



HAL
open science

Seemingly irrelevant information? The impact of legal team size on third party perceptions

Gilles Grolleau, Murat Mungan, Naoufel Mzoughi

► To cite this version:

Gilles Grolleau, Murat Mungan, Naoufel Mzoughi. Seemingly irrelevant information? The impact of legal team size on third party perceptions. *International Review of Law and Economics*, 2022, 71, pp.106068. <10.1016/j.irle.2022.106068>. <hal-03843247>

HAL Id: hal-03843247

<https://hal.inrae.fr/hal-03843247v1>

Submitted on 12 Sep 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

Seemingly Irrelevant Information?

The Impact of Legal Team Size on Third Party Perceptions

Abstract: People often appear to use irrelevant information in forming judgments about others. Using survey experiments, we show that seemingly irrelevant facts may actually be informative of actors' choices, which third parties can use to update their beliefs. Specifically, we show that subjects' perceived severity and recommended punishment for offenses are significantly increasing in the number of lawyers representing defendants. However, once subjects are informed that the defendant was randomly assigned a specific number of lawyers, the significant relationship between the perceived seriousness of the offense and the number of lawyers largely vanishes. Thus, third parties in our benchmark analysis may be using the defendant's legal team size as a proxy to update their beliefs regarding the nature of the offense committed, as opposed to being affected by irrelevant factors in forming judgments. This is because randomization makes it impossible for third parties to draw inferences regarding the nature of the offense committed by the defendant based on the number of lawyers. However, for some offenses, we find that increasing the number of lawyers raises third parties' recommended sanctions even when the number of lawyers is randomly determined, which is consistent with a psychological phenomenon called 'luck envy'.

Keywords: lawyers; third party perceptions; moral judgment; social intuitionism; luck envy.

JEL: F23; C91.

1. Introduction

The fairness and equity principles prevalent in most justice systems suggest that similar offences ought to receive the same judgment and punishment regardless of irrelevant or incidental details. Nevertheless, an expanding literature finds that wrongdoings that violate the same ethical principles and generate equivalent amounts of harm are sometimes judged differently. In particular, judgments can be impacted by factors such as the origin (Spamann and Klöhn, 2016) and the moral character of the victim (Reich et al., 2020), the language in which observers are solicited (Hayakawa et al., 2017), and the creativity displayed by the wrongdoings of offenders (Wiltermuth et al., 2017). There is also some evidence of a relationship between the group of lawyers representing a defendant and judgments (Szmer et al. 2007).

One problem in this type of research is the difficulty in separating truly irrelevant factors from seemingly irrelevant factors. This is because some information may appear irrelevant to a researcher, even when subjects may be (implicitly or explicitly) using the information in a subtle way to update their beliefs about the choices of actors that they are judging. This creates the risk of over-attributing variations in people's judgments to irrelevant factors. Here, we consider a setting to illustrate this possibility. Specifically, we investigate how the number of lawyers representing an accused person influences third parties' judgments. Following recent scholarship in law and economics (e.g. Ben-Shahar and Chilton, 2016; Puig and Strezhnev, 2017; Olsen et al., 2018; Metcalf et al., 2020) and management (Croson et al., 2007), we conduct an experimental survey to investigate whether third parties form different judgments about people who engage in similar wrongdoings (i.e. violating the same ethical principles and generating identical harms for identical victims) based on the number of lawyers that represent them.

In our experimental survey,¹ we present subjects with scenarios which are silent on how the number of lawyers representing a person accused of committing tax evasion and fraud offenses is determined. Responses to this survey indicate a significant and positive relationship between the number of lawyers representing an individual and subjects' perceived severity of the offense as well as the punishment they deem appropriate for the offense. These results hold true in both the fraud and tax evasion contexts that we study.

Without further analysis, one may interpret this relationship as suggesting that third parties' judgments are influenced by factors which are irrelevant to the nature of the offense committed. It is possible, however, that third parties use the number of lawyers representing a person as information that is relevant to their beliefs about some other unobservable facts which are related to the nature of the offense. Specifically, we hypothesize that third parties may assume that the accused individual may choose to hire different numbers of lawyers depending on the seriousness of their offenses. Because some third parties may not be well informed of the seriousness of the offense described, they may rely on the number of lawyers hired as a proxy of the seriousness of the offense and update their beliefs accordingly.

To investigate this possibility, we conduct additional treatments wherein subjects are informed that the number of lawyers representing a person has been randomly chosen. This makes it impossible for them to make inferences regarding the seriousness of the offense based on the number of lawyers. We keep all other aspects of these additional treatments

¹ In France, where the study was conducted, lay people are selected to sit as jurors in some cases in a criminal court (*Cour d'assises*) and participate in determining whether the accused is guilty and the sentence that must be applied. Nevertheless, our observations are related to issues that go beyond the opinion formation processes of jurors. We note that formal determinations made by courts frequently impact the opinions of third parties (Jacobs and Shapira 2022). In fact, third party opinions can, in many cases, diverge from the formal determinations made by courts and can deeply affect the accused afterwards. For instance, a defendant may avoid imprisonment through an acquittal, but may nevertheless face serious stigmatization, such as diminished employment prospects or social ostracism. The presence of these social pressures is an important motivation for studying the views of lay people on this topic (see also Philippe, 2017).

identical to those in our benchmark analysis. This small change causes the significant relationship between the number of lawyers and the severity of the offense to largely vanish. This lends support to the theory that third parties, in fact, use the number of lawyers chosen by the accused as a proxy for the seriousness of the offense about which they are imperfectly informed.

Results obtained from our supplementary treatments highlight an additional interesting finding. Although the significant relationship between the number of lawyers and third parties' recommended punishment for the offense of tax evasion vanishes, it persists to some extent in the context of fraud offenses. This type of behavior is consistent with a phenomenon identified in prior studies which we call 'luck envy' wherein people's degree of altruism or spite towards another is influenced by whether that person has achieved his/her position through mere chance. Thus, our finding suggests that luck envy may be activated only in some circumstances and not others (e.g. when the person commits fraud, but not tax evasion). Because the two offenses that we presented to subjects differ from each other in more than one aspect, which we discuss below, we can only speculate which of these aspects may be responsible for activating luck envy. Thus, our results indicate the need for studying luck envy more closely to identify the conditions under which it is likely to be a more prevalent behavioral phenomenon.

The broader implication of our finding is that some factors which affect third party judgments of accused people which are seemingly unrelated to the wrongdoing may, in fact, be used by third parties in updating their beliefs about the nature of the offense. Our specific finding with respect to the number of lawyers is important for defendants, particularly those who have a significant interest in protecting their reputation. Specifically, when evaluating the returns from legal expenditures, one must bear in mind a potential off-setting effect that may

come about from inferences by juries as well as third parties regarding the seriousness of the underlying offense.

In the next section, we provide a brief review of the literature most relevant to our study and explain our main hypotheses. In sections 3 and 4, we describe our main and additional experimental surveys, and present their analyses. In section 5, we discuss the implications of our findings as well as the differences between two contexts (fraud and tax evasion) which may be responsible for the activation of luck envy. We provide concluding remarks in section 6.

2. Literature review and main hypotheses

From a rational viewpoint, similar offenses causing the same harm on identical victims being associated with similar judgments seems natural. In the rationalist perspective, moral judgments and conclusions are the results of conscious moral reasoning. The process involves careful, rational deliberation and the consistent application of general moral rules or principles, and allows the removal of irrelevant factors from the judgment process (Kohlberg, 1969).

In the social intuitionist perspective, on the other hand, moral judgment is frequently made not on the basis of a conscious reasoning, but is the result of automatic, intuitive, affective evaluations (intuitions). This perspective posits that moral reasoning is usually a post hoc construction generated after a judgment has been made (Haidt, 2001). These intuitions and the subsequent moral judgment are strongly influenced by a number of factors that can a priori seem irrelevant or incidental (e.g., Wiltermuth et al., 2017; Hayakawa et al., 2017; Reich et al., 2020). As a result, third parties' judgments and preferences for punishment often diverge from those actually rendered by official authorities (e.g. judges).

We posit that the number of lawyers representing a defendant has the potential of affecting the judgments of third parties regarding both the seriousness of the offense committed by an individual as well as the appropriate punishment for the offense in question. However, a relationship of this type, if it exists, is a priori consistent with both the social intuitionist and the rationalist perspectives. From the social intuitionist perspective, many factors can influence the judgment of third parties. Observers may, for instance, perceive defendants who are represented by a larger number of lawyers as having an unfair advantage over other individuals who have inferior representation. From the rationalist perspective, third parties may perceive the number of lawyers hired by defendants as a proxy of the seriousness of their offenses, and may recommend more severe punishments for offenses they perceive to be more serious. In this latter mechanism, third parties relate the defendant's legal team choice to the defendant's perceived seriousness of the offense, which they use to update their uninformed beliefs regarding the seriousness of the offense.

The rationale behind these dynamics can be illustrated by formalizing a simple decision faced by the defendant who chooses his/her legal expenditures, e , to minimize his/her total expected costs of $p(e)L(s) + e$. Here $p(e)$ with $p' < 0 < p''$ is the probability of being found liable and $L(s)$ with $L' > 0$ are damages incurred upon being found liable as a function of the severity of the offense s . Thus, the legal expenditures $e^*(s)$ chosen by the defendant is characterized by $p'(e)L(s) + 1=0$. Therefore, the marginal impact of the seriousness of the offense on the defendant's legal expenditures are given by $\frac{de^*(s)}{ds} = -\frac{p'(e)L'(s)}{p''(e)L(s)} > 0$, which implies that the defendant's legal expenditures are increasing in the severity of the offense.

Based on the above observation, rational third parties who are imperfectly informed of the seriousness of the offense, may associate greater expenditures, which may be manifested in the form of bigger legal teams, with more severe offenses. This can cause them to

recommend more serious punishments when they observe defendants represented by more lawyers.

This inference, however, hinges on the number of lawyers representing a defendant being determined by the defendant's choice. Therefore, we would not expect this type of inference to be operational, for instance, when legal teams are exogenously (and randomly) assigned to defendants. Thus, assuming that the type of rational reasoning described above has some effect on third parties' judgments, we would expect random lawyer assignments to defendants to mitigate the relationship between the number of defense lawyers and the judgments of third parties.

Based on the above reasoning, we formulate the following hypotheses:

Hypothesis 1. A high number of defense lawyers increases the perceived seriousness of the act committed by the accused individual.

Hypothesis 2. A high number of defense lawyers increases the severity of the recommended punishment applied to the accused individual.

Hypothesis 3. The impact of the number of defense lawyers on the perceived seriousness of the offense is mitigated when subjects are informed that the number of lawyers is not related to the defendant's characteristics.

Hypothesis 4. The impact of the number of defense lawyers on the recommended punishment is mitigated when subjects are informed that the number of lawyers is not related to the defendant's characteristics.

To test these hypotheses, we conducted an experimental survey which we briefly described in the introduction and explain in further detail, below. We also briefly note the relationship between our article and the sizeable law and economics work focusing on reputational incentives. A strand of this literature focuses on documenting and measuring the

reputational effects associated with adverse legal judgments (e.g. Karpoff and Lott, 1993; Karpoff et al., 2002). On the other hand, numerous theoretical analyses of reputational incentives question how formal incentives ought to be designed given that people are also motivated by reputational incentives (e.g. Bénabou and Tirole, 2006; Cooter and Porat, 2001; Mungan, 2016; Deffains and Fluet, 2020). We complement this literature by identifying a factor, namely legal representation, which may affect the reputational impacts of adverse judgments.

3. The experimental survey

3.1. Design and participants

To investigate the effect of the number of lawyers on the moral judgment of an accused, we designed a simple survey experiment (Weber, 1992; Croson et al., 2007; Ben-Shahar and Chilton, 2016; Puig and Strezhnev, 2017) with a between-subjects design where participants are confronted with two scenarios. The offenses committed in the two scenarios share common features, in that they violate similar ethical principles and create similar harms. Following Grolleau et al. (2020), in the first scenario we describe a financial advisor who has committed fraud against his/her clients, as follows:

By using a sophisticated illegal method, a financial advisor managed to divert 300,000 € (three hundred thousand euros) to his/her personal gain. These 300,000 € represent a direct loss for his/her customers.

The second one describes an executive who has committed tax evasion, as follows: A business manager made false declarations to the financial and stock exchange authorities, which allowed him/her to illegally gain 10,000,000€ (ten million euros) over 10 years. These 10,000,000 € should have been paid to the tax department if he/she had honestly declared all his/her income to the authorities.

The scenarios were presented to the participants in a fixed order, i.e., fraud case first and tax evasion case second. This fixed order can raise the concerns that reading the fraud case first may have impacted in some way the answers to the tax evasion case that came second. However, because the same order (fraud case first, tax evasion second) was followed in all the treatments, there is no good reason to believe that the possible effect due to a fixed order will differ across treatments. It is also worth noting that participants were informed that the two scenarios were independent. Given that we are interested in differences across treatments, we do not believe that the fixed ordering raises an important issue as noted by several studies (e.g., Schönhage and Geys, 2021; Clot et al., 2022; see also Düval and Hinz, 2020 for a more general discussion of order issues in factorial survey experiments). Moreover, while randomizing the order in which the participants saw each case would allow us to rule out the possibility of cross contamination, the fixed order of scenarios can also reflect some real-world presentations where the situations are described in a given order.

The full translation of the survey (which was originally conducted in French) can be found in the online Appendix. We note that our objective is not to compare responses across these two scenarios, but to compare variations within treatments in each scenario, which we explain, below.

Participants were asked to indicate the seriousness of each offense on a Likert scale ranging from 1 (not serious at all) to 10 (extremely serious) and the sanction they would recommend. In the first scenario the sanctions that subjects could choose were monetary (up to 5 times the amount of the fraud) and in the second scenario they consisted of imprisonment up to ten years.

For each scenario, we used three treatments where the only difference across treatments was the number of lawyers representing the defendant (1, 3, and 9). We refer to these treatments as T0, T1, and T2.

The experimental survey was conducted in November 2020 among a convenience sample (Mullinix et al., 2015; Underhill, 2019) of 225 French individuals (75 per treatment; 53% female; Average age: 30 years old; 56% with 4 years or more of university studies; and 60% earning more than 1,001€ per month). Participants (which included students, colleagues, families, friends, and acquaintances) received an e-mail inviting them to participate in an online survey, by clicking on a link. Participants were not paid and were randomly assigned to treatments. In the two scenarios proposed (fraud, tax evasion), the number of lawyers (either 1 or 3 or 9 lawyers) was manipulated between-subjects. The survey was pre-tested among a small number of acquaintances not included in the final survey to improve its readability.

3.2. Results

Responses for perceived seriousness and recommended sanctions by treatment and scenario are presented in Figure 1. We observe a positive relationship between the number of lawyers and the perceived seriousness of the offense as well as subjects' recommended sanctions.

Insert Figure 1 around here

To test the statistical significance of the differences of responses to treatments, we controlled for multiple hypotheses testing using the Stata *MHTEXP* module (List et al., 2019). This correction procedure allows comparison of all treatments simultaneously and provides adjusted p-values. The hypotheses tests are two-sided and given that the null-hypotheses are the equivalence of means, we test whether the differences of means are above or below a certain threshold value. With multiple hypothesis testing, the p-value is adjusted to minimize

Type-I and Type-II errors (see List et al., 2019 for more details). The results of these tests reveal that the observed increases are all significant (all adjusted p-values are significant at the 5% level). In other words, these findings support hypotheses 1 and 2.

Moreover, to account for potential individual heterogeneity, and since each participant answered two questions for each scenario (seriousness and punishment), we also examined the effect of the number of lawyers using a multivariate regression (Table 1). Interestingly, the previous results (i.e. MHT) are robust when controlling for individuals' age, gender, education level and earnings. However, the individuals' socio-demographic characteristics are poor predictors of the perceived seriousness and sanction.

Insert Table 1 around here

4. Additional treatments

4.1. Design and participants

In the additional treatments, we used exactly the same design as in the benchmark analysis, except that we informed the subjects that the number of lawyers was randomly determined and is independent of the defendant's preferences and resources (see the online Appendix).

Our formulation (*“Randomly and independently of his/her preferences and resources, a specialized lawyer was assigned to defend this financial advisor's interests during the legal proceedings against him/her”*) may seem somewhat abrupt to some. While we attempted several formulations in pretest versions (e.g., justifying the random number by a pilot experience to ensure a fairer justice system) to introduce a higher degree of realism, we discarded them. Indeed, these seemingly more realistic formulations come at the price of introducing some additional factors (e.g., a pilot study, references to a fairer justice system)

which create ‘noise’ in the experimental design. To avoid these problems, we implemented the cleaner and somewhat abrupt formulation.

The survey was conducted among other 225 French individuals (75 per treatment; 47.5% female; Average age: 32 years old; 61% with 4 years or more of university studies; and 59% earning more than 1,001€ per month) who received an e-mail inviting them to participate in an online survey as in the main treatments. In administering the surveys, we also followed the same procedure.²

4.4. Results

Responses for perceived seriousness and the recommended sanction by treatment and scenario are presented in Figure 2. We observe a small and positive increase in perceived seriousness when the number of lawyers is increased from a single or 3 to 9 lawyers in both scenarios. Moreover, we observe a positive relationship between the number of lawyers and respondents’ recommended punishments in the two examined scenarios.

Insert Figure 2 around here

However, when controlling for multiple hypothesis testing, the differences in the means of seriousness are insignificant across all treatments with the exception of the comparison between T1 vs. T2 in scenario 2 (which has an adjusted p-value of 0.049, placing it slightly below the conventional 5% significance level).

² We note that the design could have been cleaner, notably if all participants were randomly assigned at the same time to the different treatments (main and additional). Noteworthy, it is frequent in experimental studies to conduct additional treatments after the first treatments have been conducted (Zizzo, 2013). Although the main and additional treatments were not run simultaneously, we are unaware of any large-scale events that could have affected participants’ attitudes towards legal representation to a degree that could affect results. Moreover, the subsample characteristics show that they are comparable in terms of gender, age, education, and income.

These results suggest that when subjects were informed that the number of lawyers is independent of the defendant's resources and preferences, the hypothesis that a high number of defense lawyers increases the perceived seriousness of the act committed by the defendant is largely rejected. This observation, combined with the benchmark analysis, lends support to the hypothesis that the relationship between the number of lawyers and the perceived seriousness of the offenses is smaller when the former does not reflect anything about the characteristics of the defendant (i.e. hypothesis 3).

As for the recommended punishment, in four comparisons out of six, the differences in means are not significant. In the remaining two cases (T0 vs. T1 and T0 vs. T2 in Scenario 1), the respondents recommend a punishment that is increasing (and statistically significant) with the number of lawyers, despite insignificant differences in their perceived seriousness of the offense. Thus, although people's perceptions of the seriousness of the crime is similar across these treatments, they still exhibit a preference towards punishing the defendants represented by more lawyers more severely.

As in our main analysis, we also examined the effect of the number of lawyers in a multivariate regression (Table 2). The results are robust when controlling for individuals' age, gender, education level and earnings. When increases in the number of lawyers is independent of the defendant's resources and preferences, the perceived seriousness of the wrongdoing and recommended punishment generally do not increase, except for the sanction in Scenario 1. In other words, except for the latter, the effect of the number of lawyers vanishes when individuals are informed that this number is set independently of the defendant's preferences and resources. Again, the individuals' socio-demographic characteristics are poor predictors of the perceived seriousness and sanction. In other words, hypothesis 4 is supported only for tax evasion and not for fraud. We comment on this asymmetry and how it might relate to the concept of 'luck envy' in the next section.

Insert Table 2 around here

5. Discussion and implications

Understanding how the moral judgment of individuals is influenced by team size is an important issue. An a priori irrelevant contextual but highly visible factor can unduly shape people's perceptions in the direction that is the opposite of that pursued, since a larger team is presumably employed to increase the odds of being judged more leniently. This perception effect can influence formal judgments rendered by courts as well as third parties' judgments. The latter judgment can matter more in many circumstances than the formal one.

Our main results support the hypothesis that third parties, who are uninformed about a defendants' reasons for hiring a particular legal team, are in fact likely to judge defendants represented by bigger legal teams more harshly. These observations suggest that employing a larger defense team is not always and systematically 'good news' and can have impacts that partially or fully offset the increased quality of representation that one hopes to obtain by doing so. The public image of a large defense team can inadvertently reinforce the perceived culpability of the accused individual and this issue can be strategically used by opponents, for instance, to make someone appear more guilty than s/he actually is. Therefore, a small or even single defense attorney can be more useful than a large legal team, *ceteris paribus*.

It is important to note that our observations are related to the *visible* legal team size. This can be an important distinction given the WYSIATI (What You See Is All There Is) bias emphasized by Kahneman (2011). If the team size is not visible and operates behind the scenes, it is possible that the bias will not be activated. In short, beyond the objective content of the alleged facts, our findings suggest that decision makers should also consider automatic processes used by human beings that can shape judgments.

In addition to emphasizing these irrelevant factors and their possible consequences, we briefly suggest some ways to debias individuals or at least to attenuate this effect (Lilienfeld et al., 2009). It is important to note that debiasing can sometimes be effective, but a sizeable literature reports mixed evidence and shows that there is substantial variability with some debiasing interventions that are completely ineffective (see e.g., Weinstein and Klein, 1995 about optimistic biases; Sana et al., 2002 about hindsight bias). Recent evidence on other biases suggest that informing and training people are not always effective as expected. For instance, Burns et al. (2017) argued that reductions of implicit bias through counter stereotype training do not lead to reduced stereotype application. Nevertheless, we propose some avenues that need empirical testing to assess whether their effectiveness potential can be achieved under plausible circumstances.

It might make sense to inform observers about this influence before they form their opinion and judgment to see whether this warning cancels or at least reduces the magnitude of this effect. A related approach can be to train observers to not give too much weight to the legal team size by performing simple experimental manipulations with observers as participants and exposing them to their own bias. Nevertheless, findings regarding the effectiveness of information and training interventions in debiasing people counsels in favor of caution. Alternatively, observers can be invited to consider a counterfactual with the same elements, but wherein the defendant is represented by a single lawyer. This counterfactual can be presented in a joint (rather than separate or successive) evaluation (see Bohnet et al., 2016). Another possibility might be to activate a counterfactual mindset, before letting the individuals form their judgment by exposing them to pre-task scenarios in which the salience of counterfactual thoughts is manipulated. Indeed, Kray and Galinsky (2003) found that this debiasing intervention, i.e. activating a counterfactual mindset, is effective to correct the failure of groups to seek disconfirming information to test an initial hypothesis.

We note that many observers of trials are often not specifically informed about why a particular defendant is represented by the legal team that, in fact, represents him/her. In these cases, the dynamics behind the benchmark analysis are likely to be more relevant than those underlying the additional treatments. We conducted the latter to gain a better understanding of the mechanisms which may be responsible for the relationship between legal team size and third party perceptions. Nevertheless, there may be limited circumstances in which the dynamics behind this study might be important, too. Most notably, in the context of interest group litigation, the relationship between the size of the legal team representing the defendant and the underlying offense can be quite random. This is because many interest groups determine whom to represent not based on the merits of the case, but the constitutional principles underlying it. The results of the additional treatments, in these circumstances, can have independent value.

Beyond the implications of the additional treatments in the legal representation context, we note its broader relevance for a psychological phenomenon. Specifically, our findings in this supplementary study indicate that people may not associate a greater legal team with more serious offenses and yet recommend greater punishment for people represented by more lawyers. This is consistent with what we call “luck envy”, a concept identified in prior literature. Some authors have documented that people envy lucky people who have privileges that are perceived as undeserved or advantages that seem unrelated to their efforts. For instance, Ben Ze’ev (2000) argued that individuals “often envy lucky people” (see also Kristjansson, 2002; Ferreira and Botelho, 2021). In being envious toward these people, individuals do not accuse them of behaving criminally or immorally, but rather consider themselves as occupying an undeserved inferior situation (Ben-Ze’ev, 2000). This envy is labelled as invidious by Kristjansson (2002) and can lead envious individuals to

exhibit a greater degree of spite towards people who obtain an undeserved superior situation such as greater wealth or a better position, through mere chance.

Interestingly, this phenomenon seems more likely to be present in the fraud scenario than in the tax evasion scenario that we presented in our study. Indeed, the tax evasion and fraud offenses we described, even though they both belong to the financial domain, contained some important differences, which could be responsible for the asymmetries in our results. First, it is presumably simpler to commit tax evasion rather than fraud without any specific intent, and while tax evasion is an offense whose harms are diffusely shared by a large group of people (all tax payers), fraud impacts a smaller group of individuals (e.g. the entity from which money is embezzled). Thus, people may be better able to empathize with a person who may have inadvertently committed a tax evasion offense, and may also be more concerned with the larger per-person harm inflicted by the fraud offense. Second, taxes are perceived by some individuals as legal theft by public authorities (McGee, 1994), and thus some third parties may not find the behavior of the tax evader enviable. Third, the applicable sanctions presented in the fraud scenario were monetary (up to 5 times the amount of the fraud) whereas they consisted of imprisonment in the tax evasion scenario. Given the perceived severity of imprisonment sentences, people may exercise more care in not letting factors unrelated to the severity of the offense –whether this be luck envy or any other factor— affect their judgements. Overall, the asymmetries between the two scenarios may have caused some subjects to be envious or spiteful towards the defendant only in the fraud context, and may have caused them exercise more caution in rendering judgements in the tax evasion context.

We note that our observations regarding the circumstances which may activate luck envy cannot, unfortunately, go beyond speculation. This is because our studies were not meant to isolate the impact of each of the dimensions that we described above (i.e. diffuse versus concentrated harm, the possibility of accidental commission of the offense, the

perceived moral blameworthiness of the offense, and the type of punishment the offense triggers). We stress the importance of not over-relying on our conjectures and comments regarding luck envy. Further research isolating the marginal impact of each asymmetry we have identified is required, and this would necessitate several additional experiments. Nevertheless, we point out these differences with the hope that they may inspire future research questions.

6. Conclusion

By utilizing an experimental survey, we investigated whether an a priori irrelevant contextual factor, that is, the number of lawyers, influences third parties' moral judgments and recommended punishments of an accused individual. The findings of our benchmark analysis across two scenarios clearly indicate a detrimental effect: when the number of lawyers increases, the perceived seriousness and recommended sanction increase. We also showed that this effect is substantially reduced and even largely vanishes when participants are informed that the number of lawyers was randomly determined and is independent of the defendant's preferences and resources. Our findings echo a substantial literature supporting how irrelevant factors impact judicial decisions. Rather than over-viewing this rich body of work, we mentioned some extraneous variables that should have no bearing on legal decisions that can sway judicial rulings, such as defendant gender (Pozzulo et al., 2010), gender diversity in panels of judges (Bourreau-Dubois et al., 2020), facial features of the defendant (Johnson and King, 2017), or even judges' hungriness (Danziger et al., 2011), to quote a few.

Nevertheless, a higher number of lawyers generates some benefits and we noted the importance of analyzing the trade-off between these benefits and costs arising from the previously-mentioned effects. Simply stated, the costs of a bigger and visible legal team size should also include the indirect effect on the perceptions and moral judgments of third parties.

Interestingly, another advantage of having a large observable legal team is related to plea negotiations (which account for most criminal sanctions, at least in the US). Indeed, a large defense team can be a way to signal to the prosecution that this will be a long and costly case. This, in turn, could motivate prosecutors to attempt to secure a plea quickly through a low penalty. Exploring this additional consideration can yield interesting insights and help in getting a more complete understanding of the unexplored effects of a party's legal team size.

Our investigation has some limitations that suggest promising extensions. We are well aware that the whole figure is more complex than what was tested in our experiment. We did not investigate the whole range of numbers of defense lawyers by considering a non-linear relationship such as a (inverted-)U curve between the legal team size and moral judgments and recommended punishments. Moreover, we examined these effects in the financial domain, but it is worth examining whether they remain robust across domains (e.g., environmental wrongdoings, corruption) and whether some domains are more prone to this effect. Furthermore, it is possible that there may be gender effects when the defendant is a male or a female. Indeed, recent empirical evidence demonstrated the existence of a "gender punishment gap": "following an incident of misconduct, female advisers are 20% more likely to lose their jobs and 30% less likely to find new jobs relative to male advisers" (Egan et al., 2022). Similarly, the effect may differ when the defendant is an organization (e.g., company, non-profit organization) rather than an individual, given prior studies suggesting that the judgment of individuals and organizations may differ (e.g. Jago and Pfeffer, 2019; Jago and Laurin, 2017). Regarding organizations, as for individuals, the lawyer team size can be used as a gross indicator of the organization's means (e.g., size, revenues, number of employees, power) and infraction seriousness. Moreover, given that Jago and Pfeffer (2019) showed that people judge an identical misconduct by an organization as more unethical than that by an individual, it makes sense to check whether this effect is amplified by a larger legal team.

This topic can be further analyzed by exploring additional factors that are likely to interact with the legal team size bias. A natural candidate is to extend our analysis of team size to the prosecution side. Indeed, our paper only focuses on the side of the defense, but the prosecution could also try to exploit this bias. Do more prosecutors on the team cause penalties to rise or fall? While our focus was on criminal adjudication because of the features of the French judicial system, it makes sense, especially in the US to consider civil litigation, where sentencing is not in the hands of lay people, but setting damages, at least partially, is. Exploring these issues both with respect to the defense side and the plaintiff could be very illuminating. In addition, including a “control” with another irrelevant factor (e.g., race, gender) would give readers a sense of the scale of the effect size. For example, by altering the race or gender of the defendant, researchers can reveal what doubling the number of lawyers does in comparison to switching the defendant’s race³.

Finally, we note that the effect we identified can potentially be moderated by the adversarial parties’ legal team size (e.g. the plaintiff in the civil law context) or another feature that is perceived as putting the defendant at a disadvantage (e.g. racial discrimination). We note these open questions with the hope that they may lead to follow-up studies that better elucidate the role of additional factors.

References

- Bohnet, I., van Geen, A. and Bazerman, M. 2016. When Performance Trumps Gender Bias: Joint Versus Separate Evaluation. *Management Science* 62(5): 1225–1234.
- Bénabou, R., and Tirole, J. 2006. Incentives and Prosocial Behavior. *American Economic Review* 96: 1652–1678.

³ Noteworthy, while introducing a natural “control” with a Black versus White defendant makes sense in the USA, it is a very controversial and hot issue in France, especially in a presidential election period. Moreover, it can even raise legal issues.

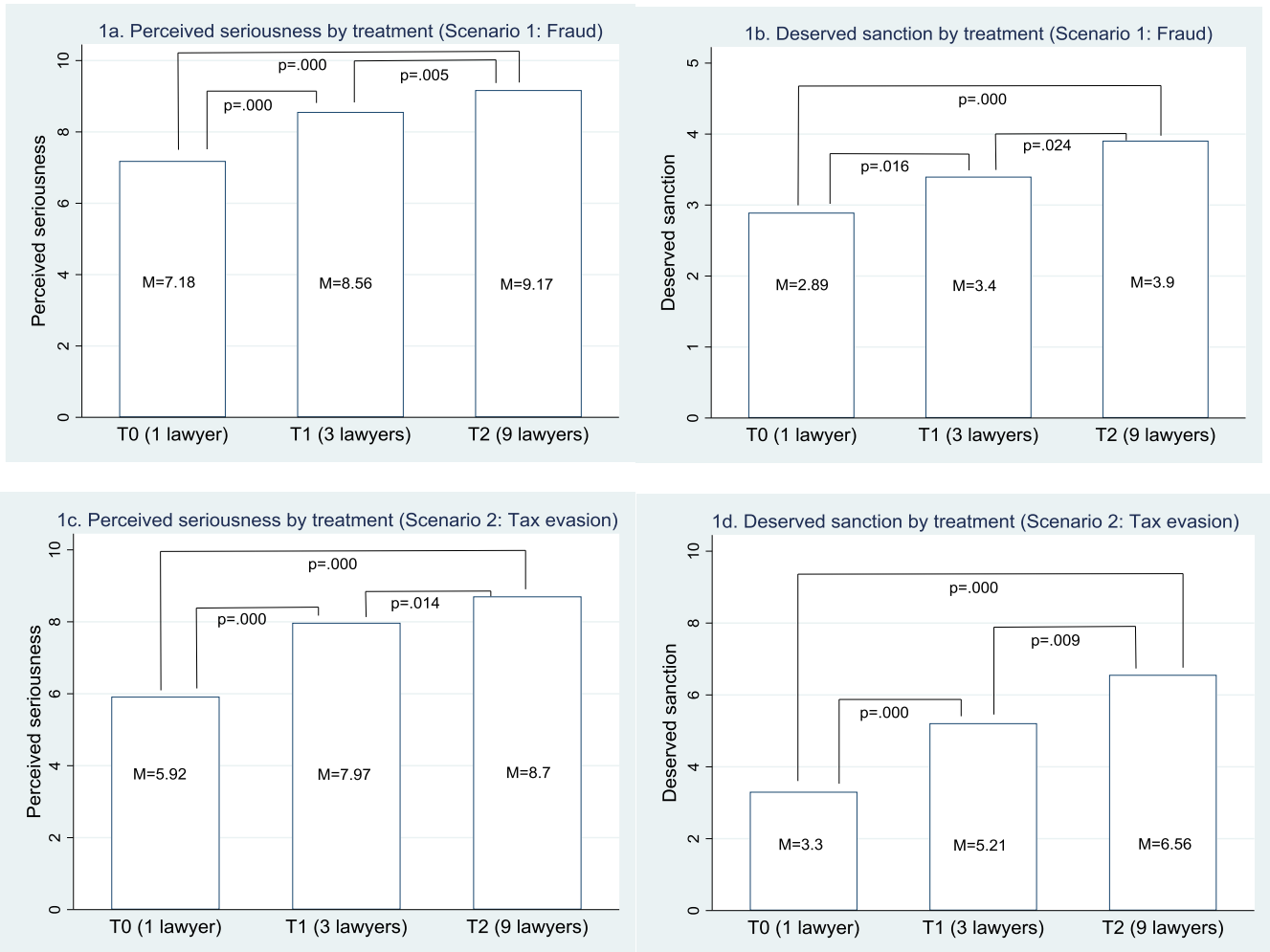
- Ben-Shahar, O., and Chilton, A. S. 2016. Simplification of Privacy Disclosures: An Experimental Test. *Journal of Legal Studies* 45: S41–S67.
- Ben-Ze'ev, A. 2000. *The Subtlety of Emotions*. Cambridge, MA: The MIT Press.
- Bourreau-Dubois, C., Doriat-Duban, M., Jeandidier, B., and Ray, J. C. 2020. Does Gender Diversity in Panels of Judges Matter? Evidence from French Child Support Cases. *International Review of Law and Economics*, 63: 105929.
- Clot, S., Grolleau, G., and Ibanez, L. 2022. Projection Bias in Environmental Beliefs and Behavioural Intentions - An Application to Solar Panels and Eco-friendly Transport. *Energy Policy* 160: 112645.
- Croson, R., Anand, J., and Agarwal, R. 2007. Using Experiments in Corporate Strategy Research. *European Management Review* 4: 173–181.
- Cooter, R., and Porat, A. 2001. Should Courts Deduct Nonlegal Sanctions from Damages? *Journal of Legal Studies* 30: 401–422.
- Danziger, S., Levav, J., and Avnaim-Pesso, L. 2011. Extraneous factors in judicial decisions. *Proceedings of the National Academy of Sciences* 108(17): 6889–6892.
- Deffains, B., and Fluet, C. 2020. Social Norms and Legal Design. *Journal of Law, Economics, and Organization* 36: 139–169.
- Düval, S., and Hinz, T. 2020. Different Order, Different Results? The Effects of Dimension Order in Factorial Survey Experiments. *Field Methods* 32(1): 23–37.
- Egan, M. L., Matvos, G., and Seru A. 2022. When Harry Fired Sally: The Double Standard in Punishing Misconduct, *Journal of Political Economy*, Forthcoming.
- Ferreira, K., and Botelho, D. 2021. (Un)Deservingness Distinctions Impact Envy Subtypes: Implications for Brand Attitude and Choice. *Journal of Business Research* 125: 89–102.

- Grolleau, G., Ibanez, L., and Mzoughi, N. 2020. Is a ‘Bad Individual’ more Condemnable than Several ‘Bad Individuals’? Examining the Scope-Severity Paradox. *Review of Law and Economics* 16: 20190017.
- Haidt, J. 2001. The Emotional Dog and Its Rational Tail: A Social Intuitionist Approach to Moral Judgment. *Psychological Review* 108, 814–834.
- Hayakawa, S., Tannenbaum, D., Costa, A., Corey, J. D., and Keysar, B. 2017. Thinking More or Feeling Less? Explaining the Foreign Language Effect on Moral Judgment. *Psychological Science* 28: 1387–1397.
- Jacob, A., and Shapira, R. 2022. An Information-Production Theory of Liability Rules. *Chicago Law Review*, Forthcoming.
- Jago, A. S., and Pfeffer, J. 2019. Organizations Appear More Unethical than Individuals. *Journal of Business Ethics* 160: 71–87.
- Jago, A. S., and Laurin, K. 2017. Corporate personhood: Lay perceptions and ethical consequences. *Journal of Experimental Psychology: Applied*, 23(1): 100–113.
- Johnson, B. D., and King, R. D. 2017. Facial profiling: Race, physical appearance, and punishment. *Criminology* 55(3): 520-547.
- Kahneman, D. 2011. *Thinking Fast and Slow*. New York. Farrar, Strauss and Giroux.
- Karpoff, J. M., and Lott Jr, J. R. 1993. The Reputational Penalty Firms Bear from Committing Criminal Fraud. *Journal of Law and Economics* 36: 757–802.
- Karpoff, J. M., Lott Jr, J. R., and Wehrly, E. W. 2005. The Reputational Penalties for Environmental Violations: Empirical Evidence. *Journal of Law and Economics* 48: 653–675.
- Kohlberg, L. 1969. Stage and Sequence: The Cognitive-Developmental Approach to Socialization. Pp. 347–480 in *Handbook of Socialization Theory and Research*, edited by David A. Goslin. Rand McNally.

- Kray, L. J., and Galinsky, A. D. 2003. The Debiasing Effect of Counterfactual Mind-sets: Increasing the Search for Disconfirmatory Information in Group Decisions. *Organizational Behavior and Human Decision Processes* 91(1): 69-81.
- Kristjansson, K. 2002. *Justifying Emotions: Pride and Jealousy*. New York: Routledge.
- Lilienfeld, S. O., Ammirati, R., and Landfield., K. 2009. Giving Debiasing Away: Can Psychological Research on Correcting Cognitive Errors Promote Human Welfare? *Perspectives on Psychological Science* 4 (4): 390–98.
- List, J. A., Shaikh, A. M., and Xu, Y. 2019. Multiple Hypotheses Testing in Experimental Economics. *Experimental Economics* 22: 773–793.
- McGee, R. W. 1994. Is Tax Evasion Unethical? *University of Kansas Law Review* 42: 411–435.
- Metcalf, C., Satterthwaite, E. A., Dillbary, J. S., and Stoddard, B. 2020. Is a Fine Still a Price? Replication as Robustness in Empirical Legal Studies. *International Review of Law and Economics*, 63: 105906.
- Mullinix, K. J., Leeper, T. J., Druckman, J. N., and Freese, J. 2015. The Generalizability of Survey Experiments. *Journal of Experimental Political Science* 2: 109–138.
- Mungan, M. C. 2016. A Generalized Model for Reputational Sanctions and the (Ir)Relevance of the Interactions between Legal and Reputational Sanctions. *International Review of Law and Economics* 46: 86–92.
- Olsen, J., Kasper, M., Enachescu, J., Benk, S., Budak, T., and Kirchler, E. 2018. Emotions and Tax Compliance Among Small Business Owners: An Experimental Survey. *International Review of Law and Economics*, 56: 42–52.
- Philippe, A. 2017. Does Introducing Lay People in Criminal Courts Affect Judicial Decisions? Evidence from French Reform. *International Review of Law and Economics*, 52: 1–15.

- Pozzulo, J. D., Dempsey, J., Maeder, E., and Allen, L. 2010. The Effects of Victim Gender, Defendant Gender, and Defendant Age on Juror Decision Making. *Criminal Justice and Behavior* 37(1): 47-63.
- Puig, S., and Strezhnev, A. 2017. Affiliation Bias in Arbitration: An Experimental Approach. *Journal of Legal Studies* 46: 371–398.
- Reich, B. J., Campbell, T., and Madrigal, R. 2020. Who Deserves Faulty Products? How Blaming the Victim Prevents Consumer Punitive Action. *Journal of Consumer Psychology* 30: 60–76.
- Sanna, L. J., Schwarz, N., and Stocker, S. L. 2002. When Debiasing Backfires: Accessible Content and Accessibility Experiences in Debiasing Hindsight. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 28(3): 497–502.
- Schönhage, N. L., and Geys, B. 2021. Partisan Bias in Politicians' Perception of Scandals. *Party Politics*. <https://doi.org/10.1177/1354068821998024>
- Spamann, H., and Klöhn, L. 2016. Justice Is Less Blind, and Less Legalistic, than We Thought: Evidence from an Experiment with Real Judges. *Journal of Legal Studies* 45: 255–280.
- Szmer, J., Johnson, S. W., and Sarver, T. A. 2007. Does the Lawyer Matter? Influencing Outcomes on the Supreme Court of Canada. *Law and Society Review* 41: 279–303.
- Underhill, K. 2019. Price and Prejudice: An Empirical Test of Financial Incentives, Altruism, and Racial Bias. *Journal of Legal Studies* 48: 245–274.
- Weber, J. 1992. Scenarios in Business Ethics Research: Review, Critical Assessment, and Recommendations. *Business Ethics Quarterly* 2: 137–159.
- Weinstein, N. D., and Klein, W. M. 1995. Resistance of Personal Risk Perceptions to Debiasing Interventions. *Health Psychology* 14(2): 132–140.

- Wiltermuth, S. S., Vincent, L. C., and Gino, F. 2017. Creativity in Unethical Behavior Attenuates Condemnation and Breeds Social Contagion When Transgressions Seem to Create Little Harm. *Organizational Behavior and Human Decision Processes* 139: 106–126.
- Zizzo, D. J. 2013. Claims and Confounds in Economic Experiments. *Journal of Economic Behavior and Organization* 93: 186–195.



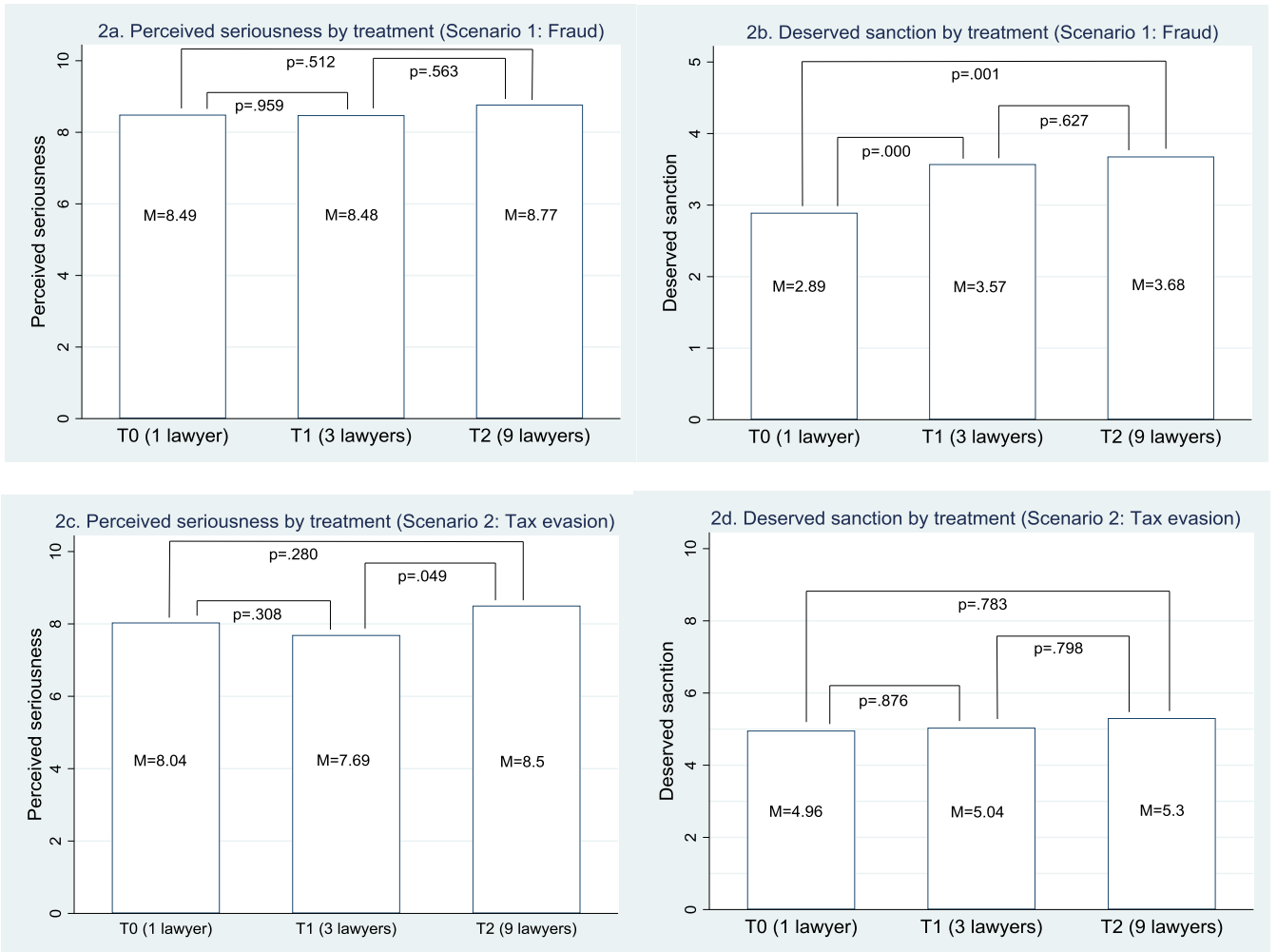
* 75 observations per treatment.

Figure 1. Perceived seriousness and deserved sanction by treatment and scenario (Benchmark analysis)

Table 1. Effect of treatments on perceived seriousness and sanctions (Benchmark analysis)

	Scenario 1 (Fraud)		Scenario 2 (Tax evasion)	
	Perceived seriousness	Sanction	Perceived seriousness	Sanction
One lawyer (Ref)	-	-	-	-
Three lawyers	1.393***	0.544**	1.950***	1.825***
Nine lawyers	1.877***	1.062***	2.673***	3.281***
Age (continuous)	0.017*	0.005	0.005	-0.006
Gender (=1 if female)	-0.076	-0.130	-0.103	-0.296
Education	Cat. 1 (Ref)	-	-	-
	Cat. 2	0.057	-0.254	0.099
	Cat. 3	0.251	0.046	0.623
Earnings	Cat. 1 (Ref)	-	-	-
	Cat. 2	0.374	0.001	0.402
	Cat. 3	0.237	-0.418*	0.164
Constant	6.386***	2.942***	5.319***	3.108***
Observations	225	225	225	225
F	9.44***	3.80***	11.42***	6.06***
R2	0.2591	0.1235	0.2973	0.1835
	Multivariate test, F		Multivariate test, F	
Wilks' lambda	0.6874***		0.6769***	
Pillai's trace	0.3235***		0.3277***	

For the variable Education, Cat. 1 to 3 refer to French baccalaureate or less, between 1 and 3 years of university studies, and 4 years or more of university studies, respectively. For the variable Earnings, Cat. 1 to 3 refer to $\leq 1000\text{€}/\text{month}$, between 1001€ and $2000\text{€}/\text{month}$, and $> 2000\text{€}/\text{month}$, respectively. ***, ** and * refer to significance at the levels of 1%, 5% and 10% respectively.



* 75 observations per treatment.

Figure 2. Perceived seriousness and deserved sanction by treatment and scenario (Additional treatments)

Table 2. Effect of treatments on perceived seriousness and sanctions (Additional treatments)

		Scenario 1 (Fraud)		Scenario 2 (Tax evasion)	
		Perceived seriousness	Sanction	Perceived seriousness	Sanction
One lawyer (Ref)		-	-	-	-
Three lawyers		0.205	0.720***	-0.129	0.181
Nine lawyers		0.218	0.648***	0.320	0.166
Age (continuous)		0.035***	0.019**	0.032**	0.035*
Gender (=1 if female)		0.528**	0.034	0.408	0.207
Education	Cat. 1 (Ref)	-	-	-	-
	Cat. 2	-0.473	-0.051	0.377	1.015
	Cat. 3	0.124	0.247	1.136***	1.275**
Earnings	Cat. 1 (Ref)	-	-	-	-
	Cat. 2	-0.135	-0.013	-0.387	-0.697
	Cat. 3	-0.446	-0.627**	-0.285	-0.445
Constant		7.237***	2.314***	6.203***	3.100***
Observations		225	225	225	225
F		4.06***	3.41***	4.09***	1.322
R2		0.1308	0.1123	0.1317	0.0467
		Multivariate test, F		Multivariate test, F	
Wilks' lambda		0.7874***		0.8541***	
Pillai's trace		0.2234***		0.1480***	

For the variable Education, Cat. 1 to 3 refer to French baccalaureate or less, between 1 and 3 years of university studies, and 4 years or more of university studies, respectively. For the variable Earnings, Cat. 1 to 3 refer to $\leq 1000\text{€}/\text{month}$, between 1001€ and $2000\text{€}/\text{month}$, and $> 2000\text{€}/\text{month}$, respectively. ***, ** and * refer to significance at the levels of 1%, 5% and 10% respectively.