Moral judgment of environmental harm caused by a single versus multiple wrongdoers: A survey experiment

Gilles Grolleau, Lisette Ibanez, Naoufel Mzoughi

Moral judgment of environmental harm caused by a single versus multiple wrongdoers:

A survey experiment

Gilles Grolleau

CEREN EA 7477, Burgundy School of Business, Université Bourgogne Franche-Comté, Dijon, France

CEE-M, Université de Montpellier, CNRS, INRAE, Supagro, Montpellier, France

grolleau@supagro.inra.fr

Lisette Ibanez

CEE-M, Université de Montpellier, CNRS, INRAE, Supagro, Montpellier, Montpellier, France

lisette.ibanez@inrae.fr

Naoufel Mzoughi (Corresponding author)

ECODEVELOPPEMENT, INRAE, 84000, Avignon, France

naoufel.mzoughi@inrae.fr

Abstract: Some experimental and archival studies have found empirical support for the scope-severity paradox (SSP), according to which the perceived harm of the same crime or wrongdoing decreases when the number of victims is greater. In the context of environmental wrongdoing, we investigate whether the SSP applies when the number of *perpetrators* of a crime or wrongdoing increases. Using a survey experiment with two scenarios and five treatments (variations of the number of perpetrators and the individual and total harms committed), we test whether the perceived severity and punishment recommendation for perpetrators of an environmental wrongdoing decrease as the number of perpetrators increases, independent of the total environmental harm committed. Unlike the studies that look at the SSP phenomenon as regards number of victims, we do not find direct support for the existence of a SSP effect regarding number of perpetrators. We do find, however, that participants evaluating the one-perpetrator treatments are more likely to judge with the highest severity. We also provide some collateral insights such as the insensitiveness of participants to the individual pollution level, once the environmental damage exceeds a

certain threshold. Our results extend previous SSP studies in important directions and suggest some policy implications, and avenues for further research.

Keywords: environmental wrongdoing; pollution; punishment; scope-severity paradox.

JEL numbers: C91; D91.

Moral judgments of an environmental harm caused by a single versus multiple

wrongdoers: A survey experiment

1. Introduction

The popular wisdom frequently goes as follows: harming many people is more serious than harming a few individuals. As such, it would stand to reason that the former would be judged and punished more severely than the latter. Nevertheless, the validity of this reasoning has been challenged in the literature. Using three studies (two laboratory experiments and an archival data study) in the financial and environmental domains, Nordgren and McDonnell (2011; see also Konis et al, 2016) found that the perceived severity of a wrongdoing was negatively correlated to the number of people victimized. The authors found that an increase in the number of victims actually decreased the perceived severity of the same wrongdoing, causing external evaluators or judges to recommend less severe punishment for individuals that victimized more people.

This counterintuitive phenomenon is known as the scope-severity paradox (SSP), which is notably due to the identifiability bias (Jenni and Loewenstein, 1997; Small and Loewenstein, 2006; Daniels, 2012; Lewinsohn-Zamir et al., 2017), according to which people tend to identify more emotionally with one well-identified individual victim than with numerous statistical ones. This bias was well-summarized by the Nobel Peace Prize recipient, Mother Teresa who stated: *"If I look at the mass I will never act. If I look at the one, I will"* (Slovic, 2007). Slovic (2007) argues that this effect comes from the fact that it is easier for people to put themselves in the shoes of one person than in the shoes of many. In addition, people's reactions are more cognition-driven when deciding about groups and more emotion-driven when deciding about individuals (Kogut, 2011). Recently, several studies found empirical support that describing a

well-identified needy individual is a more effective way to raise donations than evoking a statistically large number of victims (Kogut and Ritov, 2005; Small et al., 2007; see Lee and Feeley, 2016 for a meta-review). Interestingly, this bias has been notably studied regarding environmental issues and there is increasing support that it is likely to affect numerous environment-related decisions (e.g., Hsu, 2008; Markowitz et al. 2013; Thomas-Walters and Raihani, 2016) even if some studies are more skeptical (e.g., Pellegrin et al., 2018). Interestingly, people may be more likely to help others, e.g., by donating more to large numbers of victims (both human and non-human) if the victims are perceived as an entity (Smith et al., 2013).

Rather than investigating how the number of victims may impact the perception of harm in the environmental realm, a subject that has already benefited from a sizeable amount of academic attention, we extend the previous literature by examining whether the scope-severity paradox holds when the number of *perpetrators is greater*. For example, will citizens, consumers or judges award harsher sanctions against a single company that has polluted the environment, than against two or more similar companies that have each committed the same wrongdoing? Investigating this question is not only a matter of academic curiosity. Indeed, some companies are frequently suspected of environmental wrongdoings either at the individual company level or at the more collective level. Pollution events such as oil spills can be evoked by using a case involving either a unique perpetrator such as the BP Deepwater Horizon rig in the Gulf of Mexico, a small number of perpetrators, or a large number of perpetrators such as a list of oil spills involving several companies over the last decade. In another context, in September 2018, the European Union launched an in-depth investigation into alleged collusion of the so-called "circle of five" German automotive brands (BMW, Daimler, Volkswagen, Audi and Porsche)

in denying consumers the opportunity to buy less polluting cars, despite the technology being available to the manufacturers of vehicles (Keating, 2018).

Investigating the scope-severity paradox as regards the perpetrator extends the results of Kogut (2011). Kogut (2011) stresses that "taking the injured side increases willingness to punish mostly identified wrongdoers" (Kogut, 2011, p. 1495). She also indicates that the results of two studies presented suggest complex relationships between the identifiability of the wrongdoer and the willingness to punish. In line with previous research on the identifiability effect, "the identifiable wrongdoer evokes more intense emotions (positive and negative) than an unidentified one" (p. 1497). We go a step further by considering whether an increase in the number of wrongdoers (which can also be a way to decrease identifiability) leads to less severe judgment and punishment by outside observers. This question is especially important nowadays, when spillovers involving one wrongdoer can rapidly transform into a situation where several wrongdoers appear to be involved (Roehm and Tybout, 2006; Lee, 2017). In addition to a better understanding of outside observers' reactions, our study can also indicate whether wrongdoers are likely to unduly influence social and formal judgment thanks to the manipulation of the perceived number of wrongdoers.

It is noteworthy that several non-mutually exclusive mechanisms used to explain the identifiability bias (Jenni and Loewenstein, 1997) can also be applied to the number of perpetrators, notably the vividness and reference group effects. The vividness effect suggests that the colorful description of an individual perpetrator (or a small group of perpetrators) is more likely to evoke powerful emotional responses than a higher number of unidentified, statistical or not fully identified perpetrators (see Genevsky et al., 2013 for a development of this argument in the case of identified versus statistical victims). The reference group effect is

related to the tendency of people to feel greater concern toward some specific perpetrators as the reference group people are part of grows smaller. For instance, sanctioning 10 polluters out of a group of 10 farmers in a given area can be perceived as significant while sanctioning 10 polluters out of a group of thousands of farmers is likely to be perceived as much less noticeable.

Identifiability-related mechanisms are not the only ones that can explain the SSP in terms of the judgment of perpetrators. For instance, in cases with many wrongdoers, the social norms of what is accepted behavior can change. Indeed, a greater number of wrongdoers can (inadvertently) indicate a descriptive social norm (Farrow et al., 2017) that can be counterproductive to reach the pursued objective. Consistent with our reasoning, Hauge (2015) found that moral opinions depend on what other people actually do. Consequently, the fact that there are multiple wrongdoers may indicate the presence of a social norm that leads people to judge with less severity and recommend weaker punishment.

As far as we know, no study has examined whether an increase in the number of perpetrators of an environmental crime or wrongdoing decreases the perceived severity and recommended punishment of each perpetrator. Using a survey experiment, we therefore explore a new hypothesis, by testing whether *a higher number of perpetrators in an environmental crime leads to a more lenient judgment*. The originality of our research is at least twofold. First, we test the validity of the scope-severity paradox in terms of the perpetrator by varying the number of perpetrators in a between-subjects design. Second, we test the SSP in the environmental domain, which differs from other types of wrongdoings such as financial crimes. In an environmental crime or wrongdoing, the damage involves natural entities rather than money. As stressed above, this study will allow us to reach a better understanding of external observers' reactions to identical environmental wrongdoings as well as to anticipate whether wrongdoers

are likely to anticipate external observers' reactions and attempt to manipulate the perceived or real number of wrongdoers.

The remainder of our paper is organized as follows. Section 2 describes our empirical strategy. Section 3 exposes our main findings and discusses them. Section 4 concludes and suggests possibilities for further research.

2. Design and empirical strategy

In order to test our hypothesis, we conducted a survey experiment[1] among a convenience sample of the French population using the Foule Factory platform, which is considered as the French equivalent of Amazon Mechanical Turk (AMT).[2] Like AMT and comparable platforms, Foule Factory offers several major elements useful to conducting research, such as an integrated participant compensation system that respects French minimum wage laws, a large participant pool and a streamlined process of study design, participant recruitment, and data collection (Buhrmester et al., 2011). Consequently, using Foule Factory workers is a convenient way to obtain participation from a wide variety of individuals and seemed appropriate for our survey experiment.

Participants were invited to read two hypothetical scenarios related to two different environmental wrongdoings. The first scenario described the pollution of a river by a given

---

[1] Following Akay et al (2012, see also Graf et al., 2012), our use of the term "survey experiment" is motivated by the fact that our survey uses a between-subjects design in which participants are randomly assigned to various conditions.

[2] https://www.foulefactory.com/.

quantity of whey[3] thrown into the stream by a company, which caused a visible biodiversity loss, an increase in fish mortality and unpleasant odors in the area. The second scenario involved illegal deforestation by a real estate company of several hectares of a forest area (see Appendix 1 for the details of the scenarios and instructions). This second scenario indicated a clear delineation of the harm caused, that is, the number of hectares of forest illegally destroyed. In each scenario, participants were asked to indicate the severity of the committed crime on a Likert scale ranging from 1 (not severe at all) to 10 (extremely severe) and the number of years they felt the perpetrators should spend in jail, with a maximum threshold of 10 years in the river pollution scenario and 15 years in the deforestation scenario.

The scope-severity paradox was tested by varying the number of perpetrators with a between-subjects design, that is, one perpetrator (1) versus a small number of perpetrators (2) versus a large number of perpetrators (20). These numbers were directly inspired from those used in the original study of Nordgren and McDonnell (2011), except that we added a treatment with just one perpetrator. In order to control for the individual vs. collective levels of harm caused, our experimental design is composed of five treatments. In some of these treatments, the harm caused per individual (Pi) is maintained constant, while in other treatments, the total harm (Pg) is maintained constant (see Table 1). Participants were randomly assigned to treatments. To avoid any order effect, half of the participants read the river pollution scenario first and the other half read the deforestation scenario first.

---

[3] According the Merriam-Webster dictionary (https://www.merriam-webster.com) whey is "the watery part of milk that is separated from the coagulable part or curd especially in the process of making cheese and that is rich in lactose, minerals, and vitamins and contains lactalbumin and traces of fat".

[Insert Table 1 around here]

In the first three treatments (T1, T2, T20), the total harm to the environment increases with the number of perpetrators, while the damage caused per individual remains the same. The two other treatments (T1*, T2*), however, capture situations where the individual damage caused changes while the total harm remains constant. When keeping the *per* person harm caused constant (Pi) and increasing the number of perpetrators, the total harm caused (Pg) increases with the number of perpetrators. If severity judgments (Si) are based on individual behavior (regardless of behavior by others), the severity judgment will be identical for all treatments where the individual harm remains constant (i.e. Si1=Si2=Si20). If respondents consider the total harm inflicted on the environment, this could lead to an increase in perceived severity. Nevertheless, the social norm effect and other identifiability-related mechanisms may lead people to judge the same wrongdoing less harshly when perpetrators are numerous. In sum, we consider that there is evidence of the scope-severity paradox if the perceived severity (Si) and/or the punishment (Yi) decrease(s) with the number of perpetrators, when the harm inflicted per individual remains the same (i.e. Si1>Si2>Si20 and/or Yi1>Yi2>Yi20).

In order to capture some collateral effects, we also investigated participants' reactions to various combinations of total and per individual harm caused to understand whether or not the severity of participants' judgment was motivated by the overall harm caused. In a given wrongdoing committed by multiple individuals, each perpetrator is arguably responsible for less harm than one perpetrator would be responsible for if they committed the same wrongdoing alone. In addition, as the one perpetrator has solely caused this harm the crime might be viewed by external observers as more unusual and thus be judged more severely. For instance, there could be a social norm that judged such behavior as unacceptable. These two effects could potentially

9

go in the same direction: the wrongdoing might be perceived as more severe when there is a single perpetrator than when there are several perpetrators. In other words, our main hypothesis can be reformulated as follows:

H1: When the environmental damage caused per individual is fixed, a higher number of perpetrators leads to a more lenient judgment and sanction: $S_{i1}>S_{i2}>S_{i20}$ and $Y_{i1}>Y_{i2}>Y_{i20}$.[4]

Moreover, our experimental design also allows us to test some secondary effects related to the level of environmental damage caused:

H2: When the number of perpetrators is fixed, greater environmental damage caused per individual leads to a more severe judgment and sanction: $S_{i1}<S_{i1}^{*}$, $S_{i2}<S_{i2}^{*}$ and $Y_{i1}<Y_{1}^{*}$, $Y_{i2}<Y_{i2}^{*}$.

This hypothesis is motivated by the incurred environmental cost and the positive correlation between the level of this cost and the willingness to pay for pollution abatement. Indeed, various studies (Barriero et al., 2005; Andersson et al., 2010) show a strong positive relationship between the level of pollution (e.g., air pollution, noise) and economic valuation.

---

[4] Getting only one (or more) relationship(s) that holds can be interpreted as partial support for H1. In case of partial support for H1, an important issue that is beyond the scope of our contribution is to examine whether the relationship between the number of perpetrators and the severity of the judgment (or recommended punishment) exhibits discontinuities, leading to a more complex figure.

We also considered an additional related insight that can be considered as a complementary and possibly competing hypothesis compared to H2.[5] *Above a certain (unknown) level of environmental damage caused, participants become almost insensitive to numbers and similarly judge and punish perpetrators irrespective of their individual contribution to the environmental damage:* $S_{i1}^* \cong S_{i2}^* \cong S_{i20}$ and $Y_{i1}^* \cong Y_{i2}^* \cong Y_{i20}$.

This insight echoes the findings of some studies that illustrate that willingness to pay (WTP) to protect some endangered species is capped beyond certain limits. For instance, Desvousges et al. (1993) found that the WTP to support the protection of migratory birds from exposure to oil ponds increased only slightly when the number of affected birds was increased from 2,000 to 200,000 (see also Kahneman and Ritov, 1994). Carlsson and Johansson-Stenman (2000) also showed an insensitivity-to-scope effect in the WTP for improved air quality for people exposed to serious air pollution. Nevertheless, our design does not allow us to fully test this insight.

In order to test the effect of the number of perpetrators on the severity of a participant's judgment and recommended sanction, we use both a Student test, comparing mean individual responses across treatments, and multiple hypotheses testing using the MHTEXP module developed by List et al (2019). In addition to the Bonferroni and Holm adjusted p-values, MHTEXP provides another adjusted p-value that we call "LSX", which has been shown to be an improvement of both other statistics and is referred to as Thm3_1 in both the MHTEXP

---

[5] These two hypotheses are not necessarily mutually exclusive. For instance, one can claim that the assumption that greater environmental damage caused per individual leads to a harsher judgment and sanction (H2) is likely to apply below a given damage threshold, while above this threshold, participants become almost insensitive to numbers and judge and punish perpetrators similarly irrespective of their individual contribution to the environmental damage.

module and the original work of the authors (List et al., 2019). The tests of the hypotheses are two-sided. The null-hypotheses being the equivalence of means, we test whether differences of means are above or below a certain threshold value. With the multiple hypothesis testing, the p-value is adjusted to minimize the Type-I and Type-II errors. Moreover, in order to more generally analyze our results and take into account the potential heterogeneity of individuals, we use two regressions, controlling for age, gender, education level and earnings.

The relationship between the number of perpetrators and the perceived severity of the committed wrongdoing is analyzed using a probit regression (Greene, 2003). More formally, we create a dummy variable for extreme severity D_i from the observed variable S_i – corresponding to perceived severity – defined by:

$$D_i = \begin{cases} 0 & \text{if} \quad S_i \leq u_9 \\ 1 & \text{if} \quad S_i = u_{10} \end{cases}$$

We thus consider the probit model defined by $D_i = X_i\beta + \varepsilon_i$, where X_i is the vector of exogenous variables, β represents slope coefficients to estimated, ε_i is the disturbance term, which is assumed to be normally distributed with zero mean, and u9 to u10 correspond to threshold parameters.

Nevertheless, given that the number of years of a prison is a count variable, the effect of the number of perpetrators on recommended punishment is studied using a Poisson regression (Cameron and Trivedi, 1998). Let $Y_i (i = 1, 2, \ldots, N)$ be the dependent variable, which represents the number of years in prison recommended by each individual. Assume that Y_i is independently Poisson-distributed. The conditional mean is specified as: $E(Y_i|X_i) = exp(X_i'\beta)$, where X_i and β are the vectors of explanatory variables and parameters, respectively.

Last but not least, besides socio-economic variables, we also collected additional data to explore the main drivers of participants' judgments and recommended punishments of the environmental crimes. More precisely, participants indicated on a 7-point Likert scale the importance of each of the following factors in explaining their responses: the violation of an ethical standard ("breaking a law"), the environmental damage caused by the considered perpetrator ("individual harm to the environment") or the total harm caused by all the perpetrators ('unacceptable state of environment"). (See Appendix 1 for the precise questions.)

3. Results and discussion

For our experiment, we enlisted the participation of 774 Foule Factory workers, which we considered to be an acceptable sample size. Indeed, using a statistical analysis tool called the G*Power Program[6] we found that only 300 responses (around 60 subjects per group) are required to get a statistical power score of 0.95 with a 0.25 effect size. Our sample size is largely above this threshold. Some descriptive statistics regarding our sample are provided in Table 2. A Kruskal-Wallis test confirms that our sample is relatively well-balanced between treatments for socio-economic variables.[7] Moreover, before analyzing the responses regarding severity and recommended sanctions across treatments, let us mention some statistics about the main drivers of the participants' choices (Table 3). One can see that participants cited environmental

---

[6] http://www.psychologie.hhu.de/arbeitsgruppen/allgemeine-psychologie-und-arbeitspsychologie/gpower.html.

[7] In our sample, age categories are representative of the French population. Moreover, concerning the income level, 50% of our sample have monthy earnings above the average monthly revenue in France, which is 1,500€. However, in our sample, men and highly-educated people are over-represented (58% of men vs 49% in the French population; 73% have a Bac+2 level vs 20% of the overall French population). Consequently, we caution the reader not to over-generalize our results to the French population.

harm caused per individual as the most important factor in the severity of their judgment and recommended punishments, closely followed by the total harm to the environment.

[Insert Tables 2 and 3 around here]

Now, let us analyze mean individuals' responses in severity and recommended number of years of prison sentences. As mentioned in the previous section, we report both the results of a student test (Table 4) and multiple hypotheses testing (i.e., all treatments are simultaneously compared to each other) using the MHTEXP module (Table 5). For ease of exposition, we report in Table 4 only the most relevant comparisons, that is, those closely related to our hypotheses. The complete results are available from the authors upon request. In general, the results obtained by the two tests are rather consistent, despite some differences. Below, we analyze the results in relation to each hypothesis, considering scenario A followed by scenario B.

[Insert Tables 4 and 5 around here]

Regarding Scenario A, Table 4 indicates that there is no statistically-significant difference between the mean responses in severity of perceived harm for treatments T1, T2 and T20, i.e., $Si1$, $Si2$, and $Si20$ are not significantly different. There is also no statistically significant difference in terms of the recommended sanction between T1 and T2 (i.e., no significant difference between $Yi1$ and $Yi2$). However, the recommended sanction is significantly higher for T20 compared to T1 ($Yi20>Yi1$). In other words, when the number of perpetrators is 20, the recommended punishment is higher, compared to the case where there is one perpetrator. Moreover, when correcting for MHT (Table 5), the results remain quite similar, except that the difference between the recommended sanctions in T1 and T20 does not reach the conventional

level of statistical significance (LSX p-value = 0.112). Hence, for scenario A, the results do not support our hypothesis H1, and even suggest an inverse relationship in some cases (Yi20>Yi1).

Regarding Scenario B, Table 4 indicates that there are no statistically significant differences in terms of perceived severity of harm across the treatments T1, T2 and T20, i.e., between Si1, Si2 and Si20. There is also no significant difference in terms of the recommended sanction between T1 and T2 (that is, no difference between Yi1 and Yi2). However, the recommended sanction is significantly higher for T20 compared to both T1 and T2 (Yi20>Yi1; Yi20>Yi2). When correcting for MHT (Table 5), the results are similar, except that the difference between T2 and T20 is only close to reaching the conventional level of statistical significance (LSX p-value = 0.101). Thus, for scenario B as well, the results do not support our hypothesis H1 and even suggest an inverse relationship in some cases (Yi20>Yi1).

Overall, regarding H1, our results do not support the existence of an SSP effect for environmental wrongdoing when the number of perpetrators increases in terms of either the severity of judgment or recommended punishment. In other words, when the individual environmental damage is fixed, a greater number of perpetrators does not lead to a more lenient average judgment and sanction.

Let us now examine the results in relation to our H2 hypothesis which states that when the number of perpetrators is fixed, greater environmental damage caused per individual leads to a harsher judgment and sanction. First, for scenario A, Table 4 indicates that there is no significant difference in terms of perceived severity of harm between treatments T1 and T1* or between treatments T2 and T2*. Regarding the recommended sanction, we found that Yi1* was significantly higher than Yi1. In other words, when there was only one perpetrator, we found

that greater environmental damage caused per individual led to a harsher sanction. The previous result, however, does not hold when comparing treatments T2 and T2* (i.e., Yi2 and Yi2* are not significantly different). Interestingly, when correcting for MHT (Table 5), we obtain the same result, where the only significant difference is between Yi1* and Yi1. In other words, with the exception of this last case of a harsher recommended sanction in the one-perpetrator treatments in scenario A, the results do not support our H2 hypothesis.

Regarding scenario B, Table 4 shows that when there is only one perpetrator (treatments T1 and T1*), Si1* and Yi1* are significantly higher than Si1 and Yi1, respectively. In other words, in the one-perpetrator treatments, an increase of the environmental damage caused per individual is associated to more severe judgment and punishment. However, when there are two perpetrators (treatments T2 and T2*), there is no significant difference either in terms of the perceived severity or the recommended sanctions, since Si2* and Yi2* are not statistically different from Si2 and Yi2, respectively. In addition, when controlling for MHT (Table 5), we found no significant difference between T1 and T1*, on one hand, and between T2 and T2*, on the other. Moreover, these results hold for both the perceived severity (Si1 vs. Si1* and Si2 vs. Si2*) and the recommended sanctions (Yi1 vs. Yi1* and Yi2 vs. Yi2*). Hence, , we do not find evidence to support our H2 hypothesis in our scenario B results.

In sum, we do not find evidence in our scenario A or scenario B results to support our H2 hypothesis. When the number of perpetrators is fixed, greater environmental damage caused per individual does not necessarily lead to a harsher judgment or sanction.

A possible explanation for these findings may be related to the insight developed in Section 2 suggesting that there may exist a kind of (unknown) threshold beyond which onlookers react

similarly, irrespective of the harm caused per individual. Yet, regarding scenario A, Table 4 indicates that there are no significant differences in terms of the severity of judgment or sanctions between the treatments corresponding to greater environmental damage caused, that is, T1*, T2* and T20. Moreover, these findings hold when controlling for MHT, as reported in Table 5 (LSX p-values are all not significant when simultaneously comparing T1*, T2* and T20). For scenario B, we find the same trend, although Table 4 suggests that the recommended sanction is significantly less severe in treatment T2* (7.97) compared to T20 (9) at the 10% level. However, the latter is not robust when controlling for MHT (Table 5). In other words, our results suggest (without proving it) that on average, participants may have similarly judged and punished perpetrators for environmental damage caused beyond a certain level, irrespective of a perpetrator's individual contribution to the environmental damage. It is possible that the examined damage was perceived as so serious, that a threshold was exceeded, causing respondents to condemn all implied perpetrators similarly, even if their individual contribution differed across treatments.[8]

To analyze further, we also considered people choosing the highest levels of perceived severity or sanction. As mentioned in Section 2, we use a binary variable (1 if the perceived severity was 10, 0 otherwise) and run a probit regression to investigate what factors contribute to a participant choosing this highest grade (Table 6). Interestingly, the probability of choosing the

---

[8] Unfortunately, our data does not allow us to draw a categorical conclusion. We do not have a clear correspondence between severity judgments and support for punishment within the different conditions, and the data does not necessarily imply that it is the overall increase in severity of the damage that eliminates differences in judgment between groups. Moreover, this likely pattern can be seen as inconsistent with the results reported in Table 6 regarding the lower likelihood of ratings of maximum severity in cases with twenty perpetrators than in cases with one or two perpetrators.

harshest judgment increases when there are only one or two wrongdoers in a scenario A, compared to a treatment with 20 perpetrators. In other words, the most severe judgment is expressed when there is (are) one (or two) versus many perpetrators. The same tendency is observed in scenario B even though the parameters do not reach conventional statistical significance. This result suggests that participants are more likely to judge a wrongdoing committed by a few perpetrators more harshly than a wrongdoing committed by many perpetrators. Another interesting finding is related to the relative contribution of the main drivers of judgment. Indeed, the probability of a participant selecting the harshest level of perceived severity is more sensitive to environmental harm caused per individual ("harm to the environment") than to the overall harm ("unacceptable state of environment"). Thus, participants who selected the harshest judgments appear to be more prone to the SSP effect. This result does not seem to be driven by the level of total harm.

[Insert Table 6 around here]

Punishment, on the other hand, increases with the number of perpetrators as well as with the level of environmental damage. This result is confirmed for both scenarios (Table 7). An explanation can be found in the non-linear dynamics of pollution. Indeed, nature does not necessarily respond to gradual changes in a smooth way (Scheffer et al., 2001) and pollution accumulation may lead to irreversibility and threshold effects.[9] Moreover, given that groups might be considered more rational than individuals (Charness and Sutter, 2012), one may argue

---

[9] An interesting insight was related to the fact that natural entities (e.g., animals, plants, ecosystems) can be perceived as more helpless than human victims. Consequently, external 'judges' may decide that perpetrators deserve more severe sanctions because of the helplessness status of natural entities. Nevertheless, our design does not allow us to investigate this dimension by comparing an environmental crime with a non-environmental one.

that a wrongdoing committed by a group (or participants perceived as such) is the result of deliberate reflection, while the same crime committed by one individual is more likely to be the result of an emotional reaction, leading outside observers to judge groups more severely than individuals. We also observe some individuals heterogeneity towards punishment. Students and older people proposed more lenient sanctions.

[Insert Table 7 around here]

Our findings are preliminary but do not support the existence of an SSP effect regarding perpetrators at the aggregated level. Nevertheless, aggregated levels can hide subsamples, such as very concerned respondents, who may be more subject to the SSP. Moreover, it may make sense to further distinguish between judgment and punishment. Judgment can be seen as a moral opinion that does not impact behavior of the wrongdoers whereas punishment can be viewed as a means to change behavior. It can also be used as a deterrent for other potential wrongdoers. Our results also indicate that the social and/or ethical costs of environmental wrongdoing do not similarly influence the severity of judgment and punishment. The relative weight of "having broken the law" respective to the "environmental harm caused per individual" appears to have had a greater impact on the recommended punishment than on the perceived severity of the crime or wrongdoing. This discrepancy may indicate that people believe that punishment, rather than moral judgment, affects behavioral decisions.

It is noteworthy that our results show a high level of consistency across scenarios both for participants' judgments and their recommended punishments. However, to test the robustness of SSP insights there is a need to investigate different domains (e.g., environmental wrongdoings versus non-environmental wrongdoings), involved parties (perpetrators versus

victims) and subpopulations (concerned versus unconcerned respondents) to see whether similar effects are observed.

4. Conclusion

Using a survey experiment regarding two different scenarios involving environmental wrongdoings, our findings do not support the existence of an SSP effect regarding perpetrators at the aggregated level. At best, we found only that more often participants judged a wrongdoing most harshly judgment when there were one or two versus many perpetrators. Conversely, in several cases, our findings indicate that increasing the number of perpetrators increased the level of punishment. Even if this point is speculative and deserves a specific investigation, a greater number of perpetrators may implicitly suggest (even if not objectively proven) that there is a kind of collusion between offenders. While this is unlikely when considering the SSP in terms of victims, it seems intuitive in the case of similar perpetrators and can lead to harsher punishments.

Our results seem to indicate that the recommended punishments are not always commensurate with the perceived severity of a wrongdoing. A bias may thus arise if formal sanctions (punishment) have a greater power to trigger more ethical behavior than moral judgment (Romaniuc et al. 2016). Accordingly, an outside observer may recommend more severe sanctions for a crime committed by a large number of perpetrators than he/she might recommend for the same crime committed by a single perpetrator. If this extra punishment is morally undeserved, an important issue is to recognize this and learn how to mitigate it. A practical suggestion that remains to be tested might be to caution people in judging capacities about this potential bias in order to investigate whether an information provision at an

appropriate time (presumably, shortly before the decision making) is likely to debias perceptions and decisions.

As our study involves hypothetical decisions in delineated scenarios, it involves some limitations. These limitations indicate directions for further research, e.g., considering real decisions through archival data or field experiments, robustness studies across domains, subpopulations and perpetrators (other types of wrongdoings such as air pollution or financial crimes). Another dimension is to consider that the scope-severity paradox is a non-linear relationship, likely to follow a U-shaped curve where too few or too many perpetrators lead to harsher judgments. Rather than basing important decisions on results drawn from one-study contributions, we advocate in favor of replications and robustness checks (Hamermesh, 2007; Schmidt, 2009; Clemens, 2017). Last but not least, we did not investigate how increasing the identifiability of the unique wrongdoer (e.g., providing a company name and photo) is likely to contribute to the emergence of the SSP. In sum, rather than closing the debate, our contribution aims at expanding the scope of SSP-related research.

References

Akay, A., Martinsson, P., Medhin, H. 2012. Does positional concern matter in poor societies? Evidence from a survey experiment in rural Ethiopia. *World Development*, 40, 428–435.

Andersson, H., Jonsson, L., Ögren, M. 2010. Property prices and exposure to multiple noise sources: Hedonic regression with road and railway noise. *Environmental and Resource Economics*, 45(1), 73-89.

Barreiro, J., Sánchez, M., Viladrich-Grau, M. 2005. How much are people willing to pay for silence? A contingent valuation study. *Applied Economics,* 37, 1233-1246

Buhrmester, M., Kwang, T., Gosling, S.D. 2011. Amazon's Mechanical Turk: A new source of inexpensive, yet high-quality, data? *Perspectives on Psychological Science,* 6, 3–5.

Cameron, A.C., Trivedi, P.K. 1998. Regression Analysis of Count Data. New York: Cambridge University Press.

Carlsson, F., Johansson-Stenman, O. 2000. Willingness to pay for improved air quality in Sweden. A*pplied Economics*, 32, 661-669.

Charness, G., Sutter, M. 2012. Groups make better self-interested decisions. *Journal of Economic Perspectives*, 26(3), 157-176.

Clemens, M. 2017. The meaning of failed replications: A review and proposal. *Journal of Economic Surveys,* 31, 326–342.

Daniels, N. 2012. Reasonable disagreement about identified vs. statistical victims. *Hastings Center Report*, 42, 35–46.

Desvousges, W.H., Johnson, F.R., Dunford, R.W., Hudson, S.P., Wilson, K.N., Boyle, K.J., 1993. *Measuring Natural Resource Damages with Contingent Valuation: Tests of Validity and Reliability, Contingent Valuation: A critical Assessment.*

Farrow, K., Grolleau, G., Ibanez, L. 2017. Social norms and pro-environmental behavior: A review of the evidence. *Ecological Economics*, 140, 1–13.

Genevsky, A., Vastfjall, D., Slovic, P., Knutson, B. 2013. Neural underpinnings of the identifiable victim effect: Affect shifts preferences for giving. *Journal of Neuroscience*, 33, 17188–17196.

Graf, L., König, A., Enders, A., Hungenberg, H. 2012. Debiasing competitive irrationality: How managers can be prevented from trading off absolute for relative profit. *European Management Journal*, 30, 386–403.

Graff -Zivin, J. and M. Neidell, M. 2018. Air pollution's hidden impacts. *Science*, 359, 39–40.

Hamermesh, D.S. 2007. Replication in economics. *Canadian Journal of Economics*, 40, 715–733.

Hsu, S.L., 2008. The identifiability bias in environmental law. *Florida State University Law Review*, 35, 433-504.

Greene, W. 2003. *Econometric analysis*. Prentice-Hall, Upper Saddle River, NJ.

Hauge, K. 2015. Moral opinions are conditional on the behavior of others. *Review of Social Economy*, 73, 154–175.

Jenni, K.E., Loewenstein, G.F. 1997. Explaining the "identifiable victim" effect. *Journal of Risk and Uncertainty*, 14, 235–257.

Kahneman, D., Ritov, I., 1994. Determinants of stated willingness to pay for public goods: A study in the headline method. *Journal of Risk and Uncertainty,* 9, 5–37.

Keating, D. 2018. No respite for German automakers as dieselgate's third anniversary brings fresh cartel probe. *Forbes* (09/18), https://www.forbes.com/sites/davekeating/2018/09/18/no-respite-for-german-automakers-as-dieselgates-third-anniversary-brings-fresh-cartel-probe/#5737d3d57c9d.

Kogut, T. 2011. The role of perspective taking and emotions in punishing identified and unidentified wrongdoers. *Cognition and Emotion*, 25, 1491–1499.

Konis, D., Haran, U., Saporta, K., Ayal, S. 2016. A sorrow shared is a sorrow halved: Moral judgments of harm to single versus multiple victims. *Frontiers in Psychology*, 7, article 1142.

Lee, F.L. 2017. The spillover effects of political scandals: The moderating role of cynicism and social media communications. *Journalism & Mass Communication Quarterly*, 10, 1–20.

Lee, S., Feeley, T.H., 2016. The identifiable victim effect: a meta-analytic review. *Social Influence,* 11, 199–215.

Lewinsohn-Zamir, D., Ritov, I., Kogut, T. 2017. Law and identifiability. *Indiana Law Journal*, 92, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2785510.

List, J.A., Shaikh, A.M., Xu, Y. 2019. Multiple hypotheses testing in experimental economics. *Experimental Economics*. https://doi.org/10.1007/s10683-018-09597-5.

Lu, J.G., Lee, J.J., Gino, F., Galinsky, A.D. 2018. Polluted morality: Air pollution predicts criminal activity and unethical behavior. *Psychological Science*, 29, 340–355.

Markowitz, E.M., Slovic, P., Västfjäll, D., Hodges, S.D. 2013. Compassion fade and the challenge of environmental conservation. *Judgment and Decision Making*, 8, 397–406.

Nordgren, L.F., McDonnell, M-H. M. 2011. The scope-severity paradox: Why doing more harm is judged to be less harmful. *Social Psychological and Personality Science*, 2, 97–102.

Pellegrin, C., Grolleau, G., Mzoughi, N., Napoleone, C. 2018. Does the identifiable victim effect matter for plants? Results from a quasi-experimental survey of French farmers. *Ecological Economics*, 151, 106-113.

Piff, P.K., Stancato, D.M., Côté, S., Mendoza-Denton, R., Keltner, D. 2012. Higher social class predicts increased unethical behavior. *PNAS,* 109, 4086-4091.

Roehm, M.L., Tybout, A.M. 2006. When will a brand scandal spill over, and how should competitors respond? *Journal of Marketing Research*, 43, 366–73.

Romaniuc, R., Farrow, K., Ibanez, L., Marciano, A. 2016. The perils of government enforcement. *Public Choice*, 166 (1-2), 161-182.

Scheffer, M., Carpenter, S., Foley, J. A, Folke, C., Walker, B. 2001. Catastrophic shifts in ecosystems. *Nature*, 413, 591–596.

Schmidt, S. 2009. Shall we really do it again? The powerful concept of replication is neglected in the social sciences. *Review of General Psychology*, 13, 90–100.

Slovic, P. 2007. "If I look at the mass I will never act": Psychic numbing and genocide. *Judgment and Decision Making*, 2, 79–95.

Small, D., Loewenstein, G., Slovic, P. 2007. Sympathy and callousness: The impact of deliberative thought on donations to identifiable and statistical victims. *Organizational Behavior and Human Decision Processes*, 102, 143-153.

Small, D.A., Loewenstein, G. 2006. The devil you know: The effects of identifiability on punishment. *Journal of Behavioral Decision Making*, 18, 311–318.

Smith, R.W., Faro, D., Burson, K.A. 2013. More for the many: The influence of entitativity on charitable giving. *Journal of Consumer Research*, 39, 961-976.

Thomas-Walters, L., Raihani, N.J. 2016. Supporting conservation: The roles of flagship species and identifiable victims. *Conservation Letters*, 10.1111/conl.12319.

Table 1: Treatments used in the experiment

	Damage per individual is constant (Pi1=Pi2=Pi20)			Total damage is constant (Pg1*=Pg2*=Pg20*)		
Treatment and number of perpetrators (N):	**T1** (N=1)	**T2** (N=2)	**T20** (N=20)	**T1*** (N=1)	**T2*** (N=2)	**T20*** (N=20)
Harm per individual (Pi) = *Water pollution* *Deforestation*	**Pi1 =** 1500 l 20 ha	**Pi2 =** 1500 l 20 ha	**Pi20 =** 1500 l 20 ha	**Pi1* =** 30 000 l 400 ha	**Pi2* =** 15 000 l 200 ha	**Pi20* =** 1500 l 20 ha
Total harm (Pg) = *Water pollution* *Deforestation*	**Pg1 =** 1500 l 20ha	**Pg2 =** 3000 l 40 ha	**Pg20 =** 30 000 l 400 ha	**Pg1* =** 30 000 l 400 ha	**Pg2* =** 30 000 l 400 ha	**Pg20* =** 30 000 l 400 ha
Severity (Si) and **sanction** (Yi)	Si1 Yi1	Si2 Yi2	Si20 Yi20	Si1* Yi1*	Si2* Yi2*	Si20* Yi20*

The treatment T20 and T20 are exactly the same.*

Table 2: Socio-economic sample characteristics (N=774)

Variable name	Mean	SD	Min	Max
Student	0.07	0.255	0	1
Age	39.367	11.86	17	77
Gender (=1 if male)	0.581	0.494	0	1
High Education (Bac+2 and more)	0.727	0.446	0	1
Income (Categorical)	3.168	1.021	1	4

Table 3: Importance of main drivers of participants' judgments and punishments of the environmental crimes (N=774)

	Variable name	**Mean**	**SD**	**Min**	**Max**
Scenario A	Harm to the environment	6.655	0.718	2	7
	Breaking the law	5.935	1.311	1	7
	Unacceptable state of environment	6.523	0.883	1	7
Scenario B	Harm to the environment	6.597	0.768	2	7
	Breaking the law	5.866	1.338	1	7
	Unacceptable state of environment	6.355	1.074	1	7

Table 4: Mean responses for the severity and recommended number of years in prison and t-test for equality of means (N=774)

	T1	T2	T20	T1*	T2*	T1/T2	T1/T1*	T1/T20	T2/T20	T2/T2*	T1*/T2*	T1*/T20	T2*/T20
	Mean **(SD)**	**Mean** **(SD)**	**Mean** **(SD)**	**Mean** **(SD)**	**Mean** **(SD)**								
Scenario A													
Severity (Si)	9.49 (.96)	9.47 (1.09)	9.48 (0.83)	9.53 (0.94)	9.35 (1.29)	NS	NS	NS	NS	NS	NS	NS	NS
Years in prison (Yi)	5.92 (3.49)	6.57 (3.56)	6.83 (3.25)	6.8 (3.32)	6.44 (3.28)	NS	**	**	NS	NS	NS	NS	NS
	N=163	*N=139*	*N=149*	*N=158*	*N=165*								
Scenario B													
Severity (Si)	9.2 (1.21)	9.27 (1.26)	9.29 (1.17)	9.46 (1.04)	9.32 (1.08)	NS	**	NS	NS	NS	NS	NS	NS
Years in prison (Yi)	6.96 (4.8)	7.53 (5.22)	9 (5.04)	8.09 (5.28)	7.97 (5.17)	NS	**	***	**	NS	NS	NS	*
	N=166	*N=141*	*N=144*	*N=158*	*N=165*								

*, ** and *** refer to the significance of a t-test comparing mean individual responses at the 10%, 5% and 1% levels, respectively. NS stands for not significant.

Table 5: Results of the multiple hypotheses testing (simultaneous comparison of all treatments)

Scenario A

Compared treatments	Difference in means	Severity (Si) Unadjusted	LSX	Bonferroni	Holm	Compared treatments	Difference in means	Recommended years in prison (Yi) Unadjusted	LSX	Bonferroni	Holm
T20 vs. T1	0.007	0.947	0.947	1	0.947	*T20 vs. T1*	0.911	*0.016*	0.112	0.163	0.163
T20 vs. T2	0.015	0.898	0.988	1	1	T20 vs. T2	0.263	0.475	0.843	1	1
T20 vs. T1*	0.042	0.671	0.969	1	1	T20 vs. T1*	0.028	0.939	0.939	1	0.939
T20 vs. T2*	0.137	0.246	0.751	1	1	T20 vs. T2*	0.389	0.29	0.711	1	1
T1 vs. T2	0.023	0.860	0.984	1	1	*T1 vs. T2*	0.648	*0.098*	0.419	0.983	0.786
T1 vs. T1*	0.034	0.736	0.981	1	1	*T1 vs. T1* *	0.883	*0.025*	0.154	0.256	0.231
T1 vs. T2*	0.145	0.264	0.760	1	1	T1 vs. T2*	0.522	0.163	0.559	1	1
T2 vs. T1*	0.057	0.626	0.960	1	1	T2 vs. T1*	0.235	0.532	0.867	1	1
T2 vs. T2*	0.122	0.376	0.868	1	1	T2 vs. T2*	0.125	0.744	0.933	1	1
T1* vs. T2*	0.179	0.153	0.598	1	1	T1* vs. T2*	0.361	0.323	0.724	1	1

Scenario B

Compared treatments	Difference in means	Severity (Si) Unadjusted	LSX	Bonferroni	Holm	Compared treatments	Difference in means	Recommended years in prison (Yi) Unadjusted	LSX	Bonferroni	Holm
T20 vs. T1	0.092	0.487	0.879	1	1	*T20 vs. T1*	2.049	*0.000*	*0.000*	*0.003*	*0.003*
T20 vs. T2	0.022	0.862	0.862	1	0.862	*T20 vs. T2*	1.475	*0.015*	*0.101*	0.153	0.138
T20 vs. T1*	0.164	0.198	0.663	1	1	T20 vs. T1*	0.918	0.125	0.399	1	0.625
T20 vs. T2*	0.023	0.846	0.974	1	1	*T20 vs. T2* *	1.037	*0.073*	0.296	0.733	0.44
T1 vs. T2	0.070	0.624	0.943	1	1	T1 vs. T2	0.574	0.315	0.665	1	1
T1 vs. T1 *	0.256	*0.035*	0.211	0.35	0.35	*T1 vs. T1* *	1.130	*0.042*	0.222	0.42	0.336
T1 vs. T2*	0.116	0.341	0.777	1	1	*T1 vs. T2* *	1.011	*0.068*	0.303	0.68	0.476
T2 vs. T1*	0.186	0.173	0.632	1	1	T2 vs. T1*	0.556	0.350	0.619	1	1
T2 vs. T2*	0.045	0.723	0.933	1	1	T2 vs. T2*	0.437	0.465	0.691	1	0.930
T1* vs. T2*	0.140	0.243	0.708	1	1	T1* vs. T2*	0.118	0.845	0.845	1	0.845

Table 6: Factors explaining the selection of the highest severity in judgment (probit regression)

	Scenario A	Scenario B
One perpetrator	*0.289**	0.214
Two perpetrators	*0.327***	0.123
Twenty perpetrators (Ref)	-	-
High level of overall damage	0.036	0.049
Maximum fine	*0.807****	*1.208****
Harm to the environment	*0.756****	*0.933****
Breaking the law	*0.133****	*0.207****
Unacceptable state of the environment	*0.151**	*0.263****
Student	0.104	0.062
Age	0.007	*0.009**
Gender	-0.092	*-0.213**
High education	-0.042	0.027
Income	0.023	-0.035
Observations	774	774
Pseudo R2	0.2777	0.3608
Log likelihood	-337.62231	-319.57157
LR Chi2	259.63***	360.73***

***, ** and * refer to significance at the levels of 1%, 5% and 10%, respectively.

Table 7: Effect of the number of perpetrators on recommended years in prison (Poisson regression)

	Scenario A	Scenario B
One perpetrator	*-0.084**	*-0.143****
Two perpetrators	-0.034	*-0.105****
Twenty perpetrators (Ref)	-	-
High level of overall damage	*0.061**	*0.064***
Severity=10	*0.312****	*0.485****
Harm to the environment	*0.121****	*0.058***
Breaking the law	*0.046****	*0.055****
Unacceptable state of the environment	0.013	0.022
Student	*-0.133**	*-0.273****
Age	*-0.004****	*-0.002**
Gender	0.028	-0.006
High education	0.007	0.035
Income	-0.002	*-0.031***
Observations	774	774
Log likelihood	-2041.3028	-2580.7582
LR Chi2	242.56***	524.62***

***, ** and * refer to significance at the levels of 1%, 5% and 10%, respectively.

Appendix 1: Scenarios used in the survey instrument [adapted from Nordgren and McDonnell (2011) and translated from French]

For ease of exposition, we only present the instrument for treatments T1, T2 and T20 that focus on the scope-severity paradox.

Scenario A:

[1/2/20] top executive(s) of a food factory (has/have each) asked (his/their) employees to throw [1500/15000/30000] liters of whey in a nearby river. Throwing whey into water streams is prohibited by law because it constitutes a source of pollution. Whey causes fermentation of organic matter and a decrease of the dissolved oxygen content of the water, leading to fish mortality, a loss of biodiversity, and unpleasant odors. However, no harm to humans has been found.

Please indicate how you judge the severity of the act committed by this/each of these top executive(s) on the scale below [knowing that there is no coordination or collusion between them]:

1 Not at all severe	2	3	4	5	6	7	8	9	10 Extremely severe

Suppose guilt for this type of behavior can result in sentences of up to 10 years in prison for each convicted individual.

What jail sentence would you recommend for this [each of these] top executive[s] of this [these 2/20]

company[ies]? _____ years

Regarding your previous answers [i.e., severity and sanction], please indicate the level of importance of the following reasons:

Reasons	Not im-Portant						Very im-portant
The individual harmed the natural environment (regardless of the respect of regulations)	1	2	3	4	5	6	7
The individual broke a law (regardless of the impact on the natural environment)	1	2	3	4	5	6	7
The natural environment is in an unacceptable state	1	2	3	4	5	6	7

Scenario B:

A [2/20] real estate developer[s] has [have each] obtained a legal authorization to deforest a fixed area of forest in an area very well located in terms of real estate valuation. In order to meet the demand of some customers, this [each] promoter voluntarily and illegally deforested 20 [200/400] ha more. These [20/200/400] ha of forest contained outstanding trees, which were irreversibly destroyed.

Please indicate how you judge the severity of the act committed by this [each one of these] real estate developer[s] on the scale below [knowing that there is no coordination or collusion between them]:

1 Not at all severe	2	3	4	5	6	7	8	9	10 Extremely severe

Suppose guilt for this type of behavior can result in sentences of up to 15 years in prison for each convicted individual.

What jail sentence would you recommend for this [each of these 2/20] real estate developer[s]? _____ years

Regarding your previous answers [i.e. severity and sanction], please indicate the level of importance of the following reasons:

Reasons	Not im-Portant						Very im-portant
The individual harmed the natural environment (regardless of the respect of regulations)	1	2	3	4	5	6	7
The individual broke a law (regardless of the impact on the natural environment)	1	2	3	4	5	6	7
The natural environment is in an unacceptable state	1	2	3	4	5	6	7

Please, indicate the following information:

1. Age: ___ years old	2. Gender: Male ☐ Female ☐	3. Education: (French) Bac + ___ years. Less than Bac ☐
4. Status: Professional ☐ Student ☐ Other (Please, precise) _____ ☐		
5. Your net income/month (€): < 500€☐ Between 500€ and 1000€☐ Between 1001€ and 1500€☐ > 1500€☐		