2017-08

Heterogeneous guilt sensitivities and incentive effects

Charles Bellemare
Alexander Sebald
Sigrid Suetens

Août / August 2017

Centre de recherche sur les risques
les enjeux économiques et les politiques publiques

www.crrep.ca

Abstract

Psychological games of guilt aversion assume that preferences depend on (beliefs about) beliefs and on the guilt sensitivity of the decision-maker. We present an experiment designed to measure guilt sensitivities at the individual level for various stake sizes. We use the data to estimate a structural choice model that allows for heterogeneity, and permits that guilt sensitivities depend on stake size. We find substantial heterogeneity of guilt sensitivities in our population, with 60% of decision makers displaying stake-dependent guilt sensitivity. For these decision makers, we find that average guilt sensitivities are significantly different from zero for all stakes considered, while significantly decreasing with the level of stakes.

JEL Codes: A13, C91
Keywords: guilt sensitivity, psychological game theory, heterogeneity, stakes, dictator game

Bellemare: Département d'économique, Université Laval, CRREP, - cbellemare@ecn.ulaval.ca

Sebald: Department of economics, University of Copenhagen - Alexander.Sebald@econ.ku.dk

Suetens: Tilburg University, CentER, and TILEC - S.Suetens@uvt.nl

The authors thank audiences of seminars at the universities of Amsterdam, Copenhagen, Frankfurt, Gothenborg, Heidelberg, Innsbruck, Mannheim, Munich, Pennsylvania, and Tilburg, at the Max Planck Institute of Economics in Jena, at the Nordic Conference of Behavioral and Experimental Economics in Stockholm, and at IMEBE in Madrid for helpful comments. Furthermore, we would like to thank the editor and two anonymous referees for very helpful comments. Suetens acknowledges financial support from the Netherlands Organization for Scientific Research (NWO) through the VIDI program (project number 452-11-012).

1 Introduction

In recent years, game theorists have incorporated insights from psychology into models of strategic behavior. One of these insights is that people are sensitive to guilt, and refrain from actions that harm others or are deemed immoral (Baumeister, Stillwell, and Heatherton (1994)). In psychological game theory, guilt-sensitive agents are motivated by a propensity to avoid 'letting down' others. They form beliefs about what others believe, in order to infer how much their decisions would let others down.[2] They then trade off their own material benefit from a decision against the extent to which they believe others are let down by it (see Charness and Dufwenberg (2006) and Battigalli and Dufwenberg (2007)). The higher one's 'guilt sensitivity', the more one wishes to avoid letting down others.

Two aspects of these belief-dependent models of guilt aversion are noteworthy. First, the trade-off between people's regards for own payoff and letting down others is captured in their guilt sensitivity. Despite the fact that this distribution might crucially influence optimal behavior in economic exchanges in which individuals do not know each other's propensity to feel guilt, little is known about the distribution of guilt sensitivities among individuals. For example, the decision to trust another party, or to undertake costly screening and search activities to find a right partner, may strongly depend on how heterogeneous the potential trading partners are in their propensities to feel guilt (see Attanasi, Battigalli, and Manzoni (2015) for a theoretical model). Second, potential feelings of guilt are exclusively linked to the (higher-order) beliefs and the stakes involved. Little is known, however, about whether the trade-off between people's preferences for their own payoff and their sensitivity to guilt is independent of the stakes involved. For example, some individuals, who are sensitive to guilt when stakes are low, may become less sensitive when stakes are high. One psychological justification for this behavior is relative payoff saliency – some subjects may progressively focus more on their own payoff relative to concerns about letting down others as stakes increase. This justification is in line with findings in Smith and Walker (1993) where behavior across 31 published studies is

[2]In general, psychological game theory takes into account that people's utilities do not only depend on material payoffs, but also on their beliefs about others' behavior, as well as beliefs about the beliefs of others. See Geanakoplos, Pierce, and Stacchetti (1989) and Battigalli and Dufwenberg (2009) for general frameworks of games with belief-dependent preferences.

shown to be consistent with a model where decision-making is subject to cognitive costs which induce deviations from payoff maximization, deviations which are dampened by increasing stake sizes. Conversely, some may instead simply attach a greater importance to not letting down others when stakes increase.[3]

The effects of stake sizes have recently been studied in the context of models of risk preferences (see Holt and Laury (2002)), outcome-based social preferences (see Bellemare, Kröger, and van Soest (2008) and Yang, Onderstal, and Schram (2012)), and preferences for truthfulness (see Gibson, Tanner, and Wagner (2013)). In many cases, preferences have been shown to depend on the stakes involved.[4] Understanding how the distribution of guilt sensitivities relates to the size of stakes is thus essential for understanding the conditions under which guilt and aversion to let down others can be expected to play a substantial role in economic exchanges.

We experimentally investigate the relationship between (the distribution of) guilt sensitivities and stake size in the context of Battigalli and Dufwenberg (2007)'s model of simple guilt that assumes that individuals prefer to avoid guilt. We elicit behavior in binary-choice ('mini') dictator games for three stake levels. Binary-choice dictator games have the advantage that highly focal norms such as equal split (Andreoni and Bernheim (2009)) and reciprocity (Bicchieri, Xiao, and Muldoon (2011)) can be excluded as drivers of behavior.[5] Furthermore, the games make the interpretation of belief-dependent behavior elicited through the below-described 'menu approach' particularly straightforward.

The menu approach we use is akin to the strategy method due to Selten (1967). More specifically, choices of dictators are measured conditional on an exogenously provided sequence of possible first-order beliefs for the passive player. We show that this menu approach allows us to identify tight bounds on the individual guilt sensitivities of each dictator and, hence, the

[3]Self-serving bias on the other hand has not shown to be related to stake sizes. See Babcock and Loewenstein (1997) for a discussion.

[4]Holt and Laury (2002) find that the degree of risk aversion increases with stakes. Bellemare, Kröger, and van Soest (2008) find that the marginal disutility of disadvantageous inequality is decreasing with stakes while the marginal disutility of advantageous inequality is decreasing with stakes for young, highly educated people (see also Yang, Onderstal, and Schram (2012)). Gibson, Tanner, and Wagner (2013) find that the percentage of truth-tellers decreases with the costs of truthfulness. On the other hand, Carpenter, Verhoogen, and Burks report that multiplying the stakes by 10 does not have a statistically significant effect on the share a dictator allocates to a passive player. For general surveys, see Camerer and Hogarth (1999) and Gneezy, Meier, and Rey-Biel (2011).

[5]See Krupka and Weber (2013) for an experimental study that points out the relevance of dictator games for understanding how social norms influence behavior.

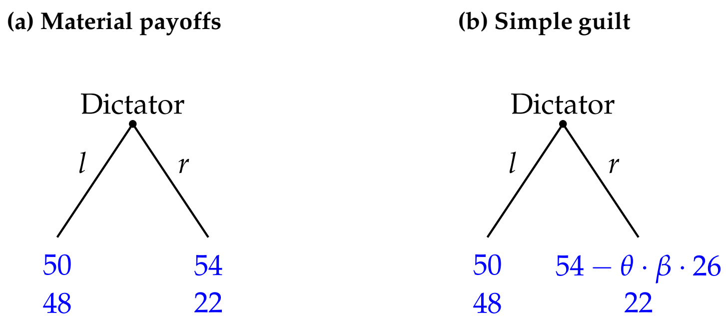
distribution of guilt sensitivities in our population. In addition, our menu approach allows to separate behavior of guilt-sensitive dictators from those motivated solely by distributional or efficiency concerns. This seperation exploits the fact that choices by dictators in our experiment who are motivated solely by distributional or efficiency concerns do not depend on the passive players' first-order beliefs.[6] Other studies use a menu approach to elicit guilt sensitivities. Attanasi, Battigalli and Nagel (2013) elicit belief-dependent preferences of trustees in a trust game through a pre-game questionnaire by letting trustees condition on first-order beliefs of trustors. They find that sensitivity to guilt is the prevalent psychological motivation.[7] Khalmetski, Ockenfels, and Werner (2015) allow for dictators to be surprise-seeking next to guilt averse. They find that guilt aversion is more prevalent than surprise seeking. Third, in our companion paper Bellemare, Sebald, and Suetens (2017), we show that choices elicited with the menu approach are similar to choices of players who are asked to self-report second-order beliefs instead.

We use our data to estimate a structural model that builds upon Battigalli and Dufwenberg (2007)'s framework (henceforth, BD). Our model allows each dictator to belong to one of three classes. The first two classes contain guilt averse dictators whose guilt sensitivity depend or not on the level of stakes. Within each class, we allow for heterogeneity due to differences in guilt sensitivities as well as heterogeneity due to other unobservable factors not taken into account in the model. The final class contains dictators motivated by distributional or efficiency concerns. We find that guilt sensitivities are heterogeneous and vary significantly with stakes for approximately 60% of dictators in our population. For this group, guilt sensitivities on average decrease significantly as stakes are increased, but remain significantly different from zero for all stakes considered. Roughly 32% of dictators are classified as stake-independent dictators. Interestingly, dictators whose guilt sensitivity depends on the stakes are generally more sensitive to guilt than dictators with stake-independent guilt sensitivity. Finally, the remaining dictators (7.9%) are classified as motivated by distributional or efficiency concerns.

[6] The method circumvents the problem of possible spurious correlation between beliefs and behavior by letting the dictator make choices conditional on first-order beliefs of the passive player. See Vanberg (2008), Reuben, Sapienza, and Zingales (2009), Ellingsen, Johannesson, Tjøtta, and Torsvik (2010), and Bellemare, Sebald, and Strobel (2011) for other methods that circumvent the problem.

[7] They model belief-dependent preferences as a combination of guilt sensitivity and intention-based reciprocity, with trustees' willingness to share being increasing (decreasing) in his second-order belief if guilt sensitivity (intention-based reciprocity) prevails in the questionnaire.

Figure 1: A mini dictator game



Notes: Numbers at the top refer to payoffs of the dictator while numbers at the bottom refer to the payoffs of the passive player. In panel (b) θ refers to the guilt sensitivity of the dictator and β to the dictator's average belief about the probability that the passive player assigns to the event that the dictator chooses option l.

The organization of our paper is as follows. In section 2 we present our experimental design and procedures. We present our experimental results in section 3. Section 4 concludes.

2 Experimental game and methods

2.1 The 'mini' dictator game

Consider the mini dictator game shown in Figure 1(a). The dictator must choose between alternatives *left* (l) and *right* (r). If the dictator chooses l, the passive player receives a monetary payoff of 48 and the dictator receives 50. If the dictator chooses r, the passive player receives a monetary payoff of 22 and the dictator receives 54.[8]

If the dictator is only concerned about his own monetary payoff in this decision situation, he chooses r. If the dictator is sensitive to guilt as modeled by BD, he may choose l depending on his belief of what the passive player expects. In particular, the dictator will choose l if he believes sufficiently strongly that he would be 'letting down' the passive player when choosing r. Figure 1(b) depicts our decision situation including the dictator's belief-dependent payoff.

[8] Note that unlike in 'typical' dictator games, in our experiments, dictators can condition their choice on the first-order belief of the matched passive player.

It can easily be seen that a guilt-sensitive dictator prefers option l over option r if

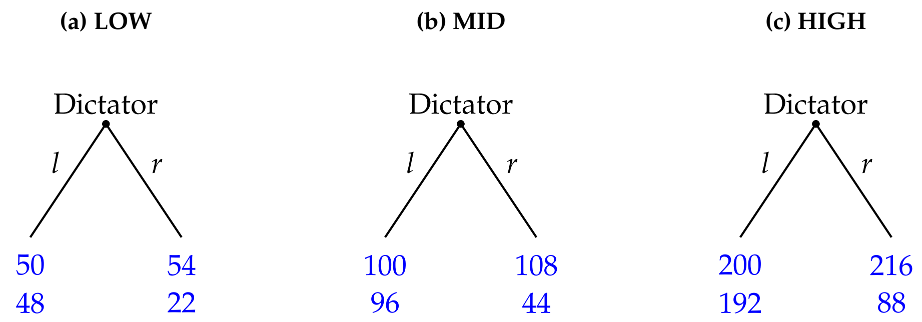
$$\begin{aligned} 50 &\geq 54 - \theta \cdot \beta \cdot 26 && (1) \\ \beta &\geq \frac{2}{13 \cdot \theta}. && (2) \end{aligned}$$

There are three main insights captured by condition (2). First, conditional on the dictator's sensitivity to guilt θ, second-order beliefs β play a key role in determining whether the dictator chooses l or r. Let $\frac{2}{13 \cdot \theta}$ denote a dictator's decision threshold. Consider dictators for whom $\theta \geq \frac{2}{13}$. It follows that there exist values of $\beta \in [0, 1]$ for which these dictators choose l (when (2) holds) and values of β for which they choose r (when (2) does not hold). The model of simple guilt thus predicts that these dictators will switch once from playing r to playing l when their second-order belief β surpasses their decision threshold. Dictators with $\theta < \frac{2}{13}$, however, will always choose r irrespective of their expectations β. Second, the level of β at which a dictator decides to switch from l to r provides information on that dictator's guilt sensitivity parameter θ. Third, multiplying all payoffs by a common factor does not affect condition (2). In other words, in the model with simple guilt, individual guilt sensitivities should be independent of the level of stakes in the game.[9]

Our experimental design exploits these insights in the following way. We ask dictators to choose between l and r for various possible values of the matched passive player's first-order belief, which consequently equals the dictator's second-order belief β. We use the level of β where the switch occurs to get information about the dictator's individual guilt sensitivity θ. Moreover, we vary the stakes in the game *within* subjects – the same subjects play games with different stake levels – to assess whether the β at which dictators switch and the corresponding estimates of θ are sensitive to stakes. The next section describes the experiment in detail.

[9] Following Battigalli and Dufwenberg (2007)'s model of simple guilt, we assume that utility functions are linear in the stake level throughout the paper. However, these properties also hold for utility functions that are power functions of the stake level.

Figure 2: Three mini dictator games



Notes: The figure shows the mini dictator games used in the experiment with material payoffs expressed in experimental points (10 points = 4 DKK).

2.2 Experimental design

Our experimental design allows to set-identify the preference parameter θ of each dictator and to assess how this set varies with the level of stakes in the game. The identified set contains all values of θ consistent with the choices of a given dictator for a given level of stakes.

Each dictator in our experiment was asked to make decisions in three games: one game in which the material payoffs are those shown in Figure 1(a), another game in which the material payoffs are doubled and another game in which material payoffs are quadrupled. We refer to these three games as LOW, MID and HIGH, respectively. Figure 2 shows material payoffs in the three mini dictator games used in our experiment.

We identify the guilt sensitivity of each dictator (and for each stake level) by using a menu approach. Specifically, we elicited the first-order beliefs of passive players by asking them to indicate how many dictators out of 10 he/she thinks will choose l.[10] Denote the first-order beliefs of a passive player by $\alpha \in \{\frac{0}{10}, \frac{1}{10}, \frac{2}{10}, ..., \frac{10}{10}\}$. Figure 3(a) presents the decision screen of the passive player. Figure 3(b) presents the decision screen of the dictator for a given stake level.[11] Each dictator in our setting was presented with 11 decisions, one for each of the first-order beliefs α the passive player could state. Each dictator could choose l independently of α,

[10] Trautmann and van de Kuilen (2015) compare the effect of different belief elicitation methods and find that subjects do not behave differently depending on whether beliefs are incentivized or not (see also Armantier and Treich (2013)). Therefore, we decided not to incentivize the measurement of first-order beliefs.

[11] Appendix A.1 includes detailed instructions. Passive players and dictators were respectively referred to as Players A and B during the experiment.

Figure 3: Dictator and passive player choice screens

(a) Passive player

Out of 10 dictators how many do you think will choose Left?

◯ 0
◯ 1
◯ 2
◯ 3
◯ 4
◯ 5
◯ 6
◯ 7
◯ 8
◯ 9
◯ 10

(b) Dictator

Suppose that the passive player believes that in the above-described decision situation...

... 0 out of 10 dictators choose Left. What do you choose? Left ◯ ◯ Right.
... 1 out of 10 dictators choose Left. What do you choose? Left ◯ ◯ Right.
... 2 out of 10 dictators choose Left. What do you choose? Left ◯ ◯ Right.
... 3 out of 10 dictators choose Left. What do you choose? Left ◯ ◯ Right.
... 4 out of 10 dictators choose Left. What do you choose? Left ◯ ◯ Right.
... 5 out of 10 dictators choose Left. What do you choose? Left ◯ ◯ Right.
... 6 out of 10 dictators choose Left. What do you choose? Left ◯ ◯ Right.
... 7 out of 10 dictators choose Left. What do you choose? Left ◯ ◯ Right.
... 8 out of 10 dictators choose Left. What do you choose? Left ◯ ◯ Right.
... 9 out of 10 dictators choose Left. What do you choose? Left ◯ ◯ Right.
.. 10 out of 10 dictators choose Left. What do you choose? Left ◯ ◯ Right.

choose r independently of α, or switch from choosing r to choosing l at any value of $\alpha > 0$.[12]

2.3 Experimental procedures

We ran the experiment in February and September 2012 in the Laboratory of the Center of Experimental Economics at the University of Copenhagen. The experiment was programmed in z-Tree (Fischbacher (2007)) and participants were recruited through ORSEE (Greiner (2004)). We ran 12 sessions with in total 284 participants (so 142 dictators).[13]

[12] This method to elicit switchpoints corresponds to the 'minimum acceptable offer' procedure often used to elicit the strategy of responders in ultimatum games (see e.g., Güth et al. (1982) and Schotter and Sopher (2007)). In another experiment using a mini trust game where players could switch back and forth between r and l as much as they liked (reported in Bellemare et al. (2017)), we found that the vast majority of players did not switch more than once (see also Attanasi, Battigalli, and Nagel (2013)). Furthermore, for those few players that switched more than once or switched in the 'wrong' way, no systematic tendency could be discovered in their behavior.

[13] Due to an error in the computer program, the data from 2 dictators were not usable, leaving us with 140 dictators.

At the beginning of each session participants were randomly allocated to a computer terminal. Once seated, they received instructions explaining that they were matched in pairs, and that they were randomly allocated a role (player A or player B). The instructions showed an example of a mini dictator game (with payoffs different from those in Figure 2) and explained that participants would be confronted with a number of such decision situations that only differed in the corresponding payoffs (see also Appendix A.1). Both players were informed that the payoff in relation to each decision situation was determined by the dictator's choice of l or r.

In order to avoid strategic reporting of beliefs, passive players were not informed that dictators would make choices conditional on their first-order beliefs. Furthermore, it was not explicitly pointed out to the dictators that the passive players were not informed. Note, however, that our instructions were not deceptive in the sense that they made false claims. We merely left out this information as we judged it would be more harmful to provide it than to not provide it. For example, telling dictators that passive players were deliberately not informed could have induced dictators to wonder whether there are things going on *they* (i.e. the dictators) are not told about.[14]

To avoid ordering effects, participants went through the three games (LOW, MID, and HIGH) in a random order. That is, participants in the role of player A (passive players) were asked about their first-order beliefs in the three games and participants in the role of player B (dictators) were asked to choose either l or r for all possible levels of first-order beliefs in the three games. Neither A- nor B-players received any feedback about outcomes (choice of matched partner or payoff) in between these games.

Only after both players had entered their beliefs/choices in all three games, their respective choices and beliefs were used to determine the payoffs of both players and payoffs were summed across the three stake levels (LOW, MID, and HIGH).

After the experimental games had finished, participants were asked to fill in a post-experimental questionnaire, learned about their payoffs, were paid and dismissed. The post-experimental

[14]Another important reason for not informing the dictators was that we tried to keep instructions as similar as possible to those in other dictator treatments (reported in Bellemare, Sebald, and Suetens (2017)).

questionnaire consisted of a gender question and a questionnaire often used in the psychology literature to elicit guilt proneness (results not reported here).

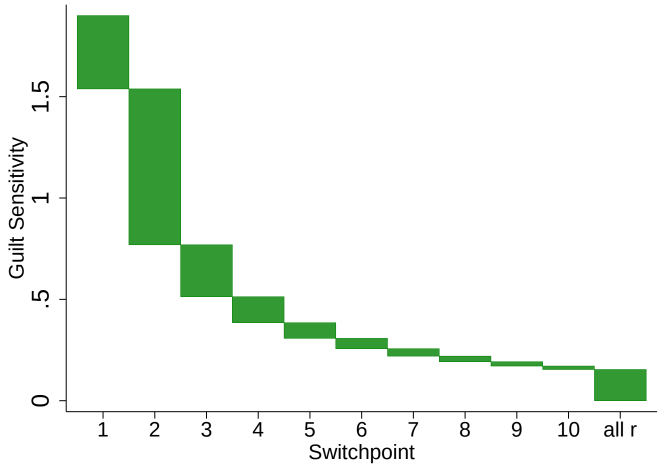
2.4 Behavioral predictions

Guilt-sensitive dictators are predicted to select the selfish option r when $\beta = 0$ as there is no letting down possible ($G(0) = 0$), and then switch to l for a second-order belief $\beta > 0$, or not switch at all in the case of weak guilt sensitivity. Consider a dictator choosing r at $\beta = 0$ and l at all $\beta \geq \frac{1}{10}$. Using condition (2) it follows that $\beta < \frac{2}{13\theta}$ for $\beta = 0$, but $\beta \geq \frac{2}{13\theta}$ for all $\beta \geq \frac{1}{10}$. Solving $\frac{1}{10} \geq \frac{2}{13 \cdot \theta}$ with respect to θ produces $\theta \geq \frac{20}{13} \approx 1.54$, so that there exists a lower bound on the dictator's guilt sensitivity. If the dictator's guilt sensitivity were lower than this lower bound, he would not switch to l at $\beta = \frac{1}{10}$, but at some $\beta > \frac{1}{10}$ (or not switch at all). We next consider a dictator choosing r for $\beta \leq \frac{1}{10}$ and l for $\beta \geq \frac{2}{10}$. On the one hand, it must be true in this case that $\theta^{max} < \frac{20}{13} \approx 1.54$ otherwise the dictator would choose l at $\beta = \frac{1}{10}$ (as in the first example). However, on the other hand, it follows from condition (2) that $\frac{2}{10} \geq \frac{2}{13\theta}$ and hence $\theta^{min} \geq \frac{10}{13} \approx 0.77$. If θ were lower than this lower bound, a dictator who wishes to avoid guilt would switch from r to l at some $\beta > \frac{2}{10}$ (or not switch at all). This reasoning can be extended to players switching from r to l for higher values of β. For a dictator who chooses r for all β, e.g., because he is selfish, $\theta^{max} < 0.15$ and $\theta^{min} = 0$.

Figure 4 presents the identification regions of θ corresponding to each switchpoint (see also the second column of Table A1 in the Appendix). We see that the width of the identification region of θ decreases with the switchpoint. Also, for the majority of the switchpoints the identification regions are very narrow. Our experimental design thus allows to make sharp estimations on the guilt sensitivities.

Dictators motivated exclusively by distributional concerns (or with a guilt sensitivity above 1.5) are insensitive to β and are predicted to always choose l or r for all values of β for a given stake level. To illustrate, take the model of Fehr and Schmidt (1999) and let $s \in \{1, 2, 4\}$ denote the payoff scale factor relative to stakes in the LOW stakes condition. Stakes under MID and HIGH are characterized with $s = 2$ and $s = 4$ respectively. The utility of choosing the selfish option r for a dictator with distributional concerns can be expressed as $54s - \gamma \cdot 32s$ where

Figure 4: The identification regions of guilt sensitivity θ



Notes: The figure shows the identification regions of guilt sensitivity θ corresponding to each switchpoint. The switchpoint is defined as the second-order belief β at which the dictator switches from the selfish choice r to the kind choice l. It is equal to k if the dictator chooses r for $\beta \leq \frac{k}{10}$ and chooses l for $\beta > \frac{k}{10}$.

$32s$ represents the difference between the dictator's and the passive player's payoffs in the experiment for stakes s, and γ represents the preference parameter measuring the strength of advantageous inequality aversion. Similarly, let $50s - \gamma \cdot 2s$ capture the utility from choosing the kind option l. The utilities of options r and l for LOW stakes ($s = 1$) are given by $54 - \gamma \cdot 32$ and $50 - \gamma \cdot 2$. It follows that inequality averse players who have a $\gamma \geq \frac{2}{15} \approx 0.133$ will play l independent of the stake level s, and r otherwise.[15,16]

Along the same lines, dictators that trade off concerns for their own payoff with concerns for efficiency will also tend to choose l irrespective of the passive player's first order belief since l maximizes the sum of the dictator's and the passive player's payoffs. Furthermore,

[15] In a similar vein, it is possible to show predictions under non-linear inequality aversion. Suppose that the utility of choosing the selfish option r is equal to $54s - \gamma \cdot m(32s)$ where the function $m()$ captures the disutility from advantageous inequality aversion. It can easily be seen that dictators with quadratic inequality aversion $m(a) = a^2$ choose l for all s if $\gamma \geq 0.0039$ and r otherwise. Moreover, dictators with square-root inequality aversion $m(a) = \sqrt{a}$ choose l for $s = 1, 2, 4$ if $\gamma \geq 0.943$. They choose r for $s = 1$ and l for $s = 2, 4$ if $\gamma \geq 0.666$, and choose r for $1, 2$ and l for $s = 4$ if $\gamma \geq 0.472$.

[16] Our identification regions are based on the model of simple guilt aversion by Battigalli and Dufwenberg (2007). They should be considered as conservative lower bounds in case players are motivated by simple guilt aversion as well as distributional concerns. Assuming that dictators are also motivated by Fehr and Schmidt (1999)'s model of inequality aversion implies that they should choose l if $\beta \geq \frac{2-15\gamma}{13\theta}$. The higher the dictator's sensitivity to advantageous inequality, the lower the threshold. For γ bigger or equal to 2/15 (0.13333), the dictator should choose l independent of β. Fehr and Schmidt (1999) report an average sensitivity to advantageous inequality of 0.315 (in Table III on page 844). Using this coefficient implies that all dictators in our experiment that are motivated both by guilt aversion and inequality aversion should have chosen the nice option l independent of β.

this tendency increases as stakes increase. To illustrate, let γ now be the dictator's sensitivity towards efficiency and let $m(l) = (50+48)s$ and $m(r) = (54+22)s$ with $s \in \{1,2,4\}$ such that the dictator's utility from choosing l and r is $50 + \gamma \cdot (50+48)s$ and $54 + \gamma \cdot (54+22)s$, respectively. It can be seen that a dictator who is concerned about efficiency will choose l for all s whenever $\gamma \geq \frac{4}{22s}$. If $\gamma \geq \frac{4}{44}$, the dictator chooses r for $s = 1$, and chooses l for $s = 2, 4$. Lastly, if $\gamma \geq \frac{4}{88}$, the dictator chooses r for $s = 1, 2$ and chooses l for $s = 4$. This is a similar pattern as the pattern predicted under concave inequality aversion (see footnote 15).

The key point underlying all of these examples is that decisions of dictators motivated solely by distributional or efficiency concerns do not vary with beliefs β for a given level of stakes.

3 Results

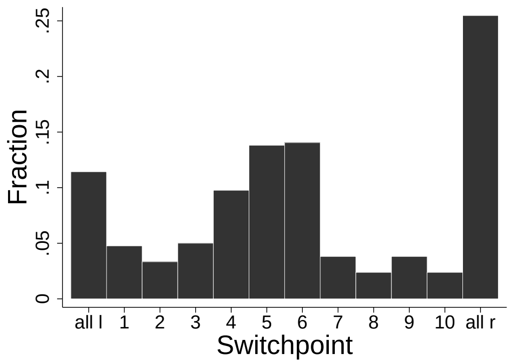
In this section we report our experimental results. In subsection 3.1 we present descriptive statistics and in subsection 3.2 we present results from structural estimations using an extension of the model of simple guilt, that allows for guilt sensitivity to be heterogeneous among individuals and variable in stake size.

3.1 Descriptive statistics

In the aggregate, dictators choose the kind option l 44% of the time. The average switchpoint across all stake levels is 6.03, meaning that the average dictator in our experiment has a guilt sensitivity in the interval $[0.26, 0.31]$. Figure 5 shows the overall distribution of switchpoints, aggregated across the three stake levels. The figure shows that dictators are heterogeneous in their guilt sensitivities. Overall, the mode (corresponding to about 25% of the dictators) is to always choose r for the three stake levels, corresponding to a guilt sensitivity below .15. About 36% of the dictators have a switchpoint between 4 and 6, so a guilt sensitivity between .31 and .38, and about 10% always choose the kind option l, given the stake level.

If we slice up the switchpoints by level of stakes we find that the average switchpoint is 5.77 in LOW, 6.02 in MID, and 6.29 in HIGH. Frequency distributions of switchpoints by stake

Figure 5: Distribution of switchpoints

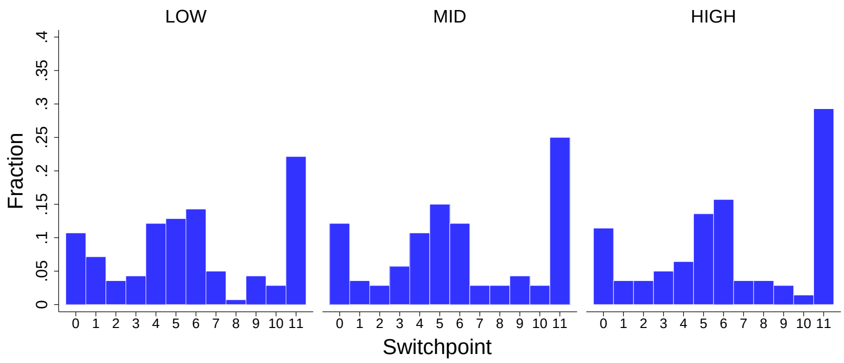


Notes: The figure shows a histogram of the switchpoints . The switchpoint is defined as in Figure 4. The figure is based on observations from 140 dictators. Panel (a) is based on all dictators ($N = 140$) and panel (b) on dictators that have been classified as having a stake-dependent guilt sensitivity with a posterior probability of at least 80% in the model estimated in Section 3.2 (see Figure 7).

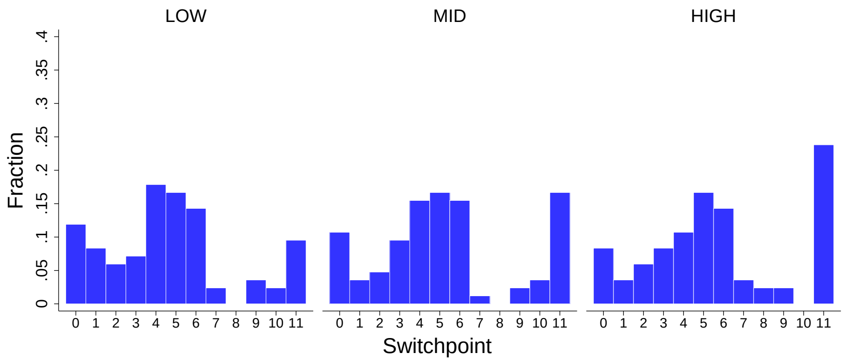
level using all observations are shown in panel (a) of Figure 6. The distributions shift more to the right as stakes increase. In particular, the share of 'all *r*' (a guilt sensitivity below 0.15) goes up as stakes increase. Table 1 reports results from exploratory regressions documenting the effects of increasing stakes. The first column presents parameter estimates of a random effects ordered logit model, using the switchpoint as the dependent variable. Threshold parameters common in parametric ordered response models were estimated but are not presented in the table to simplify exposition. The second columns presents estimates of a random effects logit model where the dependent variable takes a value of 1 when subjects chooses 'all *r*, and 0 otherwise. Explanatory variables include dummies referring to MID and HIGH, as well as two dummy variables referring to the order of play. All standard errors are clustered at the subject level. Diminishing guilt sensitivity translates to positive parameter estimates for MID and HIGH variables in both models – subjects are more prone to have higher switchpoints (in the ordered logit model) and more prone to play 'all *r* independent of the level of beliefs as stakes increase. Estimates from both models are broadly consistent with diminishing guilt sensitivity. We find that the binary variable HIGH enters positively and significantly in both models (10% and 5% respectively), suggesting lower guilt sensitivity in HIGH relative to LOW stakes environments. These descriptive results remain crude as they ignore potential heterogeneity in how subjects respond to stake levels and do not control for alternative factors including con-

Figure 6: Distributions of switchpoints by stake level

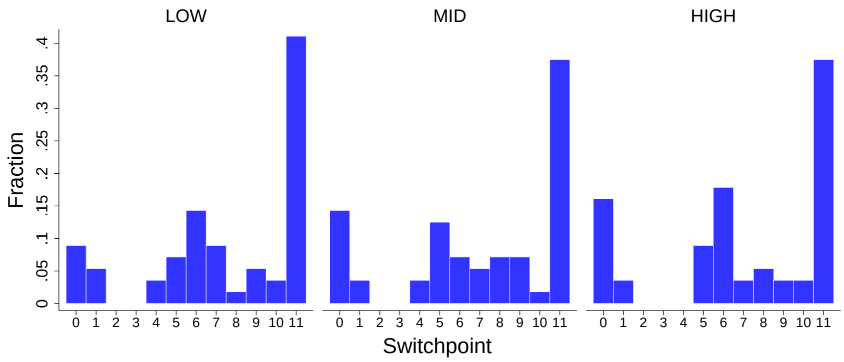
(a) All dictators



(b) Dictators with stake-dependent guilt sensitivity



(c) Other dictators



Notes: The figure shows histograms of switchpoints by stake level . The switchpoint is defined as in Figure 4. Panel (a) is based on all dictators ($N = 140$). Panel (b) is based on dictators that have been classified as having a stake-dependent guilt sensitivity with a posterior probability of at least 80% in the model estimated in Section 3.2 ($N = 56$). Panel (c) is based on other dictators ($N = 84$).

cerns for efficiency or inequity. The model presented in the next section aims to bring these factors together.

Table 1: Regression results

Model :	Ordered logit	Logit
MID	0.122 (0.183)	0.287 (0.367)
HIGH	0.395 (0.216)*	0.739 (0.356)**
Period = 2	0.221 (0.228)	0.062 (0.425)
Period = 3	0.243 (0.227)	0.407 (0.406)
N	420	420
Clusters	140	140

Notes: Estimated parameters of ordered and binary logit models with random effects, standard errors clustered at the subject level. The ordered logit model uses the switchpoint as the dependent variable. The dependent variable used for the logit model is set to 1 for subjects playing 'all r' (yes/no) in a given treatment, 0 otherwise. '*','**' denote significance at the 10%, and 5% levels respectively. Threshold parameters of the ordered logit model are not presented in the table.

3.2 Structural estimation

3.2.1 A three-class model

The data described in section 3.1 reveals important heterogeneity in the sensitivity to guilt across players and stake levels. In this section we estimate a simple choice model allowing for three classes of dictators. The first class contains dictators whose sensitivity to guilt depends on the stakes involved, while the second class contains dictators with stake-independent guilt sensitivity. Preferences are allowed to be heterogeneous within these two classes. This feature allows us to identify the fraction of dictators within the first class whose sensitivity to guilt varies both monotonically and non-monotonically with the stakes. The final class contains dictators whose decisions in the experiment are consistent with models of distributional and efficiency concerns. We make no attempts to separate distributional and efficiency concerns. This reflects the fact that our experiments and the current paper are focused on measuring heterogeneity of guilt sensitivities. Separating alternative concerns would involve different experiments and data.

Our analysis focuses on guilt from playing r defined in section 2.1,

$$G\left(\beta_j\right) = \beta_j \cdot 26 \tag{3}$$

where β_j denotes the probability that the passive player assigns to the event that the dictator will play r. Guilt from playing l is set to zero for all decisions. For each stake level players made j=1,2,...,11 binary decisions for values of β_j, ranging from 0 to 1. Let $s \in \{1,2,4\}$ denote a variable capturing the scale of the payoffs relative to the LOW stakes scenario. Let y_{ij}^s denote decision j of dictator i for stakes s, where $y_{ij}^s = 1$ when the dictator chose to play right, and 0 otherwise.

Stake-dependent class

Let $\omega_{1i} \in [0,1]$ denote the (prior) probability that a dictator's guilt sensitivity is stake-dependent. For dictators in this class, let the utility of player i from choosing either r and l be given by

$$\begin{aligned} U_{r,ij}^s &= 54 \cdot s - \sum_{\forall s} \theta_i^s \cdot G\left(\beta_j\right) \cdot s \cdot d_i^s \\ U_{l,ij}^s &= 50 \cdot s \\ \Delta U_{ij}^s\left(\theta_i^s\right) &= U_{r,ij}^s - U_{l,ij}^s \end{aligned}$$

where d_i^s is a binary indicator taking a value of 1 when stakes s occur, 0 otherwise. The term $G\left(\beta_j\right) \cdot s$ represents the level of guilt under stakes s associated with playing r. This level corresponds to the level of guilt defined in (3) scaled by s. The dictator's sensitivity to guilt is captured by the preference parameters θ_i^s. We allow θ_i^s to vary among subjects and the level of stakes (payoffs) in the game. In particular, we assume that

$$\theta_i^s = \theta^s + \eta_i^s$$

where η_i^s denotes the unobserved part of the sensitivity to guilt for stakes s. We assume that the distribution of the triplet $\left(\eta_i^1, \eta_i^2, \eta_i^4\right)$ is $h\left(\eta_i^1, \eta_i^2, \eta_i^4\right)$, a multivariate normal distribution with mean vector $(0,0,0)$ and the 3×3 covariance matrix Ω. Given $\left(\theta^1, \theta^2, \theta^4\right)$ and Ω, it is possible to compute the following probabilities

$$\Pr\left(\theta_i^1 < \theta_i^2 < \theta_i^4\right) \tag{4}$$
$$\Pr\left(\theta_i^1 > \theta_i^2 > \theta_i^4\right) \tag{5}$$
$$\Pr\left(\theta_i^1 < \theta_i^2 > \theta_i^4\right) \tag{6}$$
$$\Pr\left(\theta_i^1 > \theta_i^2 < \theta_i^4\right) \tag{7}$$

(4) and (5) represent the proportion of dictators whose sensitivity to guilt is respectively increasing and decreasing with the level of stakes. (6) and (7) represent the proportion of dictators whose sensitivity to guilt varies non-monotonically with the stakes (respectively inverted-U shape and U-shape). The proportions of these four stake-dependent types – types which are present in the raw data discussed in section 3.1 – are thus identified in the model.

Finally, we allow for the possibility that dictators make possibly suboptimal decisions and maximize $\Delta U_{ij}\left(\theta_i^s\right) + \lambda \varepsilon_{ij}^s$ where ε_{ij}^s denotes errors assumed to follow a logistic distribution and where λ denotes a scale parameter capturing the level of sub-optimal play. The error term ε_{ij} can alternatively be interpreted as capturing unobserved preferences not captured by the model.

Stake-independent class

Let ω_{2i} denote the (prior) probability that a dictator's guilt sensitivity is independent of stakes. In particular, let the utilities of choosing l and r be given by

$$\begin{aligned} U_{r,ij}^s &= 54 \cdot s - \theta_i^I \cdot G\left(\beta_j\right) \cdot s \\ U_{l,ij}^s &= 50 \cdot s \\ \Delta U_{ij}^s\left(\theta_i^I\right) &= U_{r,ij}^s - U_{l,ij}^s \end{aligned}$$

where

$$\theta_i^I = \theta^I + \eta_i^I$$

and such that that distribution of η_i^I is $g\left(\eta_i^I\right)$, a normal distribution with mean zero and variance σ^2. This specification imposes the restriction that the sensitivity to guilt is constant for a given dictator across stakes, but the value of θ_i^I can vary among dictators in this class (because of η_i^I when $\sigma^2 > 0$). Finally, we allow for the possibility that dictators make possibly suboptimal decisions and maximize $\Delta U_{ij}\left(\theta_i^I\right) + \lambda^I \varepsilon_{ij}^I$ where ε_{ij}^I denotes errors assumed to follow a logistic distribution and where λ^I denotes a scale parameter capturing the level of sub-optimal play. Note that this class also subsumes all purely selfish dictators.

Distributional and efficiency concern class

Finally, let ω_{3i} denote the probability that a dictator is motivated by distributional or efficiency concerns. Section 2.4 presented behavioral predictions for these dictators. Our data suggests that 8 dictators behave in a way which can be perfectly predicted by models of distributional and efficiency concerns. These models predict dictators will always choose the kind option l for LOW, MID, and HIGH stakes, consistent with linear or quadratic inequality aversion. Moreover, 1 dictator always choose l for MID and HIGH stakes while 2 dictators always chose l only for HIGH stakes. Behavior of these last 3 dictators is consistent with different degrees of square-root inequality aversion and efficiency concerns. These dictators are too few to properly estimate the distribution of aversion to inequality in the population. Moreover, the data do not allow to separately identify linear and quadratic inequality aversion. Our strategy is to model the likelihood of behaving in a way consistent with models of distributional and efficiency concerns and abstracting from specific functional forms. In particular, let d_i denote a binary variable taking a value of 1 when the decisions of dictator i are consistent with models of distribution and efficiency concerns, and 0 otherwise. The variable d_i takes a value of 1 for the 11 dictators discussed above. It follows that d_i can be treated as a Bernoulli random variable with probability of success equal to ω_{3i}.

Observable heterogeneity

Finally, we allow the class distribution to depend on gender and age of dictators, both of which were measured in the post-experimental questionnaire. We introduce this observable heterogeneity by specifying

$$\omega_{1i} = \frac{\exp\left(\omega_1^0 + \omega_1^1 age_i + \omega_1^2 gender_i\right)}{\exp\left(\omega_1^0 + \omega_1^1 age + \omega_1^2 gender\right) + \exp\left(\omega_2^0 + \omega_2^1 age_i + \omega_2^2 gender_i\right) + 1} \quad (8)$$

$$\omega_{2i} = \frac{\exp\left(\omega_2^0 + \omega_2^1 age_i + \omega_2^2 gender_i\right)}{\exp\left(\omega_1^0 + \omega_1^1 age_i + \omega_1^2 gender_i\right) + \exp\left(\omega_2^0 + \omega_2^1 age_i + \omega_2^2 gender_i\right) + 1} \quad (9)$$

$$\omega_{3i} = 1 - \omega_{1i} - \omega_{2i} \quad (10)$$

where $\{(\omega_k^0, \omega_k^1, \omega_k^2) : k = 1, 2\}$ are estimated parameters, age_i is measured in years (sample average of 25.1), and $gender_i$ is equal to 1 for male dictators (0.378% of dictators), and 0 otherwise.

Likelihood function

The likelihood contribution of a given dictator, given the model described above, is

$$\begin{aligned} L_i &= \omega_{1i} \int_{R^3} \left[\prod_{\forall s} \prod_{j=1}^{11} l_{ij} \left(y_{ij}^s; \theta_i^s, \lambda \right) \right] h\left(\eta_i^1, \eta_i^2, \eta_i^4; \Omega\right) d\eta_i^1 d\eta_i^2 d\eta_i^4 \\ &\quad + \omega_{2i} \int \left[\prod_{\forall s} \prod_{j=1}^{11} l_{ij} \left(y_{ij}^s; \theta_i^I, \lambda^I \right) \right] g\left(\eta_i^I; \sigma^2\right) d\eta_i^I + \omega_{3i} d_i \end{aligned}$$

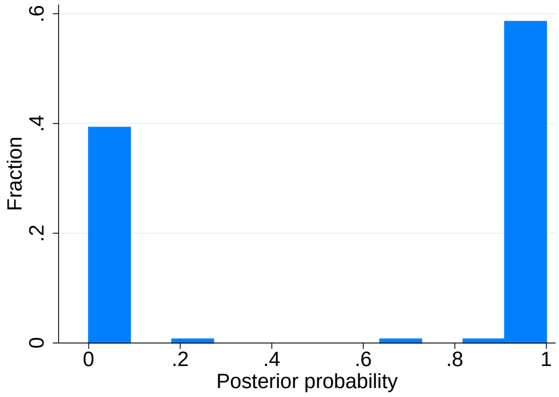
where $\omega_{1i} + \omega_{2i} + \omega_{3i} = 1$ and $l_{ij}\left(y_{ij}^s; a, b\right) = y_{ij}^s \exp(\Delta U_{ij}^s(a)/b)/(\exp(\Delta U_{ij}^s(a)/b) + 1) + \left(1 - y_{ij}^s\right)/(\exp(\Delta U_{ij}^s(a)/b) + 1)$. The sample log-likelihood of the model is given by $\sum_{i=1}^N \log(L_i)$. We estimate the model parameters $\left(\theta^1, \theta^2, \theta^4, \theta^I, \lambda, \lambda^I, \Omega, \sigma^2, \omega\right)$ using Simulated Maximum Likelihood (there are 6 free parameters in Ω). In particular, we maximize $\sum_{i=1}^N \log\left(\widetilde{L}_i\right)$ where

$$\widetilde{L}_i = \omega_{1i} \frac{1}{R} \sum_{r=1}^R \left[\prod_{\forall s} \prod_{j=1}^{11} l_{ij} \left(y_{ij}^s; \theta_i^{s,r}, \lambda \right) \right] + \omega_{2i} \frac{1}{R} \sum_{r=1}^R \left[\prod_{\forall s} \prod_{j=1}^{11} l_{ij} \left(y_{ij}^s; \theta_i^{I,r}, \lambda^I \right) \right] + \omega_{3i} d_i \quad (11)$$

represents the simulated likelihood contribution of player i. This function is computed by drawing R vectors $\left(\eta_i^{1,r}, \eta_i^{2,r}, \eta_i^{4,r}\right)$ from $h\left(\eta_i^1, \eta_i^2, \eta_i^4\right)$ and computing $\theta_i^{s,r} = \theta^s + \eta_i^{s,r}$ for all s and by drawing R values $\eta_i^{I,r}$ from $g\left(\eta_i^I\right)$ and computing $\theta_i^{I,r} = \theta^I + \eta_i^{I,r}$. We use Halton sequences to generate all draws.[17]

[17] We experimented with a specification which modeled the distribution of $(\theta^1, \theta^2, \theta^4)$ as a discrete distribution : $(\theta_k^1, \theta_k^2, \theta_k^4)$ with probability ω_k for k = 1,2,...,7. This specification led to a lower log-likelihood function (when evaluated at the solution) than our preferred specification despite the higher number of model parameters of the discrete specification (the discrete and our preferred specifications model the joint distribution of $(\theta^1, \theta^2, \theta^4)$ using

Figure 7: Posterior probabilities of belonging to stake-dependent class



Notes: The figure presents an histogram of the predicted individual posterior probabilities of belonging to the stake-dependent class, conditional on the choices of each dictator.

3.2.2 Estimation results

Table 2 presents the estimated model parameters. Estimated values of the class share parameters can be used to predicted class shares. We find that the average predicted (a priori) share of stake-dependent and stake-independent dictators is 0.599 and 0.322 respectively, with a residual share of 0.079 dictators predicted to belong to the distributional and efficiency concern class.[18] Observable heterogeneity is introduced in our model by allowing class shares to vary across age and gender. A formal test of the null hypothesis that of a lack of observable heterogeneity (no age or gender effects) translates to testing whether $\{(\omega_k^1, \omega_k^2) : k = 1, 2\}$ are jointly equal to zero. We cannot reject this hypothesis at contentional levels of significance (LR test, p-value = 0.173), suggesting little observable heterogeneity in the distribution of dictator types in the experiment.

Figure 7 presents a histogram of the predicted individual posterior probabilities of belonging to the stake-dependent class, conditional on the choices of each dictator.[19] We find that the

respectively 27 and 11 parameters). Numerically unstable solutions were obtained when we further increased the number of mass points of the discrete distribution.

[18]These numbers are computed using $\frac{1}{N}\sum_{i=1}^{N}\widehat{\omega}_{ji}$ for each class j, where $\widehat{\omega}_{ji}$ are computed using equations (8)-(10) evaluated at the estimated parameter values reported in Table 2.

[19]The predicted conditional posterior probability of player i is obtained by evaluating $\left(\omega_{1i}\frac{1}{R}\sum_{r=1}^{R}\left[\prod_{\forall s}\prod_{j=1}^{11} l_{ij}\left(y_{ij}^{s};\theta_i^{s,r},\lambda\right)\right]\right)/\widetilde{L}_i$ at the estimated values of the model parameters, where $\widetilde{L}_i$ is given in (11).

Table 2: Estimated model parameters

Stake-dependent class ($\widehat{\omega_1} = xx$)			Stake-insensitive class ($\widehat{\omega_2} = xx$)		
	Param.	Std. err.		Param.	Std. err.
θ^1	1.034***	0.141	θ^I	0.145***	0.023
θ^2	0.592***	0.066	$\mathbf{V}\left(\eta_i^I\right)$	0.020**	0.010
θ^4	0.445***	0.042	λ^I	0.765***	0.053
$\mathbf{V}\left(\eta_i^1\right)$	1.038***	0.268			
$\mathbf{V}\left(\eta_i^2\right)$	0.458**	0.119	ω_2^0	4.817*	2.805
$\mathbf{V}\left(\eta_i^4\right)$	0.251***	0.063	ω_2^1	0.013	0.804
$\mathbf{Cor}\left(\eta_i^1, \eta_i^2\right)$	0.183**	0.059	ω_2^2	-0.132	0.103
$\mathbf{Cor}\left(\eta_i^1, \eta_i^4\right)$	-0.071	0.053			
$\mathbf{Cor}\left(\eta_i^2, \eta_i^4\right)$	-0.128	0.073			
λ	6.000***	0.400			
ω_1^0	4.923*	2.962			
ω_1^1	0.723	0.753			
ω_1^2	-0.121	0.111			
$\Pr\left(\theta_i^1 < \theta_i^2 < \theta_i^4\right)$	0.101	$[0.065, 0.144]$			
$\Pr\left(\theta_i^1 > \theta_i^2 > \theta_i^4\right)$	0.317	$[0.233, 0.393]$			
$\Pr\left(\theta_i^1 < \theta_i^2 > \theta_i^4\right)$	0.248	$[0.164, 0.346]$			
$\Pr\left(\theta_i^1 > \theta_i^2 < \theta_i^4\right)$	0.334	$[0.282, 0.390]$			

Notes: '*','**','***' denote significance at the 10%, 5%, and 1% level respectively. The column Std. err. reports the 95% confidence interval for the four predicted proportions at the bottom of the table.

posterior probabilities of most dictators are predicted to be very close to 0 or 1. As a result, few players cannot be unambiguously assigned to one of the three classes we consider. This provides additional evidence that our model is able to cleanly separate dictators with stake-dependent and stake-independent guilt sensitivity.

Within the stake-dependent class we find that the estimated values of θ^1, θ^2, and θ^4 are all positive and significantly different from zero. We also find that the estimated value of θ^j decreases with the stakes, suggesting a negative relationship between average guilt sensitivity and stakes. We also find significant heterogeneity around the estimated averages. In particular, the estimated variances of all three unobserved heterogeneity terms η_i^s are significantly different from zero. The estimated correlation between η_i^1 and η_i^2 is positive and significant, suggesting that dictators with higher guilt sensitivity under LOW stakes also have higher guilt sensitivity under MED stakes. The other two estimated correlation coefficients are small and not significantly different from zero at the 5% level. Finally, the estimated value of λ is positive and significant.

The bottom of Table 2 presents for the stake-dependent dictators the four predicted proportions given in equations (4) to (7). The 95% confidence interval for each predicted proportion appears in brackets. Our model predicts that 10.1% of dictators in this class have a sensitivity to guilt which is increasing with the stakes ($\theta_i^1 < \theta_i^2 < \theta_i^4$). This is the smallest of the four predicted proportions. The prediction proportions for the other three types are higher and relatively similar. We find that the estimated value of $\Pr(\theta_i^1 > \theta_i^2 > \theta_i^4)$ is 31.7%. These dictators have a sensitivity to guilt which is strictly decreasing with the stakes. 24.8% of stake-dependent dictators are estimated to have preferences such that $\theta_i^1 < \theta_i^2 > \theta_i^4$. This non-monotonic relationship captures sensitivity to guilt which initially increases with the stakes (from LOW to MID) before decreasing for higher stakes (from MID to HIGH). Finally, 33.4% of dictators are estimated to have a sensitivity to guilt which initially decreases with the stakes (from LOW to MID) before increasing from MID to HIGH stakes.

We find that the average sensitivity to guilt for stake independent dictators is 0.145. This value is significantly lower than the estimated values of θ^1, θ^2, and θ^4 for stake-dependent dictators (p-values of all three pairwise comparisons are below 0.05). This suggests that the average level of guilt sensitivity of dictators whose guilt sensitivity does not vary with stakes is lower than the corresponding levels measured for dictators with a stake-dependent guilt sensitivity. The relatively lower levels of guilt sensitivity in the stake insensitive class partially reflects the fact that selfish dictators will be assigned to this class. Notwithstanding the presence of selfish dictators, the average sensitivity to guilt of the stake insensitive group remains significantly different from zero. Furthermore, the estimated variance of η_i^I is significantly different from zero but small in magnitude, suggesting significant but small heterogeneity around the estimated average sensitivity to guilt in this class. This heterogeneity is quantitatively and significantly smaller than the corresponding heterogeneity measured within the stake sensitive class, indicating that dictators with a constant guilt sensitivity have more homogenous levels of guilt sensitivity. Finally, the estimated value of λ^I is 0.765 and significantly smaller than the corresponding estimate for the stake sensitive class ($p = 0.000$).

In line with Battigalli and Dufwenberg (2007), the preceding analysis assumes that own payoffs enter people's utility linearly. A natural question is whether the negative relationship between average guilt sensitivities and stakes is robust when moving from linear utility in own payoffs to concave utility in own payoffs, which is empirically more relevant. Concavity in own payoffs means that condition (2) becomes $\beta \geq \frac{u(54 \cdot s) - u(50 \cdot s)}{\theta \cdot s(48-22)}$ with $u(54 \cdot s) - u(50 \cdot s) < 54 - 50$. Moving from linear utility in own payoffs to concave utility in own payoffs thus means that the switchpoint decreases as stakes increase. With exponential concave utility, for example, individuals switch more easily – they switch for a lower second-order belief β – from the 'selfish' to the 'nice' action as stakes increase for a given guilt sensitivity θ.[20] The effect of stakes on the switchpoint that we observe in our data would thus lead to a stronger effect of stakes on estimated guilt sensitivities if we were to use an exponential concave utility instead of a linear function in our estimations. A similar conclusion holds if we were to use concave expo-power functions. Back of the envelope calculations of one-parameter and two-parameter expo-power functions using parameter ranges estimated by Abdellaoui et al. (2007) and Noussair et al. (2014) show that the threshold for the second-order belief at which guilt-sensitive individuals switch from 'selfish' to the 'nice' decreases as stakes increase.[21]

4 Discussion

Guilt averse individuals make decisions by trading off their self-interest with their aversion to letting down others. In this paper, we studied whether this trade-off varies between individuals and across stake sizes. Our experimental results reveal considerable heterogeneity, with 60 percent of our population having stake-dependent guilt sensitivities. For the later group, guilt sensitivities are on average significantly lower when stakes are high than when stakes are low. Moreover, the guilt sensitivities of individuals who do not respond to stakes are estimated to be significantly lower than those of individuals who do respond to stakes. In summary, these results suggest that preferences for guilt aversion are non-separable in intrinsic preferences

[20] This is because the threshold for the second-order belief at which guilt-sensitive individuals switch from *right* to *left* (condition (2)) decreases as stakes increase. Intuitively, increasing stakes has the same effect as increasing ρ when the utility function is $u(x) = 1 - e^{-\rho x}$.

[21] Calculations available upon request.

and economic incentives. It remains to establish whether measured guilt sensitivities could be completely crowded out (on average) by increasing stakes beyond the levels considered in this paper.

The model of guilt aversion we estimated interpreted beliefs a 'descriptive' expectation of another players' expectations rather than a 'normative' expectation (social norm) of the expectations of a broader group of individuals (see e.g. Bicchieri (2006); Bicchieri and Xiao (2009)). Our design cannot fully exclude that dictators perceived beliefs as a social norm and attempted to live up to this norm. The distinction is certainly relevant in order to gain a deeper understanding of the underlying motivations for behavior. In real life, there is often overlap between both; people do not like to let down others because that would violate an expectation of mutual concern (e.g. Baumeister et al. (1994)). Disentangling both motivations is left for future research.

A Appendix (For Online Publication)

A.1 Instructions

A.1.1 General part

You are participating in an experiment on economic decision making and will be asked to make a number of decisions. Please follow the instructions carefully. At the end of the experiment, you will be paid your earnings in private and in cash. You are not allowed to communicate with other participants. If you have a question, raise your hand and one of us will help you.

The experiment is strictly anonymous: that is, your identity will not be revealed to others and the identity of others will not be revealed to you.

Payoffs in the experiment are specified in points. At the end of the experiment the points will be exchanged into DKK at the following exchange rate: **10 points = 4 DKK**.

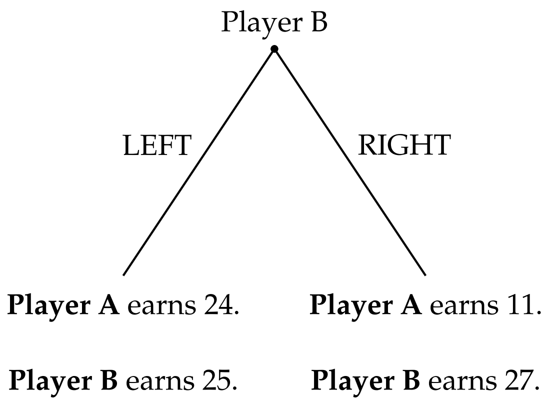
In the experiment, participants are divided into pairs. In each pair, one participant is randomly assigned to the role of "player A", and the other participant to the role of "player B".

A.1.2 Instructions player A

Your role will be Player A.

In the experiment you will be confronted with a number of decision situations like the following:

Figure A1: Example of a decision situation



That is, player B will get the chance to decide between LEFT and RIGHT. The only difference between the decision situation depicted above (in Figure 1) and the situations you will be confronted with during the experiment are the payoffs connected to player B's choices LEFT and RIGHT.

What are the decisions that have to be taken during the experiment?

Choice of player A: In each decision situation that you will be confronted with, you will be asked the following question:

- Out of 10 B-players, how many do you believe will choose LEFT?

We call the answer to this question your belief.

Choice of player B: Player B will be asked to choose LEFT or RIGHT.

How are payoffs calculated?

Assume that you are confronted with the decision situation as shown in Figure 1.

The earnings of you and player B in this decision situation depend on player B's choice. If player B chooses LEFT, you earn 24 points and player B earns 25 points. If player B chooses RIGHT, you earn 11 points and B earns 27 points.

At the end of the experiment the payoffs from the different decision situations will be summed and you and Player B will be paid accordingly.

Following these decisions there will be a questionnaire.

A.1.3 Instructions player B

Your role will be Player B.

In the experiment you will be confronted with a number of decision situations like the following:

[Figure A1 is shown]

That is, you will get the chance to decide between LEFT and RIGHT. The only difference between the decision situation depicted above (in Figure 1) and the situations you will be confronted with during the experiment are the payoffs connected to your choices LEFT and RIGHT.

What are the decisions that have to be taken during the experiment?

Choice of player A: In each decision situation, player A is informed that you can choose LEFT or RIGHT, and about the payoffs connected to these choices. Player A will be asked the following question: Out of 10 B-players, how many do you believe will choose LEFT?

We call the answer to this question player A's belief.

Choice of player B: You will be asked to choose LEFT or RIGHT. More specifically, you will be asked the following questions:

- Suppose player A believes that 0 out of 10 B-players choose LEFT, what do you choose LEFT or RIGHT?
- Suppose player A believes that 1 out of 10 B-players choose LEFT, what do you choose LEFT or RIGHT?
- Suppose player A believes that 2 out of 10 B-players choose LEFT, what do you choose LEFT or RIGHT?
- Suppose player A believes that 3 out of 10 B-players choose LEFT, what do you choose LEFT or RIGHT?
- ...
- ...
- ...
- Suppose player A believes that 10 out of 10 B-players choose LEFT, what do you choose LEFT or RIGHT?

How are payoffs calculated?

Assume that you are confronted with the decision situation as shown in Figure 1.

The earnings of player A and you in this decision situation depend on player A's belief and your choice. If player A's belief and your choice are such that you choose LEFT, A earns 24 points and you earn 25 points. If player A's belief and your choice are such that you choose RIGHT, A earns 11 points and you earn 27 points.

Example: Suppose that A believes that 8 out of 10 B-players will choose LEFT. Suppose further that you choose LEFT, if A believes that more than 4 B-players choose LEFT and that you choose RIGHT, if A believes that 4 or less B-players choose LEFT. In this case, the outcome will be (IN, LEFT) which implies that A earns 24 and you earn 25.

At the end of the experiment the payoffs from the different decision situations will be summed and player A and you will be paid accordingly.

Following these decisions there will be a questionnaire.

A.2 Supplementary tables

Table A1: Distribution of switchpoints and guilt sensitivities for stake-dependent and other players

		Stake-dependent (N=56)				Other players (N=84)
Switchpoint	θ	LOW	MID	HIGH		
all l	$--$	0.119	0.107	0.083		0.131
1	$[1.54, +\infty)$	0.083	0.036	0.036		0.042
2	$[0.77, 1.53]$	0.060	0.048	0.060		0
3	$[0.51, 0.77]$	0.071	0.095	0.083		0
4	$[0.38, 0.51]$	0.179	0.155	0.107		0.024
5	$[0.31, 0.38]$	0.167	0.167	0.167		0.095
6	$[0.26, 0.31]$	0.143	0.155	0.143		0.131
7	$[0.22, 0.26]$	0.024	0.015	0.036		0.060
8	$[0.19, 0.22]$	0	0		0.048	
9	$[0.17, 0.19]$	0.036	0	0.024		0.054
10	$[0.15, 0.17]$	0.024	0.036	0		0.030
all r	$[0, 0.15]$	0.095	0.167	0.238		0.387

Notes: The table shows the distribution of switchpoints for stake-dependent (N=84) and other (N=56) players. Dictators are labeled stake-dependent if they are classified as stake-dependent with a probability of at least 60% in the model in Section 3.2 (see Figure 7). Column θ shows the range of the guilt sensitivity parameter as a function of the switch point.

References

Abdellaoui, M., Barrios, C., and Wakker, P. P. (2007). Reconciling introspective utility with revealed preference: Experimental arguments based on prospect theory. *Journal of Econometrics*, 138:356–378.

Andreoni, J. and Bernheim, B. D. (2009). Social image and the 50-50 norm: A theoretical and experimental analysis of audience effects. *Econometrica*, 77:1607–1636.

Armantier, O. and Treich, N. (2013). Eliciting beliefs: Proper scoring rules, incentives, stakes and hedging. *European Economic Review*, 62:17–40.

Attanasi, G., Battigalli, P., and Manzoni, E. (2015). Incomplete information models of guilt aversion in the trust game. *Management Science*, Forthcoming.

Attanasi, G., Battigalli, P., and Nagel, R. (2013). Disclosure of belief-dependent preferences in a trust game. *Mimeo.*

Babcock, L. and Loewenstein, G. (1997). Explaining bargaining impasse: The role of self-serving biases. *The Journal of Economic Perspectives*, 11(1):109–126.

Battigalli, P. and Dufwenberg, M. (2007). Guilt in games. *American Economic Review Papers and Proceedings*, 97:170–176.

Battigalli, P. and Dufwenberg, M. (2009). Dynamic psychological games. *Journal of Economic Theory*, 144:1–35.

Baumeister, R., Stillwell, A., and Heatherton, T. (1994). Guilt: An interpersonal approach. *Psychological Bulletin*, 115:243–267.

Bellemare, C., Kröger, S., and van Soest, A. (2008). Measuring inequity aversion in a heterogeneous population using experimental decisions and subjective probabilities. *Econometrica*, 76:815–839.

Bellemare, C., Sebald, A., and Strobel, M. (2011). Measuring the willingness to pay to avoid guilt: Estimation using equilibrium and stated belief models. *Journal of Applied Econometrics*, 26:437–453.

Bellemare, C., Sebald, A., and Suetens, S. (2017). A note on testing guilt aversion. *Games and Economic Behavior*, 102:233 – 239.

Bicchieri, C. (2006). *The grammar of society: The nature and dynamics of social norms*.

Bicchieri, C. and Xiao, E. (2009). Do the right thing: but only if others do so. *Journal of Behavioral Decision Making*, 22(2):191–208.

Bicchieri, C., Xiao, E., and Muldoon, R. (2011). Trustworthiness is a social norm, but trusting is not. *Politics, Philosophy, & Economics*, 10:170–187.

Camerer, C. and Hogarth, R. (1999). The effects of financial incentives in experiments: A review and capital-labor-production framework. *Journal of Risk and Uncertainty*, 19:7–42.

Carpenter, J., Verhoogen, E., and Burks, S. (2005). The effect of stakes in distribution experiments. *Economics Letters*, 86:393–398.

Charness, G. and Dufwenberg, M. (2006). Promises and partnerships. *Econometrica*, 74:1579–1601.

Ellingsen, T., Johannesson, M., Tjötta, S., and Torsvik, G. (2010). Testing guilt aversion. *Games and Economic Behavior*, 68:95–107.

Fehr, E. and Schmidt, K. (1999). A theory of fairness, competition and cooperation. *Quarterly Journal of Economics*, 114:817–868.

Fischbacher, U. (2007). z-Tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics*, 10:171–178.

Geanakoplos, J., Pearce, D., and Stacchetti, E. (1989). Psychological games and sequential rationality. *Games and Economic Behavior*, 1:60–79.

Gibson, R., Tanner, C., and Wagner, A. (2013). Preferences for truthfulness: Heterogeneity among and within individuals. *American Economic Review*, 103:532–548.

Gneezy, U., Meier, S., and Rey-Biel, P. (2011). When and why incentives (don't) work to modify behavior. *Journal of Economic Perspective*, 25:191–210.

Greiner, B. (2004). The online recruitment system ORSEE 2.0 - A guide for the organization of experiments in economics. Technical report, University of Cologne, Working Paper Series in Economics 10.

Güth, W., Schmittberger, R., and Schwarze, B. (1982). An experimental analysis of ultimatum bargaining. *Journal of Economic Behavior & Organization*, 3(4):367–388.

Holt, C. and Laury, S. (2002). Risk aversion and incentive effects. *American Economic Review*, 92:1644–1655.

Khalmetski, K., Ockenfels, A., and Werner, P. (2015). Surprising gifts: Theory and laboratory evidence. *Journal of Economic Theory*, 159:163–208.

Krupka, E. L. and Weber, R. A. (2013). Identifying social norms using coordination games: Why does dictator game sharing vary? *Journal of the European Economic Association*, 11:495–524.

Noussair, C. N., Trautmann, S. T., and van de Kuilen, G. (2014). Higher order risk attitudes, demographics, and financial decisions. *Review of Economic Studies*, 81:325–355.

Reuben, E., Sapienza, P., and Zingales, L. (2009). Is mistrust self-fulfilling? *Economics Letters*, 104:89–91.

Schotter, A. and Sopher, B. (2007). Advice and behavior in intergenerational ultimatum games: An experimental approach. *Games and Economic Behavior*, 58(2):365–393.

Selten, R. (1967). Die strategiemethode zur erforschung des eingeschränkt rationalen verhaltens im rahmen eines oligopolexperiments. *in: H. Sauermann (ed.), Beiträge zur experimentellen Wirtschaftsforschung, Vol. I*, pages 136–168.

Smith, V. and Walker, J. (1993). Monetary rewards and decision costs in experimental economics. *Economic Inquiry*, 31.

Trautmann, S. T. and van de Kuilen, G. (2015). Belief elicitation: A horse race among truth serums. *Economic Journal*, Forthcoming.

Vanberg, C. (2008). Why do people keep their promises? An experimental test of two explanations. *Econometrica*, 76:1467–1480.

Yang, Y., Onderstal, S., and Schram, A. (2012). Inequity aversion revisited. *Working paper, University of Amsterdam*.