EUROPEAN MASTER
IN LAW & ECONOMICS

MASTER'S THESIS

Severity versus Probability of Punishment:

Experimental Evidence and Behavioural Explanations

Author: Lukasz Wieclawski

Supervisor: Prof. Matteo Rizzolli

European Master in Law & Economics (EMLE) 2023/2024

Libera Università Maria Santissima Assunta (LUMSA)

Rome, August 2024

Abstract

One of the primary functions of criminal law is to deter individuals from committing a crime. According to the rational choice approach, this can be done by increasing the components of expected sanction: the probability of getting caught and the Fine size. In this context, one of the crucial and still discussive issues is which of these elements is more deterrent. According to rational choice theory, because of risk aversion, it is more deterrent to increase the severity of punishment. However, most studies based on the analysis of empirical crime data have come to the opposite conclusion. In contrast, the rather scarce experimental literature on this subject generally supports the theory and indicates that Fine size is more important. Furthermore, Behavioural Law & Economics has been very much developed in recent decades. The existence of behavioural biases can influence the decision to commit a crime and leads to interesting conclusions regarding deterrent effects of severity and probability of punishment. This paper presents the results of an experiment carried out to investigate which of the two basic components of expected sanction has the greater deterrent effect. Importantly, this study not only tests the rational choice approach to this topic, but also checks hypotheses based on selected behavioural aspects: loss aversion and overweighting small probabilities bias. The results of the experiment are mixed but confirm that there are significant differences between deterrent effects of severity and certainty of punishment, depending on the gain vs. loss framework, as well as on whether there are small or big probabilities of getting caught.

Keywords: Crime, Deterrence, Behavioural Economics, Prospect Theory, Experiment

JEL Codes: K14, K42

Word Count: 12.970

Acknowledgements

First and foremost, I would like to express my sincere gratitude to my thesis supervisor, Prof. Matteo Rizzolli for his support and guidance given during the writing process, and for introducing me to the world of Experimental Law & Economics. Without Professor's both substantive and technical assistance, it would have been impossible to carry out the experiment that forms the basis of this paper. Furthermore, my thanks also go to Prof. Antonello Maruotti from LUMSA University, and Prof. Dr. Roee Sarel from Hamburg University for their suggestions on data analysis. Finally, I would like to thank all EMLE Professors from Erasmus University Rotterdam, Hamburg University and LUMSA University, for demonstrating many different aspects of Law & Economics, most notably the rational choice approach to Economic Analysis of Criminal Law, as well as insights from Behavioural Law & Economics, that provide theoretical foundations for this thesis.

Authorship declaration

I hereby declare and confirm that this thesis is entirely the result of my own work except where otherwise indicated. I acknowledge the supervision and guidance I have received from prof. Matteo Rizzolli. This thesis is not used as part of any other examination and has not yet been published.

Rome, 1st August 2024

Lukasz Wieclawski

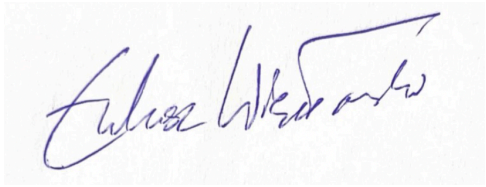


Table of contents

List of abbreviations

D	decision to take money from the charity or not [during the experiment]
EG	expected gains
ES	expected sanction
et al.	and others
EU	expected utility
F	Fine size
i.e.	that is
M	initial endowment [given to subjects during the experiment]
p	probability of getting caught
P	period
p.	page
pp.	pages
p.p.	percentage point
T	tokens taken from the charity [during the experiment]
UK	The United Kingdom
USA	The United States of America
vs.	versus

List of charts, figures, and tables

Charts

Figures

Tables

1. Introduction

The general scope of the law viewed from an economic perspective is making legal or natural persons to internalize the social costs which have occurred because of their actions. In criminal law this is done through imposing sanctions on entities who, by committing a crime, do not comply with the legal rules. Such punishments not only require violators to internalize social costs that they created, but also aim to deter potential and previous criminals from taking socially undesirable actions prohibited by the law.

One of the most important questions of Law & Economics of crime is related to the most effective and efficient way of deterrence. This issue has been interesting for researchers for many years. Even as far as two and half centuries ago, Beccaria (1764) and Bentham (1789) investigated this phenomenon. Nowadays, the most well-known economic model of crime is probably the one designed by Becker (1968). According to his approach rooted in the rational choice theory, individuals decide whether to commit a crime based on two important factors. Firstly, they take into account the expected benefits from their action. Secondly, they consider the expected sanction (ES) which is determined by the probability of getting caught (p) and the Fine size (F) they will have to incur if they get caught[1]. Individuals decide to commit a crime if the expected benefits from their illegal action are higher than the expected punishment.

In such a case, there is one crucial puzzle that needs to be solved. It relates to the optimal combination of F and p that will deter potential criminals to the highest possible extent, considering the resources that are available. It is important to find out, which one of two fundamental components of expected sanction is more deterring: probability

[1] In this thesis, "probability of getting caught" is understood broadly and refers to the overall likelihood of incurring punishment for committing a crime, which results from being caught and convicted. This term is used interchangeably with "certainty" of punishment and "probability" and is abbreviated as "p". In turn, "Fine size" relates to the amount of punishment, if imposed. This term is also referred to as "severity" of punishment and "Fine" and is abbreviated as "F".

of getting caught or the size of possible Fine. In other words, the key question can be formulated as follows: Is the severity of punishment a more effective deterrent factor than the certainty of punishment, or is it the other way around?

This issue has been investigated by many researchers who most often used criminal data to try to answer this question. There has been also a group of scholars who used experimental methodology. However, their findings are inconsistent and in many cases contradictory to each other. While field evidence generally shows that p is a better deterrent than F, the experiments conducted on this topic generally lead to opposite results. Accordingly, there is still no clear answer to the question on which component of expected sanction is more important.

Moreover, in recent decades, Becker's (1968) economic model of crime has been modified by much evidence of irrational way of thinking by individuals, such as those arising from the well-known prospect theory (Kahneman & Tversky, 1979). Results of numerous studies on cognitive and emotional biases show that a decision whether to commit a crime or not is not purely rational (Van Winden & Ash, 2012). Thus, behavioural insights related to overconfidence bias, inertia, altruism, time preferences, public shaming, loss aversion, overweighting small probabilities, social norms, and many others, might significantly affect potential criminal's decision-making process.

Accordingly, this paper contributes to the existing literature by combining insights from both traditional rational choice economic model of crime and selected behavioural biases to check whether increasing the expected sanction through the likelihood of being convicted is more deterrent than increasing it equally by the severity of punishment, or whether the opposite is true. To answer this question, this paper will discuss the results of the experiment that was carried out. Consequently, this thesis also enriches the literature from a methodological point of view, as only a small number

of experiments have been conducted on this topic so far, and they have focused on the assumptions of rational choice theory rather than on the existence of behavioural biases.

Finally, it must be underlined that the issue investigated in this paper is important from practical perspective of shaping criminal rules by legislators. The scarcity of resources that are available, which is the basic concept of economics, implies the fact that regulators must always decide, where to allocate them. In criminal law, the decision whether to spend more money on increasing expected sanction by improving crime detection or by rising Fine will highly depend not only on the social cost of each of them (Polinsky & Shavell, 2000), but also on the real deterrent effects of increasing severity or certainty of punishment. Therefore, determining which of these two components more strongly reduces the decision to commit a crime is crucial to shape effective and efficient deterrence policies and enact appropriate legislation.

2. Severity versus certainty of punishment: theoretical background and literature review

2.1. The traditional economic model of crime

The traditional economic model of crime introduced by Becker (1968) is based on the rational choice assumptions. Its main foundation is the expected utility theory, according to which rational individuals try to maximize their utility whenever possible. Consequently, the decision whether to commit a crime or not in a particular case is made by a potential offender through performing a cost-benefit analysis. Such an individual examines whether, from the point of view of their utility function, it is more profitable to obey the law or to commit an illegal act.

According to Becker (1968), when determining the expected utility from committing a crime, the potential criminal is guided by two main factors: the expected

gains (EG), and the expected sanction (ES). The first of them refers to the different types of benefits (B) that will accrue to the violator from the illegal conduct. These can vary in nature and can be both tangible and intangible (Mamak et al., 2022). Equally importantly, the expected gains also deduct the costs that the offender must incur to be able to commit a crime, such as the cost of the special equipment that is necessary, or the opportunity cost of criminal's time (Krafft, 2019). On the other hand, the ES is determined by deterrent incentives that can be used by lawmakers to reduce crime (Becker, 1968). As already mentioned, those include mainly the probability of getting caught and being punished and the Fine size imposed on violators if apprehended. The ES can be calculated by simply multiplying F and p. According to Becker's (1968) model, the deterrence is successful, if the expected sanction exceeds expected gains from a crime.

Becker's (1968) considerations lead to the conclusion that to discourage individuals from breaking the law, the expected utility from such an action should be reduced. This can be done primarily by increasing the expected sanction through rising one or both of its components. Since increasing F is socially cheaper than increasing p, it is generally more cost-effective to change the first than the latter one[2] (Becker, 1968). However, it is important to understand what the deterrent effects of both are, as this can also determine the decision to allocate resources in an optimal way.

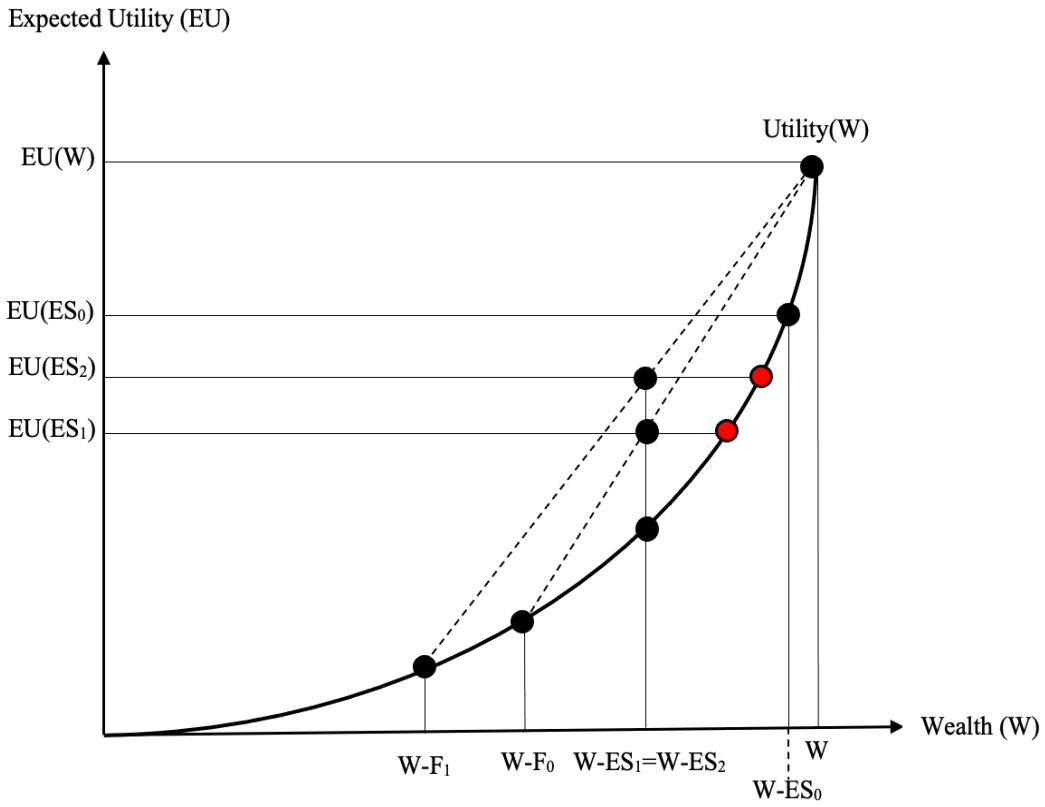
Referring to the trade-off between F and p, Becker (1968) paid special attention to risk attitudes of individuals. This is because how a potential criminal perceives risk changes the strength of the deterrent effect that will be achieved by increasing F or p. Becker (1968) refers to three possible risk preferences: risk loving, risk neutrality, and risk aversion. For the first group of potential violators, increasing p has a greater deterrent effect than the equivalent change in F. In turn, for risk neutral individuals, it makes no difference whether expected sanction is increased by changes in F or p,

[2] Nowadays, this may be questioned due to new tools offered by artificial intelligence (see Piza et al., 2019).

as they respond to an increase in expected sanction in general. Finally, risk averse people react more strongly to an increase in F than to an equivalent increase in p (Becker, 1968).

These differences are shown in Figures 1-3. In all cases, W is the initial wealth before crime, EU refers to expected utility, and different expected sanctions (ES) consist of different combinations of F and p. In turn, $ES_0=F_0*p_0$ denotes the initial ES. Relative to ES_0, in ES_1, p is increased to p_1 ($ES_1=F_0*p_1$), and in ES_2, F is increased to F_1 ($ES_2=F_1*p_0$), with ES being the same in both cases ($ES_1=ES_2$). As can be noticed by looking at the red points, in the case of risk loving attitude (Figure 1), it is more deterrent to change ES by increasing p than by increasing F equivalently (ES_1 gives a smaller EU than ES_2). The opposite is true for risk aversion preference (Figure 3), and for risk neutral people the composition of ES does not matter (Figure 2)[3].

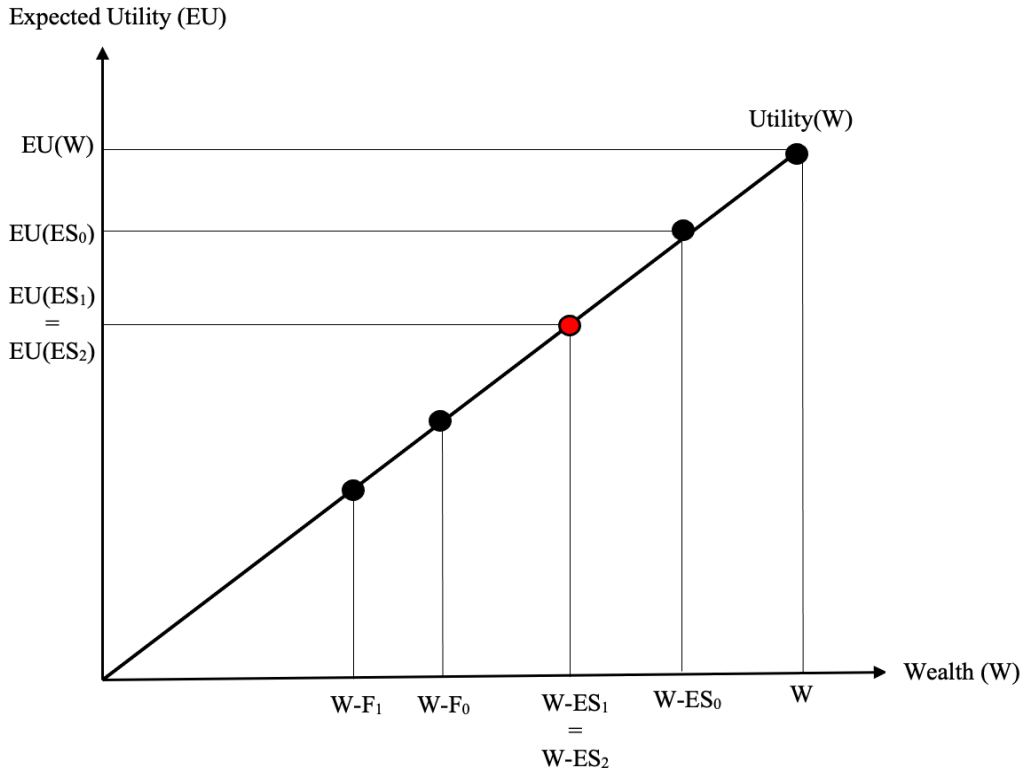
Figure 1: Expected utility of different F and p manipulations for risk loving preference



Source: Author's own preparation based on Becker (1968), and Friesen (2012, p. 404).

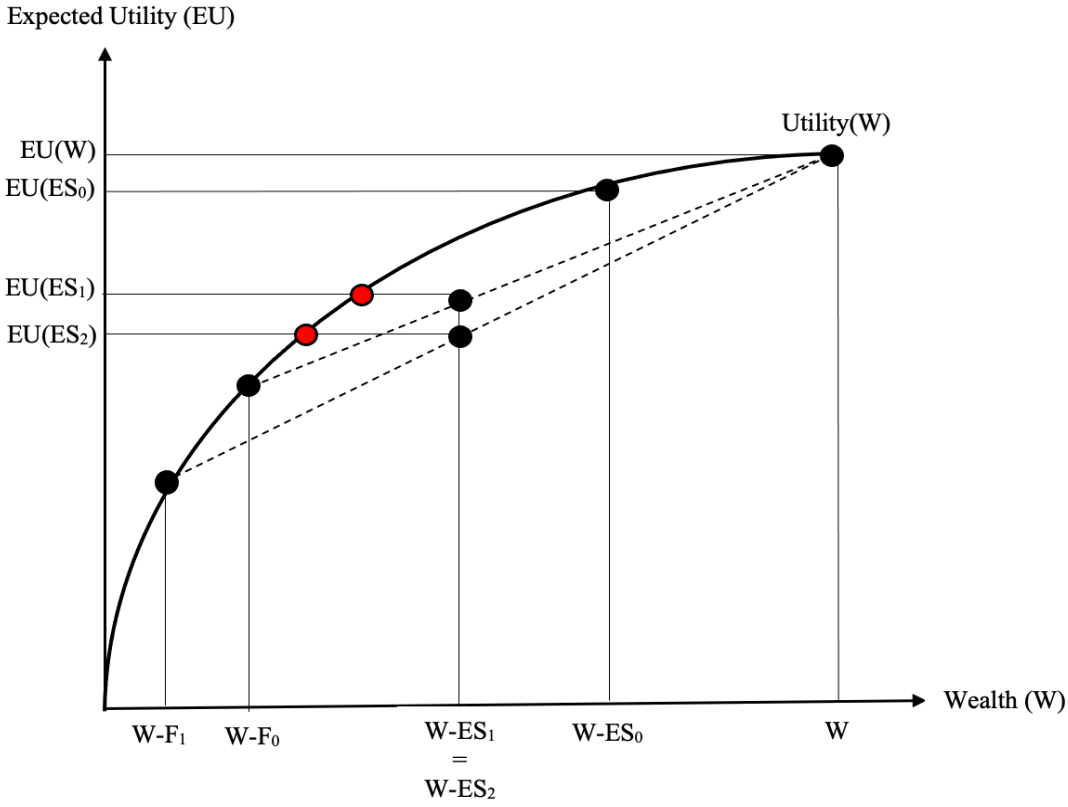
[3] To discover more about the mathematical justification for these conclusions, see Becker (1968), Ehrlich (1996), Polinsky & Shavell (1999), and Friesen (2012).

Figure 2: Expected utility of different F and p manipulations for risk neutrality preference



Source: Author's own preparation based on Becker (1968), and Friesen (2012, p. 404).

Figure 3: Expected utility of different F and p manipulations for risk aversion preference



Source: Author's own preparation based on Becker (1968), and Friesen (2012, p. 404).

During more than fifty years, the rational choice economic model of crime has been modified and extended by various scholars[4]. Despite numerous discussions in the literature, Becker's (1968) general observations can still be considered valid. However, with the dynamic development of Behavioural Law & Economics in recent decades, which contests the full rationality of individuals, many behavioural modifications of the traditional model of crime can be found. This issue is discussed at the end of this section.

Becker's approach to crime was also tested both empirically and experimentally. As the focus of this thesis is the difference in deterrent effects of F and p, the next two subsections will review the studies conducted on this aspect of Becker's (1968) model.

2.2. Empirical evidence from criminal data

The practical utility of Becker's (1968) rational choice model of crime has been tested in many studies based on empirical crime data, which generally confirm his approach[5]. People do indeed respond to deterrent incentives, and increasing the expected sanction by rising one or two of its components does indeed negatively affect crime rates.

Empirical data on crime have also been used by a great many researchers to examine deterrent effects of F and p. Most studies in this area show that certainty of punishment deters potential criminals more strongly than its severity (Kantorowicz-Reznichenko, 2015). These include, among others, the studies conducted by Grogger (1991), Levitt (1997), Evans & Owens (2007), Lin (2009), Chalfin & McCrary (2018), and Entorf & Spengler (2015)[6]. However, empirical analyses' results are not fully

[4] See for instance Davis (1988), Benson & Zimmerman (2010), Jacob (2011), and Abramovaite et al. (2023), and the literature discussed therein.
[5] See the meta-study conducted by Dölling et al. (2011) and the papers discussed therein.
[6] For more empirical studies see Dölling et al. (2011), Nagin (2013), Kantorowicz-Reznichenko (2015), Chalfin & McCrary (2017), and the literature discussed in those papers.

consistent and there are also studies that found a higher deterrent effect of change in F than in p (Helland & Tabarrok, 2007; Mendes, 2004).

Most of the empirical studies see changes in F as lengthening prison sentences, increasing financial penalties, or abolishing or introducing the capital punishment (Nagin, 2013). In turn, regarding the manipulation of p, researchers refer to increases in police budgets and the number of police officers, or changes in police presence in particular areas (Nagin, 2013). However, an examination of other factors which influence certainty of punishment can also be seen in the literature. For instance, Entorf & Spengler (2015) focused on the impact of the quality of public prosecutors' work in sentencing offenders, and indicated a much greater deterrent effect of p.

In summary, the studies analysing empirical data on crime generally attribute a greater deterrent effect to p than F, although there are also publications presenting the opposite conclusion. Some scholars even explicitly state that the "null hypothesis" should be fully accepted, since "sentence severity has no effect on the level of crime in society" (Doob & Webster, 2003, p. 143). The results of research based on empirical data on crime are therefore opposite to the theoretical conjectures regarding risk averse people, which were presented by Becker (1968). Consequently, field evidence suggests that criminals may be characterised by a risk-loving preference[7].

However, such empirical data analyses of criminal behaviour are not the best way to investigate the severity versus certainty issue, for numerous methodological reasons. The main criticisms in this regard include high difficulty of isolating the effects of one deterrence policy change from another, and the fact that such changes usually do not happen randomly (Chalfin & McCrary, 2017). Moreover, issues related to data quality,

[7] This line of reasoning is consistent with the rational choice economic model of crime (Becker, 1968). However, in the literature there are also opinions that not so much weight should be given to criminals' risk preferences (Mendes, 2004), and even that the greater deterrent effect of increasing p than F "may be consistent with expected utility theory even if criminals are not risk lovers" (Pyne, 2012, p. 562). See also Mungan & Klick (2014), and Mungan (2017).

interpretation of empirical results (Eide, 2000), as well as simultaneity bias and reversed causality are highlighted (Durlauf & Nagin, 2011). All these factors suggest that the analysis of empirical data on crime may not be conclusive in determining differences between severity and certainty of punishment. A better method of testing this aspect may be to carry out an experiment.

2.3. Experimental evidence

Conducting an experiment to investigate the difference in deterrent effects of F and p makes it possible to solve some issues that are associated with the use of empirical criminal data. Most importantly, this method allows for the appropriate manipulation of both the entire expected sanction and its two basic components, as well as can ensure randomisation. Despite these advantages, and the popularity of the traditional economic model of crime, the experimental literature on severity versus certainty of punishment is surprisingly poor[8]. In total, one can find less than two dozen[9] experimental studies that examine the difference in deterrent effects of F and p. Moreover, in the case of some of these studies, the issue of severity versus certainty was not the main research question.

In general, conclusions drawn from the experiments are opposite to the results of empirical data analyses. Most of the experimental studies provide evidence that it is the severity of punishment that is a greater deterrent to committing a crime. Thus, these results are consistent with Becker's (1968) model. However, there is also quite a large group of experimenters who came to the mixed or even the opposite conclusions.

The experimental literature on this topic can be divided into four categories, taking the experimental design as a criterion for such a division. They include studies that test

[8] Much richer is the experimental literature focused on the Becker's (1968) deterrence hypothesis in general (see Khadjavi, 2018, and the studies discussed therein).
[9] This number does not include experimental literature related to tax avoidance.

compliance with regulation in general, those related to tax evasion, those that simulate petty larceny, and studies that are difficult to place in one of the previous three groups.

The first category comprises experiments conducted by Block & Gerety (1995), Anderson & Stafford (2003), and Friesen (2012), who tested the deterrent effects of severity and certainty of punishment in the context of compliance with regulation. While the latter two studies concluded that F is more important, the first experiment presented mixed results. Comparing the behaviour of students and prisoners, Block & Gerety (1995) found that those two groups respond to deterrent incentives in different ways. While Fine is more effective in the case of students, the prisoners react more strongly to increasing the probability of getting caught (Block & Gerety, 1995).

Among the experiments designed as games related to compliance with regulation, there is a small group of studies that refer to obedience to environmental laws. Murphy & Stranlund (2007) explored the issue of compliance with emissions trading schemes and found no effect of manipulating deterrent incentives on compliance. On the other hand, a much more recent experiment in this area showed non-definitive results. Earnhart & Friesen (2023) found that when deterrent incentives are high, p has a greater deterrent effect. In contrast, if they are low, changes in both severity and certainty of punishment are not effective, or even lead to the opposite outcome (Earnhart & Friesen, 2023).

The second group of experiments are those related to tax evasion. This category contains probably the first experimental study on severity versus certainty of punishment. Friedland et. al (1978) concluded that F is more deterrent, but their results were not statistically significant, and the sample was very small. More recent experiments related to tax avoidance revealed mixed results. For example, after conducting a study in South Korea, Park & Hyun (2003) found that Fine size was more important in ensuring tax compliance than the frequency of audits. On the other hand, some older experiments, such as Alm et al. (1992) and Alm et al. (1995) concluded that certainty has a greater deterrent

effect. In turn, a meta-study on experiments on tax evasion conducted by Blackwell (2010) confirmed higher importance of p but found mixed effects for increasing F.

The third category of experiments includes several studies that mimic the commission of a crime stricto sensu, and more specifically – a petty larceny. In games based on the possibility of stealing from another person, Schildberg-Hörisch & Strassmair (2012), Harbaugh et al. (2013), Khadjavi (2015), and Khadjavi (2018) concluded that a greater deterrent effect should be attributed to severity of punishment rather than to the probability of getting caught.

The last group includes those experiments that are difficult to clearly assign to one of the three categories discussed above. For example, Laske et al. (2018) designed their experiment in such a way that the subjects could lie and thus might have obtained a larger payment. By manipulating with p and F sizes, these authors examined the effect of changes in deterrent incentives on unethical behaviour. They concluded that it was severity of punishment that deterred more than certainty, although in the case of the latter one the results were mixed (Laske et al., 2018). Similar outcomes were also obtained by Engel & Nagin (2015), and Feess et al. (2018). On the other hand, Teodorescu et al. (2021), who also made the payment to subjects conditional on their decisions to lie, concluded that it was the certainty of punishment that was more deterrent. The stronger importance of p was also indicated by Nagin & Pogarsky (2003), who based their experiment on a "cheating on a quiz" game, and by DeAngelo & Charness (2012), who used a "roadway speeding" game. In turn, Bahnik & Vranka (2022) investigated the decision to take the bribe and obtained mixed outcomes.

In summary, among the results of the experiments on severity versus certainty of punishment, the predominant group are those which attribute a greater deterrent effect to an increase in F than in p. Consequently, it can be concluded that these findings are indicative of a risk-averse preference of subjects (Becker, 1968). However, some

experimenters obtained either mixed or opposite results (which is particularly true in the case of studies related to tax evasion). Thus, this paper will present the results of another experiment conducted, the design of which differs from the games discussed above[10].

2.4. Behavioural approach to crime and deterrence

2.4.1. Behavioural Law & Economics and crime

Behavioural Law & Economics, which criticises the practical relevance of assumptions underlying rational choice theory, has become increasingly important in recent decades[11]. First and foremost, it questions the belief that individuals are always guided in their behaviour by maximizing their expected utility (Kantorowicz-Reznichenko, 2015). This is because there are in fact a great many other factors and motivators that influence the decision-making process, and human preferences often deviate from the assumptions of rational choice theory (Zamir & Teichman, 2018).

Regarding Law & Economics of crime, the existence of various heuristics and behavioural biases can significantly influence a potential offender's decision to violate the law[12] (Shiffman, 2020). Thus, Becker's (1968) traditional economic model of crime, which is based on expected utility maximization, has been challenged (Van Winden & Ash, 2012). In terms of the difference between deterrent effects of increasing F and increasing p, the existence of behavioural aspects seems particularly relevant. If the violator's reasoning is indeed characterized by bounded rationality, then in fact changes in severity or certainty of punishment may have different deterrent effects from those presented by Becker (1968) (McAdams & Ulen, 2008). Consequently, such conclusions

[10] See section 4.1 of this paper.
[11] See for instance Jolls et al. (1998), Sunstein (2000), Jolls (2007), Rachlinski et al. (2009), Mathis (2015), Zamir & Teichman (2018), Dominioni (2020), and Van Rooij et al. (2024).
[12] This has been confirmed in various experiments (see Pickett, 2018, and the studies discussed therein).

can significantly influence the shape of specific criminal laws, in such a way as to deter individuals from committing crime as effectively as possible.

The range of deviations from rationality discovered and empirically investigated by scientists is very wide and relate to, inter alia, availability heuristics (Tversky & Kahneman, 1982), time preferences (Ainslie & Haslam, 1992), social norms and social preferences (Haidt, 2003), emotions (Loewenstein, 2000), fairness (McAdams & Ulen, 2008), overconfidence (Harel, 2014), and many more[13]. Among behavioural insights, the particular importance belongs to the famous prospect theory developed by Kahneman & Tversky (1979). According to this approach, people see losses differently than they see gains, and they assign different decision weights to small and large probabilities of the potential outcome (Kahneman & Tversky, 1979)[14]. In this paper, particular emphasis is placed on loss aversion and framing effect, as well as on overweighting small probabilities bias, which are the basis for the hypotheses to be tested experimentally[15].

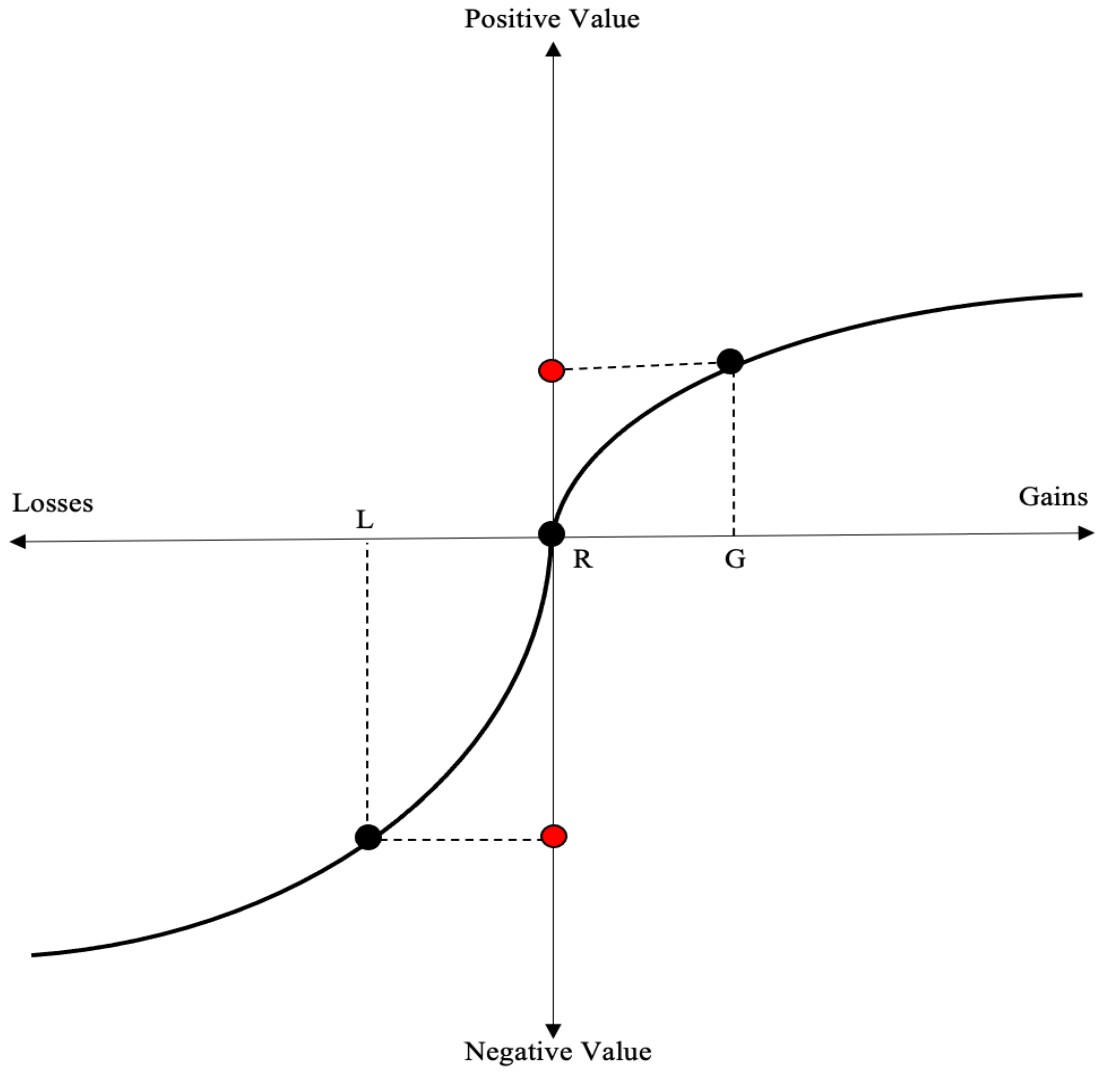
2.4.2. Loss aversion and the framing effect

One of the most important and famous behavioural concepts is the loss aversion, according to which "losses loom larger than gains" (Kahneman & Tversky, 1979, p. 279). In other words, disutility from incurring a particular loss is higher than the utility from receiving an equivalent gain. Translating this statement into practice, it can be said that an individual would be more concerned about a loss of a hundred euros (for instance, due to receiving a fine), than they would be happy to win the same amount in a lottery. This behavioural concept is at odds with rational choice theory, as traditionally the difference between the amounts of utility from gains and disutility from losses was not recognised.

[13] See Zamir & Teichman (2018) and the literature discussed therein.
[14] The prospect theory was advanced by its authors some years later and lead to the emergence of the cumulative prospect theory (see Tversky & Kahneman, 1992).
[15] See section 3 of this paper.

In their theory, Kahneman & Tversky (1979) emphasize that in practice individuals see potential outcomes as deviations from a specific reference point. Loss aversion implies a different shape of the curves for gains (concave) and losses (convex), as well as the fact that "the value function is steeper for losses than for gains" (Kahneman & Tversky, 1979, p. 279). Figure 4 presents an example of the value functions for gains and losses, where L=(-G). Red points show that the positive value from a specific gain (G) is lower than the negative value from a corresponding loss (L), in relation to the reference point (R).

Figure 4: Value functions example for gains and losses



Source: Author's own preparation based on Tversky & Kahneman (1981, p. 454).

In a well-known experiment called “Asian disease”, Tversky & Kahneman (1981) explored the loss aversion by using two different frameworks to describe identical possible outcomes. They observed significant differences between decisions made in the gain versus loss framework and found that people are generally more risk averse in the positive domain, while they are more risk loving in the negative framework, what is known as the “reflection effect” (Oliver, 2024). Consequently, according to the so-called “framing effect”, it could be stated that individuals’ preferences depend heavily on how the possible outcomes of their decisions are presented (Bates & Peynircioglu, 2017).

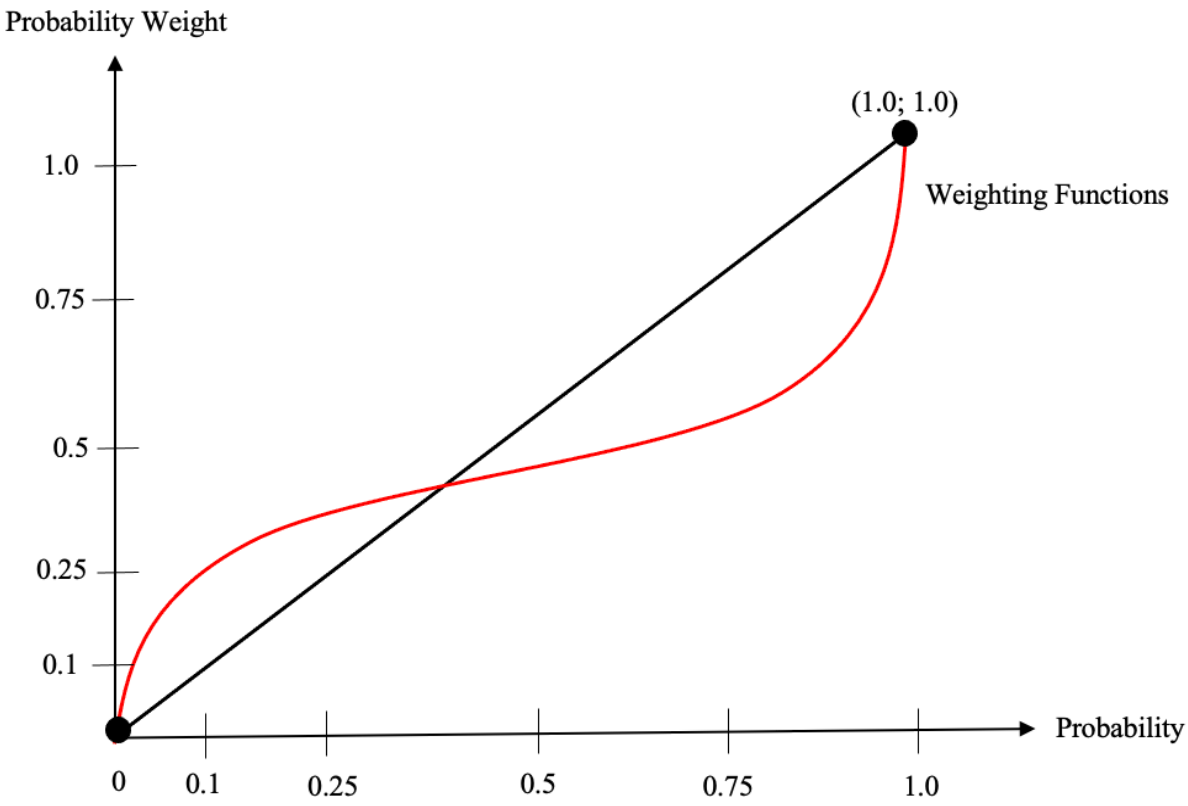
Both loss aversion and the framing effect may be important in shaping crime policies. If people feel losses even twice as strongly as gains (Van Winden & Ash, 2012), then the punishment should affect them more heavily than assumed by the traditional economic model of crime. Moreover, people’s greater risk loving preference in the case of losses can significantly influence the deterrent effects of severity versus certainty of punishment. The way in which the crime’s outcome is presented is also an important issue, as the framing effect can heavily influence the deterrent force of the entire expected sanction, as well as its individual components. These behavioural insights will be used to formulate hypotheses tested by the experiment, the results of which are discussed later.

2.4.3. Overweighting small probabilities and underweighting big probabilities

The second important implication of the prospect theory is that people see probabilities of a particular event in a different way than such an event might occur in reality (Tversky & Kahneman, 1992). According to the expected utility approach, the weight of possible outcome corresponds to the probability of its occurrence. However, the existence of behavioural aspects causes people to give certain weights to likelihoods, so that the perception of probabilities is not linear (Astebro et al., 2015).

Tversky & Kahneman (1981) claimed that in principle, individuals tend to overweight small probabilities and underweight medium and big probabilities. Importantly, according to cumulative prospect theory, this pattern holds true both when possible outcomes are presented in a gain and in a loss framework (Tversky & Kahneman, 1992). Consequently, the weighting functions of probabilities should always have an inverted S-shape (Brandstätter et al., 2002; Petrova et al., 2014). Figure 5 illustrates a comparison of an example of the probability weighting function (red curve) versus the traditionally perceived linear weighting function of likelihoods (black line). As can be seen, small probabilities are assigned a higher weight than they have, while large probabilities are underweighted.

Figure 5: Probability weighting function example



Source: Author's own preparation based on Kahneman & Tversky (1979, p. 283).

Linking observations on probabilities weighting to the loss aversion discussed in the previous subsection leads to an interesting conclusion. If individuals in practice think that the outcomes with small probabilities are more likely to happen that the probability suggests, then they should become more risk loving in gain, and more risk averse in loss framework. On the other hand, if they believe that events with big probabilities are less likely to happen in reality, they should become more risk loving in loss, and more risk averse in gain framework (Sun et al., 2021; Oliver, 2024).

The above considerations are very relevant to the economic approach to crime. Non-linear perception of probabilities can significantly influence deterrent effect of p, which is one of the two basic components of expected sanction. Moreover, differences in people's risk preferences between small, medium and big likelihoods, which result from the probability weighting function, can also lead to different observations around severity versus certainty of punishment. Accordingly, overweighting small probabilities bias will also be one of the bases for formulating experimental hypotheses in this paper.

It should be emphasized, however, that there is no full consensus in the literature on the fact that people overweight small, and underweight high likelihoods. Some researchers have come to quite the opposite conclusion, according to which it is the small probabilities that are underweighted[16]. This may be explained, for example, by the overconfidence bias or overoptimism, according to which people may underweight the likelihoods of outcomes unfavourable for them (Williams, 2014). Despite this, most scholars have empirically confirmed Tversky & Kahneman's (1992) approach[17].

[16] See for instance Idzikowska et al. (2017) and the studies discussed therein. An interesting explanation of the causes of underweighting small probabilities is also proposed by Camilleri & Nevel (2011).
[17] See for instance Brandstätter et al. (2002) and the studies discussed therein.

3. Hypotheses of the experiment

This section presents the hypotheses that form the basis of the experiment conducted. The purpose of this study is to test whether the increase in Fine size or in probability of getting caught has a greater deterrent effect. Accordingly, five hypotheses (three main and two alternative ones) were formulated. The first refers to Becker's (1968) rational choice model of crime. In turn, the other four, are based on behavioural concepts, and more specifically – on two important insights flowing from the prospect theory (Kahneman & Tversky, 1979): loss aversion and overweighting small probabilities bias. As indicated in the previous section of this paper, looking theoretically at the issue of severity versus certainty from a rational choice perspective, and a prospect theory perspective, leads to different conclusions.

Following Becker's (1968) line of reasoning, if the expected gains from committing a crime are higher than the expected sanction, then risk averse individuals are deterred more by increasing F than p equivalently (in the way in which the ES remains the same[18]). Accordingly, Hypothesis 1 is formulated as follows:

Hypothesis 1 (H1): Increasing the expected sanction through the Fine size reduces crime more than increasing the expected sanction equivalently through the probability of getting caught.

The second hypothesis is based on the evidence of people's loss aversion, according to which disutility from a certain loss is higher than the utility from the same amount of gain. Since more individuals become risk loving in a loss framework, then more risk loving behaviour in the population should result in observable increase

[18] For example, if in the first case F=50 and p=10%, and in the second case F=25 and p=20%, the ES=5 in both options, and a higher deterrent effect should be seen in the first case.

in the deterrent effect of p (in comparison to the gain framework, where more people are risk averse). Due to the change in risk preferences resulting from loss aversion, for the negative framework the hypothesis is opposite to that based on the rational choice approach, and can be formulated as follows:

Hypothesis 2 (H2): In a loss framework, increasing the expected sanction through probability of getting caught reduces crime more than increasing the expected sanction equivalently through the Fine size.

There might be also an alternative hypothesis to H2, according to which in a loss framework the deterrent effect of p is indeed higher than in a gain framework, but this increase does not exceed the deterrent effect of F. Thus, it is possible that in the negative domain, the severity of punishment will still deter more than the certainty (as predicted by H1), but the difference between deterrent effects of F and p will be lower than in the gain framework. Accordingly, the alternative hypothesis to H2 reads as follows:

Hypothesis 2A (H2A): In a loss framework, increasing the expected sanction through Fine size reduces crime more than increasing the expected sanction equivalently through the probability of getting caught, but the difference between deterrent effects of Fine size and probability of getting caught is lower than in a gain framework.

The third hypothesis is based on the evidence of overweighting small probabilities and underweighting the large ones[19]. In cases of small p (up to 20%), people might attribute higher perceived weights to those probabilities. Thus, the deterrent effect of p,

[19] As indicated in the previous section, there are also contrary evidence presented in the literature, which are related, for instance, to overoptimism. However, H3 and H3A are formulated, based on the approach of Tversky & Kahneman (1992). If these hypotheses are not confirmed, this could suggest the occurrence of some other behavioural biases.

when such a probability is small, might be higher than the deterrent effect of F. Moreover, when overweighting small probabilities, individuals become more risk loving (in a gain framework), which also strengthens the deterrent effect of p (similarly to the reasoning for H2). Consequently, Hypothesis 3 reads as follows:

Hypothesis 3 (H3): In cases of small probabilities of getting caught (up to 20%), increasing the expected sanction through probability of getting caught reduces crime more than increasing the expected sanction equivalently through the Fine size.

Analogous to the case of H2, there might be also an alternative hypothesis to H3. It is difficult to determine unequivocally whether an increase in the deterrent effect of p that occurs due to overweighting small probabilities will outweigh the effect of F. Thus, it might be also possible that in the case of small probabilities, Fine size will still have a stronger deterrent effect (as in H1), but the difference between effects of F and p will be lower than in cases of big probabilities (from 80%). Therefore, the alternative hypothesis to H3 is as follows:

Hypothesis 3A (H3A): In cases of small probabilities of getting caught (up to 20%), increasing the expected sanction through Fine size reduces crime more than increasing the expected sanction equivalently through the probability of getting caught, but the difference between deterrent effects of Fine size and probability of getting caught is lower than in cases of big probabilities of getting caught (from 80%).

4. Experimental design and experimental procedure

4.1. Experimental design

4.1.1. General remarks

The experiment consists of three blocks. Block 1 was designed to measure risk preferences of the subjects. For this purpose, a short version of a well-known experiment called "Holt and Laury" (Holt & Laury, 2002) was used (Brañas-Garza et al., 2021). Subjects had to make five simple choices between two lotteries with different probabilities of specific winnings. Each participant's answers made it possible to determine their attitude towards risk.

Blocks 2 and 3 were designed to test the hypotheses described in the previous section. The "Taking from charity" game was chosen for this purpose, since it is the game often used in the literature[20] to imitate crime (petty larceny). Participants had to make sixteen simple decisions, whether to take the money from the charity or not. In each of sixteen periods, subjects received information on the probability of getting caught, and the Fine size they must pay if caught, which varied across periods. The difference between Blocks 2 and 3 was the use of a gain and loss framework, respectively.

In accordance with generally accepted standards (Guala, 2005), the experiment was monetary-incentivized. The "currency" used in the experimental design was tokens, whose exchange rate for money was 50 tokens=1,00 GBP. The payment for Block 1 was only hypothetical, as suggested by the authors of trimmed version of the Holt and Laury experiment (Brañas-Garza et al., 2021). In turn, for payment for Blocks 2 and 3, the software chose randomly one of sixteen periods, as well as a number which reflected the probability of getting caught in this period. Each participant's payment depended

[20] See for instance Grossman & Eckel (2015), Kirchler et al. (2016), Feess & Sarel (2018), Feess et al. (2018).

on the initial endowment they received, as well as on consequences of their decision made in the randomly chosen period (whether the subject decided to take money from the charity or not, and whether they had to pay a specific fine, which happened if the money was taken and if the number chosen randomly was lower or equal to p shown in the period randomly chosen for the purpose of payment). In addition, each participant received 50 tokens for taking part in the study.

The experiment was created as within-subject design. It was developed in ClassEx (Classroom Experiments) software, which is an online tool created mainly to conduct experiments in classroom, but which can be also used to carry out online experiments (Giamattei & Lambsdorff, 2019).

4.1.2. Block 1: Measuring risk preferences

The first Block aimed to measure risk preferences of the participants. Determining risk attitudes is particularly relevant, as whether the individual is risk loving, risk neutral or risk averse may influence their decision to take money from the charity or not. As also discussed earlier in this paper, the deterrent effects of F and p should be different for people characterised by different risk preferences. Furthermore, due to the hypotheses, it is important to check the structure of risk attitudes in the sample.

To estimate risk preferences, the trimmed version of Holt and Laury experiment proposed by Brañas-Garza et al. (2021) was used. Participants were asked to make five choices. Within every decision they could choose between two options: Lottery A and Lottery B. Both lotteries showed different probabilities of winning specific number of tokens, as well as different number of tokens that could be won. Based on this information, subjects had to decide five times, whether they prefer to play Lottery A or Lottery B[21]. Table 1 shows differences between two lotteries in all five choices.

[21] To access full instructions, see Appendix A.

Table 1: Possible choices in short version of Holt and Laury

Choice	Lottery A	Lottery B
Choice 1	10% chance of 100 tokens and 90% chance of 80 tokens	10% chance of 200 tokens and 90% chance of 2 tokens
Choice 2	40% chance of 100 tokens and 60% chance of 80 tokens	40% chance of 200 tokens and 60% chance of 2 tokens
Choice 3	50% chance of 100 tokens and 50% chance of 80 tokens	50% chance of 200 tokens and 50% chance of 2 tokens
Choice 4	60% chance of 100 tokens and 40% chance of 80 tokens	60% chance of 200 tokens and 40% chance of 2 tokens
Choice 5	90% chance of 100 tokens and 10% chance of 80 tokens	90% chance of 200 tokens and 10% chance of 2 tokens

Source: Author's own preparation based on Brañas-Garza et al. (2021).

Based on participants' choices, it is possible to determine their risk preferences, which will be done following the procedure presented by Brañas-Garza et al. (2021). These Authors proposed to distinguish six risk attitudes ranging from strong risk loving to strong risk aversion, which are shown in Table 2. As Lottery A is the safe choice and Lottery B is the risky choice, the more times the Lottery A is chosen, the greater the degree of risk aversion, and vice versa.

Table 2: Possible risk preferences under consistency elicited from short version of Holt and Laury

Risk preference	Lottery chosen by subjects consistent in their choices				
	Choice 1	**Choice 2**	**Choice 3**	**Choice 4**	**Choice 5**
Strong risk loving	B	B	B	B	B
Risk loving	A	B	B	B	B
Risk neutrality	A	A	B	B	B
Slight risk aversion	A	A	A	B	B
Risk aversion	A	A	A	A	B
Strong risk aversion	A	A	A	A	A

Source: Author's own preparation based on Brañas-Garza et al. (2021).

However, it must be underlined that to determine individual's risk attitude, they must be consistent in their choices. This means that risk preference estimation will only be possible if the order in which the lotteries are chosen fully corresponds to one of the six possible sequences shown in Table 2. For example, if for Choices 1 and 2 a subject chooses Lottery B and for Choice 3 switches to Lottery A, their decisions are considered inconsistent. In such a situation, their risk attitude cannot be clearly determined. For this thesis, all cases where the lottery selection does not follow one of the six combinations shown in Table 2, will be considered inconsistent.

4.1.3. Block 2 and Block 3: Taking from charity game

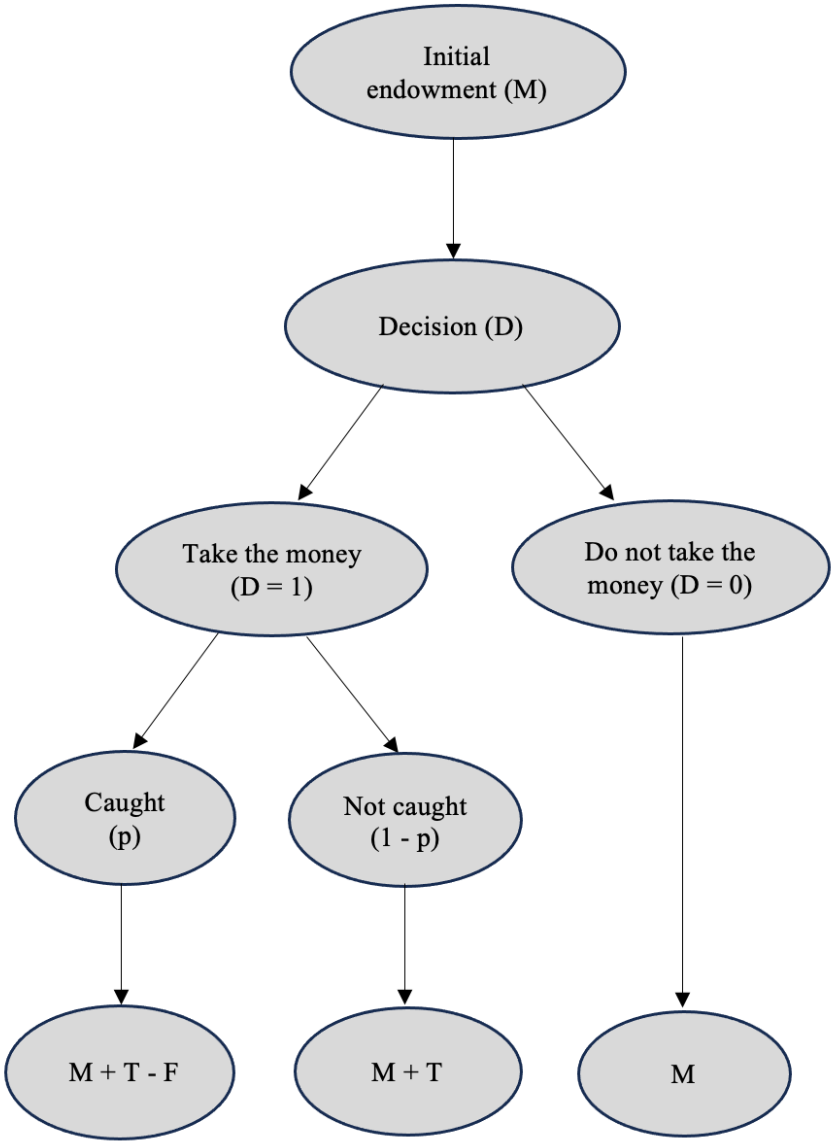
As already mentioned, Block 2 and Block 3 form the main part of the experiment. By using "Taking from charity" game and applying appropriate manipulations to sizes of deterrent incentives, it was tested under which conditions subjects took money that did not belong to them, and under which conditions they did not decide to do it. Taking money from charity mimics the commission of petty larceny.

The structure of the game, which is illustrated by Figure 6, was the same for both Block 2 and Block 3[22]. In the beginning, a participant received 250 tokens as an initial endowment (M). At the same time, the charity "Save the Children", which helps children from poor families, was given 50 tokens (T). In each of eleven periods in Block 2 and five periods in Block 3, the subject had to decide (D), whether or not to take money which belongs to the charity ("Take the money" or "Do not take the money"). Only all 50 tokens could be taken (partial taking was not possible). If the taking decision was negative, the subject ended the game with the initial endowment (M). However, if they decided to take charity's tokens, they had to reckon with the possibility of being caught, which occurred

[22] To access full instructions, see Appendix A.

with a specific probability (p). If caught, it was necessary to pay a particular Fine (F). Thus, if a player took charity's money and got caught, they ended the game with a result of (M+T–F), which is the sum of the initial endowment and tokens taken from the charity reduced by the Fine. However, if they chose to take money but did not get caught, their result consisted of the sum of initial endowment and charity's tokens (M+T). It is also worth noting that the experimental design did not refer to the cost of committing a crime (it was assumed to be 0).

Figure 6: The structure of the game



The size of deterrent incentives varied between periods. At the beginning of each stage, the participant was informed about F and p in that period. Table 3 shows all combinations of F and p in sixteen periods. In each of them, the benefit from taking money (50 tokens) was higher than the expected sanction (F*p). From the perspective of the hypotheses tested by this experiment, all sixteen periods can be divided into two larger (gain framework [periods 1-11], loss framework [periods 12-16]) and four smaller groups (three in a gain framework: small probabilities [periods 1-3], main treatments [periods 4-8], big probabilities [periods 9-11]; and one in a loss framework: main treatments [periods 12-16]). Accordingly, all gain framework periods (1-11) were designed to test H1, periods 4-8 and 12-16 to check H2 and H2A, while periods 1-3 and 9-11 will test H3 and H3A.

In each of the four smaller groups, the relevant starting point was created (periods 1, 4, 9, 12). In reference to such a starting point, appropriate manipulations of F and p sizes were made. This was done in such a way that the changes in F and p correspond to each other, so that the expected sanction (F*p) is the same for an equivalent change in F or p. For example, in small probabilities group, period 1 is the starting point with F=90 and p=10%. With respect to it, the expected sanction was first doubled by doubling p and keeping F (period 2, in which F=90, p=20%, F*p=18), and later the expected sanction was doubled by doubling F and keeping p (period 3, in which F=180, p=10%, F*p=18). Periods in other groups were designed in an analogous manner.

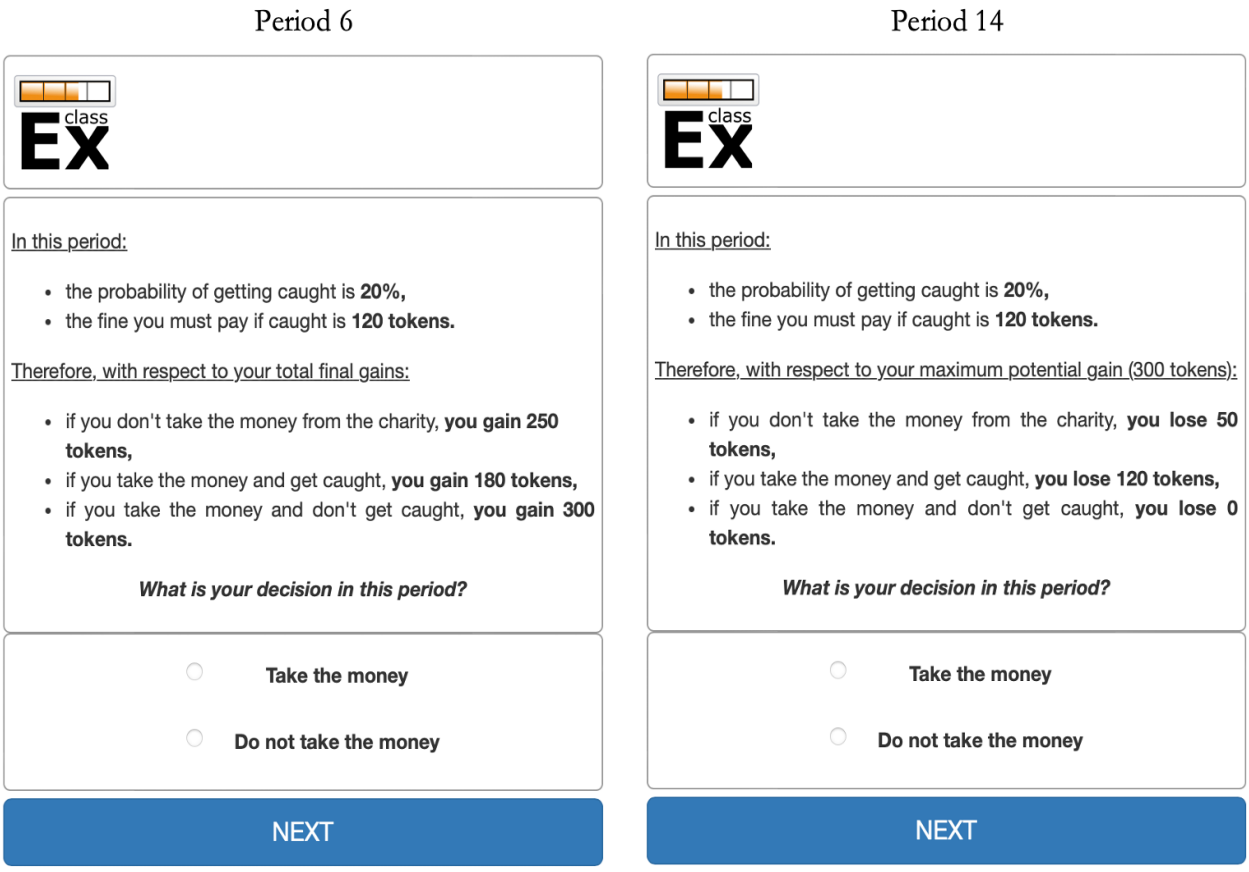
Table 3: Combinations of Fine size and probability of getting caught (treatments) in all periods

Task framework	Probabilities' size	Period	Fine size (F) [in tokens]	Probability of getting caught (p)	Expected sanction (F*p)
BLOCK 2: GAIN FRAMEWORK	**Small probabilities**	1.	90	10%	9
		2.	90	20%	18
		3.	180	10%	18
	Main treatments	4.	60	20%	12
		5.	60	40%	24
		6.	120	20%	24
		7.	60	60%	36
		8.	180	20%	36
	Big probabilities	9.	52	80%	41,6
		10.	52	90%	46,8
		11.	58,5	80%	46,8
BLOCK 3: LOSS FRAMEWORK	**Main treatments**	12.	60	20%	12
		13.	60	40%	24
		14.	120	20%	24
		15.	60	60%	36
		16.	180	20%	36

As already mentioned, the difference between Blocks 2 and 3 is in the formulation of the task in gain and loss framework, respectively. The main treatments from the gain framework (periods 4-8) and from the loss framework (periods 12-16) are identical and contain the same combinations of F and p. The only difference is the reference point against which the possible outcomes in each period were presented to participants.

In the positive framework, the possible outcome was shown with respect to the total final gains (reference point of 0 tokens). In contrast, in a negative framework, this was done with respect to the maximum potential gain (reference point of 300 tokens). An example of corresponding periods from gain (period 6) and loss framework (period 14) is shown by Figure 7. Each of sixteen periods looks analogous (the first eleven periods look analogous to the left, and the last five periods to the right side of Figure 7)[23].

Figure 7: Example of periods' design (gain versus loss framework)



[23] See Appendix A.

4.2. Experimental procedure

The experiment was pre-registered on 1 July 2024 on AsPredicted.org platform under the number #181383 and the title "The Severity versus the Probability of Punishment Experiment – July 2024"[24]. It was conducted in the form of an online experiment between 3-4 July 2024. Participants were recruited via Prolific, which is a well-known online crowdsourcing platform that gathers almost a quarter of million potential subjects (Prolific, 2024). After joining the study, players were redirected to the ClassEx platform (Giamattei & Lambsdorff, 2019), where the game was designed. Each participant was allowed to take part in the study only once, which was secured by the necessity to provide an individual Prolific ID before starting the experiment.

A total of 116 subjects took part in the study. They resided either in the UK (61,2%) or in the USA (38,8%) and most of them were female (67,2%). The age of participants ranged from 19 to 69 years old, with a median of 30. Most subjects declared White ethnicity (62,6% versus 16,5% of Black, 11,3% of Asian, and 9,6% of Mixed ethnicity)[25].

The experiment was divided into 15 sessions, the overview of which is shown in Table 4. Each session lasted on average around 25 minutes and involved from 5 to 10 participants. Due to the time difference between the UK and the USA, sessions were conducted at different times of day and night (CEST time zone). Final total payments transferred to each participant varied between sessions and ranged from 3,40 GBP to 7,00 GBP, with an average payment of 5,48 GBP. The total of all payments was 635,12 GBP.

[24] The pre-registration document is available at https://aspredicted.org/WP1_CDB.
[25] To discover more information about the participants' characteristics, see Table 10 which presents descriptive statistics of variables used to panel data regressions.

Table 4: Overview of sessions of the experiment and participants' payments

Session	Starting date and time	Duration (min:sec)	Number of subjects	Payments received by participants		
				Average payment	Minimum payment	Maximum payment
I.	03/07/2024 09:56:51	26:58	10	4,90 GBP	3,40 GBP	7,00 GBP
II.	03/07/2024 11:49:53	26:21	8	5,25 GBP	3,40 GBP	6,00 GBP
III.	03/07/2024 13:04:54	18:30	6	5,93 GBP	5,80 GBP	6,00 GBP
IV.	03/07/2024 13:52:08	19:52	8	4,98 GBP	3,40 GBP	6,00 GBP
V.	03/07/2024 15:00:32	25:55	9	6,16 GBP	3,40 GBP	7,00 GBP
VI.	03/07/2024 15:59:54	23:11	8	5,60 GBP	3,40 GBP	7,00 GBP
VII.	03/07/2024 19:32:02	21:24	10	5,95 GBP	5,83 GBP	6,00 GBP
VIII.	03/07/2024 20:28:12	28:07	8	5,28 GBP	3,40 GBP	7,00 GBP
IX.	03/07/2024 21:27:59	26:32	8	5,63 GBP	3,40 GBP	6,00 GBP
X.	04/07/2024 00:17:11	26:26	8	5,80 GBP	3,40 GBP	7,00 GBP
XI.	04/07/2024 01:25:55	18:58	7	5,60 GBP	3,40 GBP	6,00 GBP
XII.	04/07/2024 02:15:35	28:39	7	4,46 GBP	3,40 GBP	6,00 GBP
XIII.	04/07/2024 03:28:56	30:06	7	4,89 GBP	3,40 GBP	6,00 GBP
XIV.	04/07/2024 16:04:02	24:43	7	5,69 GBP	3,40 GBP	7,00 GBP
XV.	04/07/2024 22:17:23	27:33	5	6,20 GBP	6,00 GBP	7,00 GBP
Overall		**Average duration**	**Number of subjects**	**Average payment**	**Minimum payment**	**Maximum payment**
		24:53	116	5,48 GBP	3,40 GBP	7,00 GBP

On the technical side, the course of the game was as follows. Once the study began, participants saw a screen with general instructions which contained information on the game structure, payment and navigation through the experiment. They also had to provide their Prolific ID. Then, subjects saw the next screen with instructions for Block 1 and five short Holt and Laury choices to be made. The following stage contained instructions showing "Take from charity" game scenario and payment information. Each of the next eleven screens presented one decision period and looked like the example shown earlier (left side of Figure 7). Then, the screen with instructions for Block 3 was displayed (it was very similar to instructions for Block 2, with minor changes regarding the loss framework). Later, participants saw five separate screens with loss framework periods, which looked as on the right side of Figure 7. In the end, subjects saw the final screen with information on the numbers randomly chosen for the purpose of payment, as well as a summary of their payment[26]. It should be emphasized that, to avoid the influence of gambler's fallacy on participants' choices, they were not informed whether they were caught after each individual period[27]. Only at the end of the study subjects saw, whether they got caught in the period that was randomly chosen to determine the payment.

What is also important, to avoid possible influence of the order in which participants saw the periods, each session of the experiment showed a different sequence of periods, as presented in Table 5[28]. These sequences are random, and they were generated using the website random.org. Moreover, 8 sessions (I-V, XI, XII, XV) presented the standard experimental design, i.e. gain framework periods (Block 2) at the

[26] To access all the screens seen by participants, see Appendix A.

[27] This is because, for example, if a player saw immediately that they had been caught in three consecutive periods, they might think than in the fourth period they would not get caught, just because their fate would change. Thus, they might have based their decision in the fourth period only on this belief, completely ignoring the size of deterrent incentives. To learn more about gambler's fallacy see Farmer et al. (2017) and the literature discussed therein.

[28] The numbers of periods in the sequences shown in Table 5 match the numbers shown in the general "ordered" sequence of periods in Table 3.

beginning, and loss framework periods (Block 3) afterwards, while 7 sessions (VI-X, XIII, XIV) showed Block 3 before Block 2. Thus, 60 subjects made their choices in gain framework first, and 56 participants made decisions in loss framework first. In cases where Block 3 was presented before Block 2, instructions was adjusted accordingly.

Table 5: Sequence of periods shown to participants in each session

Session	**Sequence of periods**	
	Block 2	**Block 3**
I.	11, 5, 2, 1, 4, 9, 6, 3, 7, 10, 8	14, 15, 12, 16, 13
II.	6, 9, 11, 10, 5, 4, 8, 1, 7, 2, 3	15, 12, 14, 13, 16
III.	4, 2, 10, 11, 3, 5, 9, 1, 8, 6, 7	12, 14, 15, 16, 13
IV.	8, 2, 4, 10, 3, 5, 6, 9, 11, 7, 1	13, 14, 16, 15, 12
V.	1, 6, 4, 8, 7, 3, 5, 10, 11, 2, 9	13, 15, 16, 14, 12
VI.	13, 15, 12, 14, 16	7, 2, 5, 11, 6, 1, 4, 10, 8, 9, 3
VII.	14, 16, 13, 12, 15	2, 10, 1, 9, 8, 3, 11, 5, 6, 4, 7
VIII.	13, 15, 16, 12, 14	6, 5, 2, 4, 3, 1, 10, 7, 11, 8, 9
IX.	13, 14, 12, 15, 16	9, 11, 5, 3, 7, 8, 1, 6, 10, 4, 2
X.	15, 13, 16, 12, 14	10, 4, 1, 2, 7, 5, 8, 11, 6, 3, 9
XI.	7, 1, 8, 10, 5, 6, 2, 3, 4, 11, 9	13, 16, 14, 15, 12
XII.	6, 11, 8, 5, 2, 10, 3, 1, 4, 9, 7	12, 14, 13, 15, 16
XIII.	13, 16, 15, 12, 14	1, 6, 2, 8, 11, 9, 7, 3, 4, 5, 10
XIV.	15, 16, 14, 12, 13	4, 7, 10, 1, 6, 8, 11, 2, 3, 9, 5
XV.	6, 9, 5, 11, 3, 1, 10, 2, 4, 8, 7	15, 12, 14, 13, 16

5. Results of the experiment

5.1. Decisions made by the subjects and simple hypotheses testing

5.1.1. Block 1: Measuring risk preferences

The results of trimmed version of Holt and Laury experiment will be presented according to procedure suggested by Brañas-Garza et al. (2021). Of the 116 participants who took part in the study, 108 subjects responded in Block 1. The share of safe choices (Lottery A) is shown in Table 6 ("Mean" column). As can be seen, this share generally decreases with the next choice, which is consistent with the expectations. The exception is Choice 4, in which more people chose Lottery A than in Choice 3. This situation can be explained by inconsistencies of participants' choices, as discussed earlier in this thesis. In the case of this experiment, inconsistencies (at least one switch from Lottery B to A) were demonstrated by 27 subjects, who further will be treated as "Inconsistent".

Table 6: Descriptive statistics of short Holt and Laury results

Variable	Variable definition	Obs[a]	Mean	SD	Min	Max
Choice_1	Dummy variable (=1) if Lottery A chosen	108	.889	.316	0	1
Choice_2	Dummy variable (=1) if Lottery A chosen	108	.833	.374	0	1
Choice_3	Dummy variable (=1) if Lottery A chosen	108	.62	.488	0	1
Choice_4	Dummy variable (=1) if Lottery A chosen	108	.639	.483	0	1
Choice_5	Dummy variable (=1) if Lottery A chosen	108	.38	.488	0	1

Notes: [a] 8 observations for each choice are missing since 8 subjects did not make their decisions in this part of the experiment.

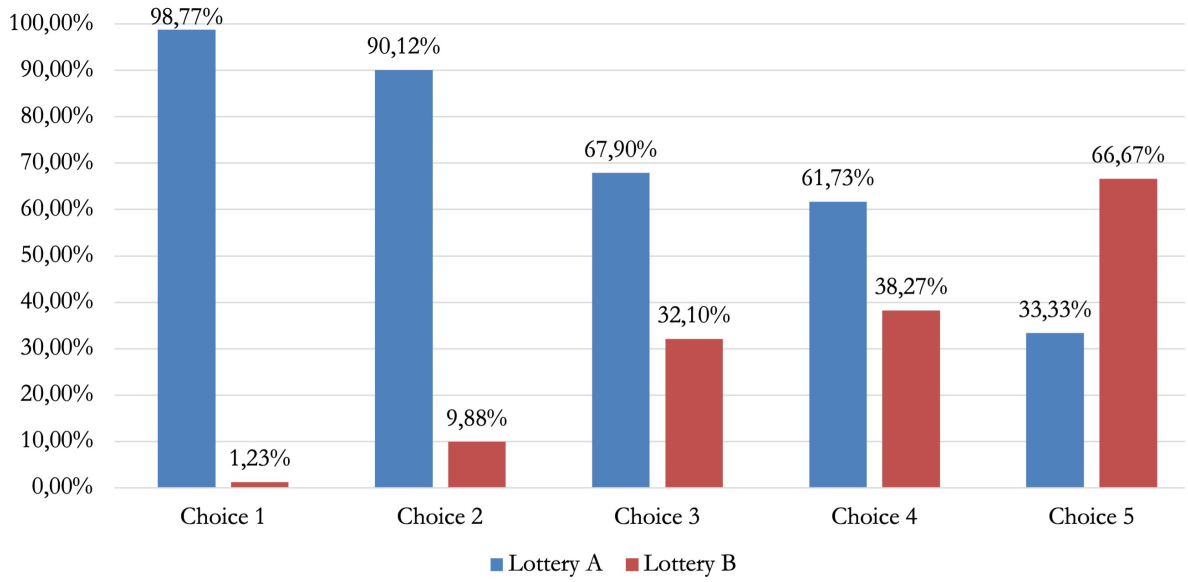
Due to the large number of inconsistencies, it is necessary to look at the decisions of the subjects for whom determining risk preference is clearly possible. Table 7 presents descriptive statistics, and Chart 1 illustrates shares of safe and risky choices for consistent participants. There is clearly a decrease in Lottery A choices and an increase in Lottery B choices with each successive choice made. This suggests that most participants are characterised by risk aversion preference.

Table 7: Descriptive statistics of short Holt and Laury results, excluding inconsistencies

Variable	Variable definition	Obs[a]	Mean	SD	Min	Max
Choice_1	Dummy variable (=1) if Lottery A chosen	81	.988	.111	0	1
Choice_2	Dummy variable (=1) if Lottery A chosen	81	.901	.3	0	1
Choice_3	Dummy variable (=1) if Lottery A chosen	81	.679	.47	0	1
Choice_4	Dummy variable (=1) if Lottery A chosen	81	.617	.489	0	1
Choice_5	Dummy variable (=1) if Lottery A chosen	81	.333	.474	0	1

Notes: [a] 35 observations for each choice are missing since 8 subjects did not make their decisions in this part of the experiment, and other 27 subjects were inconsistent in their choices.

Chart 1: Share of safe choices (Lottery A) and risky choices (Lottery B), excluding inconsistencies

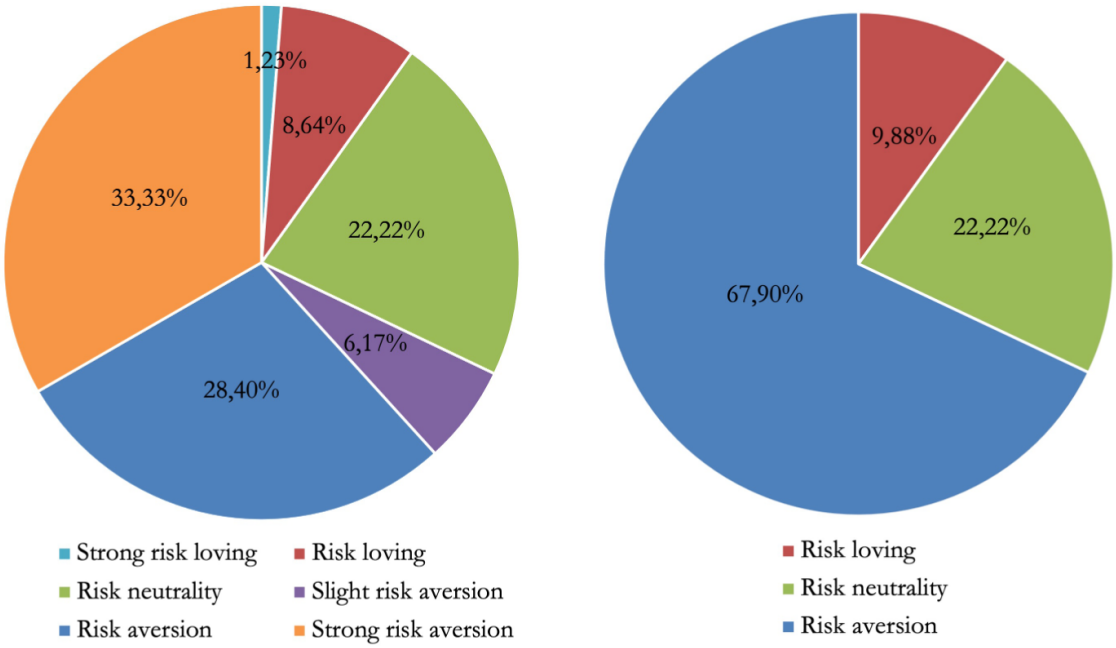


Based on decisions made by individuals consistent in their choices, their risk attitudes were determined, as shown in Table 8 and on Chart 2. The most numerous two groups include strong risk averse (33,33%) and risk averse subjects (28,40%). There are also quite a few individuals with a risk neutrality attitude (22,22%). In contrast, less than 10% of participants appeared to be risk loving (8,64%) and slightly risk averse (6,17%), while only one person is a strong risk lover (1,23%). Chart 2 also presents a simplified version of risk preferences, distinguishing only three main attitudes. Overall, almost 68% of participants can be described as risk averse, 22% as risk neutral, and 10% as risk lovers.

Table 8: Risk preferences of participants who were consistent in their choices

Number of safe choices	Risk preference	Number of risk preferences	Share	Cumulative share
0	Strong risk loving	1	1,23%	1,23%
1	Risk loving	7	8,64%	9,88%
2	Risk neutrality	18	22,22%	32,10%
3	Slight risk aversion	5	6,17%	38,27%
4	Risk aversion	23	28,40%	66,67%
5	Strong risk aversion	27	33,33%	100,00%
Overall		81	100,00%	-

Chart 2: Risk preferences (left) and simplified risk preferences (right) of subjects who were consistent in their choices



The results of Block 1 are in line with previous experimental findings[29]. Thus, there is nothing unusual happening in terms of risk attitudes in the sample of participants consistent in their choices. What is noteworthy, however, is the very high number of inconsistencies. Accordingly, the later analysis will contain "Inconsistency" as another risk attitude category to check its impact on the decision to take money from the charity.

[29] See for instance Holt & Laury (2002), Drichoutis & Koundouri (2012), Friesen (2012), and Brañas-Garza et al. (2021).

5.1.2. Block 2 and Block 3: Taking from charity game

This section presents an overview of decisions made by the subjects in the main part of the experiment, as well as the results of simple hypotheses testing obtained from two-sample test of proportions (prtest). As illustrated by Chart 3, the share of "Take the money" decisions[30] in all decisions varies from 20% (period 10) to 69% (period 12). In each of four groups of periods (see Table 3 above, "Probabilities' size" column), there were more positive taking decisions at the relevant starting points (periods 1, 4, 9, 12) than in periods in which F or p was increased in reference to the specific starting point. In 8 out of 12 cases, these differences were highly statistically significant (see Table 9 below, yellow group), similarly to 2 out of 3 differences between the starting points which had different expected sanctions (Table 9, orange group). Accordingly, this generally confirms Becker's (1968) theory, as increasing expected sanction always reduced the number of taking decisions.

What is very important for this thesis, in all cases there are observable differences between periods in which F was increased and those in which p was increased, when the expected sanction was the same (periods 2 vs. 3, 5 vs. 6, 7 vs. 8, 10 vs. 11, 13 vs. 14, and 15 vs. 16). In the main treatments in gain framework group (periods 4-8), increasing F (periods 6 and 8) had a higher deterrent effect than increasing p equivalently (periods 5 and 7), which is consistent with H1. However, these differences were found to be statistically insignificant with prtest p-values of 0.291 and 0.380 (Table 9, green group).

[30] The decision to take money from the charity will also be referred to as a "decision to commit a crime", "crime decision", or "crime choice", while the share of "Take the money" decisions in all decisions will also be called "crime rate".

Chart 3: Share of "Take the money" decisions in all decisions (all 16 periods)

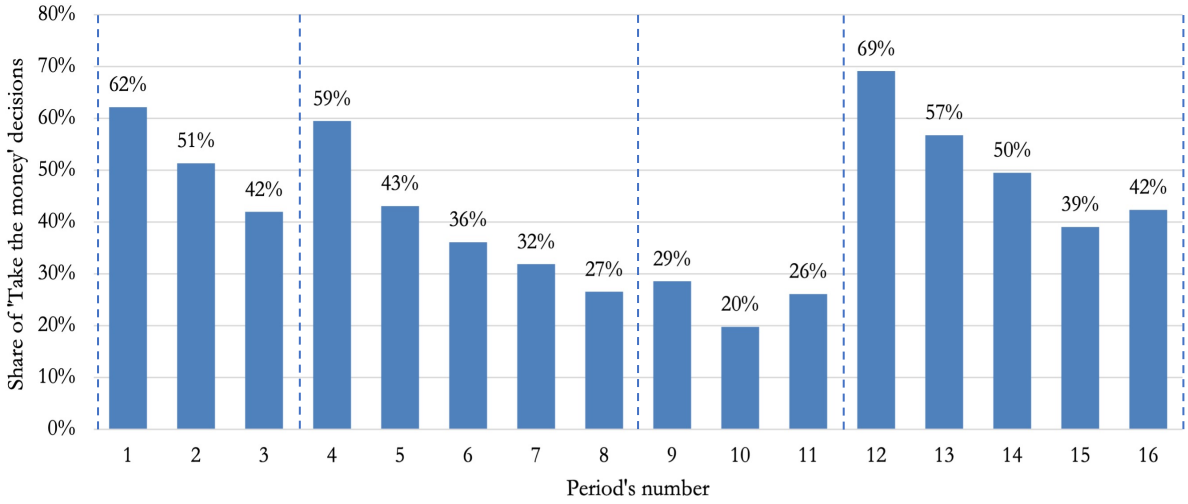


Table 9: Results (p-values) of two-sample test of proportions (prtest) between selected periods

Period[a]	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P12	P13	P15
P2	.104												
P3	.003	.160											
P4	.680												
P5				.015									
P6				.001	.291								
P7				.000									
P8				.000			.380						
P9	.000			.000									
P10		.000							.127				
P11			.013						.682	.264			
P12				.135									
P13					.043						.058		
P14						.045					.003	.282	
P15							.259				.000		
P16								.013			.000		.623

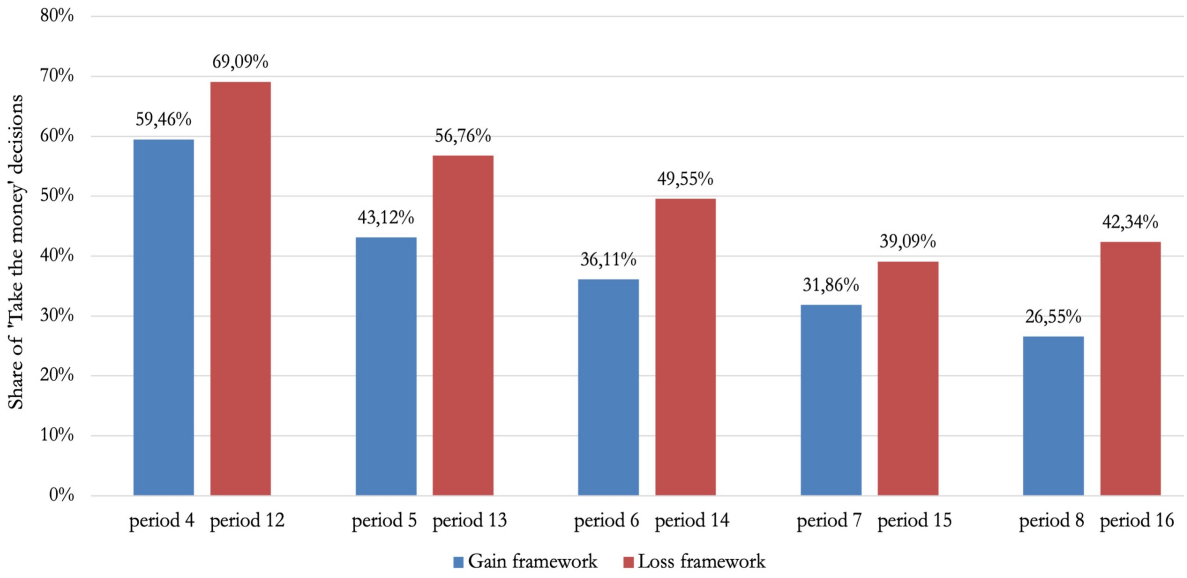
Notes: [a] The row for P1 as well as columns for P11, P14, and P16 were omitted due to not containing any values. Different colours indicate p-values for prtests between different groups of periods (yellow – prtests between periods and relevant starting points; green – between periods with the same expected sanction (F*p); orange – between different starting points; grey – between small and big probabilities periods; blue – between corresponding periods of gain and loss framework). The results were estimated with Stata 18.0.

Looking more closely at the differences between gain and loss frameworks, for the latter, there were remarkable more decisions to take money from the charity, which is illustrated by Chart 4. Three of the five differences between the corresponding periods (those with identical deterrent incentives) in the positive and negative framework are statistically significant at the 5% level with prtest p-values of 0.043, 0.045, and 0.013 (Table 9, blue group).

Regarding H2 and H2A which concern the loss framework, mixed results were obtained. For periods 13 and 14, where p and F were doubled respectively, severity appeared to have a greater deterrent effect (as in the gain framework), which is inconsistent with H2. Also, H2A was not confirmed, as the differences between deterrent effects of F and p in periods 13 and 14 and the corresponding periods from positive framework (periods 5 and 6) is almost non-existent (7,21 p.p. in loss domain vs. 7,01 p.p. in gain framework).

In contrast, when deterrent incentives were tripled (periods 15 and 16), in the loss framework the certainty of punishment was more important than its severity, which is an opposite finding to the one seen in gain framework (periods 7 and 8), and which confirms H2. These results might suggest that in the case of negative framework, unlike the positive one, the size by which deterrent incentives are increased may be important in determining the difference between deterrent effects of F and p. It should be stressed, however, that differences between periods with the same expected sanction in the loss framework (periods 13 vs. 14, and 15 vs. 16) were found to be statistically insignificant (Table 9, green group).

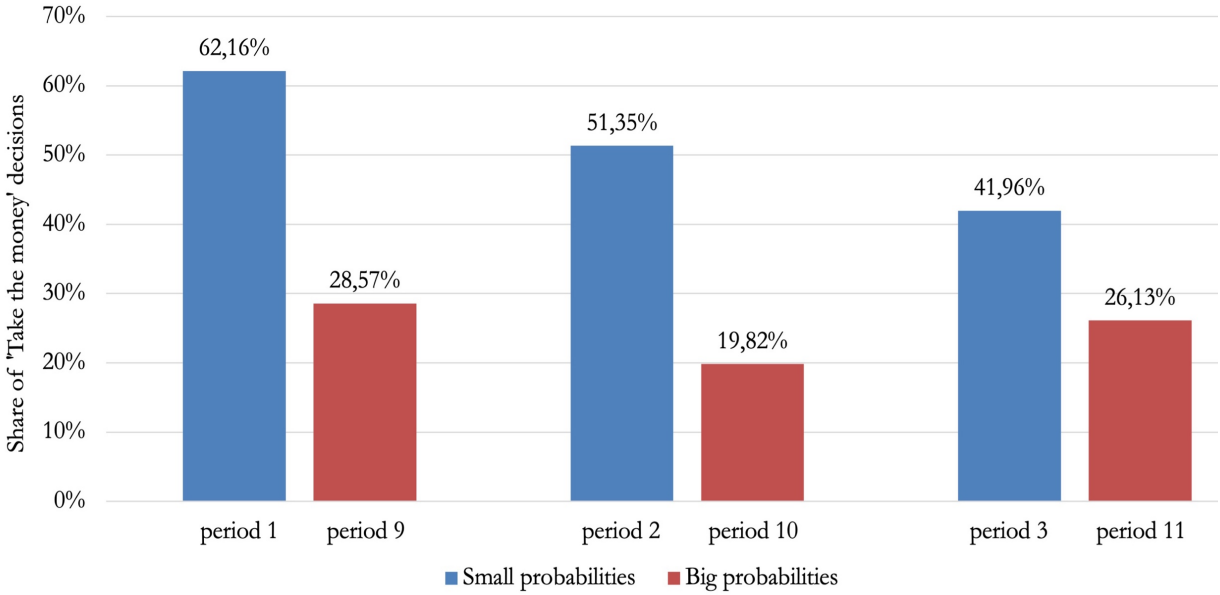
Chart 4: Share of "Take the money" decisions in all decisions (main treatments; gain versus loss framework)



When comparing periods with small probabilities and big probabilities illustrated by Chart 5, there were far fewer crime decisions in the second group and the differences were highly statistically significant (Table 9, grey group). However, this is not very meaningful, since expected sanction for periods with high likelihoods was much higher than for those with low likelihoods. Therefore, from this finding one can only conclude that Becker's (1968) deterrence hypothesis was confirmed also in this case.

However, H3 and H3A concern overweighting small probabilities. A comparison between periods 2 and 3 indicates that, contrary to H3, it was increasing expected sanction through F rather than through p that reduced taking more. In contrast, in the case of big probabilities (periods 10 and 11), where according to the theory it should be F that is more deterrent, it was p that turned out to be so, which completely negated the validity of H3A. This could suggest a full inconsistency of the results with prospect theory in terms of overweighting small and underweighting big probabilities. However, it should be again pointed out that both the differences between periods 2 and 3, and between periods 10 and 11 are not statistically significant (Table 9, green group).

Chart 5: Share of "Take the money" decisions in all decisions (small probabilities versus big probabilities)



In summary, the results of simple hypotheses testing are mixed. While H1 was suggested to be confirmed, the findings related to H2 and H2A were mixed, and H3 and H3A were not confirmed. However, all comparisons between corresponding crime rates (between periods with corresponding manipulations of F and p and with the same expected sanctions), proved to be statistically insignificant (Table 9, green group). Therefore, simple hypotheses testing generally suggests that participants' decisions were not affected by changes in F or p, and that the differences in crime rates might have been seen due to some other effects or factors. The next section will present the results of panel data regression, which examines the effects of F and p on crime decision, and which will also account for risk preference and other control variables.

5.2. Panel data regression results

5.2.1. General remarks

This section presents panel data regression results, based on probit random effects regression (xtprobit). The dataset includes decisions to take money from the charity or not in each of 16 periods for each of 116 participants, as well as treatments in each period (Fine size and probability of getting caught), the expected sanction in each period (Fine*probability), and control variables for each participant, which will be discussed below[31]. All estimations were made using Stata 18.0.

At first, the general model used for regressions, as well as descriptive statistics will be presented. The following subsections discuss estimation results for observations from periods 1-11 (to test H1), periods 4-8 and 12-16 (H2 and H2A), and periods 1-3 and 9-11 (H3 and H3A). The last subsections cover additional robustness checks that were done, and some interesting additional findings which arose from the analysis.

5.2.2. General model and descriptive statistics

The general model used to perform panel data regressions is shown by Figure 8. The dependent variable $crime$ is a dummy variable that reflects the decision to take money from the charity (1) or not (0). However, as all regressions will use a probit model with random effects (xtprobit), the variable $crime$, more specifically, reflects the propensity to take the money from charity (propensity to commit a crime). The α is a constant, β_1 and β_2 are coefficients of treatments (Fine size and probability of getting caught), β_3 refers to the coefficient of expected sanction (Fine*probability), while β_4, β_5, β_6, β_7, β_8, β_9, β_{10} refer to the coefficients of control variables. Finally, u_i reflects the individual-specific random effects.

[31] To access the full dataset, see Appendix A.

Figure 8: General model used to perform panel data regressions

$$crime_i^* = \alpha + \beta_1 Fine + \beta_2 probability + \beta_3 (Fine * probability) + \beta_4 riskpreference + \beta_5 female + \beta_6 age + \beta_7 student + \beta_8 employment + \beta_9 UK + \beta_{10} ethnicity + u_i$$

Table 10 presents descriptive statistics of variables used in the model[32]. It is worth noting that 81 observations for the dependent variable are missing because 20 subjects did not make all 16 decisions. Control variables include risk preference of an individual, which was determined from the results of Block 1 of the experiment. It is a categorical variable that includes six possible risk attitudes and *Inconsistency* for subjects for whom risk preference could not be determined. The other control variables relate to demographic factors, i.e. gender, age, country of residence, and ethnicity. In addition, there are *student* and *employment* variables, which relate to the student status and full-time employment, respectively. It can be noted, however, that a great number of observations for *student* and *employment* variables are missing, which is because the data on participants received from Prolific was incomplete[33].

The last two items in Table 10 are variables *periods_order* and *periods_answered*. The first of them refers to session dummies, which signify one of the fifteen sequences of periods that was seen by a specific participant (see Table 5 above). It will allow to check whether any order effect has occurred (for instance, whether subjects have succumbed to moral licensing). In turn, *periods_answered* refers to the number of periods in which an individual made any decision regarding taking money from the charity. These two variables are not included in the general model on Figure 8, but they will be used as control variables in additional robustness checks (see section 5.2.7).

[32] Treatments (Fine, probability, Fine*probability) are not shown since they were described in detail in section 4.1.3 of this paper.
[33] The information was coded as data_expired for 30 subjects in the case of student status, and for 51 subjects in the case of employment information.

Table 10: Descriptive statistics of variables used in panel data regressions

Variable	Variable definition	Obs	Mean	SD	Min	Max
Dependent variable						
Crime	Dummy variable (=1) if the money was taken	1775[a]	.427	.495	0	1
Risk_preference	Categorical variable (7 categories)					
Inconsistency	Inconsistent choices/ Impossibility to identify the risk preference of the individual	1856	.302	.459	0	1
Risk aversion	Risk aversion preference	1856	.198	.399	0	1
Risk loving	Risk loving preference	1856	.06	.238	0	1
Risk neutrality	Risk neutrality preference	1856	.155	.362	0	1
Slight risk aversion	Slight risk aversion preference	1856	.043	.203	0	1
Strong risk aversion	Strong risk aversion preference	1856	.233	.423	0	1
Strong risk loving	Strong risk loving preference	1856	.009	.092	0	1
Female	Dummy variable (=1) if the individual is female	1856	.672	.469	0	1
Age	Age of the individual	1856	33.716	10.983	19	69
Student	Dummy variable (=1) if the individual has a student status	1376[b]	.198	.398	0	1
Employment	Dummy variable (=1) if the individual is full-time employed	1040[c]	.4	.49	0	1
UK	Dummy variable (=1) if the individual resides in the UK	1856	.612	.487	0	1
Ethnicity	Categorical variable (4 categories)					
Asian	Asian ethnicity	1840[d]	.113	.317	0	1
Black	Black ethnicity	1840[d]	.165	.371	0	1
Mixed/Other	Mixed or Other ethnicity	1840[d]	.096	.294	0	1
White	White ethnicity	1840[d]	.626	.484	0	1
Periods_order[e]	Categorical variable (15 categories) assigning the order, in which the individual saw the periods	1856[e]	.086[e]	.281[e]	0[e]	1[e]
Periods_answered	The number of periods in which the individual made their decision	1856	15.302	2.347	1	16

Notes: [a] Taking decision is missing for 81 periods. [b] Student status information is missing for 30 subjects (480 observations). [c] Employment information is missing for 51 subjects (816 observations). [d] Ethnicity information is missing for 1 subject (16 observations). [e] Descriptive statistics of periods_order is shown only for one category which is "Order of periods in session one" (all other 14 categories differ in Mean and SD).

5.2.3. Model specifications used to check the hypotheses

For each of five groups of periods (1-11; 4-8; 12-16; 1-3; 9-11), the same six probit random effects regressions will be carried out. They are presented in Table 11. Model (1) is a simple model which contains only Fine and probability treatments, while

model (2) also includes F*p interaction term (expected sanction). Models (3) and (4), due to the large number of missing data for *student* and *employment* variables, are general models (see Figure 8 above), but without those two independent variables. Furthermore, in model (3), in contrast to model (4), there is no F*p included. In turn, model (6) contains all the independent variables presented in the general model (Figure 8), and model (5) is a similar specification, but without F*p. However, due to *student* and *employment* missing data issue discussed earlier, models (5) and (6) contain much fewer observations comparing to models (1)-(4). For the sake of simplicity, it may be also noted that odd-numbered models, unlike even-numbered models, do not contain F*p interaction term.

In the following subsections, the numbering of these specifications will be identical. To distinguish the different regressions conducted for different groups of periods, the models for periods 1-11 have only been numbered, while those for periods 4-8 also contain letter G (gains), those for periods 12-16 – letter L (loss), those for periods 1-3 – letter S (small probabilities), and those for periods 9-11 – letter B (big probabilities). For example, model (2) is a simple model with F*p for periods 1-11, while model (5L) is a general model without F*p for periods 12-16 (see Table 11).

Due to the use of probit random-effects regressions, as well as the different scales in which treatments were coded (F is between 52-180, and p is between 0,1-0,9), simple interpretations of coefficients will not be meaningful. Therefore, after each model, the elasticities (average marginal effects, eyex) for Fine and probability will also be computed. Elasticities will allow to interpret the experimental results correctly and test the hypotheses, since they will show the percentage change in the probability of taking money for a 1% change in F or p.

The results of the even-numbered models (those with F*p) for periods 4-8, 12-16, 1-3, and 9-11 will not be presented because they were identical to the regression results without F*p. This is due to Stata's omission of this interaction term, because

of collinearity. The occurrence of collinearity is not surprising, as expected sanction is determined by the treatment variables (Fine and probability) also used in the regressions. Only in one case, where more periods were used (1-11), F*p was not omitted in the regression, which will be shown in the next subsection. For this reason, in the results analysis, the focus will be generally on those from the odd-numbered models (specifications without the interaction term).

Table 11: All model specifications and their numbers

Name of the model	**Simple model**	**Simple model with F*p**	**General model without F*p, student and employment**	**General model without student and employment**	**General model without F*p**	**General model with all variables**
Treatments and control variables included in the model	• F • p	• F • p • F*p	• F • p • riskpref. • female • age • UK • ethnicity	• F • p • F*p • riskpref. • female • age • UK • ethnicity	• F • p • riskpref. • female • age • student • employment • UK • ethnicity	• F • p • F*p • riskpref. • female • age • student • employment • UK • ethnicity
Periods from which observations are used	**Number of the model**					
Periods 1-11 (H1)	(1)	(2)	(3)	(4)	(5)	(6)
Periods 4-8 (H2, H2A)	(1G)	(2G)	(3G)	(4G)	(5G)	(6G)
Periods 12-16 (H2, H2A)	(1L)	(2L)	(3L)	(4L)	(5L)	(6L)
Periods 1-3 (H3, H3A)	(1S)	(2S)	(3S)	(4S)	(5S)	(6S)
Periods 9-11 (H3, H3A)	(1B)	(2B)	(3B)	(4B)	(5B)	(6B)

5.2.4. Periods 1-11 (H1)

Observations from periods 1-11 are used to test H1, according to which severity has a greater deterrent effect than certainty. Table 12 presents full probit (random-effects) regression results for all six models, while Table 13 shows elasticities of changes in crime in relation to F and p. For each specification, the treatments' signs are consistent with theory, i.e. an increase in deterrent incentives negatively affects propensity to commit a crime. Moreover, all elasticities are highly statistically significant with p-values<0.01.

Interesting differences are observed between deterrent effects of F and p. In all even-numbered models (those which include F*p), F has a stronger effect than p. For instance, in model (2), increasing Fine by 1% decreases the propensity to take money from the charity by 1.02%, while increasing p by 1% results in 0.867% decrease in the likelihood of a positive taking decision. However, in all specifications containing expected sanction, the differences between elasticities for F and p was insignificant (Wald test p-values of 0.224, 0.265 and 0.928). Accordingly, there is no sufficient evidence to find a difference between deterrent effects of F and p in models with interaction term.

In contrast, odd-numbered specifications indicate a greater effect of certainty, and differences between elasticities for Fine and probability are statistically significant (at the 5% level in models (1) and (5), and at the 10% level in (3)). Differences between deterrent effects of p and F range from 0.153 p.p. in model (3) to 0.307 p.p. in (5), but the overall conclusion of greater strength of probability holds. Consequently, this result shows that H1 is not confirmed, and that it is p, not F which is a greater deterrent.

It is also worth noting the surprising lack of a significant effect of any of the risk preferences. In models (3) and (4) only a weak statistical significance of *Inconsistency* can be observed, which, however, fully disappears after calculating the average marginal effect (dydx) for this variable. In some specifications, significant effects of country of residence and ethnicity could be seen, which will be discussed in section 5.2.8.

Table 12: Probit random-effects regression results for periods 1-11

	Simple model	Simple model with F*p	General model without F*p, student and employment	General model without student and employment	General model without F*p	General model with all variables
	(1)	(2)	(3)	(4)	(5)	(6)
Fine	-0.009***	-0.004	-0.009***	-0.004	-0.008***	-0.003
	(0.001)	(0.002)	(0.001)	(0.002)	(0.002)	(0.003)
probability	-2.262***	-0.213	-2.236***	-0.311	-2.368***	-0.606
	(0.213)	(0.769)	(0.214)	(0.772)	(0.298)	(1.059)
c.Fine#c.probability		-0.038***		-0.036***		-0.032*
		(0.014)		(0.014)		(0.019)
Risk preferences[a]:						
Inconsistency[b]			0.504*	0.507*	0.167	0.166
			(0.304)	(0.307)	(0.377)	(0.379)
Risk aversion[b]			0.067	0.068	0.061	0.062
			(0.360)	(0.363)	(0.442)	(0.446)
Risk loving[b]			0.555	0.559	-0.370	-0.376
			(0.522)	(0.526)	(0.661)	(0.666)
Risk neutrality[b]			0.466	0.466	0.399	0.397
			(0.372)	(0.375)	(0.499)	(0.503)
Slight risk aversion[b]			0.562	0.564	0.054	0.050
			(0.582)	(0.587)	(0.730)	(0.735)
Female			0.187	0.191	-0.119	-0.115
			(0.273)	(0.275)	(0.314)	(0.316)
Age			-0.014	-0.014	-0.024	-0.024
			(0.011)	(0.011)	(0.018)	(0.018)
UK			0.582**	0.583**	0.001	-0.005
			(0.258)	(0.260)	(0.346)	(0.348)
Ethnicity:						
Asian[c]			0.611	0.617	0.355	0.347
			(0.377)	(0.380)	(0.538)	(0.552)
Black[c]			0.604*	0.612*	1.135***	1.143***
			(0.343)	(0.346)	(0.413)	(0.416)
Mixed/Other[c]			0.729	0.740*	0.460	0.466
			(0.449)	(0.452)	(0.582)	(0.586)
Student					-0.155	-0.152
					(0.473)	(0.476)
Employment					-0.019	-0.020
					(0.332)	(0.324)
Constant	1.249***	0.973***	0.710	0.445	1.790*	1.549
	(0.215)	(0.238)	(0.605)	(0.619)	(0.937)	(0.954)
Log likelihood	-640.587	-636.742	-624.704	-621.344	-324.385	-322.885
AIC	1289.173	1283.483	1279.408	1274.689	682.770	681.770
BIC	1309.606	1309.024	1355.759	1356.13	758.239	761.678
Number of observations	1222	1222	1200	1200	626	626

Notes: This table presents probit random-effects regression results for observations from periods 1-11. The dependent variable is crime (1 = Take the money, 0 = Do not take the money). Statistical significance of coefficients is reported as follows: *** p<0.01; ** p<0.05; * p<0.1. Standard errors are reported in parentheses. [a] The category Strong risk loving was omitted due to predicting failure perfectly (there was only one subject with Strong risk loving attitude). [b] The reference category is Strong risk aversion. [c] The reference category is White. The results were estimated with Stata 18.0.

Table 13: Elasticities (average marginal effects, eyex) for periods 1-11

	(1)	(2)	(3)	(4)	(5)	(6)
Fine	-0.561***	-1.020***	-0.602***	-1.065***	-0.584***	-1.058***
	(0.091)	(0.202)	(0.098)	(0.216)	(0.138)	(0.320)
probability	-0.710***	-0.867***	-0.755***	-0.913***	-0.891***	-1.039***
	(0.089)	(0.113)	(0.095)	(0.120)	(0.148)	(0.178)
Wald test p-value	0.046	0.224	0.058	0.265	0.014	0.928
Number of observations	1222	1222	1200	1200	626	626

Notes: This table presents elasticities (average marginal effects, eyex) for Fine and probability for observations from periods 1-11, as well as Wald test p-values for differences between elasticities for Fine and probability. Model VCE: OIM. Statistical significance of coefficients is reported as follows: *** p<0.01; ** p<0.05; * p<0.1. Standard errors are reported in parentheses. The results were estimated with Stata 18.0.

5.2.5. Periods 4-8 and 12-16 (H2 and H2A)

Hypotheses 2 and 2A are tested using observations for periods 4-8 (gain framework) and 12-16 (loss framework). Table 14 illustrates the results of the probit (random-effects) regression, and Table 15 shows relevant elasticities for Fine and probability. Again, the signs of all treatments' coefficients are negative, which is consistent with theory, and the elasticities for F and p are very highly statistically significant in all cases (p-values<0.01).

For the first two models for the gain framework ((1G) and (3G)) and the first two specifications for the loss framework ((1L) and (3L)), a greater deterrent effect is attributed to Fine. In contrast, for (5G) and (5L) models, which contain only almost half of observations comparing to other specifications (by including student and employment variables), the result is opposite. All differences between elasticities for F and p are highly statistically significant (see Wald test p-values in Table 15). Accordingly, these outcomes may indicate mixed results for H2 and H2A, although the small number of observations in (5G) and (5L) models casts doubt on the results of those two regressions.

Based solely on the results from models with a larger number of observations, H2 is not confirmed, as in the loss framework it is severity that has the larger (albeit very slightly larger) deterrent effect. However, H2A was confirmed, as the differences between F and p effects are significantly smaller in loss framework periods than in the corresponding gain framework periods (0.002 p.p. in (1L) vs. 0.188 p.p. in (1G), and 0.027 p.p. in (3L) vs. 0.201 p.p. in (3G)). Accordingly, there is clearly an increase in the deterrent effect of probability in the loss domain, which is consistent with the intuition. Even though in the loss framework the differences between the effects of F and p are statistically significant, their economic significance is questionable (0.002 p.p. and 0.027 p.p. are very small and rather practically irrelevant). It is also worth noting that the average marginal effects of both F and p in all models are generally larger in the gain than in the loss framework.

Similarly to models for periods 1-11, none of the risk attitudes has a statistically significant effect on the probability to take money from the charity. However, country of residence and ethnicity effects were noted in some specifications (see section 5.2.8).

Table 14: Probit random-effects regression results for periods 4-8 (gain framework) and periods 12-16 (loss framework)

	Simple model	General model without F*p, student and employment	General model without F*p	Simple model	General model without F*p, student and employment	General model without F*p
	(1G)	(3G)	(5G)	(1L)	(3L)	(5L)
Fine	-0.010***	-0.010***	-0.010***	-0.001***	-0.001***	-0.009***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
probability	-2.438***	-2.398***	-2.964***	-2.868***	-2.822***	-3.393***
	(0.501)	(0.504)	(0.727)	(0.527)	(0.531)	(0.756)
Risk preferences[a]:						
Inconsistency[b]		0.507	0.371		0.474	-0.144
		(0.318)	(0.412)		(0.340)	(0.482)
Risk aversion[b]		-0.052	-0.003		0.219	0.125
		(0.377)	(0.478)		(0.404)	(0.578)
Risk loving[b]		0.671	0.106		0.653	-0.136
		(0.549)	(0.739)		(0.583)	(0.862)
Risk neutrality[b]		0.530	0.461		0.193	0.328
		(0.384)	(0.541)		(0.408)	(0.624)
Slight risk aversion[b]		0.659	0.191		0.595	0.712
		(0.602)	(0.794)		(0.645)	(0.930)
Female		0.120	-0.172		0.104	-0.258
		(0.282)	(0.343)		(0.306)	(0.408)
Age		-0.006	-0.014		-0.003	0.014
		(0.012)	(0.019)		(0.012)	(0.022)
UK		0.567**	0.036		0.488*	0.066
		(0.273)	(0.379)		(0.286)	(0.438)
Ethnicity:						
Asian[c]		0.717*	0.331		1.269***	0.379
		(0.397)	(0.595)		(0.431)	(0.698)
Black[c]		0.636*	1.510***		0.353	0.606
		(0.360)	(0.473)		(0.382)	(0.525)
Mixed/Other[c]		0.694	0.643		0.833*	0.329
		(0.467)	(0.633)		(0.500)	(0.736)
Student			-0.215			0.856
			(0.519)			(0.617)
Employment			-0.131			-0.201
			(0.350)			(0.403)
Constant	1.349***	0.553	1.614	1.829***	1.027	1.332
	(0.307)	(0.663)	(1.056)	(0.331)	(0.718)	(1.238)
Log likelihood	-317.984	-307.160	-153.043	-322.647	-309.731	-159.331
AIC	643.967	644.321	340.085	653.294	649.461	352.663
BIC	661.236	708.805	402.058	670.555	713.918	414.695
Number of observations	554	544	283	553	543	284

Notes: This table presents probit random-effects regression results for observations from periods 4-8 (gain framework; models 1G, 3G, 5G) and periods 12-16 (loss framework; models 1L, 3L, 5L). Results from models 2G, 4G, 6G, and 2L, 4L, 6L are not presented since Fine*probability interaction term was omitted because of collinearity. The dependent variable is crime (1 = Take the money, 0 = Do not take the money). Statistical significance of coefficients is reported as follows: *** p<0.01; ** p<0.05; * p<0.1. Standard errors are reported in parentheses. [a] The category Strong risk loving was omitted due to predicting failure perfectly (there was only one subject with Strong risk loving attitude). [b] The reference category is Strong risk aversion. [c] The reference category is White. The results were estimated with Stata 18.0.

Table 15: Elasticities (average marginal effects, eyex) for periods 4-8 (gain framework) and periods 12-16 (loss framework)

	(1G)	(3G)	(5G)	(1L)	(3L)	(5L)
Fine	-0.741***	-0.779***	-0.765***	-0.505***	-0.559***	-0.458***
	(0.141)	(0.150)	(0.213)	(0.103)	(0.113)	(0.142)
probability	-0.553***	-0.578***	-0.782***	-0.503***	-0.532***	-0.631***
	(0.124)	(0.133)	(0.215)	(0.103)	(0.111)	(0.160)
Wald test p-value	0.000	0.000	0.000	0.000	0.000	0.000
Number of observations	554	544	283	553	543	284

Notes: This table presents elasticities (average marginal effects, eyex) for Fine and probability for observations from periods 4-8 (gain framework; models 1G, 3G, 5G) and periods 12-16 (loss framework; models 1L, 3L, 5L), as well as Wald test p-values for differences between elasticities for Fine and probability. Results from models 2G, 4G, 6G, and 2L, 4L, 6L are not presented since Fine*probability interaction term was omitted because of collinearity. Model VCE: OIM. Statistical significance of coefficients is reported as follows: *** p<0.01; ** p<0.05; * p<0.1. Standard errors are reported in parentheses. The results were estimated with Stata 18.0.

5.2.6. Periods 1-3 and 9-11 (H3 and H3A)

Observations for periods 1-3 and 9-11 were used to test H3 and H3A. Tables 16 and 17 present the results of random-effects probit regression and elasticities of crime in relation to Fine and probability. As in the previously discussed models, the results again confirm the general negative effect of increasing p and F on crime decisions (the negative sign). However, not all elasticities are statistically significant.

In specifications involving small probabilities, there is clearly a greater deterrent effect of Fine, which is highly statistically significant in all three models with p-values<0.01. The average marginal effects of p are not only smaller in magnitude, but also less statistically significant (at the 5% level in (1S) and (3S), and at the 10% level in (5S)). Differences between elasticities for F and p are statistically significant in all cases (Wald test p-values of 0.032, 0.032, and 0.055). These results directly contradict H3 and indicate that, in the case of small probabilities, severity rather than certainty of punishment is more deterrent.

On the other hand, in the case of regressions computed for big probabilities, elasticities for Fine were found to be statistically insignificant in all three specifications. In contrast, the effect of p in two models ((1B) and (3B)), i.e. those with more observations, proved statistically significant at the 5% level. Moreover, in models (1B) and (3B) the significant elasticities for p are very large in magnitude, as a 1% increase in p results in a decrease in propensity to commit a crime of 2.646% or 3.341%, respectively. Given that in the case of big probabilities it is certainty and not severity that has a larger deterrent effect, H3A is also not confirmed.

Again surprisingly, robust significant effects of risk preferences on propensity to commit a crime were not found. In contrast, models involving big probabilities may indicate highly significant effects of ethnicity and country of residence, as well as weakly significant effects of age and gender (see section 5.2.8).

Table 16: Probit random-effects regression results for periods 1-3 (small probabilities) and periods 9-11 (big probabilities)

	Simple model	General model without F*p, student and employment	General model without F*p	Simple model	General model without F*p, student and employment	General model without F*p
	(1S)	(3S)	(5S)	(1B)	(3B)	(5B)
Fine	-0.011***	-0.011***	-0.011***	-0.0224	-0.023	-0.046
	(0.003)	(0.003)	(0.004)	(0.042)	(0.042)	(0.055)
probability	-5.131**	-5.115**	-6.240*	-5.810**	-5.690**	-5.866
	(2.385)	(2.382)	(3.253)	(2.880)	(2.862)	(3.780)
Risk preferences[a]:						
Inconsistency[b]		0.768	0.454		0.358	-0.634
		(0.518)	(0.685)		(0.611)	(0.631)
Risk aversion[b]		0.116	0.229		0.177	0.169
		(0.599)	(0.789)		(0.728)	(0.719)
Risk loving[b]		1.583*	0.251		-1.125	-3.172**
		(0.913)	(1.169)		(1.148)	(1.443)
Risk neutrality[b]		0.660	0.696		0.405	0.400
		(0.626)	(0.917)		(0.740)	(0.791)
Slight risk aversion[b]		0.666	-0.384		0.565	0.216
		(1.025)	(1.380)		(1.124)	(1.214)
Female		0.707	0.367		-0.611	-0.968*
		(0.466)	(0.577)		(0.540)	(0.522)
Age		-0.024	-0.027		-0.036	-0.063*
		(0.019)	(0.032)		(0.025)	(0.035)
UK		0.456	-0.409		1.543***	1.000
		(0.431)	(0.621)		(0.575)	(0.649)
Ethnicity:						
Asian[c]		0.320	-0.069		0.850	0.895
		(0.620)	(0.929)		(0.780)	(0.887)
Black[c]		-0.070	0.512		1.802***	1.700**
		(0.584)	(0.767)		(0.697)	(0.718)
Mixed/Other[c]		0.284	0.839		1.664*	0.310
		(0.778)	(1.142)		(0.939)	(1.093)
Student			0.034			-0.498
			(0.829)			(0.815)
Employment			-0.515			0.818
			(0.593)			(0.577)
Constant	2.115***	1.635	2.962	4.446	4.445	7.647
	(0.613)	(1.118)	(1.822)	(3.979)	(4.109)	(5.462)
Log likelihood	-187.967	-180.806	-95.819	-146.893	-132.666	-62.870
AIC	383.935	391.613	225.638	301.786	295.332	159.741
BIC	399.179	448.508	279.047	317.030	352.227	213.248
Number of observations	334	328	171	334	328	172

Notes: This table presents random-effects probit regression results for observations from periods 1-3 (small probabilities; models 1S, 3S, 5S) and periods 9-11 (big probabilities; models 1B, 3B, 5B). Results from models 2S, 4S, 6S, and 2B, 4B, 6B are not presented since Fine*probability interaction term was omitted because of collinearity. The dependent variable is crime (1 = Take the money, 0 = Do not take the money). Statistical significance of coefficients is reported as follows: *** p<0.01; ** p<0.05; * p<0.1. Standard errors are reported in parentheses. [a] The category Strong risk loving was omitted due to predicting failure perfectly (there was only one subject with Strong risk loving attitude). [b] The reference category is Strong risk aversion. [c] The reference category is White. The results were estimated with Stata 18.0.

Table 17: Elasticities (average marginal effects, eyex) for periods 1-3 (small probabilities) and periods 9-11 (big probabilities)

	(1S)	(3S)	(5S)	(1B)	(3B)	(5B)
Fine	-0.542***	-0.594***	-0.598***	-0.662	-0.887	-2.696
	(0.145)	(0.158)	(0.221)	(1.233)	(1.583)	(3.284)
probability	-0.262**	-0.286**	-0.366*	-2.646**	-3.341**	-5.320
	(0.126)	(0.137)	(0.199)	(1.318)	(1.701)	(3.536)
Wald test p-value	0.032	0.032	0.055	0.043	0.046	0.121
Number of observations	334	328	171	334	328	172

Notes: This table presents elasticities (average marginal effects, eyex) for Fine and probability for observations from periods 1-3 (small probabilities; models 1S, 3S, 5S) and periods 9-11 (big probabilities; models 1B, 3B, 5B), as well as Wald test p-values for differences between elasticities for Fine and probability. Results from models 2S, 4S, 6S, and 2B, 4B, 6B are not presented since Fine*probability interaction term was omitted because of collinearity. Model VCE: OIM. Statistical significance of coefficients is reported as follows: *** p<0.01; ** p<0.05; * p<0.1. Standard errors are reported in parentheses. The results were estimated with Stata 18.0.

5.2.7. Additional robustness checks

In addition to using six different xtprobit models discussed above, the robustness of the results was also tested in some other specifications. Both xtlogit and probit (without random effects) regressions showed no major changes in the results. Computing all xtprobit regressions without *risk_preference* variable also did not change the outcomes. Each of all 30 specifications (6 models times 5 groups of periods) was also checked by inclusion of two additional control variables: *periods_order* which controlled for the order in which subjects saw periods (no order effect was found), and *periods_answered* which attributed to each participant the number of decisions they made (from 1 to 16)[34].

Moreover, there was another additional check made, which completely excluded from the sample observations for participants who did not make a decision in all 16 periods (i.e. for those, for whom a minimum one taking decision was missing). Again, no major changes were seen in all these specifications in terms of sign, significance, or difference in deterrent effects between F and p. Therefore, the results discussed above can generally be considered robust.

[34] See section 5.2.2 and Table 10.

5.2.8. Additional findings

In addition to the main results of the experiment related to deterrent effects of F and p, some other findings can also be noticed. Table 18 shows all average marginal effects (dydx) for control variables that proved statistically significant at either 1%, 5% or 10% level. Most of those effects appeared only in one or two of the all twelve possible models[35], which is the case of *risk loving*, *female*, and *age*, as well as *Asian* and *Mixed/Other* ethnicities. Therefore, it is difficult to consider them as robust additional findings. At this point, it should be again emphasised that surprisingly in almost every case, the effect of any risk preference on propensity to commit a crime turned out to be statistically insignificant.

The only repeated additional findings relate to ethnicity and country of residence. The results of 8 out of 12 models showed that being a subject who declares *Black* ethnicity increases the probability of taking money from the charity by 15-40 p.p. compared to subjects who declare *White* ethnicity. In turn, in half of the models, a statistically significant positive average marginal effect of *UK* variable can be noted. This suggests that living in the UK increases the probability of making a positive crime decision by over a dozen p.p., compared to residing in the USA.

However, it must be underlined, that these additional findings are not visible in all models, and in the specifications where they did occur, the average marginal effects are often significant only at the 10% level. Therefore, they should be treated as possible interesting directions for further experimental research, rather than as robust evidence.

[35] There were twelve specifications where any of control variables was used. They include models (3)-(6), (3G), (5G), (3L), (5L), (3S), (5S), (3B) and (5B).

Table 18: Statistically significant average marginal effects (dydx) of control variables, found in selected regressions

	(3)	(4)	(5)	(6)	(3G)	(5G)	(3L)	(3S)	(3B)	(5B)
Risk loving[a]								0.306 *		-0.249 ***
Female										-0.147 **
Age										-0.010 *
UK	0.140 **	0.139 **			0.144 ***		0.123 *		0.215 ***	0.151 *
Asian[b]					0.185 *		0.311 ***			
Black[b]	0.147 *	0.148 *	0.296 ***	0.296 ***	0.164 *	0.404 ***			0.283 ***	0.304 **
Mixed/Other[b]							0.211 *		0.258 *	
Number of observations	1200	1200	626	626	544	283	543	328	328	172

Notes: This table presents average marginal effects (dydx) of control variables, found in selected models (variables and models that are not presented do not reveal any significant results). Model VCE: OIM. Statistical significance of coefficients is reported as follows: *** p<0.01; ** p<0.05; * p<0.1. [a] The reference category is Strong risk aversion. [b] The reference category is White. The results were estimated with Stata 18.0.

6. Discussion and conclusions

The results of the experiment generally are mixed. At the outset, it should be emphasized that Becker's (1968) deterrence hypothesis was confirmed in almost all model specifications. An increase in any of the expected sanction components negatively affected propensity to commit a crime, and the average marginal effects (eyex) of both F and p generally are highly statistically significant with p-value<0.01 (with one exception, since in the case of big probabilities' models, the effect of F was found to be insignificant). This result is not surprising and, as discussed at the beginning of this thesis, finds extensive and robust evidence in both the empirical and experimental literature.

The main focus of this study, however, has been to investigate the differences in deterrent effects of severity versus certainty of punishment. Table 19 presents

a summary of hypotheses confirmation or disconfirmation in all xtprobit regressions, while Table 20 shows which expected sanction component appeared to be more deterrent in significant models. Even though simple hypotheses testing suggested that there were no significant differences between crime rates in periods with the same expected sanction but different components of it, the more sophisticated panel data regressions showed the existence of differences between deterrent effects of F and p.

Table 19: Summary of the results of probit random-effects regressions

Periods	**Hypotheses confirmation or disconfirmation**					
Periods 1-11 (H1)	(1)	(2)	(3)	(4)	(5)	(6)
	H1	insignificant	**H1**	insignificant	**H1**	insignificant
Periods 4-8 and 12-16 (H2, H2A)	(1G), (1L)	(2G), (2L)	(3G), (3L)	(4G), (4L)	(5G), (5L)	(6G), (6L)
	H2, **H2A**	-	**H2**, **H2A**	-	**H2**, **H2A**	-
Periods 1-3 and 9-11 (H3, H3A)	(1S), (1B)	(2S), (2B)	(3S), (3B)	(4S), (4B)	(5S), (5B)	(6S), (6B)
	H3, **H3A**	-	**H3**, **H3A**	-	**H3**, insignificant	-

Notes: Red colour indicates lack of confirmation, while green – confirmation of the hypothesis.

Table 20: Summary of the results of the experiment

Periods	**Which expected sanction component is more deterrent**					
Periods 1-11 (H1)	(1)		(3)		(5)	
	probability		probability		probability	
Periods 4-8 and 12-16 (H2, H2A)	(1G)	(1L)	(3G)	(3L)	(5G)	(5L)
	Fine	Fine	Fine	Fine	probability	probability
Periods 1-3 and 9-11 (H3, H3A)	(1S)	(1B)	(3S)	(3B)	(5S)	(5B)
	Fine	probability	Fine	probability	Fine	all insignificant

According to H1, which was based on Becker's (1968) rational choice economic model of crime, the Fine size should appear to be more deterrent. However, even though most participants (as conjectured) turned out to be risk averse, the models in which statistical significance of differences between the effects of F and p was noted, clearly indicate a greater deterrent effect of the certainty of punishment. Therefore, this result contradicts the theoretical basis of H1. At the same time, it is consistent with general findings obtained from empirical criminal data analyses, and contradictory to the results of most previous experimental studies.

Hypotheses 2 and 2A were based on prospect theory and loss aversion (Kahneman & Tversky, 1979) and predicted that there would be a noticeably greater increase in deterrent effect of p in the loss framework compared to the gain domain. H2, according to which in the loss framework the effect of p on crime was expected to even exceed the effect of F, was generally not confirmed. Its confirmation can only be seen in model (5L). However, this result is likely due to the small number of observations in that specification. This reasoning seems to be confirmed also by the corresponding model for the gain framework (5G), in which the greater deterrent effect of p was also shown. In all other models with a larger number of observations, in both gain and loss domains, it was severity that proved to be more deterrent, which contradicts H2.

On the other hand, H2A was confirmed. This is because it turned out that in the loss framework, the greater deterrent effect should be attributed to F, but the difference between the effects of F and p is noticeably smaller than in the corresponding periods from the gain framework. In the loss domain, differences between severity and certainty effects are so small that they even can be considered economically insignificant. Accordingly, this clearly shows that in the loss framework indeed p has gained in importance, compared to the gain domain.

Regarding H3 and H3A, the results obtained are contradictory to those predicted by the prospect theory and overweighting small probabilities bias (Tversky & Kahneman, 1981). In the case of small probabilities, it was severity that proved to be significantly more deterrent than certainty, which is opposite to H3. This may suggest that contradictory to theory, in periods with small probabilities people did not attribute higher perceived weights to p and became rather more risk averse than risk loving. Thus, it seems that subjects underweighted rather than overweighted small probabilities. This phenomenon may be explained, for example, by the previously mentioned overconfidence bias or overoptimism (Williams, 2014).

On the other hand, in the case of big probabilities, it was p that appeared to have a far greater deterrent effect, which directly contradicts the validity of H3A. The effect of F proved to be statistically insignificant even at the 10% level. This in turn indicates that big probabilities were overweighted. Perhaps the probabilities of getting caught equal to 80% and 90% were perceived by participants as a practical certainty of being punished, which could also be explained by other behavioural insights[36].

When analysing the above results, it is important to bear in mind the very high proportion of subjects in the sample for whom risk preference could not be determined. Approximately 30% of participants were inconsistent in their choices in Block 1 of the experiment. This exceptionally high amount may, but does not necessarily, suggest that some subjects were inaccurate in their decision-making during the study. However, the complete exclusion of observations of inconsistent participants would result in a significantly smaller sample and low representativeness of the results of the study, which after all, was not focused on risk attitudes, but on severity versus certainty of punishment. For the future, it may be suggested to use a different method

[36] See for instance Jolls (2004) or Williams (2014), and the literature discussed therein.

of determining participants' risk preferences, which makes it more difficult to be inconsistent in their choices[37].

In conclusion, the significant contribution of this study to the existing literature should be highlighted. It appears to be the first experiment of this kind which explicitly used behavioural insights to explore the differences between severity versus certainty of punishment. Obviously, the results of this study do not conclusively resolve whether it is the Fine size or the probability of getting caught that has a greater deterrent effect. They do, however, clearly indicate that there are significant differences between loss and gain frameworks, as well as between small and large probabilities of getting caught, in relation to severity versus certainty of punishment issue.

Accordingly, further experiments on this topic should not only test the robustness of the results obtained in this study, but also link other known behavioural insights to one of the key problems of economic theory of crime, which is related to severity versus certainty issue. A broader understanding of behavioural aspects in relation to deterrent effects of F and p can influence a better design of criminal policies. This will allow a more effective deterrence of potential criminals, which is one of the basic functions of criminal law.

[37] See for instance Meraner et al. (2018).

Bibliography

Abramovaite, J., Bandyopadhyay, S., Bhattacharya, S., & Cowen, N. (2023). Classical deterrence theory revisited: An empirical analysis of Police Force Areas in England and Wales, "European Journal of Criminology", 20(5), pp. 1663-1680. https://doi.org/10.1177/14773708211072415

Ainslie, G., & Haslam, N. (1992). Hyperbolic Discounting, in: Loewenstein, G., & Elster, J. (Eds.), *Choice Over Time*, Russell Sage Foundation Publications, New York, pp. 57-92.

Alm, J., Jackson, B.R., & McKee, M. (1992). Estimating the Determinants of Taxpayer Compliance with Experimental Data, "National Tax Journal", 45(1), pp. 107-114. https://doi.org/10.1086/NTJ41788949

Alm, J., Sanchez, I., & De Juan, A. (1995). Economic and Noneconomic Factors in Tax Compliance, "Kyklos", 48(1), pp. 1-18. https://doi.org/10.1111/j.1467-6435.1995.tb02312.x

Anderson, L.R., & Stafford, S.L. (2003). Punishment in a Regulatory Setting: Experimental Evidence from the VCM, "Journal of Regulatory Economics", 24(1), pp. 91-110. https://doi.org/10.1023/A:1023952115422

Astebro, T., Mata, J., & Santos-Pinto, L. (2015). Skewness Seeking: Risk Loving, Optimism or Overweighting of Small Probabilities?, "Theory and Decision", 78(2), pp. 189-208. https://doi.org/10.1007/s11238-014-9417-4

Bahnik, S., & Vranka, M.A. (2022). Experimental test of the effects of punishment probability and size on the decision to take a bribe, "Journal of Behavioral and Experimental Economics", 97(1). https://doi.org/10.1016/j.socec.2021.101813

Bates, K.B., & Peynircioglu, Z.F. (2017). Proximity to Task: A Prerequisite for Two Cognitive Biases, "The American Journal of Psychology", 130(4), pp. 477-491. https://doi.org/10.5406/amerjpsyc.130.4.0477

Beccaria, C. (1764). An Essay on Crimes and Punishments. Available at: https://oll.libertyfund.org/titles/voltaire-an-essay-on-crimes-and-punishments (access: 30th July 2024).

Becker, G.S. (1968). Crime and Punishment: An Economic Approach, "Journal of Political Economy", 76(2), pp. 169-217. https://doi.org/10.1086/259394

Benson, B.L., & Zimmerman P.R. (Eds.). 2010. Handbook on the Economics of Crime, Edward Elgar Publishing, Cheltenham. https://doi.org/10.4337/9781849806206

Bentham, J. (1789). An Introduction to the Principles of Morals and Legislation. Available at: https://oll.libertyfund.org/titles/bentham-an-introduction-to-the-principles-of-morals-and-legislation (access: 30th July 2024).

Blackwell, C. (2010). A Meta-analysis of Incentive Effects in Tax Compliance Experiments, in: Alm, J., Martinez-Vazquez, J., & Torgler, B. (Eds.), *Developing Alternative Frameworks for Explaining Tax Compliance*, Routledge International Studies in Money and Banking, New York, pp. 97-112.

Block, M.K., & Gerety, V.E. (1995). Some Experimental Evidence on Differences between Student and Prisoner Reactions to Monetary Penalties and Risk, "The Journal of Legal Studies", 24(1), pp. 123-128. https://doi.org/10.1086/467954

Brandstätter, E., Kuehberger A., & Schneider F.G. (2002). A Cognitive-Emotional Account of the Shape of the Probability Weighting Function, "Journal of Behavioral Decision Making", 15(2), pp. 79-100. https://doi.org/10.1002/bdm.404

Brañas-Garza, P., Estepa-Mohedano, L., Jorrat, D., Orozco, V., & Rascón-Ramirez, E. (2021). To pay or not to pay: Measuring risk preferences in lab and field, "Judgment and Decision Making", 16(5), pp. 1290-1313. https://doi.org/10.1017/S1930297500008433

Camilleri, A.R., & Newell, B.R. (2011). When and why rare events are underweighted: A direct comparison of the sampling, partial feedback, full feedback and description choice paradigms, "Psychonomic Bulletin & Review", 18(2), pp. 377-384. https://doi.org/10.3758/s13423-010-0040-2

Chalfin, A., & McCrary, J. (2017). Criminal Deterrence: A Review of the Literature, "Journal of Economic Literature", 55(1), pp. 5-48. https://doi.org/10.1257/jel.20141147

Chalfin, A., & McCrary, J. (2018). Are U.S. Cities Underpoliced? Theory and Evidence, "The Review of Economics and Statistics", 100(1), pp. 167-186. https://doi.org/10.1162/REST_a_00694

Davis, M.L. (1988). Time and Punishment: An Intertemporal Model of Crime, "Journal of Political Economy", 96(2), pp. 383-390. https://doi.org/10.1086/261542

DeAngelo, G., & Charness, G. (2012). Deterrence, expected cost, uncertainty and voting: Experimental evidence, "Journal of Risk and Uncertainty", 44(1), pp. 73-100. https://doi.org/10.1007/s11166-011-9131-3

Dominioni, G. (2020). Biased Trials: Insights from Behavioral Law and Economics, Springer Gabler, Wiesbaden. https://doi.org/10.1007/978-3-658-30080-7

Doob, A.N., & Webster, C.M. (2003). Sentence Severity and Crime: Accepting the Null Hypothesis, "Crime and Justice", 30(1), pp. 143-195. https://doi.org/10.1086/652230

Dölling, D., Entorf, H., Hermann, D., & Rupp, T. (2011). Meta-Analysis of Empirical Studies on Deterrence, in: Kury, H., & Shea, E. (Eds.), *Punitivity International Developments Vol. 3: Punitiveness and Punishment*, Brockmeyer, Bochum, pp. 315-378.

Drichoutis, A.C., & Koundouri, P. (2012). Estimating Risk Attitudes in Conventional and Artefactual Lab Experiments: The Importance of the Underlying Assumptions, Economics Discussion Paper No. 2012-19. https://doi.org/10.2139/ssrn.2035308

Durlauf, S.N., & Nagin, D.S. (2011). Imprisonment and Crime: Can Both Be Reduced?, "Criminology and Public Policy", 10 (1), pp. 13–54. https://doi.org/10.1111/j.1745-9133.2010.00680.x

Earnhart, D., & Friesen, L. (2023), Certainty of Punishment versus Severity of Punishment: Enforcement of Environmental Protection Laws, "Land Economics", 99(2), pp. 245-264. https://doi.org/10.3368/le.030521-0024R1

Ehrlich, I. (1996). Crime, Punishment, and the Market for Offenses, "Journal of Economic Perspectives", 10(1), pp. 43-67. https://doi.org/10.1257/jep.10.1.43

Eide, E. (2000). Economics of Criminal Behavior, in: Bouckaert, B., & De Geest, G. (Eds.), "Encyclopedia of Law and Economics", Edward Elgar Publishing, Cheltenham, Northampton, pp. 345-389.

Engel, C., & Nagin, D. (2015). Who is Afraid of the Stick? Experimentally Testing the Deterrent Effect of Sanction Certainty, "Review of Behavioral Economics", 2(4), pp. 405-434. http://doi.org/10.1561/105.00000037

Entorf, H., & Spengler, H. (2015). Crime, Prosecutors, and the Certainty of Conviction, "European Journal of Law and Economics", 39(1), pp. 167-201. https://doi.org/10.1007/s10657-012-9380-x

Evans, W.N., & Owens, E.G. (2007). COPS and crime, "Journal of Public Economics", 91(1-2), pp. 181-201. https://doi.org/10.1016/j.jpubeco.2006.05.014

Farmer, G.D., Warren, P.A., & Hahn, U. (2017). Who "Believes" in the Gambler's Fallacy and Why?, "Journal of Experimental Psychology: General", 146(1), pp. 63-76. https://doi.org/10.1037/xge0000245

Feess, E., & Sarel, R. (2018). Judicial Effort and the Appeals System: Theory and Experiment, "The Journal of Legal Studies", 47(2), pp. 269-294. http://doi.org/10.1086/699391

Feess, E., Schildberg-Hörisch, H., Schramm, M., & Wohlschlegel, A. (2018). The impact of fine size and uncertainty on punishment and deterrence: Theory and evidence from the laboratory, "Journal of Economic Behavior & Organization", 149(1), pp. 58-73. https://doi.org/10.1016/j.jebo.2018.02.021

Friedland, N., Maital, S., & Rutenberg, A. (1978). A Simulation Study of Income Tax Evasion, "Journal of Public Economic", 10(1), pp. 107-116. https://doi.org/10.1016/0047-2727(78)90008-7

Friesen, L. (2012). Certainty of Punishment versus Severity of Punishment: An Experimental Investigation, "Southern Economic Journal", 79(2), pp. 399-421. https://doi.org/10.4284/0038-4038-2011.152

Giamattei, M., & Lambsdorff, J.G. (2019). classEx – an online tool for lab-in-the-field experiments with smartphones, "Journal of Behavioral and Experimental Finance", 22(1), pp. 223-231. https://doi.org/10.1016/j.jbef.2019.04.008

Grogger, J. (1991). Certainty vs. Severity of Punishment, "Economic Inquiry", 29(2), pp. 297-309. https://doi.org/10.1111/j.1465-7295.1991.tb01272.x

Grossman, P.J., & Eckel, C.C. (2015). Giving versus taking for a cause, "Economics Letters", 132(1), pp. 28-30. https://doi.org/10.1016/j.econlet.2015.04.002

Guala, F. (2005), The Methodology of Experimental Economics, Cambridge University Press, Cambridge. https://doi.org/10.1017/CBO9780511614651

Haidt, J. (2003). The Moral Emotions, in: Davidson, R.J., Scherer, K.R., & Goldsmith, H.H. (Eds.), *Handbook of Affective Sciences*, Oxford University Press, Oxford, pp. 852-870.

Harbaugh, W.T., Mocan, N.H., & Visser, M. (2013). Theft and Deterrence, "Journal of Labor Research", 34(4), pp. 389-407. https://doi.org/10.1007/s12122-013-9169-x

Harel, A. (2014). Behavioral Analysis of Criminal Law: A Survey, in: Zamir, E., & Teichman, D. (Eds.), *The Oxford Handbook of Behavioral Economics and the Law*, Oxford University Press (Academic), Oxford, pp. 568-598. https://doi.org/10.1093/oxfordhb/9780199945474.013.0022

Helland, E., & Tabarrok, A. (2007). Does Three Strikes Deter? A Nonparametric Estimation, "The Journal of Human Resources", 42(2), pp. 309-330. https://doi.org/10.2307/40057307

Holt, C.A., & Laury, S.K. (2002). Risk Aversion and Incentive Effects, "The American Economic Review", 92(5),pp. 1644-1655. https://doi.org/10.1257/000282802762024700

Idzikowska, K., Muda, R., Kołodziej, S., & Zielonka, P. (2017). Overweighting versus underweighting of small probabilities, in: Tyszka, T., & Zielonka, P. (Eds.), *Large Risks with Low Probabilities: Perceptions and Willingness to Take Preventive Measures Against Flooding*, IWA Publishing, London, pp. 41-58. https://doi.org/10.2166/9781780408606

Jacob, A. (2011). Economic Theories of Crime and Delinquency, "Journal of Human Behavior in the Social Environment", 21(3), pp. 270-283. https://doi.org/10.1080/10911359.2011.564951

Jolls, C., Sunstein, C.R., & Thaler, R.H. (1998). A Behavioral Approach to Law and Economics, "Stanford Law Review", 50(5), pp. 1471-1550. https://doi.org/10.2307/1229304

Jolls, C. (2004). On Law Enforcement with Boundedly Rational Actors, "SSRN Electronic Journal". https://doi.org/10.2139/ssrn.631222

Jolls, C. (2007). Behavioral Law and Economics, "NBER Working Paper Series", Working Paper 12879. https://doi.org/10.3386/w12879

Kahneman, D., & Tversky, A. (1979). Prospect Theory: An Analysis of Decision under Risk, "Econometrica", 47(2), pp. 263-292. https://doi.org/10.2307/1914185

Kantorowicz-Reznichenko, E. (2015). Any-Where Any-Time: Ambiguity and the Perceived Probability of Apprehension, "UMKC Law Review", 84(1), pp. 27-59. http://doi.org/10.2139/ssrn.2202329

Khadjavi, M. (2015). On the Interaction of Deterrence and Emotions, "The Journal of Law, Economics, and Organization", 31(2), pp. 287-319. https://doi.org/10.1093/jleo/ewu012

Khadjavi, M. (2018). Deterrence works for criminals, "European Journal of Law and Economics", 46(1), pp. 165-178. https://doi.org/10.1007/s10657-015-9483-2

Kirchler, M., Huber, J., Stefan, M., & Sutter, M. (2016), Market Design and Moral Behavior, "Management Science", 62(9), pp. 2615-2625. https://doi.org/10.1287/mnsc.2015.2246

Krafft, C. (2019). The Economics of Crime, in: Krafft, C., *Economics for the Greater Good: An Introduction to Economic Thinking for Public Policy*, Minnesota Libraries Publishing Project, pp. 64-80. Available at: https://mlpp.pressbooks.pub/economicsforthegreatergood/ (access: 30th July 2024).

Laske, K., & Saccardo, S., & Gneezy, U. (2018). Do Fines Deter Unethical Behavior? The Effect of Systematically Varying the Size and Probability of Punishment, "SSRN Electronic Journal". https://doi.org/10.2139/ssrn.3157387

Levitt, S.D. (1997). Why Do Increased Arrest Rates Appear to Reduce Crime: Deterrence, Incapacitation, or Measurement Error?, "Economic Inquiry", 34(3). https://ssrn.com/abstract=10294

Lin, M.J. (2009). More Police, Less Crime: Evidence from US State Data, "International Review of Law and Economics", 29(2), pp. 73-80. https://doi.org/10.1016/j.irle.2008.12.003

Loewenstein, G. (2000). Emotions in Economic Theory and Economic Behavior, "The American Economic Review", 90(2), pp. 426-432. https://doi.org/10.1257/aer.90.2.426

Mamak, K., Barczak-Oplustil, A., Kwiatkowski, D., Małecki, M., & Zając, D. (2022). Should gains from criminal knowledge be forfeited?, „Crime, Law and Social Change", 77(1), pp. 305-320. https://doi.org/10.1007/s10611-021-09994-7

Mathis, K. (Ed.). (2015). European Perspectives on Behavioural Law and Economics, Springer International Publishing, Lucerne. https://doi.org/10.1007/978-3-319-11635-8

McAdams, R.H., & Ulen, T.S. (2008). Behavioral Criminal Law and Economics, "John M. Olin Program in Law and Economics Working Papers", Working Paper No. 440. https://doi.org/10.2139/ssrn.1299963

Mendes, S.M. (2004). Certainty, Severity, and Their Relative Deterrent Effects: Questioning the Implications of the Role of Risk in Criminal Deterrence Policy, "Policy Studies Journal", 32(1), pp. 59-74. https://doi.org/10.1111/j.0190-292X.2004.00053.x

Meraner, M., Musshoff, O., & Finger, R. (2018). Using Involvement to Reduce Inconsistencies in Risk Preference Elicitation, "Journal of Behavioral and Experimental Economics", 73(1), pp. 22-33. https://doi.org/10.1016/j.socec.2018.01.001

Mungan, M.C., & Klick, J. (2014). Forfeiture of Illegal Gains, Attempts, and Implied Risk Preferences, "The Journal of Legal Studies", 43(1), pp. 137-153. https://doi.org/10.1086/675351

Mungan, M.C. (2017). The certainty versus the severity of punishment, repeat offenders, and stigmatization, "Economics Letters", 150(C), pp. 126-129. https://doi.org/10.1016/j.econlet.2016.11.030

Murphy, J.J., & Stranlund, J.K. (2007). A laboratory investigation of compliance behavior under tradable emissions rights: Implications for targeted enforcement, "Journal of Environmental Economics and Management", 53(2), pp. 196-212. https://doi.org/10.1016/j.jeem.2006.08.003

Nagin, D.S., & Pogarsky, G. (2003). An Experimental Investigation of Deterrence: Cheating, Self-serving Bias, and Impulsivity, "Criminology", 41(1), pp. 167-194. https://doi.org/10.1111/j.1745-9125.2003.tb00985.x

Nagin, D.S. (2013). Deterrence: A Review of the Evidence by a Criminologist for Economists, "Annual Review of Economics", 5(1), pp. 83-105. https://doi.org/10.1146/annurev-economics-072412-131310

Oliver, A. (2024), Reflecting on Reflection: Prospect Theory, Our Behaviors, and Our Environment, "Behavioural Public Policy", 8(1), pp. 173–183. https://doi.org/10.1017/bpp.2021.31

Park, C.G., & Hyun, J.K. (2003). Examining the determinants of tax compliance by experimental data: A case of Korea, "Journal of Policy Modeling", 25(8), pp. 673-684. https://doi.org/10.1016/S0161-8938(03)00075-9

Petrova, D.G., Van der Pligt, J., & Garcia-Retamero, R. (2014). Feeling the Numbers: On the Interplay Between Risk, Affect, and Numeracy, "Journal of Behavioral Decision Making", 27(3), pp. 191-199. https://doi.org/10.1002/bdm.1803

Pickett, J.T. (2018). Using Behavioral Economics to Advance Deterrence Research and Improve Crime Policy: Some Illustrative Experiments, "Crime & Delinquency", 64(12), pp. 1636-1659. https://doi.org/10.1177/0011128718763136

Piza, E.L., Welsh, B.C., Farrington, D.P., & Thomas, A.L. (2019). CCTV surveillance for crime prevention: A 40-year systematic review with meta-analysis, "Criminology & Public Policy", 18(1), pp. 135–159. http://doi.org/10.1111/1745-9133.12419

Polinsky, A.M., & Shavell, S. (1999). On the Disutility and Discounting of Imprisonment and the Theory of Deterrence, "The Journal of Legal Studies", 28(1), pp. 1-16. https://doi.org/10.1086/468044

Polinsky, A.M., & Shavell, S. (2000). The Economic Theory of Public Enforcement of Law, "Journal of Economic Literature", 38(1), pp. 45-76. https://doi.org/10.1257/jel.38.1.45

Prolific (2024). We're revolutionizing how the world learns about people, so people can revolutionize the world. Available at: https://www.prolific.com/about (access: 30th July 2024).

Pyne, D. (2012). Deterrence: Increased enforcement versus harsher penalties, "Economics Letters", 117(3), pp. 561-562. https://doi.org/10.1016/j.econlet.2012.07.029

Rachlinski, J.J., Johnson, S.L., Wistrich, A.J., & Guthrie, C. (2009). Does Unconscious Racial Bias Affect Trial Judges?, "SSRN Electronic Journal", 84(3). https://doi.org/10.2139/ssrn.999490

Schildberg-Hörisch, H., & Strassmair, C. (2012). An Experimental Test of the Deterrence Hypothesis, "The Journal of Law, Economics, and Organization", 28(3), pp. 447-459. https://doi.org/10.1093/jleo/ewq015

Shiffman, G.M. (2020). The Economics of Violence: How Behavioral Science Can Transform Our View of Crime, Insurgency, and Terrorism, Cambridge University Press, Cambridge. https://doi.org/10.1017/9781316136072

Sun, Q., Lu, J., Zhang, H., & Liu, Y. (2021). Social Distance Reduces the Biases of Overweighting Small Probabilities and Underweighting Large Probabilities, "Personality and Social Psychology Bulletin", 47(8), pp. 1309–1324. https://doi.org/10.1177/0146167220969051

Sunstein, C.R. (Ed.). (2000). Behavioral Law and Economics, Cambridge University Press, Cambridge. https://doi.org/10.1017/CBO9781139175197

Teodorescu, K., Plonsky, O., Ayal, S., & Barkan, R. (2021). Frequency of enforcement is more important than the severity of punishment in reducing violation behaviors, "Proceedings of the National Academy of Sciences", 118(42), pp. 1-10. https://doi.org/10.1073/pnas.2108507118

Tversky, A., & Kahneman, D. (1981). The Framing of Decisions and the Psychology of Choice, "Science", 211(4481), pp. 453-458. https://doi.org/10.1126/science.7455683

Tversky, A., & Kahneman, D. (1982). Availability: A heuristic for judging frequency and probability, in: Kahneman, D., Slovic, P., & Tversky, A. (Eds.), *Judgment under Uncertainty: Heuristics and Biases,* Cambridge University Press, Cambridge, pp. 163-178.

Tversky, A., & Kahneman, D. (1992). Advances in Prospect Theory: Cumulative Representation of Uncertainty, "Journal of Risk and Uncertainty", 5(4), pp. 297-323. https://doi.org/10.1007/BF00122574

Van Rooij, B., Kuiper, M.E., & Piquero, A. (2024). How Punishment Affects Crime: An Integrated Understanding of the Behavioral Mechanisms of Punishment, "SSRN Electronic Journal". https://doi.org/10.2139/ssrn.4790878

Van Winden, F.A.A.M., & Ash, E. (2012). On the Behavioural Economics of Crime, "Review of Law & Economics", 8(1), pp. 181-213. https://doi.org/10.1515/1555-5879.1591

Williams, S.H. (2014). Probability Errors: Overoptimism, Ambiguity Aversion, and the Certainty Effect, in: Zamir, E., & Teichman, D. (Eds.), *The Oxford Handbook of Behavioral Economics and the Law*, Oxford University Press (Academic), Oxford, pp. 335-353. https://doi.org/10.1093/oxfordhb/9780199945474.013.0013

Zamir, E., & Teichman, D. (Eds.). (2018). Behavioral Law and Economics, Oxford University Press, Oxford. https://doi.org/10.1093/oso/9780190901349.001.0001

Appendix A

Full details of the experiment, such as full instructions, all screens seen by the participants, payments information, data on subjects' decisions and demographic data are available on the Open Science Framework (OSF) platform: https://osf.io/eu7zt/?view_only=961cc2ab3075480c82fb5fc8561547b7