)	3
Probabilistic Charity 2 (Nuclear Threat Initiative)	4.03 (2.73)	4
Probabilistic Charity 3 (The Center for Health Security)	4.69 (2.40)	4

Notes: Mean, Standard Deviation, and Median of risk and ambiguity ratings for all six charities.

We find that the data behave as generally expected, with sure-thing charities receiving higher ratings about the likelihood that their interventions will succeed as well as higher ratings for the quantifiability of their interventions, and probabilistic charities receiving lower rating correspondingly. Adding a subject’s scores for the individual charities of each bucket, we find that sure-thing charities are rated as having significantly higher probability interventions ($M=20.47$, $SD=4.52$) than the probabilistic charities ($M=13.96$, $SD=5.83$). This difference, 6.51, 95% CI [5.19, 7.82] was highly statistically significant, $t(100)=9.81$, $p<.001$. The same picture emerges with regard to the quantifiability of the interventions, with the mean of the sum of the quantifiability scores of sure-thing charities ($M=19.83$, $SD=4.60$) being significantly higher than that of the probabilistic charities ($M=12.85$, $SD=6.70$), with the difference of 6.98, 95% CI [5.62, 8.34] also being statistically significant at $t(100)=10.17$, $p<.001$. The effect sizes of these two differences, in Cohen’s d , is $d=.98$ for the probability ratings and $d=1.01$ for the quantifiability rankings. This provides strong support for our theoretically based distinction between sure-thing and probabilistic charities.

Appendix B – Robustness Checks for Main Choice

In Appendix Table 2, Model (4), we present a robustness check for Model (1) that runs the OLS model as a logit model. We find the same pattern of results throughout.

APPENDIX TABLE 2—REGRESSION RESULTS FOR MAIN CHOICE - LOGIT ROBUSTNESS CHECK
PREDICTING CHOICE BETWEEN SURE-THING AND PROBABILISTIC CHARITIES

	(4)
Risk Attitude	.002 (.044)
Ambiguity Aversion	-.033 (.037)

Numeracy	.054 (.192)
Empathy	.001 (.020)
Optimism	-.035 (.026)
Donor Type	
Warm-Glow	-.157 (.401)
Pure Altruism	-.499 (.554)
Age	.024 (.016)
Gender	.084 (.401)
Education	
Undergraduate degree	.292 (.398)
Postgraduate/Professional degree	-.176 (.453)
Religion	
Protestantism	-.673 (.579)
Catholicism	.003 (.541)
Islam	-1.074 (1.193)
Judaism	-19.179 (40192.970)
Buddhism	-19.198 (28374.585)
Hinduism	1.886 (1.651)
Religious Participation	.407 (.722)
Marriage Status	.330 (.388)
Children	-.011 (.427)
Financial Wellbeing	.067 (.195)
Employment	
Out of the workforce	-1.375* (.796)
Part-time employment	-.010 (.634)
Full-time employment	-.463 (.611)
Cox & Snell R ²	.055
Sample size	307

Notes: Log-odds and standard errors. *p<.1, **p<.05, ***p<.01, ****p<.001

We also report an additional robustness check of Model (1) in Appendix Table 3, Model (5). Specifically, we report a random effects model with the stimulus material being treated as a random effect. As we had three sure-thing and three probabilistic charities, there were nine charity pairs that participants could have been presented with. In Model (5), we treated the stimulus (i.e. the charity pair presented) as a random effect (with nine levels). Because significance levels are not unproblematic in mixed models like this, we also report 95% confidence intervals of the estimates. The results indicate that our null effect is robust to this model choice as well.

APPENDIX TABLE 3— REGRESSION RESULTS FOR MAIN CHOICE – RANDOM EFFECTS ROBUSTNESS CHECK
PREDICTING CHOICE BETWEEN SURE-THING AND PROBABILISTIC CHARITIES

		(5)
Risk Attitude	.002 (.005)	[-.008, .012]
Ambiguity Aversion	-.002 (.005)	[-.010, .006]
Numeracy	.003 (.023)	[-.042, 0.47]
Empathy	.001 (.002)	[-.005, .004]
Optimism	-.006* (.003)	[-.012, .001]
Donor Type		
Warm-Glow	-.007 (.047)	[-.099, .086]
Pure Altruism	-.040 (.059)	[-.156, .077]
Age	.001 (.002)	[-.003, .004]
Gender	-.024 (.047)	[-.115, .068]
Education		
Undergraduate degree	.041 (.048)	[-.053, .134]
Postgraduate/Professional degree	-.023 (.051)	[-.123, .078]
Religion		
Protestantism	-.043 (.062)	[-.165, .080]
Catholicism	.020 (.068)	[-.114, .154]
Islam	-.035 (.113)	[-.246, .188]
Judaism	-.006 (.354)	[-.691, .703]
Buddhism	-.060 (.265)	[-.581, .462]
Hinduism	.325 (.254)	[-.175, .825]
Religious Participation	-.004 (.083)	[-.168, .159]
Marriage Status	-.019 (.046)	[-.071, .108]
Children	.009 (.050)	[-.090, .107]
Financial Wellbeing	.009 (.022)	[-.035, .053]
Employment		
Part-time employment	.058 (.054)	[-.047, .164]
Full-time employment	.015 (.049)	[-.083, .112]
Sample size	307	

Notes: Random effects model, coefficients, and standard errors, as well as 95% CIs. *p<.1, **p<.05, ***p<.01, ****p<.001

Appendix Table 4, Model (6) presents the main results with the original exclusion criteria, i.e. we exclude variables with less than or equal to one instance of ‘1’. We find no differences in results. To maintain an acceptable level of power, all main results in the paper are reported with less strict exclusion criteria.

APPENDIX TABLE 4— REGRESSION RESULTS FOR MAIN CHOICE – STRICTER EXCLUSIONS ROBUSTNESS CHECK
PREDICTING CHOICE BETWEEN SURE-THING AND PROBABILISTIC CHARITIES

	(6)
Risk Attitude	.004 (.009)
Ambiguity Aversion	-.011 (.009)
Numeracy	.019 (.038)

Empathy	-0.002 (.004)
Optimism	-.007 (.005)
Donor Type	
Warm-Glow	.001 (.089)
Pure Altruism	-.031 (.102)
Age	.002 (.003)
Gender	.113 (.085)
Education	
Undergraduate degree	.021 (.088)
Postgraduate/Professional degree	.017 (.089)
Religion	
Protestantism	-.082 (.112)
Catholicism	-.076 (.107)
Islam	-.111 (.179)
Judaism	-
Buddhism	-
Hinduism	-
Religious Participation	.068 (.145)
Marriage Status	-.012 (.083)
Children	-.052 (.091)
Financial Wellbeing	-.012 (.040)
Employment	
Out of the workforce	-.028 (.170)
Part-time employment	.132 (.152)
Full-time employment	.097 (.141)
R ²	.084
Sample size	137

Notes: Unstandardised coefficients and standard errors. *p<.1, **p<.05, ***p<.01, ****p<.001

Appendix C – Robustness Checks for Final Choice

In Appendix Table 5, Model (7), we provide a logit model for Model (2) to check for robustness of results. The results are the same as in Model (2).

APPENDIX TABLE 5—REGRESSION RESULTS FOR FINAL CHOICE – LOGIT ROBUSTNESS CHECK
PREDICTING FREQUENCY OF DONATION

	(7)
Risk Attitude	.009 (.022)
Ambiguity Aversion	-.010 (.020)
Numeracy	.098 (.087)
Empathy	.036**** (.009)
Optimism	.016 (.013)
Donor Type	

Warm-Glow	.268 (.225)
Pure Altruism	.230 (.278)
Condition X Risk Attitude	.016 (.032)
Condition X Ambiguity Aversion	.012 (.028)
Condition X Numeracy	.003 (.099)
Condition X Empathy	-.006 (.011)
Condition X Optimism	-.017 (.017)
Donor Type	
Condition X Warm-Glow	-.097 (.322)
Condition X Pure Altruism	.897** (.400)
Age	.003 (.006)
Gender	-.272* (.156)
Education	
Undergraduate degree	-.262 (.161)
Postgraduate/Professional degree	.158 (.173)
Religion	
Protestantism	.366 (.214)
Catholicism	.246 (.238)
Islam	.863** (.428)
Judaism	.673 (.683)
Buddhism	.710 (.753)
Hinduism	.712 (.624)
Sikhism	.264 (1.273)
Religious Participation	.071 (.264)
Marriage Status	-.120 (.161)
Children	.161 (.175)
Financial Wellbeing	.116 (.076)
Employment	
Out of the workforce	.083 (.284)
Part-time employment	-.106 (.250)
Full-time employment	-.156 (.221)
Cox and Snell R ²	.075
Sample size	1177

Notes: Log-odds and standard errors. *p<.1, **p<.05, ***p<.01, ****p<.001

In Appendix Table 6, we report the four pre-registered models for Final Choice. These did not include interactions but instead split the models by type of charity presented to them. Model (8) and Model (9) have frequency of donations to sure-thing and probabilistic charities as their outcome variables respectively. Model (10) and Model (11) are of the corresponding size of donation. These results are largely in line with Model (2) and Model (3), in that empathy is predictive of donation behaviour generally, but it is the pure altruistic donor type that best predict donation choices towards probabilistic charities in Final Choice.

APPENDIX TABLE 6—REGRESSION RESULTS FOR FINAL CHOICE – PRE-REGISTERED REGRESSION MODELS
PREDICTION FREQUENCY AND SIZE OF DONATION

	(8)	(9)	(10)	(11)
Risk Attitude	.003 (.005)	.003 (.004)	.150 (.147)	.109 (.095)
Ambiguity Aversion	-.003 (.004)	<.001 (.003)	-.131 (.129)	-.019 (.077)
Numeracy	.033 (.023)	.008 (.017)	.598 (.653)	-.495 (.481)
Empathy	.008**** (.002)	.005** (.002)	.258**** (.06)	.082* (.043)
Optimism	.002 (.003)	.001 (.003)	.042 (.090)	.047 (.057)
Donor Type				
Warm-Glow	.047 (.051)	.036 (.040)	2.033 (1.485)	1.346 (.914)
Pure Altruism	.059 (.063)	.216**** (.056)	.315 (1.825)	2.806** (1.284)
Age	.001 (.002)	<.001 (.001)	.042 (.054)	.036 (.034)
Gender	-.038 (.087)	-.064* (.038)	-.255 (1.387)	-1.365 (.874)
Education				
Undergraduate degree	-.074 (.048)	-.057 (.044)	-2.564* (1.3741)	.079 (.907)
Postgraduate/Professional	.003 (.055)	.057 (.054)	-.757 (1.574)	.589 (1.007)
Religion				
Protestantism	.107 (.071)	.051 (.054)	.147 (2.042)	.589 (1.225)
Catholicism	.059 (.077)	.055 (.060)	3.103 (2.222)	.855 (1.361)
Islam	.286* (.158)	.106 (.106)	9.431** (4.550)	4.612* (2.404)
Judaism	.228 (.341)	.110 (.151)	6.908 (9.846)	2.308 (3.429)
Buddhism	.297 (.287)	.099 (.191)	15.139* (8.273)	.117 (4.338)
Hinduism	.226 (.216)	.080 (.154)	8.880 (6.223)	4.785 (3.507)
Sikhism	.199 (.343)	-.259 (.423)	-1.570 (9.899)	-5.135 (9.648)
Religious Participation	.015 (.090)	-.006 (.067)	.390 (2.592)	-1.723 (1.563)
Marriage Status	-.006 (.050)	-.039 (.040)	.387 (1.435)	-1.082 (.912)
Children	.074 (.053)	-.006 (.043)	.926 (1.534)	.737 (.978)
Financial Wellbeing	.038* (.023)	.007 (.019)	.962 (.655)	-.096 (.439)
Employment				
Out of the workforce	.093 (.092)	-.030 (.068)	1.320 (2.641)	-2.080 (1.550)
Part-time employment	.010 (.079)	-.043 (.059)	-1.060 (2.274)	-3.141** (1.353)
Full-time employment	-.014 (.069)	-.038 (.054)	.571 (1.982)	-2.985** (1.231)
R ²	.080	.054	.086	.052
Sample size	546	631	546	631

Notes: Unstandardised coefficients and standard errors. *p<.1, **p<.05, ***p<.01, ****p<.001

Models (12) and (13) in Appendix Table 7 present the results of Models (8) and (9) from Appendix Table 6 with logit models and otherwise identical specifications. We find no difference in results.

APPENDIX TABLE 7— REGRESSION RESULTS FOR FINAL CHOICE – LOGIT ROBUSTNESS CHECKS FOR PRE-REGISTERED MODELS
PREDICTING FREQUENCY OF DONATION

	(12)	(13)
Risk Attitude	.015 (.023)	.017 (.024)
Ambiguity Aversion	-.014 (.021)	.003 (.020)
Numeracy	.158 (.106)	.049 (.103)
Empathy	.038**** (.011)	.029** (.012)
Optimism	.011 (.014)	.005 (.015)
Donor Type		
Warm-Glow	.215 (.234)	.223 (.235)
Pure Altruism	.274 (.286)	1.101**** (.294)
Age	.002 (.009)	.002 (.009)
Gender	-.189 (.221)	-.369* (.225)
Education		
Undergraduate degree	-.353 (.222)	-.199 (.243)
Postgraduate/Professional degree	.010 (.247)	.291 (.306)
Religion		
Protestantism	.476 (.313)	.291 (.306)
Catholicism	.275 (.346)	.321 (.338)
Islam	1.256* (.702)	.577 (.596)
Judaism	1.018 (1.467)	.552 (.782)
Buddhism	1.290 (1.297)	.452 (.989)
Hinduism	.978 (.955)	.451 (.879)
Sikhism	.888 (1.480)	-.20.097 (40192.970)
Religious Participation	.071 (.398)	-.015 (.377)
Marriage Status	-.029 (.227)	-.228 (.236)
Children	.343 (.245)	-.048 (.253)
Financial Wellbeing	.173* (.105)	.038 (.113)
Employment		
Out of the workforce	.431 (.422)	-.161 (.394)
Part-time employment	.073 (.370)	-.250 (.348)
Full-time employment	-.040 (.327)	-.212 (.311)
Cox and Snell R ²	.079	.051
Sample size	546	631

Notes: Log-odds and standard errors. *p<.1, **p<.05, ***p<.01, ****p<.001

Appendix D – Additional Hypotheses Tests with Inconclusive Results

In this part of the appendix, we report the results regarding null hypotheses #3 and #4. Due to the low sample size, these analyses do not have the power we outlined above and are as such taken to be inconclusive. However, we still report all analyses as pre-registered.

First, we find that 83.3% of those who donate choose a sure-thing charity, and their average donation (M=53.91, SD=24.52) does not differ statistically significantly, $t(40)=.818$, $p=.418$, from those who donate to a probabilistic charity (M=45.71,

SD=22.44). In Appendix Table 8, we report the results of the context-free condition, in which the design of Main Choice was mirrored but all identifying information and context was removed from the charity pair. Model (14) reports an OLS model and Model (15) tests the robustness of these results in a logit model. Both produce the same results. For both analyses, we excluded variables that had less than or equal to one instance of ‘1’. Given the low sample size we do not interpret these results with regard to null hypothesis #3.

APPENDIX TABLE 8—REGRESSION RESULTS FOR NO CONTEXT CONDITION – OLS AND LOGIT
PREDICTING CHOICE BETWEEN SURE-THING AND PROBABILISTIC CHARITIES

	(14)	(15)
Risk Attitude	-.012 (.024)	-4.401 (913.808)
Ambiguity Aversion	.003 (.017)	.816 (1138.332)
Numeracy	.099 (.121)	89.731 (11034.812)
Empathy	.017 (.011)	4.576 (810.673)
Optimism	-.012 (.011)	-4.083 (635.613)
Donor Type		
Warm-Glow	-.043 (.223)	-9.676 (10508.857)
Pure Altruism	-.016 (.197)	40.576 (7375.105)
Age	.005 (.008)	4.758 (613.650)
Gender	.078 (.182)	54.477 (10072.358)
Education		
Undergraduate degree	.048 (.177)	64.545 (11881.462)
Postgraduate/Professional degree	-.196 (.224)	32.012 (19491.480)
Religion		
Protestantism	-.625 (.437)	-111.071 (26044.398)
Catholicism	-.101 (.286)	-108.735 (21698.451)
Islam	-	-
Judaism	-	-
Buddhism	-	-
Hinduism	-	-
Religious Participation	.080 (.362)	-72.564 (35338.129)
Marriage Status	-.056 (.189)	-92.326 (14588.695)
Children	-.068 (.221)	-6.219 (9909.734)
Financial Wellbeing	.017 (.100)	10.290 (4125.579)
Employment		
Out of the workforce	.293 (.491)	-48.198 (14982.568)
Part-time employment	-.110 (.259)	-92.719 (13479.830)
Full-time employment	.073 (.218)	-36.104 (8611.194)
R ² and Cox & Snell R ²	.394	.594
Sample size	42	42

Notes: Unstandardized coefficients and standard errors (Model 1) as well as log-odds and standard errors (Model 2). *p<.1, **p<.05, ***p<.01, ****p<.001

Second, in Appendix Table 9, we test the effects of the informational treatment. We find that participant choices are not meaningfully different between the Main Choice, the expected-value treatment, and the made-up midontic theory treatment, see Appendix Table 9. The percentage point difference in the relative frequency of donations between the Main Condition and the expected-value treatment is 0.5p.p., whereas the difference between the Main Condition and the control is 0.6 p.p. Further, the direction of the change is contra the predicted directions, as the expected-value treatment seems to have resulted in a miniscule increase in donations to sure-thing charities. Given the low sample size we do not interpret these results with regard to null hypothesis #3.

APPENDIX TABLE 9—EXPECTED VALUE TREATMENT EFFECT

	Main Condition	Expected-Value Treatment	Midontic Theory Treatment
Sure-Thing Charity	84.7%	85.2%	84.1%
Probabilistic Charity	15.3%	14.8%	15.9%

Notes: Frequency of donations to sure-thing and probabilistic charities.

Appendix E – Additional Equivalence Tests

In Appendix Table 10, we report an additional series of equivalence tests for Final Choice. In this series of tests, we focus on Table 3, Model (3) and the size of donation. As before, because we did not pre-register equivalence bounds we report a number of them to check for sensitivity of results to choice of equivalence bounds. Note though that in contrast to previous equivalence tests, this set of equivalence bounds is defined with regard to standardised regression coefficients from Model (3). Here we find that at the equivalence bounds of (-).25, risk attitudes, ambiguity aversion, and donor type are negligibly small. At the (-).5 level, numeracy and optimism are also negligibly small within these modestly large parameters. We do not have sufficient evidence in favour of a negligibly small effect even at (-).5 for empathy. We thus argue that we have relatively strong evidence in favour of risk attitudes, ambiguity aversion, and donor type as being negligibly small, while the other variables provide less strong evidence in favour of this.

APPENDIX TABLE 10—TOST FOR FINAL CHOICE COEFFICIENTS

	-1	.1	-.25	.25	-.5	.5
Condition X Risk Attitudes	1.61*	1.37*	3.85****	3.61****	7.58****	7.34****
Condition X Ambiguity Aversion	2.86***	1.68**	6.27****	5.09****	11.96****	10.77****
Condition X Numeracy	1.10	-.04	1.89**	.23	3.22****	2.07**
Condition X Empathy	-0.99	2.09**	-.16	2.92**	1.22	4.30****
Condition X Optimism	.40	1.16	1.58*	2.33**	3.53****	4.28****
Condition X Warm Glow	1.53*	2.72**	4.72****	5.91****	10.04****	11.23****
Condition X Pure Altruism	1.02	3.74****	4.56****	7.31****	10.55****	13.26****

Notes: All t-test results for TOST procedures on a variety of lower and upper equivalence bounds (in standardized coefficients) from Model (3). * $p < .1$, ** $p < .05$, *** $p < .01$, **** $p < .001$

Appendix F – Experimental Texts

Below we show the main experimental texts. See Supplemental Materials for full experimental procedure including all instructions.

Charity Descriptions

SCI Foundation (Sure-Thing Charity #1)

The SCI Foundation focuses on improving global health by combating parasitic worm infections. It works with governments or non-governmental actors worldwide to treat a number of important diseases. For example, these diseases include schistosomiasis and soil-transmitted helminthiasis, both commonly known as 'worm infections.' These infections are treated by providing deworming drugs once a year that significantly reduce the worm load and cure the present infection.

According to recent trials, multiple dose interventions have a cure-rate of 96.1%. The total cost of deworming is about \$1.03. Previous research found that deworming campaigns improve the health of the treated children (increasing weight and haemoglobin levels), prevent potentially severe outcomes (such as organ damage), as well as reduce school absenteeism, increase cognitive function, and show an improvement in labour force participation and incomes in both the treated population and the community as a whole.

Donating to the SCI Foundation contributes to this deworming effort and will reliably improve the lives of many as well as their communities. There is a yearly shortfall of funding, and donating will directly help those in need.

Against Malaria Foundation (Sure-Thing Charity #3)

The Against Malaria Foundation has as its goal to directly distribute long-lasting insecticide-treated mosquito nets in at-risk low-income communities that stop the transmission of malaria. It works with governments or non-governmental actors worldwide to address this issue of malaria deaths. Providing mosquito nets can help the recipients reduce the chance of falling ill with malaria.

In 2019, a follow-up monitoring conducted between 6 and 23 months showed that 90% of recipients that received a net had applied it properly. The cost to purchase and distribute such an insecticide treated net is \$4.95. This means that providing insecticide-treated bed nets is an effective way of reducing malaria-related deaths and thus improving the welfare of recipients and their communities.

Donating to the Against Malaria Foundation contributes to the overall effort at reducing the impact of malaria and will reliably improve the lives of many as well as their communities. There is a yearly shortfall of funding, and donating will directly help those in need.

The Center for Health Security (Probabilistic Charity #2)

The Center for Health Security has as its mission the aim of protecting global health by focusing on epidemics and disasters and by ensuring that communities are resilient to major health challenges. It works with governments or non-governmental actors worldwide to prepare for future health disasters like emerging infectious diseases. For example, they run the 'Outbreak Observatory' in which outbreaks of new diseases are monitored and which contributes to preventing and responding to these outbreaks.

Recent estimates show that the yearly chance of a severe flu outbreak is between 0.5 and 1%. Such an influenza pandemic of these proportions could result in a 5% reduction in global GDP as well as between 21 and 33 million deaths worldwide. Preparing for such rare events in advance might make them less likely to happen or at least less devastating. The charity's interventions work to reduce this chance by an unknown probability.

Donating to the Center for Health Security will help prepare humanity for those very rare but important health risks by reducing the chance of such an outbreak as well as the potential severity of it. This can be achieved by being quicker at identifying or combating novel outbreaks. Donations are needed every year to have a chance at making an impact in this area.

GiveDirectly (Sure-Thing Charity #2)

GiveDirectly has as its goal to directly send cash to people in poverty in order to allow those living in extreme poverty to spend cash on whatever it is they need most. It works with governments or non-governmental actors worldwide to address extreme poverty. These cash transfers only reach people who are living in extreme poverty and are suffering from hunger, a lack of shelter, and easily preventable diseases. Cash transfers can help the recipients alleviate these conditions.

In 2017, 90% of recipients received the payments within 70 days. For every \$1 donated, GiveDirectly delivers \$0.83 directly to those in need. Recent research investigating what recipients actually spend their transfers on found that it was spent on livestock, durable goods, savings, business expenses, health expenditures, education expenditures, and consumption overall (though alcohol and tobacco spending did not increase), overall increasing their welfare and the welfare of their communities.

Donating to GiveDirectly contributes to this poverty alleviation effort and will reliably improve the lives of many as well as their communities. There is a yearly shortfall of funding, and donating will directly help those in need.

Machine Intelligence Research Institute (Probabilistic Charity #1)

The Machine Intelligence Research Institute has as its goal to prevent or reduce the risk of artificial intelligence/machine intelligence threats. It works with governments or non-governmental actors worldwide to prepare for the potential threat that intelligent systems pose to humanity. In their work they aim to research and make safer the underlying principles and applications of artificial intelligence.

According to a recent estimation, there is a 5% risk over the next century that artificial intelligence may pose a risk to humans at the scale of human extinction (billions of lives lost). The Machine Intelligence Research Institute works towards laying the foundations for the safety of future intelligence systems to better understand the implications of such technological advances. This work promises to help reduce the threat of these rare outcomes of intelligent systems and minimise the risk that humanity is exposed to every year. The charity's interventions work to reduce this chance by an unknown probability.

Donating to the Machine Intelligence Research Institute will directly help prepare humanity for those very rare but important technological risks by reducing the chance that any such technology will end up having a highly destructive impact. This can be achieved by developing technological solutions that may make such catastrophes less likely. Donations are needed every year to have a chance at making an impact in this area.

Nuclear Threat Initiative (Probabilistic Charity #3)

The Nuclear Threat Initiative has as its goal to prevent catastrophic attacks with weapons of mass destruction such as nuclear weapons. It works with governments or non-governmental actors worldwide to prepare for the threat of nuclear war. In their work they aim to strengthen non-proliferation and disarmament agreements and advance international partnerships for nuclear disarmament verification.

According to a survey of experts, the chance of a nuclear war that kills at least 1 million people is about 0.39% every year. It has also been estimated that between 30 and 75 million people would die worldwide as a result of such a nuclear war between the US and Russia. Preparing for such rare events in advance might make them less likely to happen or at least less devastating. The charity's interventions work to reduce this chance by an unknown probability.

Donating to the Nuclear Threat Initiative will help prepare humanity for those very rare but important war risks by reducing the chance of such a conflict, as well as the potential severity of it. This can be achieved by forming new treaties or by fostering international collaboration. Donations are needed every year to have a chance at making an impact in this area.

Context-free charity texts

Charity 1 (Context-free sure-thing charity)

This charity focuses on helping people in need. Its interventions have a success rate of 96.1%. The charity's interventions improve the well-being of those treated and lead to a number of additional positive consequences. Donating to this charity reliably improves lives of many as well as their communities. For this charity, there is a yearly shortfall of funding to achieve this goal.

Charity 2 (Context-free probabilistic charity)

This charity focuses on preventing global catastrophes. These catastrophes have a chance of 0.39% to happen every year, and if they happened might cost the lives of between 30 and 75 million people worldwide. The charity's interventions work to reduce this chance by an unknown probability. Donating to this charity might make these catastrophes less likely to happen or at least less devastating. For this charity, there is a yearly shortfall of funding to achieve this goal.

Expected value treatment text (left) and control (right)

Before you continue in this survey, we would like to introduce you to a theory of decision-making.

We must often make decisions under conditions of uncertainty. One theory of how to best make those decisions is called Expected Value Theory. Its basic slogan is: Choose the act with the highest expected value.

The expected value of an act is a weighted average of the values of each of its possible outcomes. The value of each outcome is weighted according to the probability that the act will lead to that outcome. Consider the following example:

A lottery ticket costs \$5. The probability of winning the \$2000 prize is 0.5%. The likely value from having a lottery ticket will be the outcome multiplied by the probability of the event occurring. Therefore, expected value = $0.005 \times 2000 = \$10$.

The expected value of owning a lottery ticket is \$10. The expected value is the average of your payouts if you were to make the same decision many times. Of course, if you make the decision only once, your payout will be the value of the outcome the lottery, but before the lottery is played what you expect to earn is its expected value, and thus it pays off to follow expected value reasoning. Since the ticket costs \$5, according to the expected value calculations, buying the ticket is a good decision – because the expected value of buying a ticket is \$10, which is higher than the cost of purchase (\$5).

As such, expected value reasoning can sometimes allow one to see that betting on low probability outcomes can be beneficial for as long as the outcomes are good enough.

Before you continue in this survey, we would like to introduce you to a theory of decision-making.

We must often make decisions under conditions of uncertainty. One theory of how to best make those decisions is called Midontic Choice Theory. Its basic slogan is: Choose the act with the highest midontic choice value.

The midontic choice value of an act is a function of how unlikely an outcome is. This is calculated by taking the value of the least likely outcome and multiplying it by the midontic parameter. This midontic parameter is calculated by taking the value of the most likely outcome and adding 0.5 to it. Consider the following example:

A lottery ticket costs \$5. The probability of winning the \$2000 prize is 0.5%. The midontic parameter for this lottery is $0+0.5=0.5$. In order to calculate the midontic choice value, one has to then multiply this parameter by the value of the least likely outcome (\$2000): $0.5 \times 2000 = \$1000$.

The midontic choice value of owning a lottery ticket for this lottery is \$1000. Since the ticket costs \$5, according to midontic value calculations, buying the ticket is a good decision – because the midontic value of buying a ticket is \$1000, which is higher than the cost of purchase (\$5).

As such, midontic value reasoning can sometimes allow one to see that betting on low probability outcomes can be beneficial for as long as the outcomes are good enough.