

IZA DP No. 9388

**The Impact of Fine Size and Uncertainty on
Punishment and Deterrence:
Theory and Evidence from the Laboratory**

Eberhard Feess
Hannah Schildberg-Hörisch
Markus Schramm
Ansgar Wohlschlegel

September 2015

The Impact of Fine Size and Uncertainty on Punishment and Deterrence: Theory and Evidence from the Laboratory

Eberhard Feess

Frankfurt School of Finance and Management

Hannah Schildberg-Hörisch

University of Bonn and IZA

Markus Schramm

RWTH Aachen University

Ansgar Wohlschlegel

University of Portsmouth

Discussion Paper No. 9388

September 2015

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-180
E-mail: iza@iza.org

Any opinions expressed here are those of the author(s) and not those of IZA. Research published in this series may include views on policy, but the institute itself takes no institutional policy positions. The IZA research network is committed to the IZA Guiding Principles of Research Integrity.

The Institute for the Study of Labor (IZA) in Bonn is a local and virtual international research center and a place of communication between science, politics and business. IZA is an independent nonprofit organization supported by Deutsche Post Foundation. The center is associated with the University of Bonn and offers a stimulating research environment through its international network, workshops and conferences, data service, project support, research visits and doctoral program. IZA engages in (i) original and internationally competitive research in all fields of labor economics, (ii) development of policy concepts, and (iii) dissemination of research results and concepts to the interested public.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

ABSTRACT

The Impact of Fine Size and Uncertainty on Punishment and Deterrence: Theory and Evidence from the Laboratory

We develop a theoretical model to identify and compare partial and equilibrium effects of uncertainty and the magnitude of fines on punishment and deterrence. Partial effects are effects on potential violators' and punishers' decisions when the other side's behavior is exogenously given. Equilibrium effects account for the interdependency of these decisions. This interdependency is important since, in the case of legal uncertainty, higher fines may reduce the willingness to punish, which in turn reduces the deterrence effect of high fines. Using a laboratory experiment, we identify these effects empirically by means of a strategy-method design in which potential violators can condition their behavior on the behavior of potential punishers and vice versa. All our experimental findings on both partial and equilibrium effects are in line with the hypotheses derived from the theory.

JEL Classification: K14, K42, C91, D64

Keywords: deterrence, punishment, legal uncertainty, fines, laboratory experiment, partial and equilibrium effects

Corresponding author:

Hannah Schildberg-Hörisch
Department of Economics
University of Bonn
Adenauerallee 24-42
53113 Bonn
Germany
E-mail: schildberg-hoerisch@uni-bonn.de

1 Introduction

Two fundamental insights of the economic analysis of law are that deterrence is increasing in the magnitude of fines (Becker, 1968) and the accuracy of the court system (Png, 1986, Polinsky and Shavell, 2000). For the latter point, it has been noted that both higher frequencies of convicting innocent defendants (type-I errors) and higher percentages of releasing guilty defendants (type-II errors) have detrimental effects as the incentives to obey the law depend on the difference in the expected fine with and without violating the law.

The deterrence effect of higher fines is straightforward when type-I and type-II errors are exogenously given, which is usually assumed in the literature.¹ In reality, however, the frequencies of these two error types do not only depend on the evidence and the accuracy of the court system, but also on the relative weights judges and juries put on them: For a given quality of evidence, which is a noisy signal of the suspects' actual behavior, legal decision makers can reduce type-I errors at the expense of higher type-II errors and vice versa. In particular, legal decision makers' aversion against type-I errors may increase in the fine size. If this effect is strong enough and anticipated by potential violators, then higher fines may even lead to lower deterrence (Andreoni, 1991). To understand the impact of fines and uncertainty on deterrence, it is hence important to take the interdependency of the decisions by potential violators and legal decision makers into account.

We are the first to develop a theoretical model and a laboratory experiment to distinguish between the *partial* and the *equilibrium* effects of fines and uncertainty: For partial effects, we follow the overwhelming part of the literature which considers their impact on the behavior of one "market side" (i.e. violators *or* judges) under the simplifying assumption that the behavior of the other market side is exogenously given. For instance, we analyze the impact of the fine on the violation frequency when the punishment frequency is independent of the fine. Our analysis then confirms the intuition that higher fines reduce both the violation and the punishment frequency, while higher uncertainty increases the violation, but reduces the punishment frequency.

Things are less straightforward, however, when we consider equilibrium effects. In equilibrium, judges are aware that higher legal uncertainty increases the incentives to violate the law, which in turn increases the punishment incentive compared to the case in which the number of violators is exogenously given. Hence, equilibrium effects may well differ from partial effects; not only with respect to the size, but even with respect to the sign. Our theoretical model allows to capture and to compare partial and equilibrium effects. We find that, even in equilibrium, the punishment frequency of judges decreases in the fine size, and the violation frequency increases in legal uncertainty. The impacts of the fine on the violation frequency and of uncertainty on the punishment frequency, however, are theoretically ambiguous.

Based on our theoretical model, we have designed a laboratory experiment that allows us to distinguish between the partial and the equilibrium effects of legal uncertainty and the magnitude of fines on violation and punishment. Participants are randomly assigned to one of two roles, *V* ("violators") and *J* ("judges").² Next, pairs consisting of one participant *V* and one participant *J* are matched. All participants are informed that, for each pair, there is a fixed amount of money designated for donation to charity. Participant *V* can take this money for himself. If it is not taken, it may still disappear due to a random event, and this creates legal uncertainty. Each participant *J* can punish his randomly assigned participant *V* if and only if he observes that the money has disappeared.

We consider two fine levels and two degrees of uncertainty, which leads to four different decision situations. To distinguish between partial and equilibrium effects, we employ a strategy method design in which players

¹See the overview in Polinsky and Shavell (2000, Section 15).

²In order to avoid framing, the participant types were denoted *A* and *B* in the instructions, which are provided in the Appendix.

of one type can condition their behavior on the behavior of the other types. We have two treatments. In the first treatment, the strategy method is applied for players J , while players V decide unconditionally whether to take the money or not. In the second treatment, the strategy method applies to players V , and players J decide unconditionally whether to punish or not to punish. This procedure allows us to calculate partial as well as equilibrium effects: For instance, the partial effect of the fine size on the players J 's behavior is calculated by comparing the punishment frequencies that would have occurred if the violation frequencies were the same for different fine sizes. For the equilibrium effect, we compare the punishment frequencies for the fine-dependent actual violation frequencies.

Note that the frequencies of type-I and type-II errors are determined endogenously by the decisions of players V and J , so that the interdependency of violation and punishment decisions is fully captured by our experiment. In particular, for a given probability that the money disappears by a random event, the risk of making a type-I error in case of punishment is strictly decreasing in the number of actual violators. Analogously, the risk of making a type-II error in case of no punishment increases in the violation frequency. Thus, for given behavior of judges, the violation frequency in the strategy method translates each number of violators into a probability for error types.

Our experimental results are fully in line with the theoretical predictions. For partial effects, we confirm the literature that higher fines reduce both the violation and the punishment frequency. Uncertainty increases the number of violations and decreases the punishment frequency. As outlined above, equilibrium effects are theoretically less straightforward, but our empirical findings are again in line with our model: First, a higher fine leads to a significant reduction in the punishment frequency if and only if there is legal uncertainty. Second, the equilibrium impact of legal uncertainty on the punishment frequency is lower than the partial effect, and this holds for both fine levels. The intuition for this result is that the violation incentive is larger for the case with uncertainty, and this reduces players J 's incentive to punish less in response to higher uncertainty. Third, the equilibrium impact of the fine size on the violation frequency is less pronounced than the partial impact, which can be attributed to the fact that players J punish less often. Fourth, legal uncertainty leads to a pronounced incline in the violation frequency even in equilibrium, which reinforces the view that legal uncertainty is worrisome not only from a justice point of view, but also from a deterrence perspective. Last, comparing partial and equilibrium effects shows that the difference is sizable, especially for changes in the magnitude of sanctions. Thus, our results support the case for legislation to take equilibrium effects into account.

So far, most experimental papers on punishments have adopted voluntary contribution mechanisms (VCMs), in which the participants can mutually sanction non-cooperative behavior (e.g., Fehr and Gächter, 2000). There are only a few VCM experiments with punishment which assume noisy signals or vary the magnitude of fines. Grechenig, Nicklisch, and Thöni (2010) assume that participants get noisy signals on their mutual contributions in a VCM game and find that higher noise, which increases the risk of type-I errors in case of penalties, does not reduce the punishment frequency. Social welfare, however, shrinks due to retaliation of those who are innocently punished and because fines are costly. While fines are kept constant in Grechenig, Nicklisch, and Thöni (2010), Ambrus and Greiner (2012) consider different fine levels in a standard VCM game. With low fines, people do not punish often, thereby saving on punishment costs. When fines are high, there is a large deterrence effect which improves social efficiency. As both of these beneficial effects are small for intermediate levels of punishment, efficiency is U-shaped in the magnitude of fines. This is related to results in Nikiforakis and Normann (2008) and Egas and Riedl (2008) on the impact of fine sizes on punishments and contributions in VCM settings. Rege and Telle (2004) find that revealing the identities of non-contributors increases cooperation, and Nikiforakis (2008) show that the punishment frequency decreases sharply when counter-punishments are feasible.

In contrast to these VCM settings, our research question requires a setup with a clear-cut distinction

between violators and legal decision makers: each participant is assigned to just one role; as "judge" or "violate". Moreover, legal fines better resemble a situation with third-party punishment, in which those who can punish are not directly affected by the behavior of violators. Additionally, we assume that fines are costless for those who impose them, since judges themselves do not bear the social costs of punishment. Finally, VCM games are, by definition, about voluntary payments, whereas taking money designated for donation in our setting comes closer to legal infringements.

Some papers adopt third party punishment, but neither consider uncertainty nor the interdependence of violation and punishment. Fehr and Fischbacher (2004) compare third and second party punishment in response to the behavior in a dictator game. Selfish behavior is punished under both circumstances, but to a larger degree by affected second parties. Carpenter (2007) finds that punishment decreases in the cost of punishment, and similar results are reported by Anderson and Putterman (2006) for VCM-settings. Leibbrandt and López-Pérez (2011) consider third party punishment in ten variants of a game in which one party can decide on different allocations of money. Parties allocating a larger amount to themselves are frequently punished, even in cases in which the resulting allocation is efficient as it maximizes the overall amount. As in our paper, Baumann and Friehe (2015) use a setting in which potential violators may take some money and can be punished by a third party. Focusing on the equilibrium effect only, they confirm our result that potential violators are more likely to take the money when legal uncertainty is high. However, they do not find any differences in the punishment probability across groups with different levels of sanctions. Since they do not use a strategy-method approach, they cannot compare equilibrium effects with partial effects.

In order to test Becker's (1968) deterrence hypothesis in the laboratory, Harbaugh, Moanan, and Visser (2013), Khadjavi (2015), Rizzolli and Stanca (2012), and Schildberg-Hörisch and Strassmair (2012) use take game designs in which participants can "steal" from their counterparts' endowments with *exogenous punishment schemes* (fine sizes and probabilities of punishment). Overall, these papers confirm that the violation frequency is decreasing in fines and probability of punishment. However, Schildberg-Hörisch and Strassmair (2012) and Khadjavi (2015) find that small sanctions reduce deterrence compared to no sanctions which might be explained by a crowding out of intrinsic motivation for socially appreciated behavior.³ Several more papers test the effect of exogenous punishment schemes using various laboratory setups. Abbink, Irlenbusch, and Renner (2002) document the deterrent effect of probabilistic sanctions in an experiment on corruption, Nagin and Pogarsky (2003) for cheating on an incentivized quiz. Friesen (2012) uses a design in which subjects receive exogenous profits and have to decide whether to comply with a regulation at some certain cost or to violate at no immediate cost but facing the possibility of being inspected and subsequently fined. She finds that increasing the severity of punishment is a more effective deterrent than an equivalent increase in the probability of punishment. Finally, DeAngelo and Charness (2012) use a roadway speeding framing and show that committed violations decrease in the expected cost of speeding and in uncertainty about the probability of punishment. In contrast, our paper analyzes endogenously arising punishment patterns and, to the best of our knowledge, is the first one that compares the partial and the equilibrium effects of fines and uncertainty.

The famous Blackstone ratio states that it is "better that ten guilty persons escape, than that one innocent suffers". Most people share the view that avoiding type-I errors matters more, but the marginal rate of substitution between both error types varies considerably (Volokh, 1997). We argue that those who put more emphasis on type-I errors in fact punish less often.⁴ By using a take game, Rizzolli and Stanca

³Falk and Fischbacher (2002) use a take game to show that subjects "steal" more from their peers, the more their peers "steal".

⁴To account for the Blackstone ratio, some models simply put higher weight on type-I errors in analyzing optimal judgments (Miceli, 1991, Lando, 2006), and several papers explain from a rational choice perspective why type-I errors are more severe

(2012) show experimentally that the adverse effects of type-I errors on deterrence are higher than those of type-II errors. In their experimental setting, the probabilities of the two errors types are given exogenously, i.e. they do not consider the behavior of participants in their role as judges. By contrast, we focus on the interdependence of the two market sides.

In our experiment, the equilibrium deterrence effect of higher fines is reduced by the fact that players J punish less often. This effect dates back to the early legal literature on nullification pioneered by Michael and Wechsler (1937) which argues that jurors or witnesses may not be willing to participate in legal action when punishments seem unfairly high. Accordingly, theoretical models show that higher fines even reduce deterrence when the willingness to accept type-I errors is decreasing in fines to a sufficiently large extent (Andreoni, 1991, Feess and Wohlschlegel, 2009).

The legal literature has pointed out that the likelihood of type I- relative to type II-errors depends on the legal allocation of the burden of proof which is strikingly different in common and civilian law. In the US, decisions on civil claims are based on preponderance of evidence, while the proof beyond reasonable doubt-principle applied in criminal cases carries to a large degree over to civil claims in Europe (see Clermont and Sherwin, 2002 and Taruffo, 2003 for a dispute on these principles). Engel (2009) points out that the controversy about the appropriate allocation of the burden of proof needs to be related to the different goals of judicial procedures, and to the underlying assumptions on how legal decision makers actually reach their verdicts.

The remainder of the paper is organized as follows: Section 2 provides a theoretical model on the interdependency of the judges' and the potential violators' decisions. While the partial effects confirm the previous literature, we show that the impacts of the signal's noise and the magnitude of fines on deterrence and punishment are far from straightforward. In section 3, we describe the experimental setting. Section 4 provides empirical results for the partial impacts of fines and uncertainty, and section 5 extends to equilibrium effects. We conclude in section 6.

2 A model of punishment and deterrence

We first develop a model that allows to analyze the impact of fine size and uncertainty on the interplay between punishment and deterrence. In basic models on errors in court (Png, 1986, Polinsky and Shavell, 1999), it is found that both type-I and type-II errors increase the violation frequency as the difference between the probability of being punished with and without violation shrinks. A full-fledged equilibrium analysis, however, needs to take into account that higher uncertainty and higher fines may reduce the decision makers' (judges or juries) willingness to punish, so that the relative probabilities for the two errors need to be derived endogenously.

The interdependency of the behavior of judges and violators implies that we cannot treat type-I and type-II errors as exogenous - the probability of convicting an innocent (type-I error) depends on the percentage of violators in equilibrium, which in turn depends on the (anticipated) punishment behavior of judges. Consider a potential violator who, in the case of an infringement, causes an adverse outcome.⁵ In case of no infringement, an exogenous event causes the same adverse outcome with probability q .⁶ Note that this assumption implies that the adverse outcome can occur only once. In our experiment, observing the adverse outcome means that the money has disappeared, and q is the probability that this happens even in cases in

for society (see Hylton and Khanna (2007) who take up a public choice perspective, and Persson and Siven (2007) who adopt a median-voter model).

⁵In our experiment design, this corresponds to taking the amount designated for donation to charity.

⁶As in all models on errors in court, we need to assume that the facts of the case cannot be fully reconstructed ex post.

which it is not taken by the respective participant.

With ϕ as the percentage of participants who actually commit the act, the ex post probability that a violation took place after observing the loss is $\frac{\phi}{\phi+(1-\phi)q}$, and the ex post probability of no infringement is $\frac{(1-\phi)q}{\phi+(1-\phi)q}$. Thus, $\pi = \frac{(1-\phi)q}{\phi+(1-\phi)q}$ is the probability of committing a type-I error in case of punishment, while $1 - \pi = \frac{\phi}{\phi+(1-\phi)q}$ is the probability of a type-II error in case of no punishment.

For judges, we introduce the following assumptions: First, we set the utility from correct decisions to zero, that is, we take only the preference costs of misjudgments, but not the benefits of correct decisions into account. Second, we assume that judges take the consequences of their decisions for others into account. Specifically, the disutility from a type-I error is assumed to be $\alpha_i F$ where α_i is a parameter on judge i 's aversion against type-I errors, and F is the fine size. Likewise, a judge's disutility from a type-II error is $\beta_i A$ where β_i is the degree of judge i 's aversion against type-II errors, and A the severity of the infringement which we will refer to as the "amount stolen". Normalizing $\beta_i = 1$, α_i captures the ratio of the degrees of aversion against type-I and type-II errors.

Recalling the probabilities of type-I and type-II errors, it follows that a judge i who assumes a violation frequency ϕ prefers to convict a suspect if and only if

$$\frac{(1-\phi)q}{\phi+(1-\phi)q}\alpha_i F \leq \frac{\phi}{\phi+(1-\phi)q}A. \quad (1)$$

Defining $\tilde{\alpha}$ as the threshold type such that a judge prefers punishment for all $\alpha \leq \tilde{\alpha}$, we get

$$\tilde{\alpha} = \frac{A\phi}{Fq(1-\phi)} = \frac{1-\pi}{\pi} \frac{A}{F}. \quad (2)$$

The first equation in (2) shows that punishment is more likely if judges expect a higher violation frequency ϕ . The intuitive reason for this monotonic relationship between violation and punishment frequency is that the probability of making a type-I error when punishing, π (a type-II error when not punishing, $1 - \pi$) is strictly decreasing (increasing) in the violation frequency. In other words, the punishment frequency as an increasing function of the violation frequency can be transformed into a decreasing function of π , the probability of making a type-I error, which is displayed in the second equation in (2). This demonstrates that, when taking changes of the violation frequency into account, the well-known discussion of the impact of type-I and type-II errors on judges' punishment behavior becomes a discussion of the impact of violation frequencies on punishment.

For potential violators, we define m_j as a potential violator j 's intrinsic disutility from violating the law or social norm. Furthermore, violators may also have other-regarding preferences towards potential victims. Let us define a parameter $v \in [0, 1)$ such that, for $v = 0$, violators are completely selfish, and for $v = 1$, they put equal weight on the victims' and their own payoffs. Hence, if individual j assumes punishment frequency p , she steals if and only if

$$A - (1-v)p(A+F) - m_j \geq vA(1-q) - qpF(1-v). \quad (3)$$

On the left hand side, the expected benefit is the difference in the amount stolen and the expected fine, less the violator's disutility from stealing. Note that, in line with our experimental design, the expression $-(1-v)p(A+F)$ implies that in case of conviction both the amount and the fine are returned to the victim, so that F can also be interpreted as damages. All of our Propositions also hold when we assume instead that the injurer consumes A before detection and that F is not re-distributed to the victim but lost (as in the case of imprisonment, for instance). On the right hand side, the donation, which yields the partially

altruistic violator utility vA , is made with probability $1 - q$. Furthermore, qp is the probability of being punished by mistake (type-I error), in which case the fine F is taken from the alleged violator and passed on to the victim. Hence, individual j violates if and only if $m_j \leq \tilde{m}$ given by

$$\tilde{m} = A[(1 - p)(1 - v) + qv] - Fp(1 - q)(1 - v). \quad (4)$$

Equations (2) and (4) characterize the judges' and violators' optimal decisions based on their expectations on their counterparts' behavior: Inspecting these two equations yields the following Proposition:

Proposition 1 *Suppose that violators treat the punishment frequency p , and judges the violation frequency ϕ , as exogenously given. Then:*

$$(i) \quad \frac{\partial \tilde{m}}{\partial p} = -[A + F(1 - q)](1 - v) < 0, \quad \frac{\partial \tilde{m}}{\partial F} = -p(1 - q)(1 - v) < 0, \quad \text{and} \quad \frac{\partial \tilde{m}}{\partial q} = Av + Fp(1 - v) > 0 \quad \text{and}$$

$$(ii) \quad \frac{\partial \tilde{\alpha}}{\partial \phi} = \frac{A}{Fq(1 - \phi)^2} > 0, \quad \frac{\partial \tilde{\alpha}}{\partial F} = -\frac{A\phi}{F^2q(1 - \phi)} < 0, \quad \text{and} \quad \frac{\partial \tilde{\alpha}}{\partial q} = -\frac{A\phi}{Fq^2(1 - \phi)} < 0.$$

Part (i) of Proposition 1 first confirms the well-known deterrence theory: Violators are less likely to violate if they assume a higher punishment probability p and the fine size F is larger. Furthermore, the violation frequency increases in q , that is, in the probability that the loss may also be observed without infringement. This resembles the literature on the impacts of errors in court discussed in the introduction.

Part (ii) discusses judges' behavioral responses to the model parameters when they seek to minimize expected preference costs from wrong decisions: When they assume violators to break the law more frequently, when the evidence on the actual punishment is less noisy, and when fines are low, then judges will also punish for higher aversion against type-I errors, so that $\tilde{\alpha}$ increases. In a way, $\frac{\partial \tilde{m}}{\partial p}$ characterizes how violators' best responses depend on the judges' punishment frequency as, for any given p , a potential violator steals if and only if $m \leq \tilde{m}$. Similarly, $\frac{\partial \tilde{\alpha}}{\partial \phi}$ captures the effect of the assumed violating frequency on judges' best responses. While these individual best responses are intuitive, things are more involved when we take the interdependency of the decisions into account. In Bayesian Nash equilibrium, judges' and violators' expectations about their counterparts' actions coincide with their actual equilibrium choices, i.e., punishment and violation frequencies are given by the system of equations

$$p = G(\tilde{\alpha}) \quad (5)$$

$$\phi = H(\tilde{m}), \quad (6)$$

where $G(\cdot)$ and $H(\cdot)$ denote the cumulative distribution functions of judges' parameter α of relative aversion against type-I errors and potential violators' aversion against wrongdoing, provided that the thresholds $\tilde{\alpha}$ and \tilde{m} are in the supports S_α of $G(\cdot)$ and S_m of $H(\cdot)$, respectively, the intersection of which determines the Bayesian Nash equilibrium.

The following proposition summarizes the impact of noise and the fine size on equilibrium punishment and violation frequencies, where judges and violators correctly anticipate their counterparts' equilibrium choices:

Proposition 2 *Suppose that S_α and S_m are intervals. If $\tilde{\alpha} \in S_\alpha$ and $\tilde{m} \in S_m$, then the Bayesian Nash equilibrium has the following comparative static properties:*

$$(i) \quad \text{For all distributions } G(\cdot) \text{ and } H(\cdot), \quad \frac{dp^*}{dF} < 0 \quad \text{and} \quad \frac{d\phi^*}{dq} > 0.$$

(ii) *The absolute equilibrium effects of legal uncertainty on the punishment frequency, and of the fine size on the violation frequency, are smaller than the absolute partial effects derived in Proposition 1.*

Proof. See Appendix A. ■

Part (i) of the Proposition shows that, for all distributions of the judges' and the potential violators' preferences, two of the results derived for given behavior of the other market side carry over to the Bayesian Nash Equilibrium: First, the punishment frequency decreases in the magnitude of fines, $\frac{dp^*}{dF} < 0$. This is intuitive as, for all adjustments of the violators' behavior, the consequences of a type-I error are increasing in the fine size, so that judges are in equilibrium more reluctant to punish in case of high fines. Second, the standard result that higher uncertainty reduces deterrence, $\frac{d\phi^*}{dq} > 0$, also carries over from the setting with exogenously given errors in court to the Bayesian Nash Equilibrium.

However, the impacts of the fine size on the violations frequency ($\frac{d\phi^*}{dF}$) and of the noise on the punishment behavior ($\frac{dp^*}{dq}$) are ambiguous, and do hence not appear in the Proposition. The first ambiguity confirms the results of Andreoni (1991) and Feess and Wohlschlegel (2009) that higher fines may reduce deterrence if judges care sufficiently strongly about type I errors. The second ambiguity means that judges may even increase their punishment frequency in response to higher noise when the violation frequency increases sufficiently steeply in noise.

Part (ii) of the Proposition says that, whenever the equilibrium effects are ambiguous, they are less pronounced than the partial effects. This is intuitive, since the countervailing effects that make the equilibrium effects ambiguous are not present when the behavior of the other market side is treated as exogenously given.

3 Experimental design

Our objective is to analyze how judges' punishment and potential violators' decisions react to fines, legal uncertainty, and the behavior of the other market side (violators or judges, respectively). In order to test the predictions of our model in a comprehensive manner, our design has to ensure that we can compare behavior under the assumption that players take the other market side's decisions as given (partial effects, see Proposition 1) to a situation in which players anticipate the other market side's reaction to changes in fines or legal uncertainty (equilibrium effects, see Proposition 2). For that reason, we employ a strategy method design in which players of one market side can condition their behavior on the behavior of the other market side. In two treatments, we vary between subjects whether judges or violators use the strategy method, while the other market side decides unconditionally whether to violate or to punish, respectively. Moreover, we vary the level of fine and legal uncertainty within subject: each participant faces four decision situations that are characterized by different combinations of high or low levels of fines and legal uncertainty.

3.1 Punishment and violation decisions

At the beginning of the experiment, participants were randomly assigned the role of potential violators (henceforth referred to as participants V) or the role of judges (henceforth labeled participants J) with equal probability. The instructions were framed neutrally, labeling potential violators as participants of type A and judges as participants of type B, a fine as "deduction points" and a violation as "taking" instead of stealing. A translation of the complete instructions of one treatment to English can be found in Appendix B. Next, pairs consisting of one participant V and one participant J were matched randomly. Throughout the experiment, assigned roles and pairs remained constant and players acted anonymously. Each participant

V and J had an initial endowment of 2400 points.⁷ Moreover, for each pair an additional 2400 points were designated for donation to the charity "SOS-Kinderdorf" (SOS Children's Villages). The work of the charity was described as follows: "SOS-Kinderdorf supports children and families in the poorest regions of the world (medical care, school enrollment etc.), in crisis areas, but also neglected children and adolescents in Germany. SOS-Kinderdorf has been granted the DZI-certificate that attests reliability and low administrative costs."

In order to model a violation decision, each participant V had to decide whether he would like to take the donation designated for the charity for himself. In that case, there was no donation to the charity and participant V 's payoff increased by 2400 points to 4800 points. Participant J observed whether the donation was available or had disappeared. Importantly, whenever participant J observed that the donation had disappeared, this could be due to two possible reasons: first, participant V had taken the donation. Second, participant V had decided not to take the donation but it had disappeared by chance, which happened with probability q . The parameter q varies across decision situations and captures legal uncertainty. Only if participant J observed that the donation had disappeared, he had the possibility to punish V , i.e., to reduce participant V 's payoff by a prespecified fine of size F such that V 's payoff was $2400 - F$ points. If V was punished and had taken the donation, the donation was donated to the charity. If V was punished although he had not taken the donation, the donation remained unavailable and was not donated. If fine F was imposed on participant V , the F points were additionally donated to the charity. Donating the possible fine to the charity has two advantages: First, it avoids potential concerns of participant J that punishing reduces overall efficiency (defined as the sum of payments to participant V , J , and the charity). Second, it resembles the common legal practice of donating a fine to a charity whose work is related to the field of violation.⁸ If participant J decided not to punish participant V , V received his whole initial endowment of 2400 points. If V had taken the donation, the donation was added to V 's payoff, V earned 4800 points, and the charity did not receive any donation. In any case, participant J was paid 2400 points at the end of the experiment. Thus, financial self-interest did not motivate participant J 's behavior, but his decisions affected participant V 's payoff and the donation to the charity (only if there was a legal case, i.e., if the donation had disappeared).

Each participant V and J faced four decision situations that differed only in the size of the fine F ($F_{low} = 1200$ or $F_{high} = 3600$) and / or the extent of legal uncertainty q ($q_{low} = 0$ or $q_{high} = 0.5$): first, $F = 1200$ and $q = 50\%$; second, $F = 3600$ and $q = 50\%$; third, $F = 1200$ and $q = 0$; and finally, $F = 3600$ and $q = 0$ (see Table 1). Comparing behavior across decision situations enables us to isolate the effect of changes in the fine for a given level of legal uncertainty and vice versa. The instructions described the decision situation in general terms. The exact size of the fine and the level of uncertainty were only revealed on screen just before making the respective decision.⁹ One of the four decision situations was chosen randomly and the actual decisions of participants V and J in this decision situation determined their earnings and the amount donated to the charity (if any).

Finally, as shown in Table 1 in half of the sessions participants J used the strategy method for punishment decisions. For each of the four combinations of fine and legal uncertainty, they made 13 conditional, binary punishment decisions, namely whether their matched participant V should be fined if the donation had

⁷During the experiment payments were calculated in points. At the end of the experiment, points were converted into Euro at the following exchange rate: 500 points = 1 Euro.

⁸A potential disadvantage of donating the fine is that, even if J believes that V did not take the donation, he could impose the fine to redistribute from participant V to the charity if he is willing to incur the cost of "convicting an innocent". Our questionnaire data show that participants are very reluctant to do so. Using a Likert scale from 0 (not bad at all) to 9 (extremely bad), we asked participants how bad they personally feel if someone innocent is punished for something he has not done. The mean answer is 8.2, no participant chose 0 or 1.

⁹However, the instructions used the decision situation with $F = 1200$ and $q = 0.5$ as an example to illustrate participants' decision screens and carefully explained the implications of legal uncertainty, using the case with $q = 0.5$ as an example.

Table 1: Treatment Overview

Treatment	Decision situation			
	1	2	3	4
Player J strategy method (3 sessions, N=72)	$F = 1200, q = 0.5$	$F = 3600, q = 0.5$	$F = 1200, q = 0$	$F = 3600, q = 0$
Player V strategy method (3 sessions, N=72)	$F = 1200, q = 0.5$	$F = 3600, q = 0.5$	$F = 1200, q = 0$	$F = 3600, q = 0$

disappeared and x of the 12 participants V in the room had decided to take the donation, for each $x \in \{0, 12\}$. At the same time, participants V made one unconditional violation decision for each of the four decision situations, i.e., they decided whether to take the donation or not. Counting the number of violators in the current session allowed to determine which of the judges' conditional punishment decisions was actually implemented.

Similarly, in the remaining sessions participants V made all violation decisions using the strategy method. In each decision situation, they made 13 conditional, binary violation decisions, namely whether to take the donation if y of the 12 participants J in the room decided to punish if the donation had disappeared, for each $y \in \{0, 12\}$. At the same time, participants J made an unconditional punishment decision for each of the four combinations of fine and legal uncertainty, i.e., they decided whether to punish their randomly matched player V if they observed that the donation had disappeared. The number of punishing players J in the current session determined which of player V 's conditional violation decisions was implemented.

3.2 Measurement of beliefs and attitudes towards risk

Using the strategy method data, we will analyze equilibrium effects of changes in fine or legal uncertainty under the assumption that participants correctly anticipate the other market side's reaction to the change under consideration. In order to be able to judge the validity of this assumption, we additionally measure strategy method players' beliefs concerning the behavior of the other market side. For each of the four decision situations, participants J were asked to guess the number of participants V who had decided to take the donation among the 12 players V present in the room. Similarly, participants V were asked for their belief concerning the number of participants J among the 12 present participants J who decided to impose the fine if they observed that the donation had disappeared. Participants knew that their belief for one of the four decision situations would be drawn randomly and paid according to the belief's precision. More precisely, participants had the possibility to win additional 2000 points or 4 Euro. The probability of winning increased in the precision of their belief and equaled $p_{win} = 1 - 1/144$ (belief - actual number)². For those participants who were interested in the exact payment scheme, the scheme was explained on screen. This procedure of eliciting beliefs, the randomized quadratic scoring rule, follows Schlag and van der Wee (2013) and incentivizes participants to report their beliefs truthfully, without imposing assumptions on their risk preferences.

In order to measure participants' risk preferences, we used a procedure similar to the one proposed by Holt and Laury (2002). Participants had to decide whether they preferred a safe payment of 1200 points or a lottery that would either pay 3000 with probability p or 0 points with probability $(1 - p)$. In ten separate decisions, p was 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, or 1, respectively. One of the ten decisions was drawn

randomly and paid.

3.3 Procedural details and implementation

The detailed sequence of events in each session was as follows:

1. Instructions: Upon entering the lab, subjects drew a card containing a number and were asked to sit in the respective booth. They read the instructions, were encouraged to ask questions in private, and answered several control questions on the computer.
2. Four decision situations concerning punishment or violation behavior
3. Measurement of beliefs concerning the behavior of the other player type: only strategy method players
4. Measurement of risk attitudes
5. Questionnaire¹⁰

The experiment was conducted at the BonnEconLab, Germany with 144 participants, 72 in each treatment. Most participants (94%) were students and majored in various subjects. Age varied between 18 and 57 years (median age 23, 92% in the range 19–30 years). We used the software z-Tree (Fischbacher, 2007) for conducting the experiment and ORSEE (Greiner, 2004) for inviting participants from the BonnEconLab’s participant pool and recording their participation. The experiment took place in 6 sessions in June 2015.

In total, the experiment lasted about 60 minutes (including payment). On average, participants earned a total of 14.20 Euro. Participants’ overall payments consisted of the sum of up to four payments: a show-up fee of 4.00 Euro, the payment from one randomly drawn decision situation, the incentivized belief (for strategy method players only) and the payment from one randomly drawn risk-related decision. Payments were made in a separate room to ensure privacy. We donated a total of 676.80 Euro to the NGO ”SOS-Kinderdorf”, which corresponds to the overall amount that participants had decided to donate.

4 Partial impacts of fine size and uncertainty

In our model, we have derived two kinds of results. The *partial effects* summarized in Proposition 1 capture the impacts of the fine size F and legal uncertainty q on the violation frequency for *given* punishment frequencies; and vice versa. The respective experimental results are presented in section 4. By contrast, Proposition 2 refers to the *total effects* of changes in F or q in equilibrium, in which the reactions of the other market side are anticipated. Results on the tests of Proposition 2 are discussed in section 5.

4.1 Players J

Following Proposition 1, we test three hypotheses on the behavior of players J :

H1. The punishment frequency increases in the violation frequency.

H2. For identical violation frequencies and with legal uncertainty, the punishment frequency is lower for $F = 3600$ compared to $F = 1200$.

H3. For identical violation frequencies and fines, the punishment frequency is lower for $q = 50\%$ compared to $q = 0$.

¹⁰The post-experiment questionnaire measured socio-economic and demographic characteristics. Moreover, it assessed participants’ general attitudes concerning moral behavior, donations to charity, individual responsibility for own actions, aversion against legal type I and type II error, and judgments concerning determinants of the effectiveness of legal rules.

By definition of partial effects, the hypotheses are derived under the assumption that the violation frequency is unaffected by F and q . Thus, we use the strategy method for players J , i.e., for each of our four combinations of F and q we elicit players J 's behavior for any number of actual violations by players V . As all 36 players J made decisions in four different situations, we have 144 observations when defining all decisions for a specific F - q combination as a single observation. In line with H1, the Pearson's pairwise correlation coefficients between the number of violators and the number of punishing players J are significantly positive at $p < 0.001$ for all four situations. Moreover, 135 of our 144 observations are consistent in the sense that the punishment frequency is weakly increasing in the number of violators. As inconsistent behavior increases the probability of both error types, players J who understand the setting should behave consistently; irrespective of their relative aversion to type-I and type-II errors.¹¹ For the tests of Hypotheses 2 and 3, we exclude the 9 inconsistent observations from our analysis. Figure 1 displays the cumulative distribution of the punishment frequency, i.e. the share of players J who punish for each possible numbers of violators. The number of players J who never punish is shown as a thirteenth category on the horizontal axis.

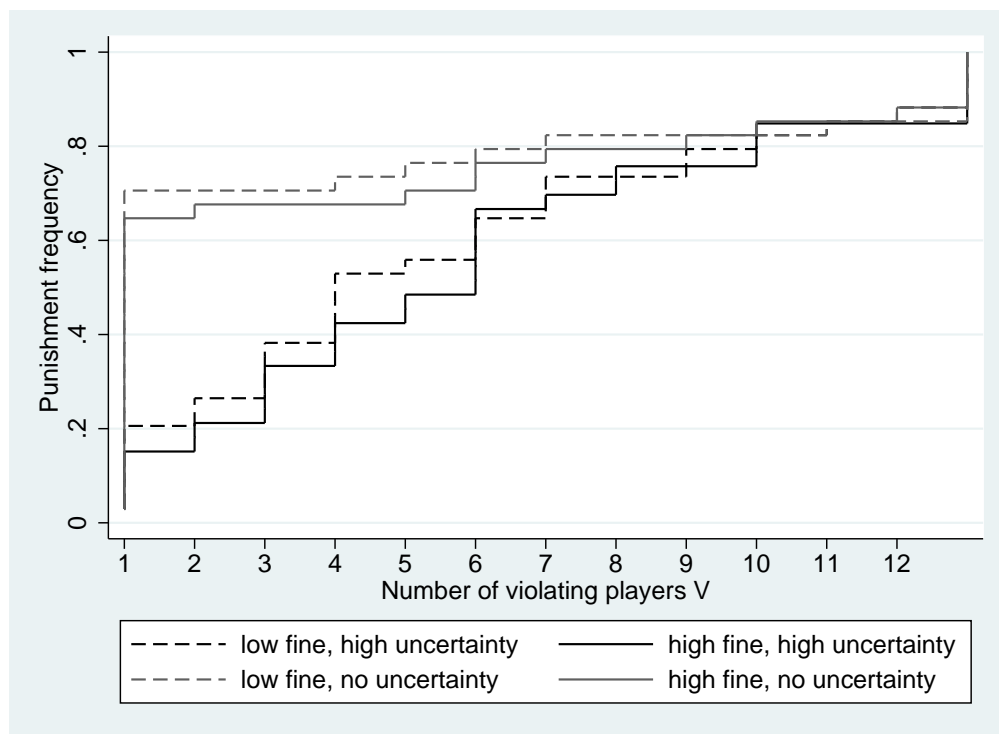


Figure 1: Punishment frequencies.

Then, denote by \tilde{V}_i a threshold such that the i -th player J punishes if and only if $V \geq \tilde{V}_i$. The cumulative distribution in figure 1 also shows the distribution of thresholds over players J : Considering the low fine ($F = 1200$) and high uncertainty ($q = 50\%$) as an example, about 40% of all players J punish for 3 violators, while more than 50% punish for 4 violators. Thus, the violation frequency threshold that triggers punishment

¹¹An example for inconsistent behavior is a player J who punished for two, but not for three violators. This player J , however, could reduce both error types by punishing for three, but not for two violators.

is 4 for 52%-40%=12% of all judges.

To test H2, we compare the thresholds for $F = 1200$ and $F = 3600$ for the case with legal uncertainty, $q = 50\%$. In line with H2, thresholds are significantly higher for $F = 3600$, with $p = 0.03$ in the Wilcoxon signed ranks test.^{12,13}

Similarly to the impact of the fine size, we test the impact of legal uncertainty by a Wilcoxon signed ranks test on the differences in the punishment thresholds, now comparing $q = 0$ and $q = 50\%$ for $F = 1200$ and for $F = 3600$, respectively. The respective p -values are both smaller than 0.001, so that H3 cannot be rejected.

Note that, for q given, any exogenously given number of violators in the strategy method translates into error types: for instance, for $q = 50\%$ and five violators (i.e. a violation frequency of $\phi = 5/12$), the probability of a type I-error in case of punishment is $\pi = \frac{(1-\phi)q}{\phi+(1-\phi)q} = \frac{7}{17} = 41.2\%$, while the probability of a type II-error in case of no-punishment is $1 - \pi = 58.8\%$. Let us next compare the risk of type-I and type-II errors that players J are willing to take. Substituting for $q = 50\%$ and taking into account that the violation frequency is $\phi = V/12$, one can easily check that the frequencies of both error types are identical for four violators; i.e. $\pi = 1 - \pi = 0.5$ for $V = 4$. For $F = 1200$, 18 players J have $V \leq 4$ and 16 players J have $V > 4$, while 14 judges have $V \leq 4$ and 19 $V > 4$ for $F = 3600$. Thus, the type-I error aversion is moderately increasing in the fine size.

4.2 Players V

For players V , we test the following three hypotheses:

H4. The violation frequency decreases in the punishment frequency.

H5. For identical punishment frequencies, the violation frequency is lower for $F = 3600$ compared to $F = 1200$.

H6. For identical punishment frequencies and fines, the violation frequency is higher for $q = 50\%$ compared to $q = 0$.

In line with H4, the Pearson's pairwise correlation coefficients between the number of punishing players J and the number of violators for our four decision situations are all significantly negative at $p < 0.001$.

For the tests of Hypotheses 5 and 6, we proceed analogously to players J by considering only those 131 out of 144 players V who behave consistently. Analogously to figure 1, figure 2 shows the share of violators for each of the possible numbers of punishing players J and for each of the four decision situations.

Following the same logic as for players J , we define a threshold \tilde{J}_i such that the i -th player V takes the money if and only if $J \leq \tilde{J}_i$. In line with H5, a comparison of these thresholds for $F = 1200$ and $F = 3600$ for the case with legal uncertainty yields $p < 0.001$ for the Wilcoxon signed ranks test. Also in line with our model, there is no significant impact of the fine size for $q = 0$ (Wilcoxon signed ranks test, $p = 0.191$). Finally, in line with H6, the violation frequency is significantly lower for $q = 0$ compared to $q = 50\%$ (Wilcoxon signed ranks tests, $p < 0.001$ for both fines). Summing up, none of the six hypotheses on the *partial* impacts of the magnitude of fines and legal uncertainty on the punishment and violation behavior is rejected by our data.

¹²Since type-I errors are excluded for $q = 0$, our theory predicts a punishment frequency of 100%. However, the legal literature discussed in the introduction has pointed out that some decision makers may prefer not to punish, even without uncertainty, when they perceive the fine as inappropriately high. This is also the case in our experiment, but the Wilcoxon signed ranks test shows that there is no significant difference in the punishment behavior between $F = 1200$ and $F = 3600$ ($p = 0.35$) for $q = 0$. Thus, in line with our model, a higher fine leads to a significant reduction in the punishment frequency only with legal uncertainty.

¹³Throughout the paper, we report two-sided p -values.

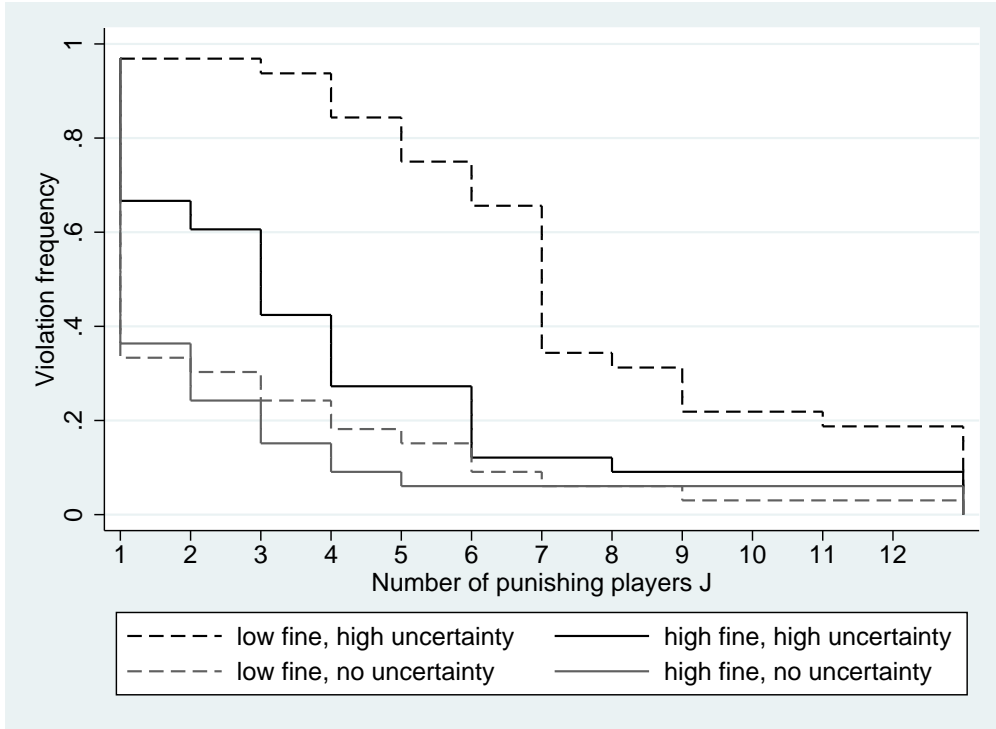


Figure 2: Violation frequencies.

5 Equilibrium impacts of fine size and uncertainty

So far, we have tested theoretical predictions that treat the behavior of the other market side as constant (*partial* effects). In equilibrium, however, players V anticipate that fewer players J punish for high fines, which is in turn anticipated by players J . Due to the aforementioned interdependencies of both players' strategies, the *total* effects of F and q are less straightforward than the partial ones (see Proposition 2). Our experimental design enables us to quantify these equilibrium effects and to compare them to partial effects.

5.1 Players J

For players J , table 2 summarizes the insights from Proposition 2.

Table 2: Theoretical results on equilibrium effects for players J .

Equilibrium effect of fine	$\frac{dp^*}{dF} < 0$	$\frac{dp^*}{dF} \geq \frac{\partial p^*}{\partial F}$
Equilibrium effect of legal uncertainty	$\frac{dp^*}{dq} \geq 0$	$\frac{dp^*}{dq} > \frac{\partial p^*}{\partial q}$

We start with analyzing the impact of the fine, for which our model yields a clear hypothesis:

H7. For the case with legal uncertainty, the punishment frequency in equilibrium is lower for $F = 3600$ compared to $F = 1200$.

By contrast, it depends on the players J 's preferences to avoid type-I relative to type-II errors whether the equilibrium effect is larger or smaller than the partial effect. Thus, we state no hypothesis on the comparison of the two effects, but it is nevertheless interesting to report and interpret our experimental findings.

Table 3 summarizes our data analysis for $q = 50\%$ and for the two fine levels. Taking the average over all sessions, 9.3 out of 12 players V took the money for $F = 1200$, but only 5.3 for $F = 3600$ (see the entries in the first row of table 3). To calculate these numbers, we consider the players V 's unconditional behavior.

Next, recall that the *partial* effect of the change in F on players J 's punishment behavior needs to be calculated under the counterfactual assumption that the violation frequency is independent of F . We hence get the partial effect by assuming that the number of violators is the same for both fine sizes. If 9.3 players V had taken the money for $F = 3600$ (as they did for $F = 1200$),¹⁴ we know from the strategy method that 75.8% of players J would punish for $F = 3600$, compared to 79.4% for $F = 1200$. The difference is the partial effect of the fine size as shown in the second row of table 3: *If* players J assumed that the higher fine had no impact on the violation frequency, then an increase in the fine size from $F = 1200$ to $F = 3600$ would lead to a decline in the punishment frequency by 3.6 percentage points.¹⁵

Table 3: Partial and equilibrium impacts of the fine size on the punishment frequency

	$F = 1200$	$F = 3600$
Actual violation frequency	9.3 (78%)	5.3 (44%)
Pun. frequency given 9 violators	79.4%	75.8%
Partial effect	-3.6% <i>points</i>	
Pun. frequency with actual violation frequency	79.4%	48.5%
Equilibrium effect	-30.9% <i>points</i>	
Violation frequency assumed by players J	6.8	4.6

The equilibrium effect is calculated by assuming that players J correctly anticipate the impact of the fine size on the violation frequencies. Thus, we compare the actual number of violations, 9.3 for $F = 1200$ and 5.3 for $F = 3600$, which yields a rather large reduction in the punishment frequency by 30.9 percentage points.¹⁶ Furthermore, table 3 shows that the equilibrium effect is more pronounced than the partial effect. The intuition is that players J who correctly anticipate players V 's behavior reduce their punishment frequency for two reasons; due to the higher fine size (direct effect) and due to the lower violation frequency (indirect effect). The last row in table 3 shows that, while players J generally underestimate the violation frequencies to some degree, they reasonably well anticipate the deterrence effect of higher fines.

We now turn to the impact of legal uncertainty. Recall from table 2 that our model yields no clear-cut prediction for the equilibrium effect itself, but for the comparison to the partial effect:

H8. The equilibrium impact of legal uncertainty on the punishment frequency is smaller than the partial effect.

¹⁴Alternatively, we could assume that the number of violators is 5.3 not only for $F = 3600$ as observed in the data but also for $F = 1200$.

¹⁵We have already considered the partial effects in section 4 where we compared the thresholds for the number of violators which triggers punishment. Here, we calculate them explicitly as to compare them to equilibrium effects.

¹⁶This is supported by a McNemar change test that compares players J 's unconditional punishment decisions for $q = 50\%$ for $F = 1200$ and $F = 3600$ (see the first row of Table 6, the difference of which is significant at $p = 0.01$).

Table 4: Partial and equilibrium impacts of legal uncertainty on the punishment frequency

$F = 1200$	$q = 50\%$	$q = 0\%$
Actual violation frequency	9.3 (78%)	2.0 (17%)
Pun. frequency given 9 violators	79.4%	82.4%
Partial effect	3.0% <i>points</i>	
Pun. frequency with actual violation frequency	79.4%	70.6%
Equilibrium effect	-8.8% <i>points</i>	
Violation frequency assumed by players J	6.8	3.4
$F = 3600$	$q = 50\%$	$q = 0\%$
Actual violation frequency	5.3 (44%)	1.0 (8%)
Pun. frequency given 5 violators	48.5%	70.6%
Partial effect	22.1% <i>points</i>	
Pun. frequency with actual violation frequency	48.5%	64.7%
Equilibrium effect	16.2% <i>points</i>	
Violation frequency assumed by players J	4.6	1.7

Let us start with the upper part of table 4 which refers to $F = 1200$. The first row shows that legal uncertainty leads to a large increase in the violation frequency. For the partial effect, we again need to assume identical violation frequencies for both levels of q . We then find that legal uncertainty reduces the punishment frequency by 3 percentage points. Interestingly, the equilibrium effect of legal uncertainty for $F = 1200$ has the opposite sign, i.e. legal uncertainty even leads to a higher punishment frequency. While this seems somewhat counter-intuitive, it does not contradict our theory which yields no clear-cut prediction on the equilibrium impact of q : Intuitively, if players J anticipate that the violation frequency increases sharply in response to uncertainty, then they may have incentives to punish more often. For $F = 3600$, the equilibrium effect goes in the opposite direction with an equilibrium effect of 16.2 percentage points.

Most importantly, however, the results are in line with H8 for both fines as the equilibrium effect is always smaller than the partial effect; -8.8 compared to 3.0 percentage points for the low fine, and 16.2 compared to 22.1 percentage points for the high fine. Finally, while players J again underestimate the overall violation frequency, they anticipate well that it is increasing in uncertainty.

Over all, our experimental evidence supports the argument that players J seek to anticipate adjustments in the behavior of players V so as to minimize their disutility from making type-I and type-II errors. A natural question is then that about the incidence of these errors. In our experiment, type II errors occurred more frequently than type I errors as hypothesized by the Blackstone ratio: When there was legal uncertainty ($q = 50\%$), 19.4% of players J made a type-II error for $F = 1200$, and 27.8% of them made a type-II error for $F = 3600$. Furthermore, the expected frequency of type-I errors is 11.8% of all cases for $F = 1200$ and 13.2% for $F = 3600$.¹⁷

¹⁷In our experiment design, the incidence of type-I errors is probabilistic. Given that player V has not violated but player J has decided to punish if the donation has disappeared, whether a type-I error occurs depends on the random event that determines whether the donation disappears exogenously. This was only drawn for the randomly chosen decision situation, but not the three further decision situations.

5.2 Players V

The test of Proposition 2 for players V proceeds along the same lines. Table 5 summarizes our theoretical results.

Table 5: Theoretical results on equilibrium effects for players V .

Equilibrium effect of fine	$\frac{d\phi^*}{dF} \geq 0$	$\frac{d\phi^*}{dF} > \frac{\partial p^*}{\partial F}$
Equilibrium effect of legal uncertainty	$\frac{d\phi^*}{dq} > 0$	$\frac{d\phi^*}{dq} \leq \frac{\partial p^*}{\partial q}$

While our model does not predict the sign for the equilibrium effect of the fine, it yields the following hypothesis:

H9. For the case with legal uncertainty, the equilibrium impact of the fine size on the violation frequency is less pronounced than the partial impact.

Table 6: Partial and equilibrium impacts of the fine size on the violation frequency

	$F = 1200$	$F = 3600$
Actual punishment frequency	7.0 (58%)	3.7 (31%)
Viol. frequency given 7 punishers	34.4%	12.1%
Partial effect	-22.3% <i>points</i>	
Viol. frequency with actual pun. frequency	34.4%	27.3%
Equilibrium effect	-7.1% <i>points</i>	
Punishment frequency assumed by players V	6.3	5.0

To measure the partial effect, we follow the same logic as for players J . For $F = 1200$ and 7 punishers, table 6 shows that the violation frequency is 34.4%. If 7 players J also punished for $F = 3600$, the violation frequency would have been 12.1%. The difference of -22.3 percentage points thus captures the partial effect. Using the actual number of violators of 3.7 for $F = 3600$ in row 3 then leads to an equilibrium reduction by 7.1 percentage points. In line with H9, this shows that the deterrence effect of a higher fine is indeed reduced when players V anticipate that players J punish less often.

While our model does not make a clear prediction on the equilibrium effect of higher fines on violation frequencies, it is still interesting to study the empirical evidence. Table 6 documents a moderate, 7 percentage points decrease in the violation frequency due to the increase in the fine.¹⁸ Row 4 shows that players V are aware of the fact that the punishment frequency decreases in the fine, but they underestimate the effect size.

Our model next predicts that, even when taking the players J 's behavior into account, higher legal uncertainty reduces deterrence:

H10. In equilibrium, the violation frequency is higher for $q = 50\%$ compared to $q = 0\%$.

¹⁸Analyzing players V 's unconditional decisions adds support to our result that the equilibrium effect of a higher fine on violation frequencies is negative: for $q = 50\%$, the first row of Table 3 shows that the violation frequency decreases from 77.8% to 44.4% as the fine increases from 1200 to 3600 (McNemar change test, $p < 0.01$).

Table 7: Partial and equilibrium impacts of legal uncertainty on the violation frequency

$F = 1200$	$q = 50\%$	$q = 0\%$
Actual punishment frequency	7.0 (58%)	10.3 (86%)
Viol. frequency given 7 punishers	34.4%	6.1%
Partial effect	-28.3% <i>points</i>	
Viol. frequency with actual pun. frequency	34.4%	3.0%
Equilibrium effect	-31.4% <i>points</i>	
Punishment frequency assumed by players V	6.3	10.0
$F = 3600$	$q = 50\%$	$q = 0\%$
Actual punishment frequency	3.7 (31%)	10.0 (83%)
Viol. frequency given 4 punishers	27.3%	9.1%
Partial effect	-18.2% <i>points</i>	
Viol. frequency with actual pun. frequency	27.3%	6.1%
Equilibrium effect	-21.2% <i>points</i>	
Punishment frequency assumed by players V	5.0	9.6

In line with H10, table 7 shows equilibrium effects of -31.4 and -21.2 percentage points for a decrease in q from 50% to 0% for the low and the high fine, respectively.¹⁹ Thus, even when players V anticipate that players J are less willing to punish in case of legal uncertainty, uncertainty still leads to more violations. This reinforces the view that legal uncertainty is worrisome not only from a justice point of view, but also from a deterrence perspective.

5.3 Probit models on unconditional behavior

As all of our hypotheses on the equilibrium effects have been tested by applying the strategy method, it is interesting to see if the main results are similar when we consider the unconditional behavior of players J and V in response to different levels of fines and uncertainty. In this case and by contrast to the strategy method, players J and V can only form expectations on the behavior of the other market side.

For the unconditional behavior, table 8 presents the marginal effects of probit models using players J 's binary punishment and players V 's binary violation decisions as dependent variables. In addition to dummy variables for the high fine and the high level of uncertainty, the explanatory variables include a measure of risk attitudes (i.e. the switching point in the Holt-Laury lottery²⁰), the subjective importance of donations for fighting poverty and the aversion against making type-I and type-II errors as elicited in our questionnaire.

Most importantly, the impacts of F and q are qualitatively the same as for the strategy method: Increasing the fine from $F = 1200$ to $F = 3600$ or introducing uncertainty reduces the punishment frequency. For violators, a higher fine increases deterrence, while introducing uncertainty increases violations. For both market sizes, the impact of uncertainty is more pronounced than the increase in the fine. In line with the assumptions of our model, players J who have a higher aversion against type-I errors punish less often, and those players V who put higher emphasis on donations take the money less often.

¹⁹Differences between partial and equilibrium effects, for which our theory yields no clear-cut prediction, are moderate.

²⁰A higher switching point implies a lower willingness to take risks.

Table 8: Probit models on players J 's punishment and V 's violation decisions

	(1) Players J	(2) Players V
$F = 3600$	-0.1477*** (0.0406)	-0.2063*** (0.0477)
$q = 50\%$	-0.3920*** (0.0466)	0.4147*** (0.0386)
Switching Point	0.0568* (0.0304)	-0.0264 (0.0220)
Donations Important	0.0325 (0.0586)	-0.1362*** (0.0449)
Type-I Error Aversion	-0.1186** (0.0559)	0.0628 (0.0617)
Type-II Error Aversion	0.0190 (0.0170)	-0.0444** (0.0182)
N	140	140

The table displays average marginal effects. Robust standard errors, which are clustered at the participant level, in parentheses. *, ** and *** denote significance at 10-percent, 5-percent and 1-percent levels, respectively.

6 Conclusion

We analyze the impact of fine size and legal uncertainty on the frequencies of punishments and legal infringements. In our theoretical model, we assume that legal decision makers have heterogeneous preferences with respect to type-I and type-II errors, and that potential violators have different preference costs from violating the law. Based on these assumptions, we first confirm theoretical results that higher uncertainty and higher fines reduce the punishment frequency when legal decision makers treat the violation frequency as exogenously given (*partial effects*). Analogously, higher uncertainty then increases, and higher fines reduce the violation frequency. Taking the interdependency between the two market sides into account in order to analyze the *equilibrium effects* of fines and uncertainty, results turn out to be less straightforward: while higher fines still reduce the punishment frequency and higher uncertainty still increases the violation frequency, the impact of the fine size on the violation frequency and the impact of uncertainty on the punishment frequency turn out to be theoretically ambiguous.

We then conduct a laboratory experiment that allows us to test our hypotheses on partial effects and equilibrium effects. To the best of our knowledge, our paper is the first one that identifies and compares the partial and equilibrium effects of fines and legal uncertainty. In the experiment, there are players V who may decide to take money that was supposed to be donated to charity, and players J who are matched with a player V whom they may punish based on an imperfect signal on whether this player V has actually taken the money. Applying the strategy-method approach, we can elicit the players J 's willingness to punish for each exogenously given number of violators. Players J 's decisions translate into the probability of type-I errors in case of punishment, and into the probability of type II-errors in case of no punishment. Partial effects can then be analyzed by using the counterfactual assumption that the violation frequency is independent of the fine size and legal uncertainty, while equilibrium effects can be studied by considering the actual violation frequencies.

Our experimental results are in line with our theoretical predictions. Assuming that the behavior of the other market side is given (partial effects), higher fines reduce the violation and the punishment frequency, while legal uncertainty increases the number of violations and decreases the punishment frequency. Also in line with our theory, a higher fine leads to a significant reduction in the punishment frequency if and only if there is legal uncertainty.

Our findings on equilibrium effects seem to be important from a legal perspective: First, they suggest that the impact of legal uncertainty on the punishment frequency is lower than the partial effect, because judges anticipate that the violation incentive increases in uncertainty. This is a countervailing effect that *ceteris paribus* reduces the risk of type I-errors, and hence enhances the punishment incentive. Second, our empirical results support previous theoretical results that the deterrence effect of higher fines is reduced by the fact that violators anticipate that judges punish less often. And third, in our experiment legal uncertainty increases the violation frequency even in equilibrium, which reinforces the view on the benefits of thorough investigations and a well-functioning legal system.

Acknowledgements

Financial support from the German Research Foundation (DfG) through SFB-TR 15 and by the Portsmouth Business School's Research Project Funds is gratefully acknowledged and did not influence study design or analysis of results. We thank Anne Mertens for programming and excellent assistance in implementing the experiment.

References

- ABBINK, K., B. IRLBUSCH, AND E. RENNER (2002): "An experimental bribery game," Journal of Law, Economics, and Organization, 18(2), 428–454.
- AMBRUS, A., AND B. GREINER (2012): "Imperfect public monitoring with costly punishment: An experimental study," American Economic Review, 102(7), 3317–3332.
- ANDERSON, C. M., AND L. PUTTERMAN (2006): "Do non-strategic sanctions obey the law of demand? The demand for punishment in the voluntary contribution mechanism," Games and Economic Behavior, 54(1), 1–24.
- ANDREONI, J. (1991): "Reasonable doubt and the optimal magnitude of fines: should the penalty fit the crime?," RAND Journal of Economics, 22(3), 385–395.
- BAUMANN, F., AND T. FRIEHE (2015): "Proof Beyond a Reasonable Doubt: Laboratory Evidence," Discussion paper.
- BECKER, G. S. (1968): "Crime and Punishment: An Economic Approach," Journal of Political Economy, 76(2), 169–217.
- CARPENTER, J. P. (2007): "The demand for punishment," Journal of Economic Behavior & Organization, 62(4), 522–542.
- CLERMONT, K. M., AND E. SHERWIN (2002): "A comparative view of standards of proof," American Journal of Comparative Law, 50(2), 243–275.

- DEANGELO, G., AND G. CHARNES (2012): “Deterrence, expected cost, uncertainty and voting: Experimental evidence,” Journal of Risk and Uncertainty, 44(1), 73–100.
- EGAS, M., AND A. RIEDL (2008): “The economics of altruistic punishment and the maintenance of cooperation,” Proceedings of the Royal Society of London B: Biological Sciences, 275(1637), 871–878.
- ENGEL, C. (2009): “Preponderance of the evidence versus intime conviction: A behavior perspective on a conflict between American and Continental European law,” Vermont Law Review, 33, 435–467.
- FALK, A., AND U. FISCHBACHER (2002): “Social norms. ‘Crime’ in the lab—detecting social interactions,” European Economic Review, 46, 859–869.
- FEES, E., AND A. WOHLSCHLEGEL (2009): “Why higher punishment may reduce deterrence,” Economics Letters, 104(2), 69–71.
- FEHR, E., AND U. FISCHBACHER (2004): “Third-party punishment and social norms,” Evolution and Human Behavior, 25(2), 63–87.
- FEHR, E., AND S. GÄCHTER (2000): “Cooperation and Punishment in Public Goods Experiments,” American Economic Review, 90(4), 980–994.
- FISCHBACHER, U. (2007): “z-Tree: Zurich toolbox for ready-made economic experiments,” Experimental Economics, 10(2), 171–178.
- FRIESEN, L. (2012): “Certainty of punishment versus severity of punishment: An experimental investigation,” Southern Economic Journal, 79(2), 399–421.
- GRECHENIG, K., A. NICKLISCH, AND C. THÖNI (2010): “Punishment despite reasonable doubt - a public goods experiment with sanctions under uncertainty,” Journal of Empirical Legal Studies, 7(4), 847–867.
- GREINER, B. (2004): “An online recruitment system for economic experiments,” Discussion paper.
- HARBAUGH, W. T., N. MOCAN, AND M. S. VISSER (2013): “Theft and deterrence,” Journal of Labor Research, 34(4), 389–407.
- HOLT, C. A., AND S. K. LAURY (2002): “Risk aversion and incentive effects,” American Economic Review, 92(5), 1644–1655.
- HYLTON, K. N., AND V. KHANNA (2007): “A public choice theory of criminal procedure,” Supreme Court Economic Review, 15(1), 61–118.
- KHADJAVI, M. (2015): “On the interaction of deterrence and emotions,” Journal of Law, Economics, and Organization, 31(2), 287–319.
- LANDO, H. (2006): “Does wrongful conviction lower deterrence?,” Journal of Legal Studies, 35(2), 327–337.
- LEIBBRANDT, A., AND R. LÓPEZ-PÉREZ (2011): “The dark side of altruistic third-party punishment,” Journal of Conflict Resolution, 55(5), 761–784.
- MICELI, T. J. (1991): “Optimal criminal procedure: Fairness and deterrence,” International Review of Law and Economics, 11(1), 3–10.

- MICHAEL, J., AND H. WECHSLER (1937): “A rationale of the law of homicide II,” Columbia Law Review, 37(8), 1261–1325.
- NAGIN, D. S., AND G. POGARSKY (2003): “An experimental investigation of deterrence: Cheating, self-serving bias, and impulsivity,” Criminology, 41(1), 167.
- NIKIFORAKIS, N. (2008): “Punishment and counter-punishment in public good games: Can we really govern ourselves?,” Journal of Public Economics, 92(1), 91–112.
- NIKIFORAKIS, N., AND H.-T. NORMANN (2008): “A comparative statics analysis of punishment in public-good experiments,” Experimental Economics, 11(4), 358–369.
- PERSSON, M., AND C.-H. SIVEN (2007): “The Becker paradox and type I versus type II errors in the Economics of crime,” International Economic Review, 48(1), 211–233.
- PNG, I. P. (1986): “Optimal subsidies and damages in the presence of judicial error,” International Review of Law and Economics, 6(1), 101–105.
- POLINSKY, A. M., AND S. SHAVELL (2000): “The Economic Theory of Public Enforcement of Law,” Journal of Economic Literature, 38(1), 45–76.
- REGE, M., AND K. TELLE (2004): “The impact of social approval and framing on cooperation in public good situations,” Journal of Public Economics, 88(7), 1625–1644.
- RIZZOLLI, M., AND L. STANCA (2012): “Judicial errors and crime deterrence: theory and experimental evidence,” Journal of Law and Economics, 55(2), 311–338.
- SCHILDBERG-HÖRISCH, H., AND C. STRASSMAIR (2012): “An experimental test of the deterrence hypothesis,” Journal of Law, Economics, and Organization, 28(3), 447–459.
- SCHLAG, K. H., AND J. J. VAN DER WEELE (2013): “Eliciting probabilities, means, medians, variances and covariances without assuming risk neutrality,” Theoretical Economics Letters, 3(1), 38–42.
- TARUFFO, M. (2003): “Rethinking the standards of proof,” American Journal of Comparative Law, 51(3), 659–677.
- VOLOKH, A. (1997): “N guilty men,” University of Pennsylvania Law Review, 146(1), 173–216.

Appendix

A Proof of Proposition 2

If S_α and S_m are intervals, $\tilde{\alpha} \in S_\alpha$ and $\tilde{m} \in S_m$, then the Bayesian Nash equilibrium p^* and ϕ^* is given by the solution to the system of equations

$$p^* = p(\phi^*; F, q) = G(\tilde{\alpha}(\phi^*; F, q)) = G\left(\frac{\phi^*}{q(1-\phi^*)} \frac{A}{F}\right) \quad (7)$$

$$\phi^* = \phi(p^*; F, q) = H(\tilde{m}(p^*; F, q)) = H(A[(1-p^*)(1-v) + qv] - Fp^*(1-q)(1-v)). \quad (8)$$

Derivatives of Best-Reply Functions. For given ϕ , the impacts of F and q on p are

$$\frac{\partial p}{\partial F} = g(\tilde{\alpha}) \frac{\partial \tilde{\alpha}}{\partial F} \quad (9)$$

$$\frac{\partial p}{\partial q} = g(\tilde{\alpha}) \frac{\partial \tilde{\alpha}}{\partial q}. \quad (10)$$

Similarly, for given p , the impacts of F and q on ϕ are

$$\frac{\partial \phi}{\partial F} = h(\tilde{m}) \frac{\partial \tilde{m}}{\partial F} \quad (11)$$

$$\frac{\partial \phi}{\partial q} = h(\tilde{m}) \frac{\partial \tilde{m}}{\partial q}. \quad (12)$$

Comparative Statics of Equilibrium Punishment and Violation Probabilities. The system of total differentials of the system (7) and (8) is

$$\begin{pmatrix} 1 & -g(\tilde{\alpha}) \frac{\partial \tilde{\alpha}}{\partial \phi} & -g(\tilde{\alpha}) \frac{\partial \tilde{\alpha}}{\partial F} & -g(\tilde{\alpha}) \frac{\partial \tilde{\alpha}}{\partial q} \\ -h(\tilde{m}) \frac{\partial \tilde{m}}{\partial p} & 1 & -h(\tilde{m}) \frac{\partial \tilde{m}}{\partial F} & -h(\tilde{m}) \frac{\partial \tilde{m}}{\partial q} \end{pmatrix} \begin{pmatrix} dp \\ d\phi \\ dF \\ dq \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \quad (13)$$

Hence, the comparative statics are:

$$\frac{d\phi^*}{dF} = \frac{h(\tilde{m}) \left(\frac{\partial \tilde{m}}{\partial F} + \frac{\partial \tilde{m}}{\partial p} g(\tilde{\alpha}) \frac{\partial \tilde{\alpha}}{\partial F} \right)}{1 - h(\tilde{m}) \frac{\partial \tilde{m}}{\partial p} g(\tilde{\alpha}) \frac{\partial \tilde{\alpha}}{\partial \phi}} \quad (14)$$

$$\frac{dp^*}{dq} = \frac{g(\tilde{\alpha}) \left(\frac{\partial \tilde{\alpha}}{\partial q} + \frac{\partial \tilde{\alpha}}{\partial \phi} h(\tilde{m}) \frac{\partial \tilde{m}}{\partial q} \right)}{1 - g(\tilde{\alpha}) \frac{\partial \tilde{\alpha}}{\partial \phi} h(\tilde{m}) \frac{\partial \tilde{m}}{\partial p}} \quad (15)$$

$$\frac{d\phi^*}{dq} = \frac{h(\tilde{m}) \left(\frac{\partial \tilde{m}}{\partial q} + \frac{\partial \tilde{m}}{\partial p} g(\tilde{\alpha}) \frac{\partial \tilde{\alpha}}{\partial q} \right)}{1 - h(\tilde{m}) \frac{\partial \tilde{m}}{\partial p} g(\tilde{\alpha}) \frac{\partial \tilde{\alpha}}{\partial \phi}} \quad (16)$$

$$\frac{dp^*}{dF} = \frac{g(\tilde{\alpha}) \left(\frac{\partial \tilde{\alpha}}{\partial F} + \frac{\partial \tilde{\alpha}}{\partial \phi} h(\tilde{m}) \frac{\partial \tilde{m}}{\partial F} \right)}{1 - g(\tilde{\alpha}) \frac{\partial \tilde{\alpha}}{\partial \phi} h(\tilde{m}) \frac{\partial \tilde{m}}{\partial p}}, \quad (17)$$

Recall by Proposition 1 that $\frac{\partial \tilde{m}}{\partial p} < 0$, $\frac{\partial \tilde{m}}{\partial F} < 0$, $\frac{\partial \tilde{m}}{\partial q} > 0$, $\frac{\partial \tilde{\alpha}}{\partial \phi} > 0$, $\frac{\partial \tilde{\alpha}}{\partial F} < 0$, and $\frac{\partial \tilde{\alpha}}{\partial q} < 0$. Hence, the denominators of all right-hand sides are positive and larger than 1, and the numerator of the right-hand side of (16) is positive, and that of (17) is negative, which proves part (i) of the proposition. Furthermore, by (9), (10), (11) and (12), part (ii) of the proposition follow immediately. ■

B Translated version of sample instructions

In the following, we present sample instructions for the treatment in which participants A ("potential violators") use the strategy method. The instructions for the treatment in which participants B ("judges") used

the strategy methods differed as minimally as possible, i.e. only concerning the decision screens and their description.

General explanations

Welcome to this economic experiment.

In the following, we will explain how you can earn money by making decisions in this experiment. Please read the instructions carefully. If you have any questions, please raise your hand and we will come to your seat.

During the whole experiment it is not allowed to talk to the other participants, to use cell phones, or to launch any programs on the computer. Disregarding any of these rules will lead to your exclusion from the experiment and from all payments.

During the experiment we do not talk about Euro but points. Your total payment will be calculated in points first. At the end of the experiment, your total points will be converted into Euro, using the following rule:

500 points = 1 Euro.

At the end of the experiment you will receive your payment in cash. Your payment in this experiment depends on your own decisions and possibly on those of the other participants. Additionally, you will get 2000 points for attending the experiment in time that will be paid at the end of the experiment, independent of your decisions.

On the next pages, we describe the exact procedures of the experiment.

The experiment

In this experiment, there are **two types of participants**, A and B. You have been randomly assigned the role of participant A.²¹ There are as many participants A as participants B. Pairs consisting of one randomly chosen **participant A** and one randomly chosen participant B are matched. You will not get any information on your assigned participant's identity, neither before nor after the experiment. Your assigned participant will not get any information on your identity, either.

At the beginning of the decision phase, all participants A and B receive an initial endowment of 2400 points **each**. Additionally to the initial endowments, a donation of 2400 points to SOS Kinderdorf is planned for each pair. SOS-Kinderdorf supports children and families in the poorest regions of the world (medical care, school enrollment etc.), in crisis areas, but also neglected children and adolescents in Germany.

²¹Instructions for players B differed only in this sentence: You have been randomly assigned the role of participant B (instead of A).

SOS-Kinderdorf has been granted the DZI-certificate that attests reliability and low administrative costs. Participant A has the possibility to take the 2400 points that are designated for the donation (“donation”) for himself. If he decides to do so, there will not be any donation to SOS-Kinderdorf and participant A’s payment will rise by 2400 points to 4800 points.

Participant B can observe whether the 2400 points that are designated for donation for SOS-Kinderdorf are available or have disappeared. **IMPORTANT:** If the donation has disappeared, there are two possible reasons why the donation has disappeared. First, participant A has taken the donation for himself. Second, if participant A has decided not to take the donation, the donation can disappear by chance. This happens with probability $q\%$. Later (but before your decisions) we will tell you the exact value of q .

Only in case participant B observes that the 2400 points for donation have disappeared, he has the possibility to reduce participant A’s payment by a pre-specified number of points F (“deduction points”). These deduction points will always additionally be donated to SOS-Kinderdorf. Later (but before your decisions) we will tell you the exact value of F .

If participant B decides to deduct points from participant A, participant A’s payment will be reduced by the deduction points **for sure** and participant A’s payment is $2400 - F$ points. If participant A had taken the donation for himself, it is still donated to SOS-Kinderdorf in case participant B decided to deduct points. If the donation had disappeared by chance, it remains lost and will not be donated.

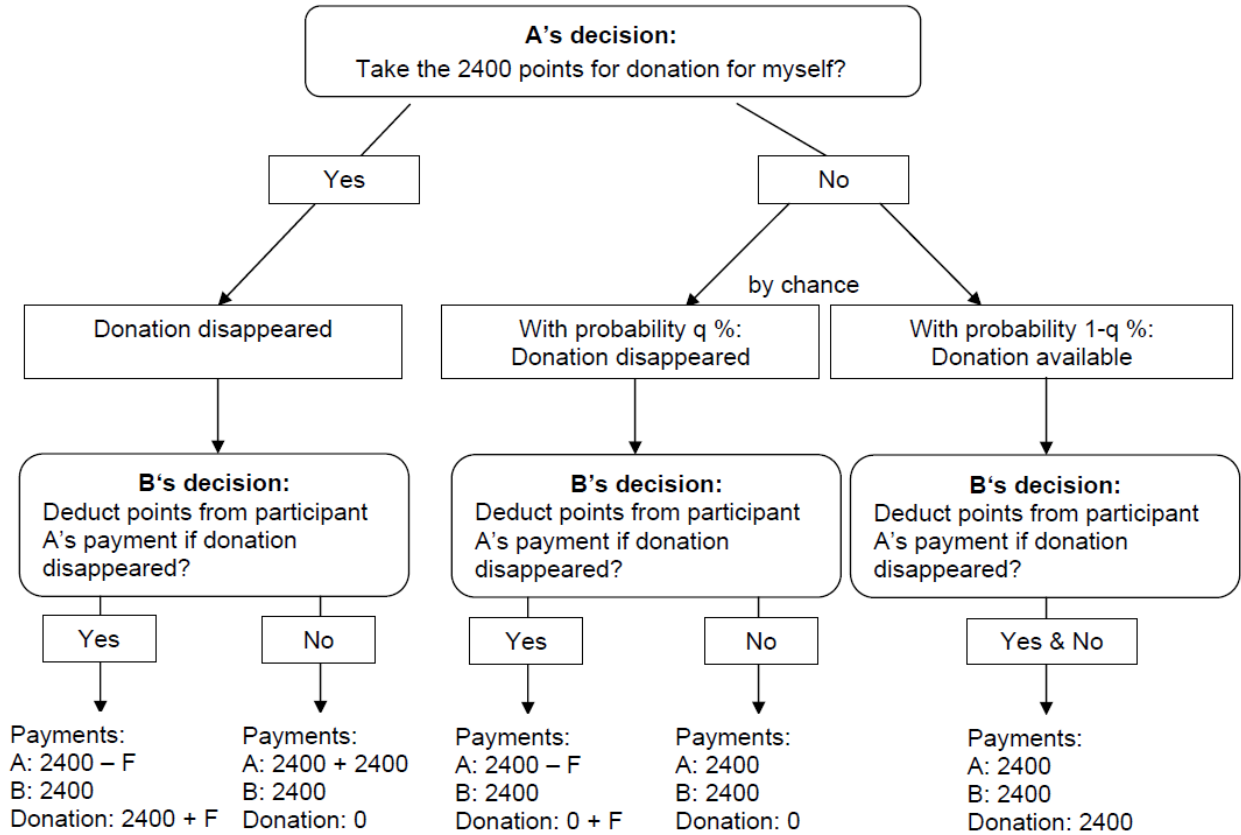
If participant B decides not to deduct any points from participant A, participant A earns his whole initial endowment of 2400 points. If participant A had taken the donation for himself, he gets this money additionally and will earn $2400 + 2400$ points donation = 4800 points. There is no donation to SOS-Kinderdorf.

To put it differently: If participant B observes that the donation has disappeared, he cannot always be sure whether it has disappeared by chance or because participant A took the donation. If participant B then decides to reduce participant A’s points, he might do it although participant A did not take the donation. If participant B decides not to reduce participant A’s points and participant A took the money, participant A is allowed to keep the money and there is no donation to SOS-Kinderdorf.

Independent of his decision, participant B always receives his initial endowment of 2400 points in cash at the end of the experiment. If the donation has disappeared, participant B’s decision influences participant A’s payment and the amount donated to SOS-Kinderdorf. Consequently, it depends on participant A’s and B’s decisions and on chance whether the donation (2400 points) will indeed be donated to SOS-Kinderdorf. **If the donation is available at the end of the experiment, it will be donated to SOS-Kinderdorf for real!**

Altogether, you will make decisions in **four different decision situations** during this experiment. All decision situations correspond to the description above. They **only differ in the amount of deduction points F and in the probability q of the donation disappearing by chance** if participant A did not take the donation. One out of the four decision situations is chosen randomly. This randomly chosen decision situation and participant A’s and participant B’s decisions in this situation determine both participants’ payments, whether there is a donation and which amount is donated to SOS-Kinderdorf. This means that all decision situations are chosen to be paid out with the same probability. Therefore, all decisions are important.

The graph on the next page summarizes all possible situations, events and payments:



You make your decisions using the computer. The example below corresponds to one out of the four decision situations.

Participants A will see the following decision screen:

Your 1. decision.
You are participant A.

The possible deduction points F are 1200 points. In case you have not taken the donation, it disappears with probability 50% by chance (therefore $q=50\%$).
Under these circumstances, would you like to take the donation of 2400 points for yourself? Please click YES or NO in all rows of the table below.

This many of the 12 participants B in this room decide to deduct points in case the donation has disappeared:

	Yes <input type="radio"/>	No <input type="radio"/>
0 (none)	Yes <input type="radio"/>	No <input type="radio"/>
1	Yes <input type="radio"/>	No <input type="radio"/>
2	Yes <input type="radio"/>	No <input type="radio"/>
3	Yes <input type="radio"/>	No <input type="radio"/>
4	Yes <input type="radio"/>	No <input type="radio"/>
5	Yes <input type="radio"/>	No <input type="radio"/>
6	Yes <input type="radio"/>	No <input type="radio"/>
7	Yes <input type="radio"/>	No <input type="radio"/>
8	Yes <input type="radio"/>	No <input type="radio"/>
9	Yes <input type="radio"/>	No <input type="radio"/>
10	Yes <input type="radio"/>	No <input type="radio"/>
11	Yes <input type="radio"/>	No <input type="radio"/>
12 (all)	Yes <input type="radio"/>	No <input type="radio"/>

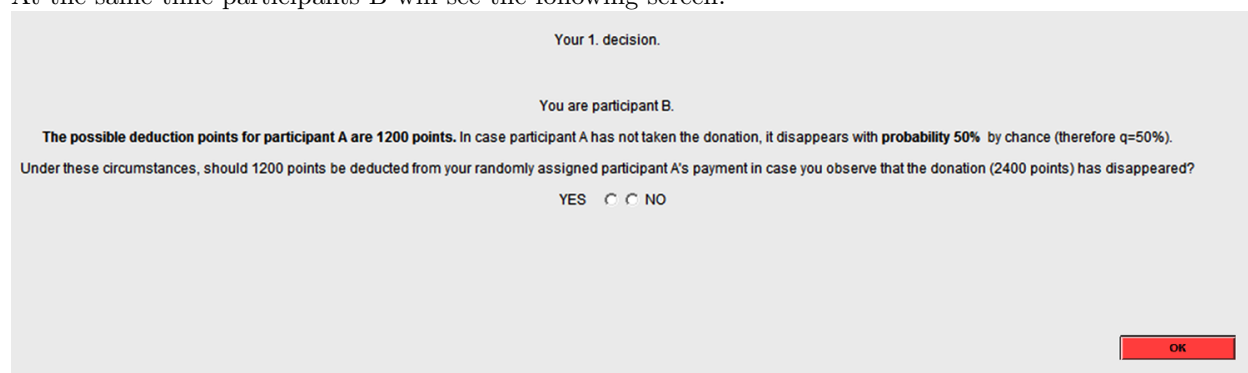
OK

The table enables participants A to condition their decision “take the donation - yes or no?” on the number of participants B who deduct F points from their participant A’s payment if they observe that the donation has disappeared. As soon as all participants B made their decision, it is determined which row of the table corresponds to the actual decisions made by the participants B in this room. Participant A’s decision (“take the donation - yes or no?”) for this row of the table is implemented.

Example: 5 out of 12 participants B in this room decide to deduct points if the donation has disappeared. Therefore, A’s decision in the row in which the 5 is written in the first column will be implemented.

For their decision, participants A click on YES or NO in all rows of the table on the corresponding screen and continue by pressing the OK-button. Afterwards, three further decisions will follow that differ in the value of possible deduction points F and / or in the probability q.

At the same time participants B will see the following screen:



Participants B click either YES or NO on the corresponding screen and continue by pressing the OK-button. Thereafter, three further decisions will follow that differ in the value of possible deduction points F and / or in the probability q.

IMPORTANT! Please note:

If $q=0\%$ and participant B observes that the donation has disappeared, he knows for sure that participant A took the donation for himself. There is no uncertainty about why the donation has disappeared because it is not possible that it has disappeared by chance ($q=0\%$).

For $q=50\%$, the following table shows how likely it is that participant B observes that the donation has disappeared if a particular number of participants A takes the donation for themselves.

For $q=50\%$: If as many participants A in this room decide to take the donation for themselves this percentage of participants B observe that the donation has disappeared.
0 (none)	50%
3 von 12	62.5%
6 von 12	75%
9 von 12	87.5%
12 (all)	100%

After all participants made their four decisions, one of these four decision situations is chosen randomly. For this decision situation, all participants will see a feedback screen with the following information:

- Randomly chosen decision situation (exact value of deduction points F and probability q)
- Decisions of all participants (A: take donation - yes or no, B: deduct points - yes or no)
- Has the donation disappeared - yes or no?
- Resulting payments for participants A and B as well as final donation

At the end of the experiment all participants A and B are **paid** their points **in cash** and the donation to SOS-Kinderdorf will be implemented for real if the money for the donation is available. Before we will show you a second feedback screen, we kindly ask you to answer several questions. Some questions can increase your earnings.

If you have any questions, please raise your hand. An experimenter will come to your seat to answer your questions. As soon as all questions are answered we will start the experiment on the computer. Before your decisions, we would like to ask you to answer several computerized questions concerning the rules of the experiment in order to ensure that you will be able to make well-informed decisions later on.