

**Applying Adam Smith's Theory of Moral Sentiments to elicited social norms:  
Giving and Taking in Dictator Games<sup>1</sup>**

Nina Serdarevic

Centre for Applied Research, Fair Insight Team, Norwegian School of Economics  
Helleveien 30, 5045, Bergen, Norway

Nina.Serdarevic@snf.no

Sigve Tjøtta

Department of Economics, University of Bergen  
Fosswinkelsgate 14, Bergen, Norway

Sigve.Tjotta@uib.no

**Abstract**

The main contribution of this paper lies in applying Adam Smith's moral theory to explain how some actions become jointly recognized as socially appropriate while others do not, and how these jointly recognized rules affect actual behavior. To illustrate the strength of Smith's theory, we apply it to the norm-elicitation method by Krupka and Weber (2013) among a representative sample to explain how different elicited social norms vary across "giving" and "taking" contexts in the dictator game.

**Keywords:** social norms, elicitation of social norms, dictator game, representative experiment

**JEL Classification:** C91, C92

---

<sup>1</sup> *Corresponding author:* Nina Serdarevic. E-mail: Nina. [Serdarevic@snf.no](mailto:Serdarevic@snf.no). Address: SNF, Helleveien 30, 5045 Bergen, Norway. We thank Robert Rebnor at BI for facilitating the experiments and Sondre V. Michelsen and Fredrik Sortland, who assisted in the data collection. We thank Småforsk for financial support. Declarations of interest: none.

## 1. INTRODUCTION

The theoretical and empirical study of social norms spans several disciplines in the social science and humanities (see review by Legros and Cislighi, 2020). A common assumption in the existing literature is that social norms capture “collective perceptions ... regarding appropriateness of different behavior” (Krupka & Weber, 2013, p. 499). Elinor Ostrom defines social norms as a “shared understanding about actions that are obligatory, permitted, or forbidden” (Ostrom, 2000, pp. 143–144). The collective perception, jointly recognized, and shared understanding feature of norms captures their “social” dimension. For a norm to exist in a certain society, it is required that most people in this society judge similar actions as morally right and specific wrongdoings as morally wrong.

In recent years, a considerable body of experimental research has emerged examining how this social dimension of norms can be measured and quantified to explain variation in behavior observed in different settings. The elicitation of social norms should consider that people act in a certain way because they believe others do the same or because they expect others to think one should act in that particular way (Bicchieri, 2006, Bicchieri & Chavez, 2010; Bicchieri & Xiao, 2009). Extending the assumption of Houser and Xiao (2011) that coordination games can be combined with economic incentives to reveal shared understandings, Krupka and Weber (2013) empirically investigate how collective perceptions regarding appropriateness relate to actual choices in dictator games.<sup>2</sup>

This paper asks *how* some types of actions become jointly recognized as socially appropriate in some contexts whereas others do not? Our primary contribution lies in the implementation of Adam Smith’s *Theory of Moral Sentiments* (1759, TMS) to inform this question. Smith’s moral theory offers an alternative process-based approach that teaches us how collective patterns of behavior emerge and how individuals learn to act with propriety within social norms.<sup>3</sup>

---

<sup>2</sup> See, elicitation of social norms using the Krupka and Weber (2013) method see for instance, Vesely (2015) for fairness, Vesely and Klöcker (2018) for environmental behavior and Barr et al., (2015) for discriminatory behavior. See also Stevens (2018) and Dathe et al., (2022) for relevant discussions on social norms and ethics.

<sup>3</sup> Other explanations of social norms and behavior are rooted in Bentham’s maximization utility approach (e.g. Benabou & Tirole 2006, 2011; Benabou et. al 2020), and the indirect evolutionary approach pioneered

The basic premise in Adam Smith's moral theory is that humans are sociable. We want to belong socially; a human "naturally desires, not only to be loved, but to be lovely ... naturally dreads, not only to be hated, but to be hateful" (Smith, 1759, III.2.1, pp. 113–114).<sup>4</sup> These incentives, the desire to belong socially and the fear of being left out, drive the process of how some actions become jointly recognized as either right or wrong in different contexts.<sup>5</sup>

In the existing literature, one frequently substitutes "the desire to belong socially" and "fear of being left out" with some intrinsic motivation for following social norms. However, this does not enlighten why people have an incentive to follow them and how different contexts elicit different social norms. The logic of Smith's theory offers a deeper and different understanding of how we study and conceive human behavior and social norms. For Smith, human interactions cannot be compared to a mechanical machinery in which the same object acts the same way under the same conditions. Rather, it is about the process that facilitates mutual agreement; his theory offers another way of understanding human conduct but also a different method of analysis through which the human discovery process can be understood. Central to this process is how individuals learn to judge others, how others judge them and how they in turn learn to judge themselves.<sup>6</sup> Out of this process arise context specific social norms.

To illustrate how Smith's theory can inform us about the process of how some actions become socially appropriate while others become inappropriate, we first present his theory of moral judgement. We then apply it to a dictator game experiment conducted with a representative sample, varying between a Give and Take framing. In the Give variant, the dictator was entitled to the

---

by Güth & Kliemt (1998). But it is beyond the scope of this paper to describe these theories and how they share similarities with and deviate from Adam Smith's TMS.

<sup>4</sup> This and all subsequent reference to *The Theory of Moral Sentiments* are to Glasgow edition, Smith (1759). References include, in this order, part (in upper case Roman numerals), section (where relevant, in lower Roman numerals), chapter (in Arabic numerals) and paragraph (in Arabic numerals).

<sup>5</sup> See Mansi et al., (2019) for a review and critique of the economics of happiness whereby the authors echo Smith's arguments that happiness and pleasure consists in more than money and depends on joint agreement, and societal and familial relations.

<sup>6</sup> This process is captured by the subtitle added in the fourth edition: An essay towards an analysis of the principles by which men naturally judge concerning the conduct and character, first of their neighbours and afterwards of themselves.

money and decided how much to give to the recipient. In the Take variant, the recipient was entitled to the money, and the dictator decided how much to take from the recipient. Both choice environments offered the dictator the same choice set and therefore also the same payoff set. We randomized half of the subjects to the commonly used Krupka and Weber (2013) norm elicitation framework in which subjects were in the role of Smithian spectators judging the propriety of the choices available to a hypothetical dictator. The other half of subjects were in the role of actual dictators making decisions in the two choice environments for which we had collected social appropriateness ratings. Our claim is that Smith offers a theory that can inform the process of *how* different context may elicit different norms and rules of conduct, not that these different contexts *generate* different rules of conduct in our experiment.

More broadly, our paper contributes to the literature on the formation of beliefs about others and ourselves, which play a crucial role for the emergence of social norms (see review by Bursztyn et al., 2021). An element that is at the core of Smith’s moral theory is the spectatorial process of moral judgement. We judge others, and we are aware that others judge us. In the imagination, we make constant efforts to adjust our sentiments, as spectators, to those of the people “principally concerned”, both actors and those acted upon. As people principally concerned, we try to adjust our feelings to a level that spectators can go along with. Importantly, Smith offers reasons for why we sometimes fail to perfectly anticipate the feelings of others, something that has also been documented experimentally in the study of Bursztyn et al. (2020) whereby individuals fail to correctly perceive what constitutes appropriate behavior in certain settings.

More specifically, our paper contributes to the conversation on how contextual changes, such as allowing dictators to take money from the recipient, affect choices in dictator games (Dreber et al., 2013; Cox et al., 2007; List, 2007; Dana et al., 2006; Bardsley, 2008; Cappelen et al., 2013; Suvoy, 2003; Eichenberger & Oberholzer-Gee, 1998; Gerlach & Jaeger, 2016). Essentially, how a choice problem is presented can be seen as constituting its frame (Tversky and Kahneman, 1981). Although many scholars hypothesize that the dictator would be relatively averse to taking, as this could be seen as violating the other’s entitlement, the literature has obtained mixed results, where some studies report differences in behavior between the giving and taking frame (e.g. List, 2007; Krupka & Weber, 2013), whereas other find no differences (e.g. Suvoy, 2003; Dreber et al., 2013).

We differ from dictator games with taking options studied by, for instance, List (2007) and Bardsley (2008), which, in addition to changing frames, also alter the choice space by including additional taking options. Our paper instead joins other research that keeps the choice space fixed and examines the effect of changing frames with associated entitlements, such that dictators may either give to or take money from the recipient (e.g. Dreber et al., 2013; Krupka & Weber, 2013; Suvoy, 2003; Eichenberger & Oberholzer-Gee, 1998). Herein, our paper and experimental design also relates to the question of endowment effects in the dictator game as the variability in the observed give-take framing effects could be due to the initial allocation of the endowments (Halvorsen, 2015; Cappelen et al., 2013). That is, whether the entire initial endowment is placed with the dictator or the recipient, could matter for the dictators' decisions and feelings of entitlement to the endowment. Related to this, Dreber et al., (2015) discuss and examine experimentally whether the taking frame may justify violating the other's entitlement, finding no effect in their setting.

Finally, by applying Smith's theory to the dictator game, we also contribute to the growing interdisciplinary literature in *humanomics* that integrates insights from Adam Smith's theory of human behavior into experimental economics (Brown, 2011a; Paganelli, 2009; Smith & Wilson, 2019; Serdarevic, 2021; Tjøtta, 2019; Wight, 2013; Young, 2009).<sup>7</sup> Our contribution to this strand of the literature is to extend and experimentally apply Smith's moral theory to dictator game and the norm elicitation framework by Krupka and Weber (2013), commonly used to elicit social norms in experimental economics. As argued by Kimbrough et al., (2014), framing effects can be understood as manipulations of the relevant social norm; "individuals import social norms from outside the lab, and the imported norm varies with induced context" (p. 12). Echoing this argument, we apply Smith's theory to dictator games and claim that different frames highlight different cues that enter into people's moral judgements and subsequently guide their behavior.

---

<sup>7</sup> See also non-experimental studies such as Bly et al., (2018), Easterly (2021), Steeds (2022), and Norman (2018) for discussions and applications of Smith's intellectual projects in *The Wealth of Nations* and *The Theory of Moral Sentiments*.

The rest of the paper proceeds as follows. Section 2 outlines elements of Smith's theory that are relevant to the analysis. Sections 3 and 4 present the experimental design and results. Section 5 concludes.

## **2. ADAM SMITH'S THEORY OF SOCIAL NORMS AND HOW THEY AFFECT CHOICES**

In this paper, we turn to Adam Smith's TMS to explain how variations in context, in particular variations in entitlement to the endowment in dictator games, may elicit different rules of conduct and choices. Smith's theory is about moral judgment: how we judge others, how others judge us, and how we in turn judge ourselves (Campbell, 1971, p. 87). His theory explains how we morally judge each other and ourselves, not how humans behave morally.<sup>8</sup> There can, however, be no sharp distinction between two such theories, as moral judgments and behavior are interlinked (Campbell, 1971, p. 87). Moreover, Smith's theory is about how we morally judge, not how we *ought* to judge. According to Smith, TMS "is not concerning a matter of right ... but concerning a matter of fact. We are not ... examining upon what principles a perfect being would approve of the punishment of bad actions; but upon what so weak an imperfect creature as man actually and in fact approves of it" (Smith, 1759, II.i.5.10, p. 77).

Central to Adam Smith's theory is that humans are sociable; it is only in and through society that we can morally judge each other and become moral beings. Without a society, there is no one whom we can morally judge and no one who is judging us, we have no experience of morally judging others or of being judged by others. Without such experience, there can be no prospect of learning, developing, and cultivating the judgment of one's own behavior. Thus, growing up in a society provides the "mirror" humans need to develop moral judgments and to sense what constitutes proper behavior (Smith, 1759, III.1.3, p.110).

Sympathy is central to Smith's theory of moral judgment. Sympathy represents social bonding; it is the sharing of any feeling of pleasure or pain (Raphael & Macfie, 1976, p. 13).

---

<sup>8</sup> It should be noted, however, that Smith used neither the words "norm" nor "social norm" in TMS, according to Haakonssen and Skinner's *Index to Works of Adam Smith* (2001).

Smith's concept of sympathy differs from the normal meaning of the word and did so even in his own time. Smith's sympathy is the correspondence between sentiments; it is about sharing of feeling between spectators and acting agents. This correspondence between sentiments is an imaginary process where we put ourselves in the shoes of the person whom we judge. If there is harmony in sentiments by our imaginative change of place and the actual acting agent, we judge the action as proper, if it is disharmony, we judge it as improper.

The sympathetic process is enabled by the imagination. However, our perceptions of what others think, and feel is always imperfect due to distance between spectators and agents. Distance is necessary because it makes room for the effort and modification of behavior that is needed until a state of mutual sympathy with one another is reached. It emphasizes that some feelings are appropriate to be felt while others are not, creating room for the constraint of our self-love so that harmony with the sentiments of another can be achieved. In the imagination, we make constant efforts to adjust our feelings; as spectators, we modify our feelings to those of the people principally concerned in a set of circumstances, and as people principally concerned, we adjust our feelings to a level with which spectators can go along. Distance is presupposed because although the imagination allows us to project ourselves into the place of other's, it cannot eradicate all our differences and create a perfect copy of others' feelings. That is, spectators can only represent an idea of others' feelings in their imagination.

As spectators, we sympathize with all persons involved in a situation; in our experiments, this includes the dictator and the recipient. It is easier for us to sympathize with pleasant emotions like benevolence than with unpleasant ones like resentment. Therefore, we tend to have "double sympathy" for actions with a beneficent tendency (Smith, 1759, I.ii.3, pp. 34–38). Furthermore, humans have something that can be viewed as a "divided sympathy" in entering a situation in which actions are rooted in unpleasant passions (Smith, 1759, I.ii.4, pp. 38–40). Our sympathy is divided between the person who shows resentment and the person who is the object of the resentment. In Smith's own words:

With regard to all such passions, *our sympathy is divided between the person who feels them, and the person who is the object of them.* The interests of these two are directly opposite. What our sympathy with the person who feels them would prompt us to wish for,

our fellow-feeling with the other would lead us to fear. (TMS I.ii.3.1, p.34, our emphasis added)

We make two kinds of moral judgments about other people's actions. First, as spectators, we judge the propriety and impropriety of an action. Judgments of propriety are backward looking and relate to the suitability or unsuitability of an action. The judgments focus on the motive of the acting agent—in our setting, the dictator. Second, we judge the merit and demerit of an action. This judgment looks forward to the consequences of an action for the agent being acted upon, in our case, the recipient. In the judgment of merit, sympathy extends to both the motive of the acting agent and the gratitude or resentment of the agent being acted upon. In what follows, we elaborate more on these two kinds of moral judgments and how they relate to the elicitation of social norms.

### *The propriety of an action*

Propriety relates to the intention of the acting agent, assessing whether it is suitable in the specific situation. As spectators, we judge the propriety of an action by imagining changing place with the acting agent, in our experimental setting the dictator. Our judgment of the propriety or impropriety of an action reflects our imaginary sympathy or lack of it. If, in this imaginary change of place, we entirely sympathize with the situation, we approve of the action. In contrast, if we do not entirely sympathize, we judge the action improper.

In the process of judging an action, we compare it with what is normally expected in the situation—what the “greater part of men have actually arrived at” (Smith, 1759, I.i.5.9, p. 26). If the passions underlying an action are too high or too low, the spectator cannot go along with them. Hence, propriety relates to the ordinary, the “mediocrity” that deserves to be approved of (Smith, 1759, I.i.intro.1, p. 27).

Moreover, Smith claims that there exists an asymmetry in the intensities of our judgments from deviating from this mediocrity, from what is proper. We find it easy to sympathize with pleasant feelings like joy and difficult to sympathize with unpleasant ones like grief. Moreover, “our propensity to sympathize with joy is much stronger than our propensity to sympathize with sorrow” (Smith, 1759, I.iii.5, p. 45). The asymmetry in our judgments of deviations from the mediocrity of propriety varies with social and unsocial passions.



Social passions like kindness, friendship, and esteem may sometimes be excessive. Our judgment of such excessive passions is of low intensity; “we blame it, we still regard it with compassion, and even with kindness and never with dislike. We are more sorry than angry” (Smith, 1759, VI.iii.15, p. 243). And, conversely, our judgement is harsher with regard to actions that are derived from social passions that are too moderate. According to Smith, such actions are called “hardness of heart ... and, by excluding him from his friendship of all the world, excludes him from the best and most comfortable of all social enjoyments” (Smith, 1759, VI.iii.15, p. 242).

This asymmetry in the intensity of judgment is reversed for actions that are derived from unsocial passions. The excess of unsocial passions like “anger, hatred, envy, malice, revenge ... renders a man wretched and miserable in his own mind, and the object of hatred, and sometimes even of horror, to other people” (Smith, 1759, VI.iii.16, p. 243). A lack of unsocial passions is, however, very seldom complained about.

As an example, consider being in a restaurant and deciding how much to tip. The propriety of the gratuity—how much it should be in each situation—varies across societies and even restaurants. Assume a gratuity of 15 to 20% as standard, whereas 25% is given for excellent service. Applying Smith’s theory to this example would suggest that we, as spectators, tend to judge deviations from this standard asymmetrically; we assess a lack of tipping more harshly—we may call it cheap, despicable, and shameful—than we assess a 50% tip, which may be judged extravagant and exaggerated. Similarly, there is an asymmetry in the propriety of the waiter’s reaction to no tips. If the waiter verbally attacks the customer, we will assess this more harshly than if the waiter does not react at all.

### *Judging the merit and demerit of actions*

The judgment of the merit or demerit of an action is related to the consequences of the actions for the one acted upon, the recipient in our experimental situation. According to Smith, our judgment of merit and demerit is a compound of two sentiments.

First, our judgment of merit arises out of our assessment of how we would feel, in an imaginary exchange of situations, if we were affected by such an action—whether we would feel gratitude or resentment. If this imaginary exchange is pleasant, gratitude is in place. Conversely, if this imaginary exchange is unpleasant, resentment is in place. By bringing the agent acted upon “home to myself, I feel gratitude arise in my own breast, I necessarily approve of the conduct of

his benefactor ... and regard it as meritorious, and the proper object of reward” (Smith, 1759, II.i.5.11, p. 78). Hence, according to Smith, we as spectators judge an action’s merit or demerit even if the agent acted upon is not able to show gratitude or resentment themselves.

In our dictator setting, this means that spectators may judge the merit and demerit of the choices toward the recipient, the agent acted upon, even though the recipient cannot, by experimental design, actually show gratitude or resentment toward the dictator. Still, from experience the spectator uses the imagination to determine what reaction would be approved of and which would not if any reaction took place. To judge merit, we put ourselves in the place of the recipients and whether we sympathize with gratitude or resentment towards the dictator in their place. If we feel gratitude, we approve of the recipient being grateful and see the dictator’s actions as being worthy of praise and reward. If we feel resentment, we approve of the recipient being resentful and see the dictator’s action as being worthy of blame and punishment.

Second, we do not judge an action as meritorious unless we also judge the motive of the acting agent as beneficent; “if in the conduct ... there appears to have been no propriety, how beneficial soever its effects,” it does not require any reward (Smith, 1759, II.i.4.1, p. 73). Similarly, we will not judge an action as morally bad unless we also judge the motive of the wrongdoer as wicked.<sup>9</sup> Hence, according to Smith, the senses of merit and demerit are compound sentiments:

we cannot ... enter ... into the gratitude of the person who receives the benefit unless we ... approve of the benefactor, so... sense of merit seems to be compound sentiment, and to be made of two distinct emotions: a direct sympathy with the sentiment of the (acting) agent, and an indirect sympathy with the gratitude of those who receive the benefit of the actions (Smith, 1759, II.i.5.2, p 74).

---

<sup>9</sup> However, Smith argues that even though we aspire to judge the intended consequence of an action, we tend to judge the actual consequence. Smith names this an irregularity in our judgement (Part II, Section II). The effect of this irregularity is first to diminish our sense of the merit or demerit of those actions which arose from the most admirable or blameable intentions when they fail of producing their proposed effects. Secondly, to increase our sense of the merit or demerit of actions, beyond what is due to the motive or affection which from which they proceed, when they accidentally give occasion either to extraordinary pleasure or pain. (97:1).

The asymmetry of our moral judgment of the deviation from the mediocrity of propriety is also present when judging the merit or demerit of an action because we, as spectators, find it easier to imaginarily sympathize with pleasant passions than with unpleasant ones. Hence, we find it easier to go along with gratitude than with resentment. Furthermore, we find it easier to go along with a beneficent motive of the acting agent than with a wicked motive.

Thus, when judging the merit or demerit of an action, we also take into account the propriety or impropriety of the acting agent's conduct; there is a close connection between judgments of propriety and judgments of merit. In our setting, in which we use the Krupka and Weber (2013) method to elicit judgments of social appropriateness of dictator choices, we cannot precisely distinguish between these two standards of judgment: the propriety of the dictator's choices and the merit with regard to the recipient's reaction as the spectator does not observe their reaction.

### *Judging ourselves*

To understand what may be guiding the dictator's actual behavior, we look to Smith's theorizing about how we judge ourselves. Smith writes that we "endeavor to examine our own conduct as we imagine any other fair and impartial spectator would examine it" (Smith, 1759, III.1.2, p. 110). In the process of judging ourselves, we place ourselves outside of ourselves and judge our conduct in the same way as a fair and impartial spectator would do. This consciousness, however, is shaped from humans' experience in society; it is not something that is inborn. In the same way as we judge others, we compare our own conduct with the standards of propriety, the mediocrity according to which most people behave, but also with

exact propriety and perfection, so far as we are each of us capable of comprehending that idea. ... The wise and virtuous man directs his principal attention to ... the idea of exact propriety and perfection. There exists in every man, an idea of this kind, gradually formed from his observations upon the character and conduct both of himself and other people. It is the slow, gradual, and progressive work of the great demigod within the breast, the great judge and arbiter of conduct" (Smith, 1759, VI.iii.23–25, p. 247).

The incentives behind the approval and disapproval mechanisms are humans' desire of praise and fear of blame but also their desire of praiseworthiness and fear of blameworthiness. Smith distinguishes between our desire for praise and desire of praiseworthiness.

The love of praiseworthiness is by no means derived altogether from the love of praise. Those two principles, though they resemble one another, though they are connected, and often are blended with one another, are yet, in many respects, distinct and independent of one another (Smith, 1759, III.2.2, p. 114).

Hence, we find satisfaction in our praiseworthy conduct even if no one would praise it.

Is the love of praiseworthiness genuine, or is it rather mere desire of praise? For Smith, the desire for praiseworthiness is genuine in the sense that our experience in the process of moral judgments provides us with a means of distinguishing between “praiseworthiness” and “mere praise,” and an incentive to do so (Griswold, 1998, p. 131). Admiring those people with a character we approve of leads us to desire to become ourselves the possessors of such agreeable character. In the process of moral judgment, we acquire a form of self-awareness or consciousness of how we feel we *ought* to behave (Campbell, 1971, p. 160). This gives us an incentive to strive for being praiseworthy even if no one would praise our actions or conduct. Therefore, an action or a person is not praiseworthy because we have praised the action or the person. The reverse is, however, true: we praise the person or action because we believe this person to be praiseworthy; “so far is love of praiseworthiness from being derived altogether from that of praise; that the love of praise seems ... to be derived from that of praiseworthiness” (Smith, 1759, III.2.3, p. 114).

In a similar way, we have an incentive to avoid being blameworthy even if no one would blame us. Our desires for praise and praiseworthiness and fear of blame and blameworthiness drive us to take actions we expect others to go along with and deter us from actions they cannot go along with.

### *Social norms as general rules of conduct*

Out of the process of the moral judgment of each other's and our own behavior emerge general rules of conduct. These general rules are abstract in the sense that all actions of a certain kind are approved or disapproved of in different circumstances.

Our continual observations of others, insensibly lead us to form for ourselves certain rules concerning what is fit and proper either to be done or to be avoided.... We do not originally approve or condemn particular actions; because, upon examination, they appear to be agreeable or inconsistent with a certain general rule...The general rule, on the contrary, is formed, by finding from experience, that all actions of a certain kind, or circumstances in a certain manner, are approved or disapproved of (Smith, 1759, III.4.7-8, p.159).

He continues to assert the importance of rules being general as

it is often impossible to accommodate to all the different shades and gradations of circumstance, character, and situation, to differences and distinctions which, though not imperceptible, are, by their nicety and delicacy, often altogether undefinable' (TMS VI.i.1.22, p. 227).

Returning to the tipping example: most of us are acquainted with tipping norms in our local communities, in which situations to tip or not, and how much tipping is expected in a particular situation. But as a visitor in another country, we may be lost, not having acquired the same sense of tipping. We therefore often spend time and money reading guidebooks and asking others with more experience to figure out the tipping norm in the place we are visiting. Thus, while rules are general, we use the spectatorial process to apply them to particular contexts and particular instances.

### *Incentives to follow general rules*

In what cases do we have an incentive to follow the general rules of conduct, and in what cases do we have a direct incentive to follow another motive? Smith formulates the question by asking:

in what cases our actions ought to rise chiefly or entirely from a sense of duty; or from a regard of general rule; and in what cases some other sentiments or affection ought to concur, and have a principal influence (Smith, 1759, III.6.1, p. 171).

Our understanding of Smith's use of the normative "ought" here is that it is similar to him arguing that our desire for praiseworthiness is a genuine incentive, as it defines what ought to be praised (e.g., Campbell, 1971, p. 13; Griswold, 1999, p. 131). In the process of morally judging others and ourselves, we acquire a consciousness of in which situations we ought to follow a general rule and in which situations we ought not to. Hence, the "incentive to follow general rules" is our understanding of what Smith means by "our actions ought to rise ... from a sense of duty; or from a regard of general rule" (Smith, 1759, III.6.1, p. 171).

According to Smith, answers to this question depend, first, on the underlying passion that prompts us to any action independent of all regard for the general rules, and, second, on the "precision and exactness or the looseness and inaccuracy, of the general rules themselves" (Smith, 1759, III.6.2, p. 171).

Smith says that the general rules of beneficence determine the responsibility

of prudence, of charity, of generosity, of gratitude, of friendship, are in many respects loose and inaccurate, admit many exceptions, and require so many modifications, that it is scarce possible to regulate conduct entirely on them ... [and they] present us rather with a general idea of the perfection we ought to aim at, than afford us any certain and infallible directions for acquiring it (Smith, 1759, III.6.9–10, pp.174–176).

Hence, actions based on social passions are unifying passions; they are based on a "doubled sympathy" between the person exhibiting these passions and the object of those passions. The social passions comprise of "generosity, humanity, kindness, compassion, mutual friendship and esteem" (Smith, 1759, I.ii.4.1, p. 38), ought to follow our heart, not general rules. Smith explains that beneficent actions

ought to proceed as much from the passions themselves than as any regard from to the general rules of conduct.... A benefactor thinks himself but ill requited, if the person upon whom he has bestowed his good offices, repay them from a cold sense of duty, and without any affection to this person (Smith, 1759, III.6.4, p. 172).

The contrary maxim takes place regarding unsocial passions like hatred and resentment that are dividing passions; our sympathy is divided between the person who feels them, and the person who is the object of them. Here we have the incentive to follow the general rule, not the passions. Smith compares the general rules of justice to grammar: they are “precise, accurate, and indispensable” (Smith, 1759, III.6.11, p. 175). The rules of justice point to what we ought not to do. We

ought always to punish with reluctance and more from sense of propriety of punishing than any savage disposition to revenge. Nothing is more graceful than the behaviour of the man who appears to resent the greatest injuries, more from sense that they deserve, and are the proper objects of resentments, than from feeling himself the furies of that disagreeable passion; who, like a judge, considers only the general rule (Smith, 1759, III.6.4, p. 172).

The selfish passions hold a middle position between the social and unsocial passions with regard to the question of whether a sense of duty or the selfish passions ought to guide our behavior. According to Smith:

The pursuit of the object private interest, in all common, little, and ordinary cases, ought to flow rather from regard to the general rules which prescribe such conduct, than from any passion for the objects themselves; but upon more important and extraordinary occasions, we should be awkward, insipid, and ungraceful, if the objects themselves did not appear to animate us with considerable degree of passion (Smith, 1759, III.6.6, pp. 172–173).

In other words, in the “common, little, and ordinary” cases, we have the incentive to follow the general rules, not our selfish passions. As Smith writes: “To be anxious, or to be laying a plot either

to gain or to save a shilling, would degrade the most vulgar tradesman in the opinion of all his neighbors” (Smith, 1759, III.6.6, p.173)

Returning to our tipping example: if we conceive of tipping as a “common, little, and ordinary” situation, we have an incentive to follow the tipping norms of a particular place and situation. This can also explain why many of us, when visiting new places, put effort into finding out what the tipping norms are.

In the “more important and extraordinary” cases, we have an incentive to follow the selfish passions, as these bear a close resemblance to ambition:

A person appears mean-spirited, who does not pursue these with some degree of earnestness of their own sake.... Those great objects of self-interest ... are the objects of the passion properly called ambition; a passion, which when it keeps within the bounds of prudence and justice, is always admired in the world (Smith, 1759, III.6.7, p. 173).

Smith explains that self-love, inspiring one to better one’s condition, is not only proper; it is admirable. He offers some examples. We should have little respect for a gentleman who did not try to pursue self-love and obtain offices “when he could acquire them without either meanness or injustice.” A parliament member who is not happy about their election gains little respect from their neighbors.

One can also think about the choice to get married or to have children, or the choice of which educational path to pursue as properly governed by selfish passions rather than rules of behavior—whether or not others think it is proper to get married or have children or to pursue a particular educational path. Nevertheless, although we might have an incentive to follow the selfish passions in the “more important” cases, in certain societies today and in the past, the choice to get married and even whom to marry is not an individual choice but is largely influenced by societal expectations, especially for women.<sup>10</sup>

---

<sup>10</sup> Ahlawat (2017) examines marriage norms and the notion of honor in rural Haryana, showing that violation of prescribed rules of marriage results in severe punishment.



### 3. EXPERIMENT

One lesson learned from Adam Smith's theory of moral judgment is that actions as well as judgments of them depend on the situation. Recall that according to Smith, social norms are abstract in the sense that "all actions of a certain kind, or circumstanced in a certain matter, are approved or disapproved of" (Smith, 1759, III.4.7–8, p. 159). To examine Smith's conjectures about actions as well as judgements, we conducted an experiment whereby subjects were placed in a dictator game situation.

As dictators, subjects were given the right to decide the allocation of some money between themselves and an anonymous fellow subject, the recipient. In the experiment, we varied whether the dictator or the recipient was entitled to the money. In the Give treatment, the dictator was given the money, and the dictator decided how much of this money to give to the recipient. In the Take treatment, the recipient was given the money, and the dictator decided how much to take from the recipient and transfer to himself.

#### *Design*

Subjects were randomly assigned to one of four treatments in a between-subject experiment. In the two propriety elicitation treatments, we implemented the Krupka and Weber (2013) method to elicit the propriety of taking and giving behavior in the dictator game.

In the Give Elicitation treatment, subjects took the role of Smithian spectators who judged the actions of a hypothetical dictator endowed with 2000 NOK (NOK = Norske kroner; 1 USD = 9.20 NOK at the time of the experiment), while the hypothetical recipient was endowed with 0 NOK.<sup>11</sup> In the scenario the spectators were judging, the dictator could decide how much to give to the recipient. The dictator could choose among five choices: giving nothing, 500 NOK, 1000 NOK, 1500 NOK, or 2000 NOK. In the Take Elicitation treatment, the hypothetical dictator was endowed with 0 NOK, while the recipient was endowed with 2000 NOK. The dictator could then decide how

---

<sup>11</sup> Erkut et al., (2015) elicit norms separately from dictator, recipient, and disinterested third party respondents in the Krupka and Weber (2013) framework and find that elicited norms are stable and insensitive to the role of the respondent.

much to take from the recipient. The dictator could choose among five decisions: take everything, take 1500 NOK, take 1000 NOK, take 500 NOK, or take 0 NOK.<sup>12</sup>

After being presented with the hypothetical scenarios, the spectators were asked to rate the appropriateness of each of the five choices available to the dictator. The subjects rated each choice as either “very socially inappropriate,” “somewhat socially inappropriate,” “somewhat socially appropriate,” or “very socially appropriate.”

For each choice, if the participant’s appropriateness rating was the same as the modal response, they would participate in a drawing to receive 500 NOK. Thus, subjects had an incentive to reveal what they perceived to be the collectively shared judgment of the propriety of each action the dictator could take and not their own personal judgment.

The two other treatments examined whether the behavior of subjects, now in the role of actual dictators, corresponded to the elicited propriety ratings in the above hypothetical situations. In the Give Choice treatment, a randomly chosen subject was given 2000 NOK and had to decide how much he or she wanted to give to a randomly chosen recipient. Subjects could choose among giving nothing (0 NOK), 500 NOK, 1000 NOK, 1500 NOK, or 2000 NOK. In the Take Choice treatment, a randomly chosen recipient was given 2000 NOK. Subjects, as dictators, decided how much they wanted to take from the recipient’s endowment and transfer to themselves, and how much to leave for the recipient. They could choose among taking everything (2000 NOK), 1500 NOK, 1000 NOK, 500 NOK, or 0 NOK.

A total of 2,085 subjects were drawn from the Norwegian Citizen Panel, an online panel that offers a cross-section of the Norwegian population aged 18–76 years (Ivarsflaten, 2017). The Norwegian Citizen Panel is a probability-based online survey offering a representative cross-section of the Norwegian population over the age of 18. The sample is recruited by post and contains individuals from various age categories, education levels, and geographical regions. The fielding company reported that the survey lasted 20 min on average where the dictator game response times were estimated to be one minute. The number of subjects in each treatment is reported in Table 1. After the experiment was finalized, a random draw decided who would get paid according. The subjects whose names were drawn were paid according to a double-blind

---

<sup>12</sup> The detailed experimental instructions can be found in Appendix A.

procedure. Four subjects were paid 500 NOK in the Give Elicitation and four in the Take Elicitation treatment. In the Give Choice treatment, two subjects were paid 500 and 1500 respectively. In the Take Choice treatment, two randomly drawn subjects were paid 1000 each.

**Table 1:** Main features of the experimental design

<b>Framing</b>	<b>Initial Entitlement</b>	<b>Elicitation</b>	<b>Choice</b>
Give to	Dictator	N=499	N=553
Take from	Recipient	N=490	N=543

*Results: Propriety elicitation*

In the Give Elicitation treatment, we replicated the findings of Krupka and Weber (2013), in which the most socially appropriate dictator choice was an even split of the money.<sup>13</sup> In particular, a total of 92% of the subjects rated giving half of the money to the recipient as either socially appropriate or very socially appropriate.

The Give Elicitation treatment is related to what we perceive to be the beneficence domain governed by social passions. Dictators' allocation of positive amounts of money to the recipient may be understood, by the spectator, as beneficent acts toward the recipient. According to Adam Smith, beneficence is always free and "cannot be exhorted by force" (Smith, 1759, II.ii.1, p. 7). Hence, neither the recipient nor the spectator, or the experimenter, for that matter, can require the dictators to give to the recipient, but it is always appreciated, nonetheless. Thus, the dictator cannot be punished for a lack of beneficence in keeping most of the money but may be blamed for it.

As can be seen from Table 2, the subjects deemed all dictators' choices on average to be socially appropriate except the choice of keeping the money for oneself. But giving nothing was deemed appropriate by 36% of the subjects.

The even split may be interpreted as Smith's "mediocrity", the proper choice in this situation. As many as 71% of the subjects deemed it "very socially appropriate." In line with

---

<sup>13</sup> We compare our Give Elicitation treatment with Krupka and Weber's (2013) "*Standard*" treatment, in which the dictator was entitled to 10 US dollars and could choose to give to the recipient. The elicited ratings in our experiment are comparable with Krupka and Weber's result, as reported in their Table 1 (p. 505).

Smith's theory, we observe an asymmetry in propriety ratings when actions deviate from what is deemed proper: deviating from even sharing by giving all the money to recipient was judged more socially appropriate compared to deviating by keeping all the money for oneself (OLS coefficient: -0.60 vs. -0.96,  $p < 0.01$ ).

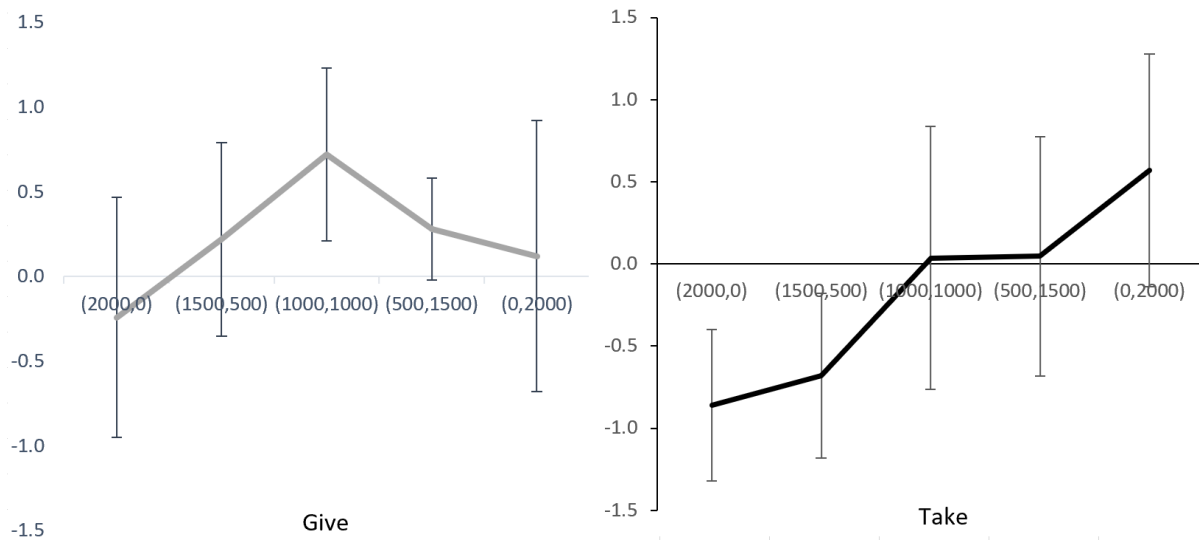
**Table 2:** Social appropriateness ratings (%) between treatments, payoff (dictator, recipient)

**Note:** Ratings are rescaled to range between -1 and +1, where -1 (--) = Very socially inappropriate, -0.33 (-) = Somewhat socially inappropriate, 0.33 (+) = Somewhat socially appropriate, and 1 (++) = Very socially appropriate.

GIVE (N = 499)							TAKE (N = 490)						
Payoff	Give	Mean	--	-	+	++	Take	Mean	--	-	+	++	Rank-sum (z)
(2000, 0)	0	-0.24	36	28	21	15	2000	-0.86	89	4	2	5	15.85***
(1500, 500)	500	0.22	7	26	44	23	1500	-0.68	65	27	5	4	19.94***
(1000, 1000)	1000	0.72	4	4	21	71	1000	0.036	29	16	26	29	14.15***
(500, 1500)	1500	0.28	9	25	31	35	500	0.046	25	17	36	23	4.65***
(0, 2000)	2000	0.12	24	20	20	36	0	0.57	14	5	13	69	9.42***

appropriate. Modal responses are shaded. All rank-sum tests are two-tailed, \*  $p < 0.1$ ; \*\*  $p < 0.05$ ; and \*\*\*  $p < 0.01$ .

**Figure 2:** Mean propriety ratings in the Give Elicitation and Take Elicitation Treatments



**Note:** The figures report the mean appropriateness and the standard deviation of the mean of each action available to the dictator. The x-axis depicts the resulting payoff (dictator, recipient) of the dictator's distribution choice. Ratings are rescaled to range between -1 and +1.

We relate the elicited ratings in the Take Elicitation treatment to the justice domain of dictator choice. Justice is abstaining from harming others. In our case, the dictator taking money from the recipient may be understood as harming the recipient and therefore as unjust. The dictators, however, were given the authority by the experimenter to take the money from the recipient's money. This "blessing" to take money may counter an incentive to judge the dictator's taking from the recipient as socially inappropriate. We show that letting the recipient keep the whole endowment was on average deemed as very socially appropriate by 69% of subjects and may be interpreted as propriety. On the other hand, as many as 93% of subjects rated taking everything from the recipient as either very or somewhat socially inappropriate. Hence, there was substantial agreement among the spectators that taking all the money from the recipient is socially inappropriate.

Comparing the appropriateness ratings across treatments, we observed that appropriateness ratings differed for fixed payoffs. Taking everything from the recipient was on average deemed more socially inappropriate than giving nothing in the Give Elicitation treatment ( $p < 0.01$ ,  $z = 15.85$ ). At the other extreme, leaving all the money with the recipient was deemed more socially appropriate in the Take Elicitation treatment than in the Give elicitation treatment ( $p < 0.01$ ,  $z = 9.42$ ). We also observed that 50/50 sharing was rated more socially appropriate in the Give than in Take Elicitation treatment ( $p < 0.01$ ,  $z = 14.15$ ). In the latter, however, subjects did not seem to agree on what was appropriate, as 45% deemed the choice socially inappropriate and 55% deemed it appropriate. Comparably, in the Give elicitation treatment, only 8% deemed it inappropriate, while as many as 92% deemed it socially appropriate.

### *Result: Actual choices*

Recall that in accordance with Adam Smith's theory, when acting, one takes into account both the "underlying passions which prompt us" to act, independently of social norms, and the "preciseness and exactness or the looseness and inaccuracy" of the social norms themselves.

We may interpret the Give Choice treatment as relating to the beneficence domain, as giving to the recipient is doing something good for the recipient. According to Smith, the social norms for beneficent actions are "loose and inaccurate," and hence beneficent actions should flow from the heart more than from social norms. However, in the dictator choice situation, there is also

a selfish passion involved, as the decision affects also the dictator's own payoff. If the subjects perceive the experimental situation to be "common, little, and ordinary", they have, according to Smith's theory, an incentive to follow the social norm, which in this case means to share the money equally.

We relate the Take Choice treatment to the justice domain and the unsocial passions, as taking money from the recipient may be perceived as an act of injustice and evoke resentment among spectators. For Smith, the same passion, resentment, could destroy a society if unconstrained, but it was necessary to get the rules of justice enforced. According to Smith, the social norms for justice are precise and exact, and in such a situation, the social norm should restrict us from harming others. As already noted, however, the dictator has been given the right, from the experimenter, to take from the recipient's money. Hence, it is not a priori clear whether subjects perceive taking money from the recipient as unjust. Still, based on the elicited appropriateness ratings, taking all the money for oneself was deemed "very socially inappropriate" by 89% of the spectators (See Table 2). Therefore, following Smith's theory, we would expect fewer subjects to take all the money from the recipient in Take Choice than in the Give treatment.

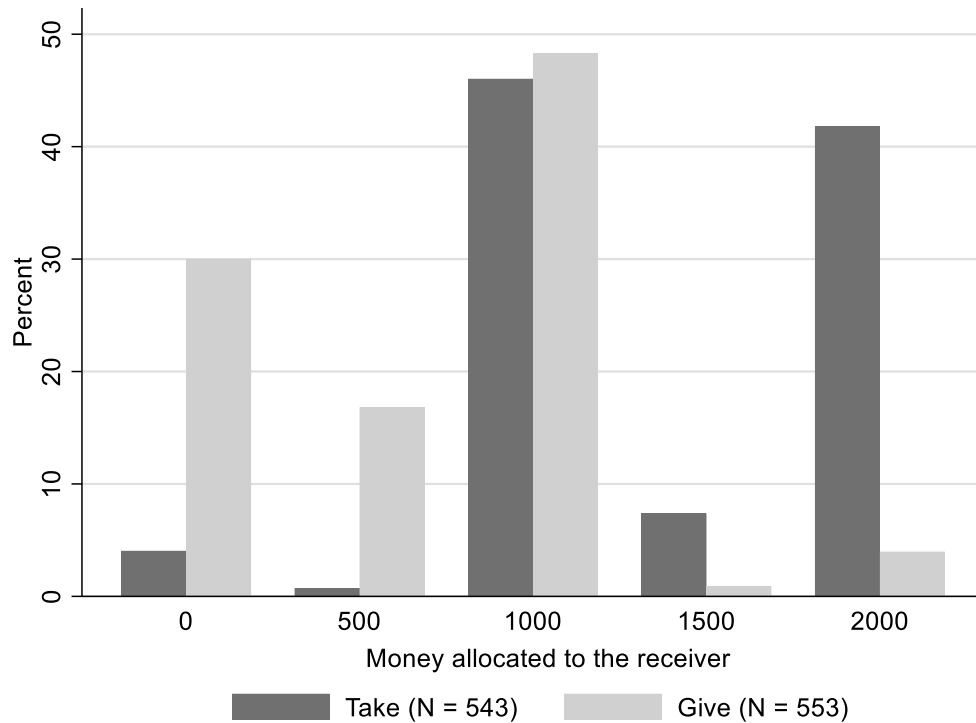
As can be seen from Figure 3, there is a remarkable shift in the distribution of the dictator's actual choices. In Give Choice, the distribution is similar to the typical distribution in the dictator game: the average distribution is 33% of the endowment, comparable to 28% in the meta study by Engel (2011, p. 588). The bimodal distribution has its first peak at the even distribution of the money and the second on the allocation leaving the dictator with all the money and nothing to the recipient. In Take Choice, however, the bimodal distribution is reversed: the first peak is still at the even distribution, but the second peak is on the allocation leaving the recipient with all the money. On average, dictators in Take Choice allocated 71% of the money to the recipient. This is significantly more than the 33% in Give Choice ( $p < 0.01$ ,  $z = 19.32$ ).

This shift in distribution is consistent with the elicited socially appropriateness ratings in the Give and Take Elicitation treatments. A larger percentage (42%) of subjects in Take Choice chose to let the recipient keep the whole endowment than in Give Choice, in which only 4% gave all the money to the recipient. This latter result is comparable to the dictator meta-study in which 5.4% of dictators gave all the money to the recipient (Engel, 2011, p. 589).

We also observed a high percentage of subjects in both treatments who chose the even split. If subjects in the Give Choice perceive the experimental situation as "common, little, and

ordinary,” they would have the incentive to follow social norms in this case. To reiterate, in the appropriateness ratings in the Give Elicitation treatment, 92% of the subjects rated this action as socially appropriate. In the Take elicited treatment, 55% did the same, and one explanation is that the subjects did not perceive it as unjust to take from the recipient’s money, as the experiment had explicitly allowed them to do so.

**Figure 3:** Dictator choices, varying according to whether dictator gives or takes.



In addition to varying the entitlements to the money, we also varied the framing of the dictator’s choices: subjects could either *give to* a recipient in the Give treatments or *take from* a recipient in the Take treatments. We chose this wording to keep the situation coherent with the actual situations the subject faced. We considered it unnatural to frame the dictators’ choices in the Take treatments as “giving to,” as the recipient had already been endowed with the money by the experimenter. Similarly, in the Give treatments, it seems unnatural to frame the dictator’s choice as “taking from,” as the dictator was endowed with the money; one does not take something that one has already been endowed with.

However, by varying both the entitlements to the endowment and the framing, our experiment may suffer from a confound in which the allocation of the endowment may influence the perceived entitlement of it. We addressed this concern in a second experiment reported in Appendix B. In this experiment, we dampened the entitlement effect by allocating the money to a common pool from which the dictator could either take or give. Thus, rather than entitling the dictator or the recipient to the entire endowment of money, we made both jointly entitled to a common pool. We show that propriety judgements of giving and taking do not vary between treatments when the endowment is placed in a common pool from which the dictator can either give or take.

#### **4. DISCUSSION**

A variety of economic behaviors have been explained by social norms. Still, to say something about how elicited norms affect behavior across different context, it is important to understand the social discovery process guiding individuals towards mutual agreement about appropriate standards of conduct in different situations. The main contribution of this paper is to apply Adam Smith's moral theory to inform this discovery process.

Whereas the heart of Smith's moral theory is that rules guide behavior, his main purpose is to explain *how* the spectatorial process of judgement moves necessarily separated individuals toward mutual agreement of what constitutes appropriate and inappropriate behavior. This argument echoes the work of Benabou and Tirole (2006) who similarly look to Smith's theory that sheds light on the channels involved in sustaining and inhibiting intrinsic motivations, emphasizing the role of deservingness and self-evaluation through the eyes of other fair and impartial spectators.

Smith explains that for a norm to exist in a certain society or a specific context, it is required that most people in this context judge similarly certain actions as morally right and specific wrongdoings as morally wrong. This assertion echoes the contemporary discourse among scholars studying norm-guided behavior. Applications of Smith's moral theory add to the conversation by explaining that to achieve this correspondence, we as spectators and people principally concerned have an incentive to make constant efforts to adjust, yet imperfectly, our feelings to a level that we imagine others going along with. It is this process of mutual adjustment that gives rise to social norms.



There is a call by the scholarship in the social sciences to better understand why individuals deviate from the standard rational-choice model (see DellaVigna, 2009: 318, 347; Bicchieri, 2006: 19). According to these models, payoff-identical actions in the dictator game should not be responsive to frames and whether the action involves giving compared to taking. But, as our, and many other papers have shown, context made salient through framing does matter for actions as well as what is considered appropriate.

To this end, social norms commonly enter as post-hoc explanations for context dependent actions. But for Smith, context is what gives meaning to outcomes (see Smith and Wilson (2019) for a discussion). Thus, one theoretical implication of applying the Smithian theory to the dictator game is that what is considered right and wrong depends upon context; one would do “the right thing” not for strategic calculations, but because that is what a real as well as an imaginary observer would judge as “the right thing” to do in that particular setting.

From a practical point of view, it is essential to understand the process whereby internal sanctions impede the violation of the appropriateness of certain behaviors. Although most economic experiments measure generosity through studying giving behavior, outside the lab actions indeed involve both giving and taking.<sup>14</sup> For instance, the design of most institutions requires individuals to overcome narrow self-interest in non-strategic situations resembling our dictator game in which the passive recipient cannot exert punishment or reward on the decision maker. Examples constitute tax compliance which consists in actions involved “taking” and charitable giving which consist in “giving”. Absent external punishment and reward, it is necessary to understand *why* people are other-regarding and *how* this depends on the context.

---

<sup>14</sup> For instance, Gächter et al., (2017) argue that social cooperation often requires collectively beneficial but individually costly restraint to provide (give) and maintain (abstain to take too much) a public good. Herein lies also the interest in understanding sustainable behaviors, which is increasing among the scholarship (see for instance, Sharma et al., 2021; Bansal et al., 2022). Studies in experimental economics have explored give-take choice framing effects in public goods games mimicking environmental and climate change settings (see Cox and Stoddard 2015 for a review).

## 5. SUMMARY REMARKS

We have argued, and shown experimentally, how this spectatorial process of judgement could be operating in two different contexts eliciting judgments and behavior consistent with the rules of justice and beneficence. Notably, we are not arguing that social norms evolve during an experiment. Our main goal was to apply Smith's theory to the norm elicitation method by Krupka and Weber (2013) and argue that actions in the dictator game which involve taking and which may be arising from Smithian unsocial passions, result in disapprobation.

We claim that this may be an expression of the sentiment spectators have for justice concerns. Actions such as giving which may be understood to arise from the social passions tend to lead to approval, expressing spectators' underlying sentiments for beneficence, instances of kindness and generosity. Finally, sympathy as a social validation process requires passions to be neither higher nor lower than what is considered proper in a given situation. We have shown that there exists an asymmetry in the intensity of spectators' judgments when actions deviate above and below this degree of propriety, and this varies with the passions that are governing actions in a particular situation.

Mentioning some limitations is due. As we consider a representative sample of the Norwegian population, replicating our study in other countries in which social norms regarding taking and giving could vary is important for the external validity of our findings. Moreover, we also have an online lab experiment. An important question is whether this would replicate in the field. Some studies on framing effects show that lab and field experiments could produce different results. Neumann (2019) reports framing effects in the lab but find no difference in the field. Grossman and Eckel (2015) show that a charity is left with 20% of the endowment irrespective of the game framing. On the other hand, Bott et al., (2020), show that situating misreporting in the tax declaration report as an action that involves "taking" from the old and vulnerable, increases tax compliance.

Still, if Smith's argument that context matters is solid, these somewhat varying results need not come entirely as a surprise. A study online versus a study in the field, increasing monetary stakes from what Smith refers to as "common, little, and ordinary" things, and making the identity

of the recipient salient, create yet another context that could affect not only actions but also elicit different perceptions of prevalent social norms.

## 6. REFERENCES

- Ahlawat, N. (2015). Marriage Norms, personal choices, and social sanctions in Haryana. *Sociological Bulletin*, 64(1), 91-103.
- Bansal, S., Jain, M., Garg, I., & Srivastava, M. (2022). Attaining circular economy through business sustainability approach: An integrative review and research agenda. *Journal of Public Affairs*, 22(1), e2319.
- Bardsley, N. (2008). Dictator Game Giving: Altruism or Artefact? *Experimental Economics*, 11(2), 122–133.
- Barr, A., Lane, T & Nosenzo, D., (2018) On the social appropriateness of discrimination. *Journal of Public Economics*, 164, 153-164.
- Bénabou, R., & Tirole, J. (2006). Incentives and prosocial behavior. *American Economic Review*, 96(5), 1652-1678.
- Benabou, R., & Tirole, J. (2011). Laws and norms (No. w17579). *National Bureau of Economic Research*.
- Bicchieri, Cristina. (2006). *The Grammar of Society: The Nature and Dynamics of Social Norms*. Cambridge: Cambridge University Press.
- Bicchieri, C., & Chavez, A. (2010). Behaving as expected: Public information and fairness norms. *Journal of Behavioral Decision Making*, 23(2), 161-178.
- Bicchieri, C., & Xiao, E. (2009). Do the right thing: but only if others do so. *Journal of Behavioral Decision Making*, 22(2), 191-208.
- Blay, A. D., Gooden, E. S., Mellon, M. J., & Stevens, D. E. (2018). The usefulness of social norm theory in empirical business ethics research: A review and suggestions for future research. *Journal of Business Ethics*, 152(1), 191-206.
- Bott, K. M., Cappelen, A. W., Sørensen, E. Ø., & Tungodden, B. (2020). You've got mail: A randomized field experiment on tax evasion. *Management Science*, 66(7), 2801-2819.
- Bursztyn, L., González, A. L., & Yanagizawa-Drott, D. (2020). Misperceived social norms: Women working outside the home in Saudi Arabia. *American Economic Review*, 110(10), 2997-3029.
- Bursztyn, L., & Yang, D. Y. (2021). *Misperceptions about others* (No. w29168). National Bureau of Economic Research.

- Brown, V. (2011a). *Intersubjectivity, The Theory of Moral Sentiments and the Prisoners' Dilemma*, The Adam Smith Review, 6: 172-90, F. Forman-Barzilai (ed.), Routledge.
- Campbell, T. (1971). *Adam Smith's Science of Morals*. Routledge.
- Cappelen, A. W., Nielsen, U. H., Sørensen, E. Ø., Tungodden, B., & Tyrann, J. R. (2013). Give and Take in Dictator Games. *Economics Letters*, 118(2), 280–283.
- Cox, C. A., & Stoddard, B. (2015). Framing and feedback in social dilemmas with partners and strangers. *Games*, 6(4), 394-412.
- Cox, J. C., Friedman, D., & Gjerstad, S. (2007). A Tractable Model of Reciprocity and Fairness. *Games and Economic Behavior*, 59(1), 17–45.
- Dana, J., Cain, D. M., & Dawes, R. M. (2006). What You Don't Know Won't Hurt Me: Costly (but quiet) Exit in Dictator Games. *Organizational Behavior and Human Decision Processes*, 100(2), 193–201.
- Dathe, T., Dathe, R., Dathe, I., & Helmold, M. (2022). Ethical Theories. In *Corporate Social Responsibility (CSR), Sustainability and Environmental Social Governance (ESG)*, 95-106.
- Dreber, A., Ellingsen, T., Johannesson, M., & Rand, D. G. (2013). Do people care about social context? Framing effects in dictator games. *Experimental Economics*, 16(3), 349-371.
- Engel, C. (2011). Dictator Games: A Meta Study. *Experimental economics*, 14(4), 583-610.
- Easterly, W. (2021). Progress by consent: Adam Smith as development economist. *The Review of Austrian Economics*, 34(2), 179-201.
- Eichenberger, R., & Oberholzer-Gee, F. (1998). Rational Moralists: The Role of Fairness in Democratic Economic Politics. *Public Choice*, 94(1–2), 191–210.
- Erkut, H., Nosenzo, D., & Sefton, M. (2015). Identifying Social Norms Using Coordination Games: Spectators vs. Stakeholders. *Economics Letters*, 130, 28-31.
- Gächter, S., Kölle, F., & Quercia, S. (2017). Reciprocity and the tragedies of maintaining and providing the commons. *Nature Human Behaviour*, 1(9), 650-656.
- Gerlach, P., & Jaeger, B. (2016). Another Frame, Another Game? Explaining Framing Effects in Economic Games. In A. Hopfensitz & E. Lori (eds.), *Proceedings of Norms, Actions, Games (NAG 2016)*. Institute for Advanced Studies.
- Grossman, P. J., & Eckel, C. C. (2015). Giving versus taking for a cause. *Economics Letters*, 132, 28-30.
- Griswold Jr., C. L. (1998). *Adam Smith and Virtues of Enlightenment*. Cambridge University Press.
- Güth, W., & Kliemt, H. (1998). The Indirect Evolutionary Approach: Bridging the Gap Between Rationality and Adaptation. *Rationality and Society*, 10(3), 377–399.
- Halvorsen, T. (2012). Moral motivation in dictator games, PhD Thesis

- Houser, D., & Xiao, E. (2011). Classification of natural language messages using a coordination game. *Experimental Economics*, 14(1), 1-14.
- Ivarsflaten, E. et al. (2017) Norsk medborgerpanel 2017 - Study Documentation. *UiB og UNI Research Rokkansenteret*.
- Kimbrough, E. O., Miller, J. B., & Vostroknutov, A. (2014). *Norms, frames and prosocial behavior in games*. Working paper, Simon Fraser University, Bocconi University, Maastricht University.
- Krupka, E. L., & Weber, R. A. (2013). Identifying Social Norms Using Coordination Games: Why Does Dictator Game Sharing Vary? *Journal of the European Economic Association*, 11(3), 495–524.
- Legros, S., & Cislighi, B. (2020). Mapping the social-norms literature: An overview of reviews. *Perspectives on Psychological Science*, 15(1), 62-80.
- List, J. A. (2007). On the Interpretation of Giving in Dictator Games. *Journal of Political Economy*, 115(3), 482–493.
- Mansi, J., Gagan, D., & Mandeep, M. (2019). Can I sustain my happiness. A review, critique, and research agenda for the economics of happiness. *Sustainability*, 11(6375), 1-36.
- Norman, J. (2018). *Adam Smith: What he thought, and why it matters*. Penguin UK.
- Ostrom, E. (2000). Collective Action and the Evolution of Social Norms. *Journal of Economic Perspectives*, 14(3), 137–158.
- Paganelli, M. P. (2009). Smithian Answers to Some Experimental Puzzles. In J. T. Young (Ed.), *Elgar companion to Adam Smith* (pp. 181-192). Edward Elgar.
- Raphael, D. D., & Macfie, A. L. (1976). Introduction to Smith (1759). Liberty Fund.
- Serdarevic, N. (2021). More or less money: The Love of Praiseworthiness and the Dread of Blameworthiness in One-player games. *Erasmus Journal for Philosophy and Economics*, 14(2), 1-24
- Smith, A. (1759). *The Theory of Moral Sentiments*. Liberty Fund (Glasgow ed. 1976).
- Skinner, A. S., & Haakonssen, K. (2001). *Index to the works of Adam Smith*. Liberty Fund.
- Smith, V., & Wilson, B. (2019). *Humanomics: Moral Sentiments and the Wealth of Nations for the Twenty-First Century*. Cambridge University Press.
- Sharma, G. D., Tiwari, A. K., Talan, G., & Jain, M. (2021). Revisiting the sustainable versus conventional investment dilemma in COVID-19 times. *Energy Policy*, 156, 112467.
- Stevens, D. E. (2018). *Social norms and the theory of the firm: A foundational approach*. Cambridge University Press.

- Steeds, L. (2022). The Social Ecology of Adam Smith: Reconsidering the Intellectual Foundations of Political Economy. *New Political Economy*, 27(1), 132-145.
- Sund, O.K.S. (2017). It is not all about money. A study of social norms based on Adam Smith's Theory of Moral Sentiment (Master thesis), Department of Economics, University of Bergen.
- Suvoy, R. (2003). The Effects of Give and Take Framing in a Dictator Game (Doctoral dissertation, University of Oregon).
- Swope, K., Cadigan, J., Schmitt, P., & Shupp, R. (2008). Social Position and Distributive Justice: Experimental Evidence. *Southern Economic Journal*, 811-818.
- Tjøtta, S. (2019). More or less money? An Experimental Study on Receiving Money. *Journal of Behavioral and Experimental Economics*, 80, 67–79.
- Tversky, Amos and Daniel Kahneman (1981). "The framing of decisions and the psychology of choice," *Science*, 211: 453–458.
- Veselý, S. (2015). Elicitation of Normative and Fairness Judgments: Do Incentives Matter?. *Judgment and Decision Making*, 10(2), 191-197
- Veselý, S. & Klöckner, A.C. (2018). Global social norms and environmental behavior. *Environment and Behavior*, 50(3), 247-272
- Wight, J. B. (2013). Sentiments and Motivations in Adam Smith and Vernon Smith. *Forum for Social Economics*, 42(4), 298–301).
- Young, J. T. (Ed.). (2009). *Elgar companion to Adam Smith*. Edward Elgar.
- Zizzo, D. J. (2010). Experimenter Demand Effects in Economic Experiments. *Experimental Economics*, 13(1), 75–98.

## 7. APPENDIX

### Appendix A: Instructions, Experiment<sup>15</sup>

#### Give Elicitation

You will now evaluate the choices available to a “person A” in a hypothetical situation. Four participants in the Norwegian Citizen Panel will be randomly selected to receive NOK 500 if they have evaluated one of the choices by person A similarly to the majority of the participants in this experiment.

The hypothetical situation is as follows:

“Person A is randomly selected to receive NOK 2,000. Person A can give their money to another randomly selected person B. Person A can choose between giving nothing (NOK 0) to B, NOK 500, NOK 1,000, NOK 1,500, or everything (NOK 2,000).”

The table below shows the choices person A can make. Please indicate how socially appropriate you think each of the five choices are. By socially appropriate is meant “to be in accordance with moral or appropriate social behaviour”.

#### EVALUATION TABLE

---

<sup>15</sup> The experiment was fielded in Norwegian. The original instructions in Norwegian are available upon request for replication purposes. Please note that the English translation provided in this paper deviates somewhat from the English translation offered by the fielding company (Skjervheim & Høgestøl, 2017).

### Take Elicitation

You will now evaluate the choices available to a “person A” in a hypothetical situation. Four participants in the Norwegian Citizen Panel will be randomly selected to receive NOK 500 if they have evaluated one of the choices by person A similarly to the majority of the participants in this experiment.

The hypothetical situation is as follows:

“Person B is randomly selected to receive NOK 2,000. Person A is randomly selected to take money from person B. Person A can choose between taking nothing (NOK 0) from person B, NOK 500, NOK 1,000, NOK 1,500 or everything (NOK 2,000)”.

The table below shows the choices person A can make. Please indicate how socially appropriate you think each of the five choices are. By socially appropriate is meant “to be in accordance with moral or appropriate social behaviour”.

## **EVALUATION TABLE**

### Give Choice

Participant A in the Norwegian Citizen Panel is randomly selected to receive NOK 2,000. Participant A can give money to another randomly selected participant, participant B. If you are randomly selected to be a participant A, you must choose how much to give to participant B. What you give is kept by participant B. You can choose between giving nothing (NOK 0) to the second participant, NOK 500, NOK 1,000, NOK 1,500, or everything (NOK 2,000).

### Take Choice

Participant B in the Norwegian Citizen Panel is randomly selected to receive NOK 2,000. Someone else, participant A, is randomly selected to take money from participant B. If you are randomly selected to be a participant A, you must choose how much to take from participant B. What you do not take is kept by participant B. You can choose between taking nothing (NOK 0) from the second participant, NOK 500, NOK 1,000, NOK 1,500, or everything (NOK 2,000).



## Appendix B: EXPERIMENT II

A possible objection to our experiment is that we varied between the Give and Take framing as well as the entitlement to the money between the dictator and recipient. A similar design feature is, however, also present in Krupka and Weber (2013). In their “*Standard*” treatment, the dictators were endowed with 10 US dollars and were asked to give to the recipient. In the “*Bully*” treatment, however, the dictator and recipient were endowed with 5 US dollars each. The dictator’s choice was framed as giving to the recipient if the dictator transferred money from their 5 dollars, and framed as taking if the dictator transferred money from the 5 dollars the recipient was given. Thus, the elicited norms and corresponding behavior may vary because of the framing, but also because of differences in sentiments of entitlements between the “*Standard*” treatment and the “*Bully*” treatment.

In Experiment II, we dampened the entitlement effect by allocating the money to a common pool from which the dictator could either take or give. Thus, rather than entitling the dictator or the recipient to the entire endowment of money, we made both jointly entitled to a common pool. In the Give treatment, the dictator chose how much to give from the common pool to the recipient. The remaining went to the dictator. In the Take treatment, the dictator chose how much to take from the common pool and transfer to himself. The remaining went to the recipient. Like in Experiment I, we elicited spectators’ appropriateness ratings and actual dictator choices in these two choice environments.

### *Design*

In the propriety elicitation experiment, subjects were asked to rate the social appropriateness of a hypothetical dictator’s choices. This dictator could allocate money from a common pool of 1000 NOK between himself and another subject (the recipient). We consider again a between-subjects design. In the Give Elicitation treatment, the hypothetical dictator could choose between giving nothing (0 NOK), 250 NOK, 500 NOK, 750 NOK, or everything (1000 NOK). In the Take Elicitation treatment, the dictator could take 0 NOK, 250 NOK, 500 NOK, 750 NOK, or everything (1000 NOK).<sup>16</sup>

---

<sup>16</sup> The detailed experimental instructions can be found in Appendix C.

After having been presented with the hypothetical situation, the subjects were asked to rate the appropriateness of each choice available to the dictator. The subject could rate each action on a four-point scale as “very socially inappropriate,” “somewhat socially inappropriate,” “somewhat socially appropriate,” or “very socially appropriate.” As in Experiment I, for each choice the hypothetical dictator could make, the spectators’ appropriateness ratings were incentivized to match the modal response in the session in order to participate in a drawing to receive 500 NOK.

A total of 78 undergraduate students participated in a classroom experiment at the Norwegian Business School in Bergen in March 2017. The students were informed that 10 of them would earn 500 NOK each.<sup>17</sup> The experiment was conducted with pen and paper. Subjects were given 5 minutes to finish their evaluation.

Another set of subjects faced the corresponding decision environments, but as actual dictators. The dictators were asked to allocate money from a common pool of 1000 NOK between themselves and a recipient. Subjects were randomly assigned one of two treatments. In Give Choice, the dictators were asked to choose how much of the common pool to give to the recipient, keeping the rest for themselves. The dictators could choose between giving nothing (0 NOK), 250 NOK, 500 NOK, 750 NOK, or everything (1000 NOK). In Take Choice, the dictators were asked to choose how much to take from the common pool to themselves, leaving the rest to the recipient. The dictator could choose between taking nothing (0 NOK), 250 NOK, 500 NOK, 750 NOK, or everything (1000 NOK).

A total of 117 undergraduate students at the University of Bergen participated in the actual dictator choices. The experiment consisted of four sessions, all conducted with pen and paper in the Citizen Lab at the University of Bergen in February 2017.

## *Results*

The results from Experiment II show that there is no significant difference between framing the dictator choice as either *give to* or *take from* when the dictator allocates money from the common

---

<sup>17</sup> A total of 157 students participated in a larger study, 78 of whom participated in our experiment, while the rest participated in an experiment not reported here. In the randomized payment, 10 subjects out of 157 were drawn as winners.

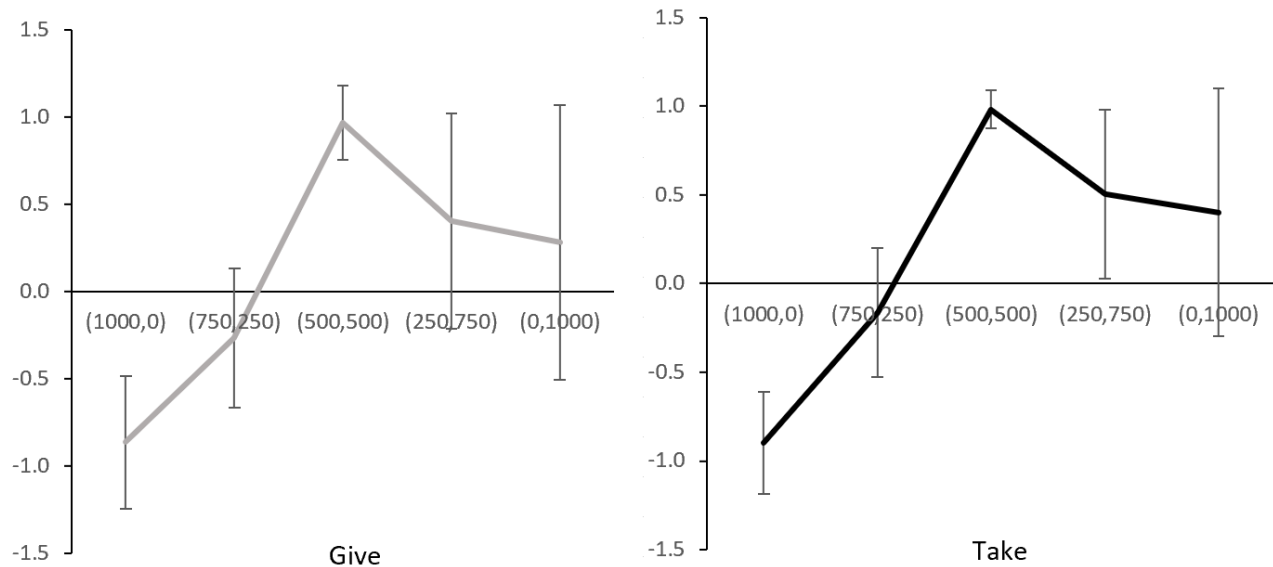
pool, neither in the propriety elicitation treatments or in the actual dictator choices; see Table B1 and Figure B1. Hence, propriety judgements of giving and taking are conditional on *whom* one is taking from. Taking from another person who was entitled to the money is different from taking from a common pool, implicitly suggesting that it is both persons' property.

**Table B1:** Social appropriateness ratings (%) between treatments, payoff (dictator, recipient)

GIVE (N = 39)							TAKE (N = 39)						
Payoff	Give	Mean	--	-	+	++	Take	Mean	--	-	+	++	Rank-sum (z)
(1000, 0)	0	-0.86	85	13	0	3	1000	-0.90	87	10	3	0	-0.323
(750, 0)	250	-0.26	13	64	23	0	750	-0.16	5	64	31	0	1.113
(500, 500)	500	0.97	0	3	0	97	500	0.98	0	0	3	97	0.018
(250, 1000)	750	0.40	5	21	32	42	250	0.50	0	15	44	41	1.514
(0, 1000)	1000	0.28	21	10	26	44	0	0.40	10	18	23	49	0.618

**Note:** Ratings are rescaled to range between -1 and +1, where -1 (--) = Very socially inappropriate, -0.33 (-) = Somewhat socially inappropriate, 0.33 (+) = Somewhat socially appropriate, and 1 (++) = Very socially appropriate. Modal responses are shaded. All rank-sum tests are two-tailed, \*  $p < 0.1$ ; \*\*  $p < 0.05$ ; and \*\*\*  $p < 0.01$ .

**Figure B1:** Mean propriety ratings in the Give Elicitation and Take Elicitation Treatments

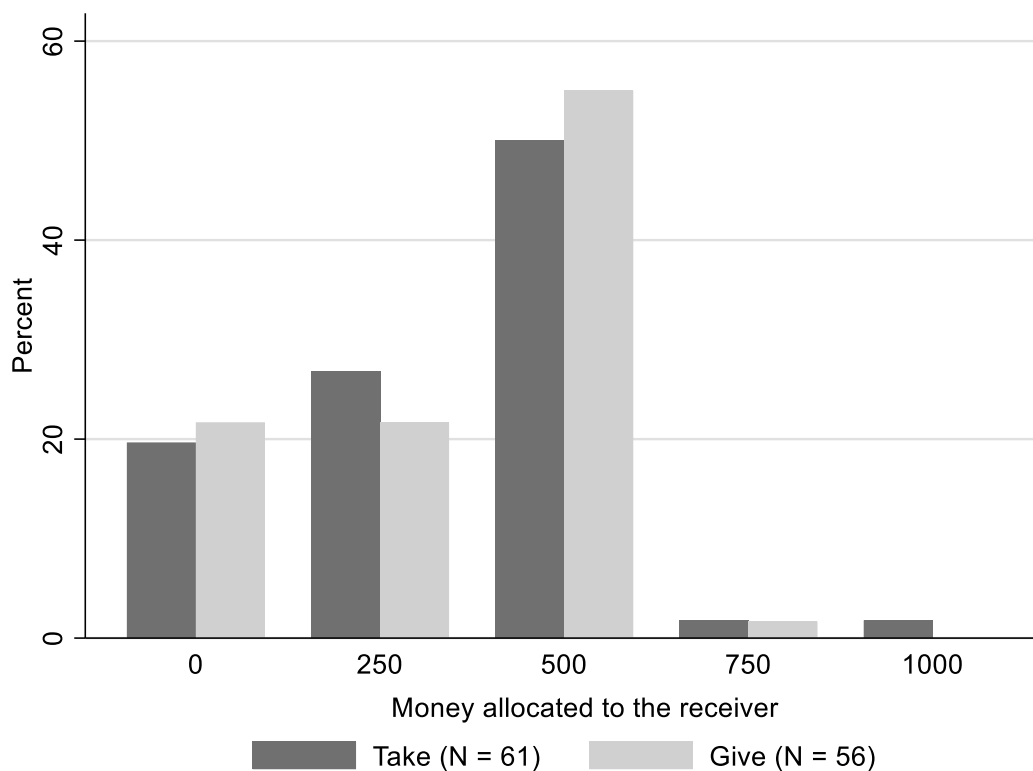


**Note:** The figures report the mean appropriateness and the standard deviation of the mean of each action available to the dictator. The x-axis depicts the resulting payoff (dictator, recipient) of the dictator's distribution choice. Ratings are rescaled to range between -1 and +1.

The proper choice seems to be equal sharing of money; taking half and giving half is judged as the most socially appropriate action by the spectators. Note the asymmetry in deviating from the “mediocrity” point, which is similar to the Give Elicitation treatment in Experiment I and previous findings in the literature (see Krupka and Weber, 2013, p. 506, footnote 15; Erkut et al., 2015). The result reveals that the asymmetry in this case seems to be that dictator sharing more than half the common pool is socially appropriate. The dictator choosing to keep more from the pool, however, is deemed socially inappropriate. As we already noted, this asymmetry is in accordance with Smith’s theory.

The choices of the dictator in the Give Choice and Take Choice treatments follow a similar pattern of elicited propriety ratings; the dictators’ actual choices are unaffected by whether the actions are framed as “to take” or “to give” when taking or giving from a common pool ( $p = 0.978$ ,  $z = -0.038$ ) as shown in Figure 5. Mirroring the elicited propriety ratings, we observe that the most common is the 50/50 split.

**Figure B2:** Dictator behavior, varying according to whether dictator gives or takes.



## **Appendix C: Instructions, Experiment 2<sup>18</sup>**

### Give Elicitation

Assume that two randomly chosen individuals, Person A and Person B, own a common pool of 1000 NOK. Person A can decide how much of the money to give to Person B. What Person A gives, Person B receives, the remaining is kept by Person A. Person A can choose between giving nothing (0 NOK), 250 NOK, 500 NOK, 750 NOK, or everything (1000 NOK).

The table below shows the different choices available to Person A. Please indicate how socially appropriate you find each of the five actions. Remember that you can be drawn to receive 500 NOK if your evaluation of the randomly selected action is equal to the evaluations of the majority of the other participants. Mark your answers with an X.

### **EVALUATION TABLE**

### Take Elicitation

Assume that two randomly chosen individuals, Person A and Person B, own a common pool of 1000 NOK. Person A can decide how much of the money he or she wants to take from the pool for themselves. What Person A takes, he or she keeps, the remaining amount is received by Person B. Person A can choose between taking nothing (0 NOK), 250 NOK, 500 NOK, 750 NOK, or everything (1000 NOK).

The table below shows the different choices available to Person A. Please indicate how socially appropriate you find each of the five actions. Remember that you can be drawn to receive 500 NOK if your evaluation of the randomly selected action is equal to the evaluations of the majority of the other participants. Mark your answers with an X.

### **EVALUATION TABLE**

---

<sup>18</sup> The original Norwegian instructions are available in Sund (2017).

### Give Choice

Two randomly chosen individuals, Person A and Person B, own a common pool of 1000 NOK. Person A decides how much of this money to give to Person B. What Person A gives is received by Person B, the remaining amount is kept by Person A.

If you are selected to be Person A, you have to choose how much from the pool you want to give to Person B. You can choose between giving nothing (0 NOK), 250 NOK, 500 NOK, 750 NOK or everything (1000 NOK).

### Take Choice

Two randomly chosen individuals, Person A and Person B, own a common pool of 1000 NOK. Person A decides how much of this money to take for themselves. What Person A takes is kept by him or her, the remaining amount is received by Person B.

If you are selected to be Person A, you have to choose how much you want to take from the pool. You can choose between taking nothing (0 NOK), 250 NOK, 500 NOK, 750 NOK or everything (1000 NOK).