

UNIVERSITY OF HELSINKI

An Outline of the Dictator Game: Room for expansion and refinement

Master's Programme in Economics, Research track
Master's thesis

Author:
Santeri Lehtinen

Supervisor(s):
Professor Klaus Kultti

27.08.2024
Helsinki

Faculty: Faculty of Social Sciences

Degree programme: Master's Programme in Economics

Study track: Research track

Author: Santeri Lehtinen

Title: An Outline of the Dictator Game: Room for expansion and refinement

Level: Master's Thesis

Month and year: August 2024

Number of pages: 59

Keywords: behavioral economics, decision making, dictator game, experimental economics

Supervisor or supervisors: Klaus Kultti

Where deposited: Helsinki University Library

Additional information:

Abstract:

This paper examines the methods, results, theoretical frameworks and shortcomings of the 40 years of dictator game research. It provides an overview of the main results, their reliability and consistency and the theoretical backing used to explain them. The game was created to test assumptions of the economic behavior of individuals in a situation where one party has anonymity, control of funds and no reciprocity, and thus no strategic or selfish reason to allocate funds to the other party. Over the years, the game has developed to test the behavior differences between groups with different characteristics, as well as a rich array of related games and variants to test behavior in a more sophisticated manner. A significant majority of studies find evidence of giving.

During research for this paper, meta-analyses and individual studies, especially those not covered by meta-analyses, were used in addition to theoretical papers. The main areas of focus in addition to the results were on the methods used, their weaknesses and the difficulties in the comparability of results caused by incomparable methods. Theoretical explanations were then considered and evaluated on the larger scope of results across different experiments.

This paper concludes that the widespread evidence of giving in the dictator game is not explainable as only caused by selfish concerns for personal benefit or simple mistakes. No single theory explains all the differences in behavior in the variety of experiments covered, and susceptibility for bias and error in human decision making is considerable. With the addition of the evidence for giving in even young children, and clear patterns in distributions emerging between many experiments, I conclude that most dictators value the utility the other party derives from their gift and that this giving has a typical distribution. Results, where giving is eliminated, can be explained by procedural factors. There is evidence that most behavior in a standard dictator game setting is rationalizable or nearly so.

I find that the standardization of protocols is lacking, and despite research indicating these differences cause strategic considerations, they have not been eliminated. Standardizing protocols and in particular making double-blind experiments the norm, as well as preventing dictators from also acting as recipients are the main recommendations of this paper. There is reasonable doubt that experiments that cover very similar or identical research questions are incomparable, making drawing larger conclusions difficult. Likewise, the applicability to life outside the lab is not simple or well-studied. The widespread use of students, especially economics students, limits how representative of the population dictator game experiments are. Finally, there is room for considerable exploration of more extreme dictator games, where very high payoffs or extreme multipliers could test the robustness of results. Sampling could also be improved in representativeness and sizes.

Table of contents

1	Introduction	5
2	Accessibility	7
2.1	Origins	7
2.2	An overview of the meta-analysis by Engel (2011)	8
2.3	Development, charitability and fairness	16
2.4	Team versus individual giving	17
2.5	Willingness to participate and delegation	18
2.6	Multi-step games	22
2.7	Earned endowment	23
2.8	Limited decision space and a pool of funds across multiple games	24
2.9	Giving and taking treatments	25
2.10	Compatibility with GARP and rationalizable preferences	26
2.11	Interactive and non-interactive games	29
2.12	Children as dictators	31
2.13	Artificial intelligence models as dictators	33
3	Theory	35
3.1	Giving as a mistake or artefact	35
3.2	Giving explained by psychological biases	36
3.3	Giving explained by rational preferences	38
3.4	Protocol differences and strategic considerations	41
3.5	Origin of giving	43
4	Discussion	44
4.1	The limitations of the use of students	45
4.2	Dictators in daily life	46
4.3	Game avoidance and explanations	46
4.4	Complexity in explanatory variables	48
5	Conclusion	51

References	55
Working papers	58
Econometrics	58

1 Introduction

The perfectly selfish, rational and aware player of game theory assumptions is one kind of monster. Compared to a human, the perfect rationality and understanding of the nature of the game is nearly machine-like; One would expect humans to make mistakes in games, and they often do. However, while proneness to mistakes of actual economic agents is well acknowledged, how selfish people are may be less clear, with a range of views being espoused throughout the history of philosophy and of economics. The dictator game is an experiment to test the cold part of that cold, calculated monster, and has for close to 40 years shed light on how people behave in anonymity, where they have the opportunity to advance their own economic interests at the expense of another, faceless and unknown person. The results have implied that the actual actors in economic affairs are not all cold, in addition to not being perfectly rational or efficient.

Behavioral research has explored the factors affecting decision-making in a variety of economic decision-making situations. In many of these situations, the standard rationality assumptions used in most economic models, most relevantly for decision experiments in game theory, are violated. In the Dictator game, where one participant chooses to allocate funds between themselves and another, powerless, party, a sizable proportion of decision-makers choose to give a portion of the pot to the other participants against their own assumed economic interests. Although a sizable proportion shares nothing (usually a minority, but over 20%), the majority typically shares some positive amount, with very few giving more than half. These results have traditionally been taken as evidence of altruism towards even anonymous individuals, although some studies have found that the effect disappears with changed framing (Bardsley 2008). Thus, the dictator game provides a useful setting to study the relation of altruism and situational factors, using framing to modify the experiment for different situations.

There is a wealth of experiments following what may be called the “standard” form of dictator games. Typical characteristics include full anonymity, where the recipient is not known, and will not find out the identity of the dictator. In the standard form, the recipient is another participant of the experiment, which often results in them being somewhat similar, such as fellow students when experimenting with university

students (as is often done). The dictator is asked to split a pot of (unearned) money, often \$10, between themselves and the anonymous other party. Their decision space is not limited, and the dictator cannot alter the total size of the pot by distributing it differently. However, a lack of solidified protocols and procedures have caused small differences between outwardly similar experiments, which as dictator game experimentation has shown, can lead to significantly different behavior.

Thus, the dictator game, after some novel initial findings, has suffered from a lack of rigor and consistency. Lack of standardization makes studies on in-name the same question not actually comparable, and the lack of setting hypotheses prior to experimentation makes results less valid. Meanwhile the connection to more generalized behavior in settings outside the lab is weakened by the lack of application and testing of theoretical frameworks in the context of the game, especially prior to recent years. Dictator game research is defined by scattered results, and more should and could be done to obtain more valid and significant results, especially as it relates to economic behavior.

This research will study the applications, results and their significance in dictator game research. An overview of the types and results of dictator games undertaken in the course of behavioral research will be given, starting with an overview by utilizing meta-analyses, such as Engel (2011) of more standard treatments and research questions, including assessing the methods, the reliability and significance of results, and possible criticisms. The volume of research limits the number of papers covered, but common trends in methodology and the theoretical outlooks proposed will be evaluated. This paper aims to outline the recent developments, as well as the possibilities yet to be fully explored. The dictator game as a tool can hopefully be wielded more precisely based on this in future use cases.

2 Accessibility

2.1 Origins

The beginnings of the dictator game (which is not really a game, as the other party does not play) were in another game, the ultimatum game (In which one party makes an offer of a pot split and the other party either accepts or rejects). The first form of the dictator game was a way to test the willingness to punish prior perceived unfairness, even with cost to self, something also relevant in the ultimatum game. The first dictator experiment had two parts. (Kahneman et al. 1986) In the first part, students were asked to split 20 dollars between themselves and an anonymous student. However, they were only given two options: \$18 to self and \$2 to the other, or an even split, \$10 to each.

This was however followed by an experiment, where a decision-maker that previously participated in the first experiment was given a pot to split between three people. If the two other participants made the same splitting decision in the prior experiment, a simple payout without decisions would occur. If they split differently, the decision-maker was given the chance to choose which of the two decision-makers received a payout in addition to themselves. The options were to get 6 dollars for the dictator and the one who chose two split the prior pot unevenly, or to get 5 dollars for the dictator and the even-splitter. The goal was to observe whether participants were willing to punish an unfair allocation. Like in many behavioral experiments, only a percentage of the participants were paid for each portion, and in the second portion only those who were not already paid could participate. 76% of dictators in part 1 chose to split the pot evenly, while 74% chose to punish the uneven distributor in part 2, even with cost to themselves.

The result of the first experiments raised further interest in testing whether there is a propensity in subjects to behave fairly, even when unobserved. With further experiments, the second part was stripped out, and the game was simplified. In the study by Forsythe et al. (1994), a group of 8 dictators were placed in a single room, while the recipients were in another room. While communication was forbidden and the dictators were not playing the game with each other, they were able to notice each other. The dictators were asked to record the splits for themselves and for the recipient on a paper sheet. There were two different pot sizes, \$5 and \$10,

experimented with on different days, and participants were paid an attendance fee of \$3 in some \$5 experiments and not in others. In most ways, this sets the standard form of the experiment, except for the fact that dictators are not typically allowed to observe each other's choices.

2.2 An overview of the meta-analysis by Engel (2011)

This standard form has been utilized rather often, with the meta-analysis by Engel (2011) covering 129 papers published between 1992-2009. This already excludes 76 experiments in that period that present a fundamentally different game. Some of the games excluded give power to the recipient, making them more similar to ultimatum games, some also keep the sanctioning, as in Kahneman et al. (1986) and some limit the options of the dictator. Some also use the dictator game to ask a different research question. Thus, there is also a sizable amount of research which do not cleanly fit into the standard model analysis. As seen later, even this sample does not necessarily only contain comparable experiments, as dictators have acted as recipients as well in some experiments.

Using the methods of both univariate and multivariate analysis, Engel (2011) estimates the significance of a list of factors commonly treated for in dictator game experiments. They find that the mean percentage of the pot given to be 28,35%, with only 6 out of 616 resulting in an average of zero, with a meta-analysis (where results of each study are treated as singular data points and weighed by their standard errors) of 445 treatments for which standard errors could be obtained resulting in a similar share. Individual give rates are concentrated in a few peaks, where the option to give none being the most chosen by far (approximately 36%). Two other significant peaks are the 50/50 split (approximately 17%) with 5,44% choosing to give the entire pot to the recipient! This result looks to rather starkly confirm that there is substantial giving in dictator games, but studies vary widely between each other. It also presents the fact that people act quite heterogeneously, which is noted in the distributions of most individual studies as well. This alone does drive a quite significant crack in the assumption that all people can be treated as acting in the same way every time when decisions are made. In univariate analysis, quite a few of the treatments are significant variables, although this changes for some variables depending on whether individual data or meta-regression are used.

Most studies generally find similar distributions: A sizable portion in the 20%-50% range gives nothing, while giving half is the most common contribution and only a small percentage gives more than half, often chalked up as an error or outlier, especially in limited samples. The main difference between results seems to be the mean of giving, with lower means associated with fewer dictators giving at all. Thus, the main thing discussed is the mean of giving, unless the distribution merits mention due to being unusual or shifted, with some choice being especially common in comparison to other experiments.

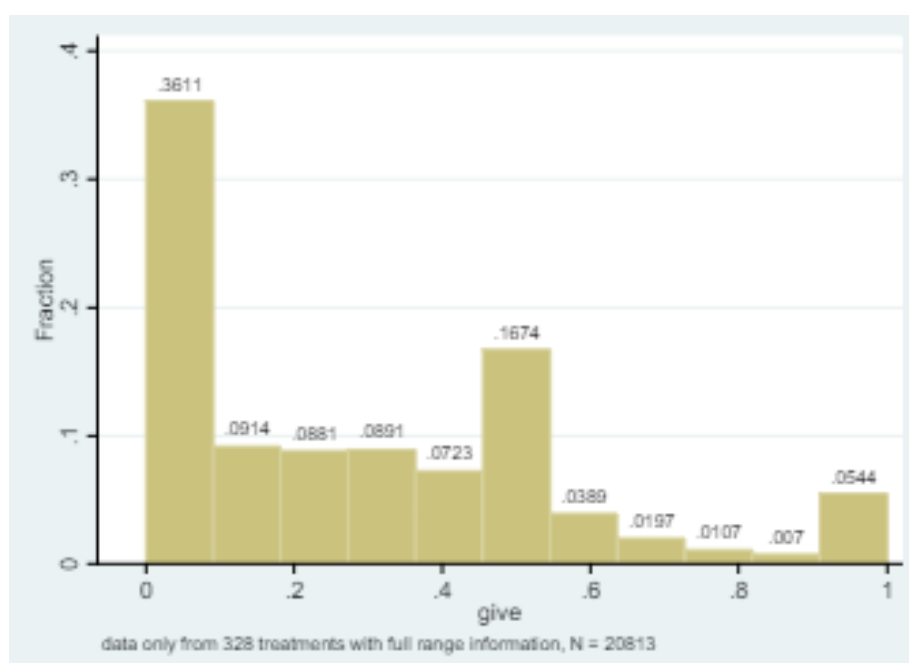


Figure 1. The distribution of individual give rates in Engel (2011), which closely resembles the distribution in most studies. Figure: Engel (2011)

The treatments presented in the Engel meta-analysis are divided into the categories of incentives, social control, distributive concerns, framing, social distance, and demographics. All the treatments are put in univariate analysis first a meta-regression and this was supplemented with individual data (her also called OLS), based on the distribution of responses from the studies. This data is secondary as it is less accurate, was in some cases imprecisely reconstructed and lacking in some experiments reported. It bears mentioning, that since the data set is so large and the factors playing into individual decision making so numerous, most significant effects are very small in magnitude, explaining typically less than 10% of the variance.

In the category of incentives, restraining the action space is insignificant in the meta-regression, but has a weakly significant positive effect in OLS. Introducing uncertainty in the payoff to the recipient also results in insignificance in the meta-regression, while in OLS it has a significant negative effect, where more uncertainty results in less giving. Having multiple dictators competing is based on a study on schoolchildren competing for the highest payoff for an extra reward. This produced no significance in the meta-regression, and at the individual level an effect that dictators contribute less in competitive conditions. In the study mentioned, this was linked to boys being less pro-social in competitive situations (Houser and Schunk 2009). Playing the game hypothetically results in no significant difference, but randomly paying only some dictators results in rather substantial and significant negative effect in OLS, when treating the prior two and the standard pay as distinct categories. Both forms of univariate analysis are insignificant for stake size (not necessarily surprising, as the name value in dollars may have rather different purchasing power in different countries and across 25 years), but in studies specifically manipulating it, increasing stake size results in a very small significant negative effect in the meta-regression. Undergoing the game multiple times has a clear and significant negative effect in the meta-regression. Making the dictator a group and not an individual results in insignificant meta-regression results, but a negative effect on willingness to give in the OLS.

A later meta-analysis by Larney et al. (2019) finds that stake size has a small, but significant negative effect on giving in the dictator game. In contrast, they do not find that effect in the ultimatum game. In this case, the high stakes for dictator game studies were not necessarily high in those participant studies that had varying stakes (median high \$20). Stakes higher than \$20 were uncommon in western studies, but developing nations allowed opportunities for endowments up to more than three months salary. These results of course may be affected by other factors related to these countries. An example of high stakes is Leibbrandt et al. (2015) which found that when both giving and taking framings (one allocates the initial endowment to the dictator, one to the other party) were used with participants from villages in Bangladesh, participants allocations for the recipient decreased substantially with higher stakes in the giving treatment, but not in the taking treatment. The high stakes here represented approximately 5 months of average income, while low stakes were equivalent to one day of average wages. It is worth noting, that in comparison to

games in the developed world, these sums are both rather large as a percentage of income. In the low stakes treatment, a giving variant left the recipient with 25% of the endowment on average, while the taking variant left them with 29,2%. In comparison, high stakes left the recipient with 3,7% of the endowment in the giving variant, while the taking variant left the recipient with 32,6%. This raises the question whether such results would be replicated with very high stakes in the developed world, and while stake size may have a negative effect of giving, it also seems to increase the “moral cost” of taking as well.

In the social control category (Engel 2011), making the recipients aware of the dictator’s identity has rather conflicting results: For the result to be significant in the meta-regression, stakes have to be controlled for, and then the effect on giving is negative. In OLS, the results match the expected positive significant effect. Giving a social cue (a rather vague category, the example provided was of three large dots arranged to evoke a face shape) had no significance, except in studies where it was the part manipulated, in which case the effect is rather large and significant in the meta-regression. Hiding the fact that the payoff is due to a dictator’s decision from the recipient has a significant negative effect on the willingness to give in the meta-regression. Using concealment methods to guarantee dictator-experimenter anonymity results in no significance, except when repeated interaction is controlled for. In that case there is a weakly significant negative effect in the willingness to give in the meta-regression.

In terms of distributive concerns, making the recipient receive the pot and then giving the dictator the option to take a fraction result in no significant effect in either individual effects or the meta-regression. To get a significant effect, Engel controls for stakes and adds the interaction term. This results in a positive and significant effect on willingness to not take in the meta-analysis. Making the recipient deserving (such as a charity etc.) produces rather clear results: There is a clear and significant positive effect to give in the meta-regression. The changes in the distribution are also significant: The portion of dictators who keep everything decreases to less than 20%, while the share who give everything increases to over 20%. This is notably one of the few cases, where giving all of the endowment gets chosen by a relatively large percentage of dictators. This is repeated by the later meta-analysis by Umer et al. (2022), along with unearned endowments resulting in higher giving in comparison to

earned endowments. The same is found by Brañas-Garza (2007): By stating at the end instructions “Note that he relies on you” (meaning the recipient), the dictators were spurred on to contribute significantly more.

When the recipient has earned the pot and the dictator has the option to take, the dictator takes less, although this result is only weakly significant in the meta-regression by Engel (2011). Making the game non-zero sum, by giving the recipient a multiplier on the share they give away results in a significant positive effect, with larger multipliers resulting in significantly more willingness to give in the meta-regression. Multiple recipients have a positive significant effect on the willingness to give to each in the meta-regression. This is especially significant, as it implies that a dictator facing two recipients should be willing to lose over twice as much money as a dictator facing one. Both the dictator having to earn the pot and the recipient also receiving an endowment have a significant negative effect on the dictator’s willingness to give in the respective meta-regressions.

In terms of framing, presenting the dictator game in a different frame in general does not have a significant effect. As frames are very varied, their effects should be as well. This is why they were not involved in Engel’s study: Each framing may not have enough experiments done with it to make it worthwhile. Framings used in studies include implicit religious priming (Gomes and McCullough 2015), presenting the game in the context of buying and selling a commodity (Hoffman et al. 1994), or presenting the game in the workplace (Carpenter et al. 2005). It should be noted that there are implicit framing effects in many studies: Children participating at school could result in different results than at another location, and the context of participating at a university could have different effects for students in comparison to other adults. The only exception that was included in the meta-analysis by Engel (2011) is handling physical notes or coins. Handling physical currency had a positive and significant effect on the willingness to give in the meta-regression. As in the case of giving to a deserving recipient, the distribution was significantly different: Dictators were less likely to give nothing, and approximately three times as likely to give everything, that is approximately 15% doing so.

The results on social distance are counter-intuitive: Being closer to the recipient had a significant negative effect on the willingness to give in the meta-regression. However, Engel rejects this result based on the different degrees of closeness being

equal, except the completely anonymous one. They deduce that the result is caused by a statistical artefact, caused by other factors inducing willingness to share. They state that this makes multivariate analysis preferable. This should also be considered as a limitation for the dictator game in general and making very broad meta-analyses in specific: Different experiments find very different willingness to give even among those exposed to the same treatments. Then some characteristic or treatment being more or less represented in subject pools with different willingness to give can result in skewed results. This is especially relevant in the final category, demographics, which are often collected even in experiments where they are not the focus.

In the demographic category in Engel (2011), the first result is especially significant: The effect of the dictator being a student is relatively large, negative and significant in the meta-regression. The distribution of the willingness to give is also significantly different, with the mode of non-students being a 50/50 split while for the students it is the dictator keeping the whole pie. Similarly, non-students are approximately 4 times more likely to give the recipient all of the funds. As Engel notes, student experiments appear to underestimate the magnitude of the deviation from expected, “rational” economic behavior of the *Homo economicus*.

Engel (2011) notes that in most papers gender is not reported, but women give significantly more in the meta-regression of 12 studies that gender was collected. The sample of studies is small, and luckily there are other meta-studies on this specific question, for example Bilén et al. (2021). In their data women are generally slightly more generous as the dictator (some 4%), but this effect is much larger when the recipient is a more deserving target, a charity (10%). Engel (2021) notes that they find no significant race effect, but that the OLS option cannot be used due to a lack in distributions. The recipient being a woman had a significant and large effect on the willingness to give in the sample of 12 studies, with this explaining 73,2 percent of the variance. Additionally, controlling for the recipient gender, the effect of women giving more was no longer significant.

As an example, Chowdhury et al. (2017) experimented with gender differences in relation to two treatments with 35 men and 35 women as dictators for each: One where the dictator receives an endowment (the standard or “giving” game) and one where the recipient receives an endowment (the “taking” game). No significant difference was found in the giving game, with both genders giving slightly over 21% of

the endowment to the recipient. While the recipient was left with approximately the same percentage overall in the taking game, women gave significantly more than in the giving game (32,6%), while men gave significantly less (10,0%). Thus, women left more than three times as much as men for the recipient in the taking game, raising further questions of differences between genders and also between framings, possible explanations for this gap in theoretical terms and its possible significance. This could also explain why framing the game as a taking or giving game was not found to be significant by Dreber et al. (2013). On the other hand, Korenok et al. (2014) find that taking is not equivalent to giving; as the amount needed to be taken to achieve a given payoff increases in their multiple different initial allocations, the payoff to the recipient increases. Thus, the negative effect of taking seems to be larger than the positive effect of giving, although one could also explain this difference as caused by the breaking of norms required to take from the recipient's endowment.

In measuring the differences between different cultures, split into "Western", "developing countries" and "primal societies", the meta-study by Engel (2011) found significantly that "The more a society is primal, the more dictators are willing to share" in the meta-regression. Treating the non-Western divisions as categorical variables caused the difference between Western and developing countries to disappear, "while the difference between developed and primal societies is all the stronger" in the meta-regression. The distributions are also significantly different: The Western distribution follows the earlier mentioned distributions with close to 40% giving nothing, an equal split being another peak at around 15% and approximately 5% giving all. In developing countries, the distribution between giving 0-50% is rather even, while giving over 50% is extremely rare. In "primal societies", an even split is the peak, which descends towards 0, while giving more than half is very uncommon. It is worth questioning, whether these categories are meaningful or overlap some other factors. Characterizing certain societies as "primal" may also evoke preconceptions which are not accurate or appropriate in research and the category should be named differently.

The study by Engel (2011) also finds that age has a substantial and significant positive effect in the meta-regression. Splitting the data into four categories (children, students, middle age and old age) results in significantly different distributions: Students give as described earlier, children are very unlikely to give more than half,

but are more evenly split between half and none, middle-aged are much more likely to split evenly, while the elderly are much more likely to give all. It should be noted that vastly more student-aged individuals have been studied overall, although Carpenter et al. (2005) also find that workers are more willing to give than students. This has the potential to lessen the deviation from selfishness in most studies, perhaps resulting in some of the outliers where giving is very limited or nonexistent. Additionally, behavioral economists would do well to remember that most people are not students, and thus the population at large may not behave like them and keep this in mind when evaluating implications at the societal level.

As univariate analysis explained only a minor part of the variance between results, multivariate analysis was undertaken by Engel (2011). The multivariate meta-regression explains almost half the variance, but the results are more complex. Quite a few of the prior treatments are insignificant, and out of the significant effects, old age is the most positively impactful, followed by having multiple recipients, the recipients having earned the pot, the recipient being more deserving and the efficiency being higher for the recipient. Meanwhile, the dictator having earned the pie is the most negative, followed the recipient receiving a separate endowment, the dictator being a child, the dictator being a student, the decision being concealed from the experimenter, the game being repeated, action space being limited, the dictator being a group and the least negative, being socially closer.

The individual OLS differs in that using real money is now significant and positive, as is the dictator being identified. Efficiency for the recipient is no longer significant, and neither is a constrained action space, a repeating game, the concealment of the dictator's decision from the experimenter, the degree of social proximity and the recipient receiving a separate endowment. Using three different logit models, a truncated regression and a tobit model provides yet again different significant effects, sometimes with outliers, where depending on model, a significant effect may be positive or negative.

Engel (2011) concludes that the share of dictators who choose to give is remarkably similar to that found by Kahneman's original experiments. Engel makes the point that most interest has been concentrated in the 63,89% that contribute, and not the fact that 36,11% do not. Furthermore, those who contribute give an average of 42,64% (resulting in a mean giving of approximately 28%) of the endowment, "exploiting the

opportunity to their advantage”. Dictator games thus display human heterogeneity, and that behavior depends on both internal and situational factors. Some reservation should be put on the effects estimated; as can be seen later in the more detailed view of individual studies, framing, sampling and the incentives of experiment design can affect each other, and thus the results, especially if certain estimated factors have coincided with differences in experimental design that has not been accounted for.

2.3 Development, charitability and fairness

In reference to the Engel (2011) results on “primal” societies, Cochard et al (2021) performed meta-analyses on the ultimatum game and the dictator game to try to analyze social preferences across different populations. The purpose was to determine how much offers differed to determine how much of an impact fairness concerns had on decision makers, because in the ultimatum game there are other, additional concerns due to the other party and the size of the dictator game offer in comparison to the ultimatum game offer can be used to estimate how much the fairness hypothesis explains of the results. In the fairness hypothesis, the initial decision maker’s perception of fairness decides the offer in both the ultimatum game and the dictator game, resulting in similar offers. Thus, the experiment could then compare the results between different regions to determine how economic development affects the applicability of the fairness hypothesis, where both games would reach the same results.

For comparability the study only included studies utilizing the standard form. These excluded studies in which the set of actions is too restricted, no money is at stake (real or fake), the recipient is a charity, the dictator earned the position, anonymity is not guaranteed, there is some sort of competition, a prior different game was played, more than two subjects or computerized subjects were implied, or a non-neutral framing was used. Thus, for a sizable number of studies, only the control treatment was included. Cochard et al. (2021) then recorded (only factors relevant to the dictator game are included) the average offer and its standard error, the year and economic indicators of development (three macroeconomic: GDP per capita, HDI and the poverty rate, as well as the ease of doing business index and bank account penetration rate) and finally a third category including: the size of the endowment, whether participants were economics students, whether the experiment was run in a

laboratory, whether a double blind procedure was used, whether the game was repeated and a dummy on whether the decision for what to do in the different roles is done before the roles are given out. The medians for all economic metrics are rather high, which is not unexpected as conducting experiments in developed nations is easier and there are some funding concerns in less developed nations.

Cochard et al. (2021) finds that in general, the mean offer for the dictator game is rather similar to Engel's findings. Concerning the development level, they also find, like Engel (2011), that higher development is correlated with lower dictator game offers. Interestingly, this difference does not apply to the ultimatum game, which leads to the fairness hypothesis not being rejected for lower income countries but being rejected for higher income ones. They note that standards of fairness vary between cultures and note that their bold assumptions of ultimatum game offers working as proxies for fairness may not be accurate, as factors such as different guesses of the minimum acceptable offers and risk aversion relating to offers not being accepted, may be different between countries.

2.4 Team versus individual giving

The dictator game has also been used to compare teams to individual decision makers to teams of decision makers. Although Cason and Mui (1998) found that groups behave more selflessly, Luhan et al. (2009) raise the point that other studies have mostly found that teams behave less altruistically than individuals. Their own experiment also found this and specifically that only the level of altruism of the least altruistic member has a significant effect on team altruism as well as that after team interaction, teams become more selfish. There are some differences in methodology (3 vs 2 team person teams, possible order effects in Luhan et al. (2009)), but it seems more likely that the findings of Cason and Mui (1998) are outliers. Luhan et al. (2009) also further explore the messages of the teams exchanged before the decision, finding mentions of acting rationally and arguments calling from participants' economic expertise. They mention that in Cason and Mui (1998) team members were more identified. One could perhaps also raise the argument, that a team diminishes personal responsibility, where support for lower charity could be found in for example Hamman et al. (2010). Additionally, it may be worth considering whether the humanization of communicating with another party makes more altruistic team

members more willing to bend to the desires of more selfish members due to accounting more for their teammates' utility than that of an anonymous and unknown other party.

2.5 Willingness to participate and delegation

Lazear et al. (2011) try to differentiate between three types of decision makers in dictator games: "willing sharers", who share a non-zero amount and seek to enter the situations to do so, "reluctant sharers", who try to avoid the sharing situation, but do share a positive amount when in it, and "non-sharers", who do not share in sharing situations. They assume that "non-sharers" are indifferent between entering and not, when the endowment and the payoff for not entering are the same. They then tested several predictions concerning the behavior of these three groups using both repeated and non-repeated dictator games, along with different sortings, where the endowment is the same, smaller than or larger than the offered alternative of not playing the game. The study found that the connection between amount given and willingness to pay to avoid the game is not linear in general, but is linearly positive among reluctant sharers. These results assume some consistency of behavior, as other studies have remarked that behavior in dictator games is remarkably inconsistent. Even so, the presence of willing sharers casts doubts on any perfectly self-interested models.

Dana et al. (2006) experiment further into willingness to exit the dictator game, with exits that are costly. They found that approximately one third of participants were willing to exit a 10-dollar dictator game and receive 9 dollars instead. The other party would not get to know that a dictator game was to be played at all, thus making the exit quiet (and avoiding any disappointment). However, when the possible payments from the dictator game were masked by another experiment (and thus the second party could not know they were or were not part of the game), almost no dictators exited. No correlation between amount given and willingness to exit was found. The limitations of the study (a small sample of 40 students from a single university that take business classes for the first experiment, and 21 for the replication treatment as well as 24 for the fully private treatment in the second experiment) means that the results do not merit sweeping conclusions.

To further examine the willingness to quit, Broberg et al. (2007) utilized an auction model (first having the dictators give a reservation price and then drawing a random payment to compare whether the dictators exit) to estimate the mean exit price (the amount they are willing to take instead of playing and below the endowment) as well as the proportion willing to quietly quit with an exit price below the endowment using 125 prospective dictators. 64% of prospective dictators were willing to exit at a payment below the endowment, with the mean payment accepted among the exiting dictators being 82% of the endowment. The mean donation in the dictator game was 27% of the endowment (however, those with a lower reservation price were likelier to be selected out in the exit phase). Broberg et al. (2007) also found that those giving more in the dictator game were more prone to lower reservation prices, although this was not significant. The study seems to confirm the willingness to exit of a sizable proportion of dictator game participants, although the result is even more striking, with the divisions of dictator types proposed by Dana et al. (2006) cast in significant doubt: Only 36% fit the behavior models of the proposed “willing sharers” and “non-sharers”. The auction model used may have impacted this, as it allowed the prospective dictators to set their reservation price first, instead of being offered a pre-determined percentage. Broberg et al. (2007) also recruited their participants from students in the Royal Institute of Technology in Stockholm, Sweden, while Dana et al. (2006) used students from Carnegie Mellon University in Pittsburgh, the United States.

Another way to investigate willingness to participate is to give the dictators the option to delegate. In the experiments conducted by Hamman et al. (2010), agents were paid a flat fee for getting chosen by a principal dictator from among three candidates. There were three experiments: One without pre-knowledge of agent behavior and the possibility to decide the allocation themselves for certain dictators, one with ex-ante signaling and one where competition among agents is addressed. The principals mostly looked to pick agents more selfish than themselves, thus reducing the amount shared. This was reinforced by repeating rounds of the experiment, where the most selfish agents got repeatedly picked, with likelihood to continue with the same agent decreasing over increased giving, in all experiments. They also found that agents in the final round gave significantly more, when there was no possible consequence for giving any amount, as agents were rewarded purely for being chosen (prior to the round).

In the first experiment by Hamman et al. (2010) it took some rounds for agents to allocate less than the baseline, with the first four rounds being similar in amount given to it. The next four rounds saw agents allocating less than the baseline on average. Principals were very likely to retain the same agent if they donated very little (90% for donations under \$1 out of a \$10 endowment), with the likelihood to retain decreasing with increasing donations. In the final four rounds, a significant number of delegators were given the option to make the split themselves: some 60% chose to do so. Those who did shared very little, like agents, which contrasts to the findings of Lazear et al. (2011) that a significant percentage of dictators would look to give a significant amount and would not avoid the dictator game. In Hamman et al. (2010) experiment one, dictators chose to share less than \$1 out of \$10 93% of the time. This could be explained by desensitization to low amounts given, as even when not given the option to delegate, dictators who had delegated to agents in previous rounds gave significantly less when forced back into the dictator role. This could also be attributed to an anchoring effect from the prior decisions. As mentioned above, agents shared even more generously in the last round than the baseline, but not significantly so.

In the second experiment by Hamman et al. (2010), the prospective agents were required to state their intentions to give before agents were chosen by the dictators, who could then use this information when choosing agents. The only way dictators had to enforce these announcements was choosing another agent in subsequent rounds. In contrast to the first experiment, amounts given through agents were less than the baseline from the first rounds, with only the last round forming the exception (Any information or announcement in the last round is truly “cheap talk” in the game theory sense, as there is no way to punish it). Very reliably, principals chose agents that had the strongest reputation for giving the least.

The final experiment by Hamman et al. (2010) eliminated competition by assigning each principal one agent and giving the principal the ability to act as a dictator after seeing the agent’s announcement of intent. Agents were paid based on how often they were selected. Results were much as in experiment 2, with the amount shared by agents decreasing over time and likelihood to delegate also increasing over time.

In total, principals who delegated gave less in all experiments, while also answering to the questionnaire that they felt that they had acted fairly more often than those who acted as dictators themselves. Hamman et al. (2010) raise the argument that

delegation can allow for responsibility to diffuse into nothing by both parties avoiding responsibility for the less equitable result. Yet, agents display less than perfect loyalty when they are beyond punishment in the final round. Using agents may not simply act to improve efficiency but can instead allow the principal to wash their hands of immoral behavior that benefits them. The results of Hamman et al. (2010) might be used to provide support to the view that dictator game-like giving is less likely to be observed in the real world due to the dictator being able to avoid the situation by delegation, especially among those who have the resources to delegate their financial management. The result of the experiment could then be tested: One would expect there to be a significant difference in charitable giving between those who delegate their asset planning to financial advisors and those who take a more personal role, among similarly wealthy groups.

Following up on Hamman et al. (2010), Ezquerra and Kujal (2020) report the findings of their modified version of the experiment. Their focus is on the difference between the baseline dictator game, endogenous delegation (the principal may make the decision themselves or delegate) with treatments where the agents knew/did not know that the dictator could choose themselves (the results were very similar), as well as compulsory delegation. Each delegation treatment featured two agents. Ezquerra and Kujal (2020) found that the amount donated between the baseline and endogenous delegation treatments did not differ significantly, but that the amount donated in the compulsory delegation treatment was significantly lower than in the other two. Those who chose not to delegate gave less on average than the baseline, as Lazear et al. (2011) also found concerning the willingness to avoid the dictator game. But they also gave less than the agents in compulsory delegation, while agents in endogenous delegation gave similar amounts to dictators in the base dictator game.

To explain this Ezquerra and Kujal (2020) also found that the reaction to switch agents when they allocate higher amounts exists in endogenous delegation, but it is small, while in compulsory delegation it is much larger. This would imply that any results in Hamman et al. (2010) concerning the ability for dictators to step in and split the pot themselves is influenced by the conditioning of the earlier rounds and thus not representative of a situation where a dictator could always choose to not delegate. However, Ezquerra and Kujal (2020) had a rather limited sample (20 principals and recipients in the baseline, 37/22/37 and 22/12/22

Principals/Agents/Recipients in the endogenous delegation treatment and 17/10/17 in compulsory delegation), and their baseline was also considerably higher than any of the Hamman et al. (2010) baselines, which may have influenced the results.

Additionally, Hamman et al.'s sample was some 10 years older and from the Eastern US (Carnegie Mellon University and the University of Pittsburgh), while Ezquerra and Kujal (2020) recruited their participants from London (Middlesex University London).

2.6 Multi-step games

A somewhat similar case, where adding an additional step causes dictators to behave more selfishly, is presented by Dana et al. (2007). They find that in their baseline treatment, where dictators choose between option A (\$6 for the self and \$1 for the recipient) and option B (\$5/\$5), most dictators (74%) choose option B. However, when the payoffs to the other party were hidden (Thus the dictator did not know whether the options were as in the baseline, or alternatively A as \$6/\$5 and B as \$5/\$1) and the dictator was given the option to see the payoffs, only slightly more than half chose to see them, resulting in a much more selfish allocation (63% chose A in the baseline payoffs, 81% in the alternate), even with the same payoffs. The high proportion in the alternate is not a surprise, as those who chose to reveal would mostly choose the outcome with higher payouts for both parties.

In the third treatment there were two dictators with just one choosing the fair option (\$5, \$5, \$5) resulting in it being implemented instead of the selfish option (\$6, \$6, \$1). In this case, 65% chose the unequal option, with it not being clear that this would adversely affect the recipient, indicating according to Dana et al. (2007) that a diffusion of responsibility could increase selfishness, as possibly also indicated by experiments on groups of dictators and the delegation experiments mentioned above. In the fourth treatment, dictators were given up to ten seconds to make a choice before being cutoff and the choice being made randomly by the computer. With a normalized distribution of the cutoff and each dictator having at least 4 seconds to make their choice, 24% were cutoff. Among those not cutoff, a slight majority (55%) chose the selfish choice of (\$6/\$1). This is significantly different from the baseline.

Dana et al. (2007) draw the conclusion that dictators do not necessarily value a more fair outcome, but instead feel compelled to give by factors like guilt. Then, relaxing

transparency allowed the “moral wiggle room” to behave selfishly, while maintaining the illusion of fairness. The percentage implementing the fair outcome approximately halves, despite the opportunity to implement it anyway without additional cost. Some caveats must be kept in mind: The numbers in the experiment were low (19, 32, 20 and 29) which can result in less reliable results, with only very few mistakes or slightly different sampling shifting the results. If a sample is 20, the mere change of action from four participants is enough to shift a share from 40% to 60%. A significant percentage still acted in the more equitable manner, like in Dana et al. (2006). However, it may be reasonable to approach “fairness” as a concept more through rules and guidelines and less through a very consistent and direct approach to implement the most equitable outcome, as mentioned by Dana et al. (2007). Norms should then be emphasized in fairness models, rather than calculations on outcomes (humans employ heuristics rather than perfect calculation when evaluating options).

2.7 Earned endowment

Cherry et al. (2002) set out to prove that giving can essentially be eliminated by adding in the dictator earning the endowment prior to making the decision to split it, in addition to the standard anonymity assumptions. They used three treatments, with an additional division between high stakes (\$40) and low stakes (\$10) versions of the experiment due to the potential to earn either amount in the earnings stage: A baseline standard dictator game, a treatment where the dictator was asked to split an amount (\$10 or \$40 depending on performance) earned by them in an earlier quiz (the dictators were informed that the other party had not had a chance to earn anything), and the earned treatment with a double-blind procedure, where the administrator of the experiment had no chance to observe the choices.

In the baseline dictator game, dictators decided to share zero only 19% and 15% of the time in the low-stakes and high-stakes variants respectively. On the other hand, in the earnings treatments the percentages were 70 and 79, and in the double-blind earnings treatment 95 and 97! Cherry et al. (2002) present this as closing the gap between results obtained in dictator game experiments and claim that with these treatments “altruism was the exception and self-interest was the rule” They also infer that strategic concerns and not fairness appears to be the reason for giving, but it is

not necessarily very clear. Fairness cannot so easily be dismissed as a factor, however. Windfall-wealth may promote giving (otherwise we would expect to see a massive amount of spontaneous and anonymous giving in the real world), but saying that fairness cannot be the reason, why earned wealth is not shared could also be wrong. People feel strongly that they deserve their earned wealth, and one would suspect that were the roles reversed so the dictator could take from another's earned wealth, they would take significantly less than 100%. The right to the fruits of one's labor would likely be described as "fair". And even if windfalls of funds are rare in day-to-day life (but do exist!), situations where all participants have a somewhat equal justification to funds (acquired because of shared effort or communal funds, for example) are not.

2.8 Limited decision space and a pool of funds across multiple games

Bolton et al. (1998) set out to test several hypotheses. These arose from the tendency for some amount of giving in dictator games, but also the result that when limited in options, dictators tend to choose the most selfish option. The first is the "I'm-no-saint hypothesis", which states:

If the dictator in the dictator game is restricted to two division choices, one being the equal split and the other being an unequal split favoring the dictator, then the percentage choosing the equal split will be equal to the percentage that would choose the equal split in the unrestricted game.

This would explain the low amount selecting the equal split by decision-makers erring on the side of their own benefit when forced to choose between two options, where they must choose between giving more or less than they would prefer.

The second hypothesis to be tested concerned multiple rounds of games, named the "rational giving hypothesis", it states that:

Let k be a fixed dollar amount, and fix dictator and recipient roles, n participants in each role. Suppose that each dictator (recipient) plays n dictator games with n distinct partners, each game having a pie of value k/n . The proportion of the pie that a dictator leaves for each recipient diminishes as n increases.

The idea is that repeated games results in less giving in comparison to a single game with an endowment equal to the smaller endowment games combined. The secondary hypothesis is that the dictator would give the same amount to each recipient. While inspiration for these hypotheses came from comparing dictator and impunity games

(in the impunity game, the recipient may walk away from their share, but the dictator is unaffected), where the impunity game has a limited decision set with two distributions, Bolton et al. (1998) set up their experiment using dictator games.

The experiment had combinations of treatments in a 2x2 grid, with dictators facing any combination of 10 or 1 consecutive games and 2 or 6 division options per game, with a total endowment of \$10 for each treatment (so the 10 games were worth 1/10 as much as the singular game). The two option games featured perfectly even and perfectly selfish choices, while the 6 option games featured additionally each option in between, in 10% increments. Bolton et al. (1998) then made two further treatments to pursue additional questions raised: an anonymity treatment, where the dictator placed a chosen amount of money from 1 \$5 bill and 5 \$1 bills into an unmarked box to be then distributed to receiving parties without informing the experimenter. This allowed additional anonymity to the game, but its result is diminished by the substantial difference in using real money, as this leads to higher giving on average (Engel 2011). A second additional treatment tested for kindness only featured \$5/\$5 and \$6/\$4 splits, or an even or slightly dictator-favored split. Each treatment featured between 24 and 33 individual dictators.

The results are mostly supportive of the “I’m no saint hypothesis”, with the percentage giving half falling when the options are reduced in the one game treatments. The percentage giving 0 significantly rises as the only other option left, as expected. On the other hand, Bolton et al. (1998) finds no support to the idea that more rounds lead to less total giving in comparison to a higher endowment shared once. Additionally, dictators do not treat the recipient in a consistent manner, with just over half appearing to behave in a random manner with their gifts varying in size. No support for the idea that anonymity reduces giving was found, but as mentioned above, this is likely not very useful due to the different way the treatment was conducted (handling real money vs payment coming after).

2.9 Giving and taking treatments

Bardsley (2008) compared giving and taking across three different experiments. Double anonymity was ensured by the experimenter not being present while the decisions were made, and then assistants, who were unaware of the experimental design and unaware of the choices, making payments. In the first experiment, after

giving both dictators and recipients £6, a treatment of the dictator game with doubled transfers, with an endowment of £6 and allowing a maximum giving of £4, is compared to one with a take option of up to £2, with taking costing twice that from the other party. Only 6/26 gave in the first treatment, and 1/29 in the second. In the second experiment, a more standard dictator game was conducted instead, with both parties receiving a £4 show-up fee and the dictator receiving a £7 endowment. All transfers were at a 1:1 rate and in the second treatment the dictator could take up to £2 from the recipient's show-up fee. 23/33 gave in the standard treatment, and 15/32 in the taking treatment. Unlike the first experiment, the difference was only significant at the 10% level. In the third experiment, dictators were informed that they had received an endowment of £10 and their partner of £5. Transfers remained at 1:1, with the dictator allowed to give up to £3 in the first treatment and only taking in the second treatment, up to £3. 17/31 chose to give, and 24/29 chose to take in the treatments.

These results are the basis for Bardsley's (2008) statistical analysis of the theories, but the results are quite aberrant. In higher efficiency scenarios, like experiment 1, Engel (2011) notes that higher efficiency correlates to higher giving, and vice versa. This raises questions of perhaps the experiment being too confusing or the result otherwise being an outlier due to sampling or some other reason, especially, with "standard" dictators in experiment 3 giving substantially more. With the lowered significance level of experiment two, and the possibility of anchoring affecting the results and especially possible influences from the non-standard design (endowments to both parties prior being known to the dictator, varied decision spaces of less than the full endowment and differently sized endowments), the rejection of social preference theories as a significant factor behind decision-making is not very convincing.

2.10 Compatibility with GARP and rationalizable preferences

In contrast to other studies, which find substantial heterogeneity and inconsistency, Andreoni and Miller (2002) test behavior for consistency and compatibility with utility maximization by testing whether behavior fits with the Generalized Axiom of Revealed Preferences (GARP), which is necessary and sufficient for well-behaved preferences. To estimate consistency, subjects were given a "menu of choices" with

different multipliers for the payoffs to the other party (they received multiplier times the amount given up by the dictator). The dictator was then asked to allocate tokens for each of these budgets, with 8 budgets in the first four sessions of the experiment, and 11 budgets in the last round. The relative price of giving (or the amount given up divide by the amount received by the other party) varied from 0.25 to 4, and the endowment was either 40, 60, 75, 80 or 100 tokens. Unusually, “subjects were encouraged to use a calculator to check their decisions”, which may explain heightened consistency and rationality. One of the budgets would then be chosen at random, carried out, and the dictators were paid out according to their tokens.

Additionally, in the last session dictators made five additional decisions on pre-allocated tokens were presented in random order. Decision-makers were presented with divisions of (10, 130), (20, 110), (50, 50), (110, 20) and (130, 10) between themselves and another party, respectively. They were instructed to assign a worth between 0 and 10 cents for each division, and then one would be carried out randomly. This was done to test for jealousy at the extremes: In each case, both parties combined would earn the most by the decision-maker always choosing 10 for themselves (or the recipient), or in other words the most unequal option.

The average given in those budget sets with a relative price of giving of one (the regular dictator game) was approximately 20%, a not unusual figure. In addition to the generalized axiom of revealed preferences (GARP), Andreoni and Miller (2002) also tested for violations of the weak (WARP) and strong axioms (SARP) (definitions quoted under the theory section), with 18 of the 176 violating one or more of the axioms, but for 14 of these the violations were relatively minor, and would only require minor budget relaxations to avoid. Only 3 subjects could not have their behavior rationalized in the form of a quasiconcave utility function. Meanwhile, simulating random allocations resulted in 78,1% of subjects violating all three axioms when 8 budgets were used as in the first four sessions, rising to 94,7% in the scenario with eleven budgets to set.

Andreoni and Miller (2002) posit that the rationalizable behaviors fit with the forms of three utility functions, with those fits divided between those who fit the strong form of those utility functions, and those who fit the weak. Of the strong fits, 22,7% of all participants behaved perfectly selfishly, where their utility function depends only on their own payoff. 14,2% of participants divided the payoffs evenly, which imply

Leontief preferences, where the lower payoff dictates the utility level. Finally, 6,2% strongly fit the perfect substitutes utility function, where the total payoff indicated their utility level, thus allocating tokens to the party who received a more efficient payoff. Thus 43% fit these strong preferences, fitting these utility functions perfectly.

Meanwhile, the rest are divided between which of these they fit best. These preferences do not result in those exact utility functions but are rationalizable.

Among these, the proportion of weakly selfish and Leontief utility functions are similar as in the strong, while the percentage resembling perfect substitutes doubles to be the same as Leontief. In the final sessions, Andreoni and Miller (2002) explain that some 23% behaved non-monotonously, and this jealousy effect was rather severe with the last two unequal divisions, leading to valuation going down by more than half. They thus note that a minority of participants make non-monotonic choices motivated by jealousy. It keeps bearing in mind that in comparison to the 170 participants for the main part of the experiment, this only concerned 35 participants.

The main takeaway is then that while dictators are heterogenous, they seem to mostly behave in ways that are rationalizable with utility functions. Xue and Ohtsubo (2020) replicated the experiment of Andreoni and Miller (2002) in Japan, using two different stake sizes across two studies: One where the stake is approximately one fifth of Andreoni and Miller's (2002), and one where it is similar in nominal terms, thus likely somewhat different due to inflation and different purchasing power. The numbers of participants were 78 and 58 respectively, of which 85% and 81%, again respectively, did not violate GARP across the eleven budgets. Xue and Ohtsubo (2020) also found that in their studies, a substantially higher percentage exhibited preferences that fit with altruistic utility functions, in particular 72,7% exhibited Leontief preferences (weak or strong) in the lower stakes study, and 50% in the higher stakes study. Thus, higher stakes did lead to more selfish preferences, but together with perfect substitutes preferences (11,7% and 17,2% respectively), non-selfish preferences were the majority, while not violating GARP.

Jakiela (2013) expanded Andreoni and Miller (2002) by adding a "giving" vs "taking" endowment component to test the consistency of the elasticity of substitution, and thus whether dictators were consistent in their "types" across different framings. Dictators were expected to divide a series of budgets of different prices (relative prices of giving being $1/3$, $1/2$, 1, 2 and 3) for giving between themselves and another,

anonymous party, with whether the endowment was earned by themselves or their partner changing between rounds of budgets. (Both players played dictator for each round, one as the “giver” and one as the “taker”, but only one was selected for payouts.) The pairings of each round were random, with the experimenter not being able to identify the decisions of any individual player. The budgets varied from 16 to 72 dollars and were represented as tokens. These budgets are on the higher end and as reported by Engel (2011) higher endowments are correlated with lower tendency to give. The experiment was conducted at the University of California, Berkeley, in 2006 using 144 undergraduate students.

The results of comparing giving and taking are not surprising: dictators allocated 12,7% of the total tokens to the other party in the giving rounds, and 19,0% in the taking rounds, which is significant. Dictators allocated equal numbers of tokens (thus reducing efficiency in exchange for equity) more in the taking rounds (9,3% vs 6,3%) and also allocated the budget evenly significantly more often in the taking rounds (13,3% vs 6,3%), but were almost equally likely to allocate 0 to their partner in both roles (59,5% in giving and 59,1% in taking). On individual decisions, Jakiela (2013) finds great heterogeneity like previous studies. As a further complication to Andreoni and Miller (2002), some players were consistent within a role, but these “types” were different! Another finding is that the relative price of sharing increasing the budget spent on the other person did not seem to depend on the role, implying that for the most part, the elasticity of substitution remains constant in aggregate. A linear regression did not lend support to the idea that the effect of “taking” role depends on the price of sharing. Unlike Andreoni and Miller (2002), Jakiela (2013) did not investigate whether participants violated GARP, but also found that the leaning was more towards Leontief preferences than perfect substitutes in those who were not perfectly selfish.

2.11 Interactive and non-interactive games

Further concerning the possible rational player types, Grech and Nax (2020) introduce the idea that dictator games can be divided into two protocols: interactive and non-interactive (In interactive protocols dictators are also recipients, see theory section for details), which have potentially different implications for rational decision makers. This may then result in different behavior in interactive games than in the

standard, non-interactive games, potentially making models designed on the assumptions of non-interactive games unusable in interactive cases. Experimentally, Grech and Nax (2020) set to test first whether the protocols are identical, and if this is rejected, whether rational choice benchmarks can explain the differences between protocols. Regardless, if the protocols are not identical, any model based on the assumptions of the non-interactive protocol is not directly applicable or testable in the interactive protocol. The experiment was identical to Andreoni and Miller (2002), except that there were two treatments: interactive and non-interactive and that budgets were fixed. The experiment was held online, with 206 dictators per protocol. A total of 20 decisions were made by each dictator, with the multiplier of redistribution varying between 0.1 and 2 in 0.1 increments. Each participant received a \$1 fee in addition to earnings based on dividing the 1000 tokens for each decision (with one randomly chosen to be paid out), with the maximum possible bonus payout being \$15.

The basic findings were unsurprising. The mean giving in the interactive protocol was 31,3% (higher than usual) and in the non-interactive 26,5%. Giving zero was only observed in 13,2% and 22,5% of the cases respectively, which is lower than most dictator games, while around 4% in both gave the full endowment (Not necessarily surprising due to possible efficiency-maximizers facing relative cost of giving under 1). Giving fell as the multiplier for the giving increased, as Jakiela (2013) also noted, but the total received by the other party increased due to increasing multiplier. In terms of the distributions or the treatments, they are significantly different at the 95% level in two different tests (Kolmogorov-Smirnov and Mann-Whitney-Wilcoxon) while the variance of the protocols is not significantly different. For individual multipliers of giving, the difference is significant mostly for multipliers under 1. The difference cannot only be explained by rational behavior, as while zero-giving changes significantly (and giving half significantly at a 90% significance level), there is no significant change in giving the full endowment, which would be expected as rational behavior would lead to extremal equilibria in the interactive protocol (see theory section for more specifics). Thus, the hypothesis that the protocols are different solely due to changes in rational incentives is rejected by Grech and Nax (2020).

An example of the use of two different protocols when comparing different groups brought up by Grech and Nax (2020) is Fisman et al. (2015), where the behavior of Yale law students in the dictator game was compared to average US citizens. The finding was that Yale students were more selfish and more efficiency-minded than the average citizen (students were found to be more selfish in the meta-study by Engel (2011) as well), but Yale students played an interactive game, while the average citizens' game was non-interactive. This raised the interest to the experiment mentioned above, where the results were that the interactive protocol was connected to higher giving. To explore this further, Grech et al. (2022) added protocol data to the meta-study by Engel (2011) to attempt to parse differences between the protocols. Following several individual experiments where differences were found, the meta-regression also found significant differences between protocols.

Grech et al. (2022) reports from the regressions on Engel (2011) data (corrected for incorrect standard errors and means) that on average interactive protocols and those "role uncertainty" protocols where all participants submit decisions, but only some are carried out, result in lower giving and more efficient-minded decision-making. This is the expected effect based on theoretical predictions, but opposite to the one found in Grech and Nax (2020). In general, Grech et al. (2022) makes the point that these results are not perfect or complete (there haven't been many studies focused on protocol effects), but that protocols have significant effects and should be studied further. Results from different protocols are not comparable, and differences in protocols may vary in size and direction.

2.12 Children as dictators

An interesting subject group for exploring the motivation behind giving are children. Children are often taught to share and other pro-social behaviors, perhaps implying that pro-social behavior is learned or cultural, and this can be tested using the dictator game. Engel (2011) found that children as a distinct group give on average similar amounts to students, and less than older groups. Compared to students, children were less likely to give zero, but also less likely to give as much as half of the endowment, resulting in a similar average. However, "children" as a category is rather broad, and to see how pervasive altruistic behavior is, looking at very young

children is likely more informative, as older children may be more similar to students, maybe only a few years away in age.

Thus, the Benenson et al. (2007) study on 4-, 6- and 9-year-olds is of special note. In addition to performing a dictator game with very young children, the experiment also compared results from areas of higher socioeconomic status to areas of lower socioeconomic status, discerned based on the percentage receiving free school lunches (a welfare program based on parent income), with less than 5% receiving them in higher socioeconomic status areas and more than 50% in the lower status areas. The lower socioeconomic status schools were also located next to public housing. Instead of money, stickers were used of a type similar to ones used to reward students, but from out of country to enhance value through exclusivity.

To obtain results, an analysis of variance (ANOVA) was performed by Benenson et al. (2007) on the number of stickers given using age, socioeconomic status, sex of the dictator and sex of the recipient as independent variables. Age level and socioeconomic status had significant positive effects, but while the difference between 4- and 9-year-olds was significant (9-year-olds donate significantly more), the differences between 4- and 6-year-olds as well as 6- and 9-year-olds were not. More importantly, the average percentage of stickers donated for each socioeconomic status and age combination varied from approximately 25% to approximately 40%, results similar to other experiments done on students or adults. In all categories, less than half give nothing. Benenson et al. (2007) do another analysis excluding 12 of the total 360 participants (3,3%) that donated all or nearly all stickers; Based on Engel (2011), these are not necessarily uncommon outliers, as Engel (2011) finds that up to 5% donate the full endowment in some experiments.

Based on the pilot testing of Benenson et al. (2007), 4-year-olds seem to be the youngest age at which children could understand the rules of the game. While older children were more generous (as were children from areas of higher socioeconomic status), even the youngest children gave at not-dissimilar rates to those found among adults. Thus, generosity may be partially learned, but seems to be present at very young ages keeping in mind that working with children produces certain difficulties and requires some changes in protocols. However, double-blind anonymity was ensured, and the value of the stickers was ensured by first asking the dictator to pick their 10 favorites among 30 presented. The translation of stickers to value is

imperfect, but the children affirmed that they liked them. Engel (2011) did find that children donate in relatively similar ways to adults so the result is not an exception; The vast majority also donated 5 or less, which would be expected if the motivations are similar to adults.

Other studies on children are varied. Gummerum et al. (2010) found similarly that even 3-year-olds give significantly (median 27,3%) in a similar study using stickers, with 4- and 5-year-olds giving more and the percentage increasing with age. They also found that girls give significantly more. This tendency for children to give more with age is replicated in the overview of Sutter et al. (2019), as is the tendency for girls to give more. Fehr et al. (2008) however, finds that children 3-4-years old are mostly selfish. The setup of the experiment is unfortunately not comparable and would also likely lead to different results compared to the standard dictator game among adults as well: Children were asked to allocate sweets between themselves and another party. The decisions were, with the allocation to the dictator listed first: (1,1) vs (1,0), (1,1) vs (1,2) and (1,1) vs (2,0). While only approximately 8% choose (1,1) over (2,0) among 3-4-year-olds, over 60% chose (1,1) over (1,0). There is simply not enough granularity, and the experiment is not similar enough to conclude that even very young children aren't willing to give to some degree. On the other hand, the tendency for older children to give more finds some support.

2.13 Artificial intelligence models as dictators

As dictator game behavior is rather well known by this point, and as preferences for certain outcomes have been identified, it can also be used to explore the behavior of human-like Large Language Models (LLMs). In fact, the dictator game has been used to compare the behavior of LLMs to human behavior in many recent papers, but most of these have not yet been peer reviewed or published in a publication, thus an overview of such results is premature. As an example of this use, Brookins et al. (2023) found that GPT-3.5 gave away half of the endowment in most cases, with the rest giving mostly 40% or 60% in contrast to much more selfish preferences found in human studies. Additionally, the rapid changes in LLMs can cause any experiments to be out of date in short order. Additionally, LLMs can be given priors to modify their behavior. This was used by Fan et al. (2024) to test for consistency of LLM given certain preferences. They found that the common preference of Self-Interest and

Equality lead to consistent choices for all three LLMs tested (GPT-3, GPT-3.5 and GPT-4), while preferences for maximizing total income and for altruism lead to sporadic (In the case of GPT-3.5 and GPT-4) or numerous (for GPT-3.0) errors in consistency. These results should not be taken as fact, but rather as instructive of the new ways the dictator game can be used to evaluate model behavior. In this way, the dictator game can help illuminate differences between human behavior (or models that simulate such behavior) and models that are made to give the most pleasing answer to the user. If the future contains a vast number of economic choices made by artificial intelligence or machine learning models, knowing how giving they will be is of great economic interest.

3 Theory

Over the last few decades, researchers have put forth many theories that would explain the behavior observed in the dictator game and other games (the ultimatum game, the impunity game and the public goods game among others) as well as tested these hypotheses using experiments, sometimes evaluating behavior using multiple experiments. The conventional selfish theory would indicate that the dictator should always keep the entire endowment to themselves, and if given the option to take they should take the entire endowment. (Similarly, in the ultimatum game the theory would imply that the proposer should offer the lowest possible positive amount, and the receiver should accept any positive amount.) These are the first theories that the first dictator game experiments contradicted, and the root of the search for other theoretical explanations.

It is important to note first that there are many ways in which humans make mistakes or have biases, some of which can be consistently reproduced in a significant part of the populace. As the wide range of giving in the dictator game implies, people are heterogenous. People have a wide variety of motivations, which should be obvious from how results from the dictator game differ widely depending on a myriad factors, like the framing, earned wealth vs unearned, deserving recipients et cetera. Thus any theoretical model should be approached from the perspective that it may only explain some percentage of behavior, and that there may be other, stronger factors that can be introduced to override the behavior predicted. This does not necessarily mean that these less powerful factors do not matter in many situations, but that they can only explain human behavior in a limited way. If a relatively simple theory could explain human behavior comprehensibly, economics would become a rather less contentious and exciting field!

3.1 Giving as a mistake or artefact

Another possible explanation to the classical selfish model, and closest to it, is simply that any giving is the result of a mistake or is an artefact. Cherry et al. (2002) present that one can drive the percentage given very low by using earned wealth and use this to argue that any giving is simply a part of selfish strategic considerations. However, other models might explain this just as well, such as models incorporating fairness as an explanation. The right to one's earned wealth is a strong part of the conception of

fairness or justice at least in western societies. Schurter and Wilson (2009) conclude that justice is the decisive factor by allocating the role of dictator by four different procedures: A quiz, which is fair and merit-based, a die roll, which is fair and merit neutral, Seniority, which is unfair but based on merit and an unannounced, merit neutral and fairness neutral procedure. Those in the merit-based allocation gave significantly less (18% and 24% vs 35% and 34%), while there was no significant difference between fair and unfair. Fairness is a subjective experience and may even induce behaviors that lead to more unequal outcomes.

3.2 Giving explained by psychological biases

Bardsley (2008) suggests, relying on Dana et al. (2007), that the giving is rather more volatile than other theories might suggest, and that theory is not directly applicable to real life situations or field experiments. They keep open the possibility that choices are explained by factors such as norms, values and tastes adding value to the other sides payoff, meaning that decision-makers would simply weigh the options, but complex inputs make generalizing a theory of giving very challenging. They note that dictators should then be willing to also give when the taking option is available, but do not. Thus the “rational” theoretical expectation is shattered, but this seems to correlate rather well with research into the anchoring effect, where prior decisions (even merely done by others) or the limits of the decision space effect the decision. O’Garra and Sisco (2020) find evidence for anchoring of the first kind, while giving the take option would give opportunity for the second kind to happen. Anchoring could help explain some other cases, such as where dictators gave less in Hamman et al. (2010) after several rounds of (delegated) lower giving.

One of the other options that Bardsley (2008) also suggests are that the behavior of giving in the dictator game is a “Hawthorne” effect, where the awareness of observation causes the subjects to modify their behavior. The specifics suggested are that dictators modify their behavior to suit what might be considered acceptable by cues given by the protocols of the experiment. They explain that the dictator game being a “giving experiment” or giving being the only action one can take normalizes giving, and by including taking, it is also presented as acceptable. The most extreme version of the effect warps the experiment into a wholly different game: If the dictator believes acting in a certain manner pleases the experimenter and/or increases their

chances to be selected as a dictator in another round or another experiment, they may be inclined to give to get to play again and thus profit. This possible effect is another reason for double blind experiments, and making sure participants know this. In some ways this is like the distortion caused by interaction (see below). Secondly, Bardsley (2008) considers stochastic models with utility measurement errors or random preferences. The first is discarded due to the completely transparent payoffs, while the second makes no difference to the result. Thus, Bardsley based on these three concludes there is no evidence for context-free pro-social behavior.

There are some weaknesses to this line of reasoning: One, there is the possibility of other models to explain behavior that have not been considered (presented later in this section), psychological factors affecting the amount given does not necessarily mean that giving is not pro-social, but merely that it depends on context, and third that choices are always made in some context, regardless of the procedures of the experiment. A ship losing its heading and ending up in the wrong port does not imply that it did not aim to get to another, specific target. It is possible that dictators do have preferences, but the inconsistencies and biases of the human mind also affect the results.

Another hypothesis from existing experiments is that dictators would rather avoid giving situations to selfishly retain funds, but cave to pressure somewhat when in those situations. The hypothesis causes some thorny questions on what constitutes altruism. It is hard to say that conditional givers are not pro-social at all, as they still choose to give (even anonymously), when put into the situation to make the choice. However, from the economic perspective exact determination may not need to be made if the behavior is: Some significant percentage of people act pro-socially when presented the situation but are willing to pay to avoid such situations. This could be explained as a hybrid theory, where giving is based on avoiding disappointment, but it is added to the baseline selfishness of maximizing own payoffs.

Bolton et al. (1998) presented the “I’m-no-saint hypothesis” which seeks to codify and explain behavior in dictator games with only two options: a more equitable one, and a more selfish one. They posit and find some support that when forced to choose between a more selfish and a more giving option compared to their preferred level of giving, dictators choose the more selfish option. They hypothesize that this may be an aspect of the rules-based behavior of dictators (instead of acts of kindness) and put

emphasis on the effect on the context of the game on results. They also find support for the rather common view that dictators are relatively heterogenous, although the fact that their experiments were not repeated on the same dictators makes it possible that differences in behavior was caused by dictator heterogeneity.

3.3 Giving explained by rational preferences

If the giving is then not considered an artifact or caused by mistakes (which the rather amount of dictator game research would imply), how could one model it rationally? Andreoni and Miller (2002) set out to answer whether individual decision-makers behave consistently and rationally across several dictator games with different multiplier for giving and found that most individual dictators behave in a manner that is consistent with the axioms of revealed preferences. The definitions Andreoni and Miller (2002) used are:

Let A, B, \dots, Z be distinct bundles of alternatives, each lying on a linear budget constraint. Then define two concepts (see Varian (1993)):

Directly Revealed Preferred: A is *directly revealed preferred* to B if B was in the choice set when A was chosen.

Indirectly Revealed Preferred: If A is directly revealed preferred to B , B is directly revealed preferred to C, \dots to Y , and Y is directly revealed preferred to Z , then A is indirectly revealed preferred to Z .

The classic revealed preference axioms are due to Samuelson (1938) and Houthakker (1950):

Weak Axiom of Revealed Preference (WARP): *If A is directly revealed preferred to B , then B is not directly revealed preferred to A .*

Strong Axiom of Revealed Preference (SARP): *If A is indirectly revealed preferred to B , then B is not directly revealed preferred to A .*

WARP is necessary and SARP is both necessary and sufficient for the existence of strictly convex preferences that could have produced the data. Varian (1982), in applying the theorems of Afriat (1967), generalized the theory to allow indifference curves that are not strictly convex:

Generalized Axiom of Revealed Preference (GARP): *If A is indirectly revealed preferred to B , then B is not strictly directly revealed preferred to A , that is, A is not strictly within the budget set when B is chosen.*

Note that if choices violate WARP they must also violate SARP, and if they violate GARP then they must also violate SARP, but the opposite is not true. As Varian shows, satisfying GARP is both a necessary and sufficient

condition for the existence of well-behaved preferences, given linear budget constraints.

The choices would then fit with one (or several) possible other-regarding (or pointedly not other-regarding in the selfish case) types of utility functions. These are the Selfish, Leontief and Perfect Substitutes types. Thus decision-makers could be both heterogenous, but also internally consistent and rational, exhibiting a certain set of preferences. This makes explaining the giving as an artifact, a mistake or because of confusion unlikely. One would then need to explain how the great differences found by altering the setup or characterization are caused. These changes would have to effect the payoffs to result in some settings close to 0 giving. These preference-based theories are also not necessarily compatible with the “rules constraining selfishness” -type explanations presented further below.

The main differences between these types could also be expressed in terms of price elasticities, where those who are inelastic favor equity and those who are more elastic in reference to price changes favoring efficiency. In Jakiela (2013) the perfectly efficiency favoring are termed utilitarians, as their preferences maximize total payoffs as is the goal of utilitarianism (even if the utility from those payoffs cannot be known in this case), and calls those perfectly favoring equitability Rawlsian after the social welfare function where the welfare level is equal to that of the lowest utility member of society. The most perfectly extreme utility functions may be represented as (given players i and j , budget (π_i, π_j) and constraint $(\pi_i + p\pi_j \leq m)$ $u_i(\pi_i, \pi_j) = \pi_i + \pi_j$ for the utilitarian function, $u_i(\pi_i, \pi_j) = \min\{\pi_i, \pi_j\}$ for the Rawlsian function and $u_i(\pi_i, \pi_j) = \pi_i$ for the selfish function (All only slightly edited from Jakiela (2013)). Of course, most preferences are not as absolute or perfect in fit and for example Andreoni and Miller (2002) finds that 14.2% behaved in a manner consistent with perfectly Rawlsian preferences and 6,2% in a manner consistent with perfectly utilitarian preferences, in addition to the 22,7% behaving perfectly selfishly.

These utility functions can be combined into a single constant elasticity of substitution (CES) utility function that also covers cases where all three factors (efficiency, equity and selfishness) have an influence, if the elasticity of substitution is constant across different experiments and framings. Jakiela (2013) finds that this is the case, and present the function as follows:

$$u_i(\pi_i, \pi_j) = \left[\alpha_i \pi_i^{\rho_i} + (1 - \alpha_i) \pi_j^{\rho_i} \right]^{1/\rho_i}$$

where $\sigma_i = 1/(\rho_i - 1)$ is the elasticity of substitution between π_i and π_j ... As α_i and ρ_i vary, the CES class of utility functions spans a wide range of social preferences. When $\alpha_i = 1$, Player i is purely self-interested and her utility function reduces to $u_i(\pi_i, \pi_j) = \pi_i$; it is impossible to estimate the willingness to tradeoff equity and efficiency among such players. When ρ_i is equal to one, the CES utility function represents generalized utilitarian preferences; as ρ_i ranges from one to negative infinity, the indifference curves of a player CES preferences first approach those of the Cobb-Douglas utility function (as $\rho_i \rightarrow 0$) and then approach the Leontief (right angle) indifference curves of a Rawlsian (as $\rho_i \rightarrow -\infty$). Thus, the range of values of ρ_i from negative infinity to one represents the spectrum from Rawlsianism to utilitarianism. If Player i 's preferences can be represented by a CES utility function with $\alpha_i \in (0, 1)$, then the optimal budget share that she spends on her partner's tokens is

$$s_i^*(p) = \frac{p^{-\rho_i/(1-\rho_i)}}{p^{-\rho_i/(1-\rho_i)} + A_i^{1/(1-\rho_i)}}$$

where $A_i = \alpha_i/(1 - \alpha_i)$. For all ρ_i less than zero, $s_i^*(p)$ is increasing in p — as in the Rawlsian case; the opposite is true for $\rho_i \in (0, 1)$. Those closer to the Rawlsian extreme are ready to pay for equity, spending more on Player j at higher prices.

In examining how to translate lab results into real world behavior, Levitt and List (2007) instead present the incorporation of non-income maximizing preferences into the utility function through adding moral costs in addition to the individual's wealth (which rises in importance as the monetary stakes rise). They divide these into three categories: financial externality on others, set of social norms broken (including laws) and the scrutiny of others. Interestingly, while the financial externalities rise with stakes, breaking norms induces costs based on the strength of that norm, with larger transgressions more costly, and scrutiny increases costs in general, with more scrutiny resulting in more costs. This model could thus incorporate wealth-maximizing behavior in morally neutral situations, such as when choosing investments, and the possible tendency for higher stakes to lessen giving, as the weight of wealth increases in comparison to moral cost. This model could be described as a hybrid model, where both concern for others' welfare and social norms, along with scrutiny, affect the level of giving. Different levels of scrutiny could

also form an explanation for different behavior in real life situations, as by their nature lab experiments are high-scrutiny situations, even with anonymity. This could also explain differences in behavior between games that are different levels of anonymous, where the scrutiny of the experimenter changes based on the level of anonymity.

3.4 Protocol differences and strategic considerations

Setting the view that clear types of utility functions can be simply discovered into question, Grech and Nax (2020) posit that theoretically, different dictator game protocols can lead to different strategic incentives for rational players. They base this on dividing dictator games into two different protocols: Non-interactive, the original protocol where participants are either only dictators or recipients and Interactive, where they are both simultaneously (in either pairs or “loops” of relationships, where they are one person’s dictator and another’s recipient, see image below). They posit that the dictator game analysis on rationally altruistic preferences relies on the non-interdependence (where each participant is in only one role and making the dictator “game” not really a game, but a decision-making experiment where strategy is eliminated from consideration) of the decisions. Alternatively, they argue that when payoffs depend on the actions of another player (the Interactive protocol) the strategic considerations may affect rational players. In fact, the games are only equivalent if the players are unaware of potential strategic considerations, which seems unlikely as players are usually directly told of the potential payoffs (the belief that these games are different is sufficient, even if they are not!). Grech and Nax (2020) then note that this difference in protocols has not been considered in the previous literature, like in Jaskiela (2013) and the meta-study by Engel (2011). It is worth noting, that giving has been observed across both protocols; Indeed, the first dictator games were specifically non-interactive. Thus, it is most relevant when discussing the theoretical models and incentives dictators have and how to test these models.

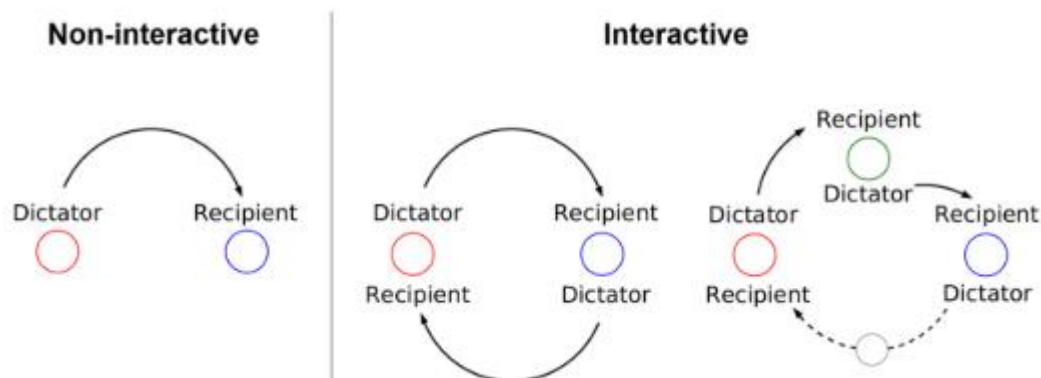


Figure 2. An illustration of the different protocols of dictator games from Grech and Nax (2020), illustrating the difference in non-interactive and interactive games which may lead to strategic considerations in the latter. Grech and Nax (2020)

While optimal decisions to maximize the utility functions presented by Jaskiela (2013) are relatively simple in the non-interactive protocol, Grech and Nax (2020) explain that for those who are not absolute in their preferences (perfectly selfish, perfectly Rawlsian, perfectly utilitarian and perfectly unselfish), their best response to depends on what other participants in their “loop” are doing. In certain cases, altruistic individuals may be led to give none: A not-perfectly selfish, (but placing greater weight on their own payoff) player may instead of giving some small amounts seek to ensure their payoff is the highest and give nothing, because they know their recipient is also a dictator and they thus assume they will allocate some for themselves. On the other hand, a more altruistic than not player may give the full endowment even if in non-interactive settings they may give less than the full amount (although this type of behavior has not been spotted much in experiments).

The experimental results do find a difference between protocols, but they cannot be fully explained by rational actors changing their decision due to different optimal play for their utility functions. Thus interactive and non-interactive dictator games are not theoretically or practically equivalent, and interactive games should be studied separately within their own theoretical framework, testing hypotheses with that specific protocol. Games where dictators are also recipient lead to dictators behaving differently, possibly evaluating the payoffs of the whole “loop” of dictators to compare to their own before making decisions. This adds to the view that the beliefs and perceptions of the participants may also be important, and that clearly even small

differences in framing have effects on behavior, but this does not necessarily help explain the core theoretical question of why dictators give.

In the further meta-regression on protocols, Grech et al. (2022) divide effect types in to three categories: population effects, effects due to framing and effects due to different incentives. Protocols fit under incentive effects which have significant effects, while population effects consist of factors like sex and age, where some have significant effects and some have little if all and finally framing effects are well attested, but complex and varied, sometimes not significant at all. The theory around incentive effects fit well into utility theory, but theory on why the old give more and what causes this has been less studied.

3.5 Origin of giving

Explanations of dictator game behavior must consider the rather widespread evidence of giving, including in studies on very young children (Benenson et al. 2008). Even though experiments with children have many shortcomings (including possibilities for misunderstandings and ethical restrictions), children have been found to contribute to anonymous parties, which makes an explanation relying on learned behavior more restricted: Children as young as 4-years old give, and thus giving must be learned prior to this, or be innate or some combination thereof. This is further supported by findings of some amount of giving in different cultures around the world. Possible explanations for the origin of giving behavior is likely best left to other fields, while the theoretical background for the volatility of giving under changing conditions is ripe for further exploration and testing using economic methods.

4 Discussion

In relation to the applicability of the research it must be asked; what level of altruism is considered significant? In addition, what would be the level we would expect in real life situations? For the most basic assumptions of agents perfectly pursuing their economic interests, even a propensity to give in very suitable conditions or circumstances can introduce some complications. However, the bar set by the standard form dictator game is very high: The recipient is completely unknown to the giver, and they have no way to gauge whether the other party will receive more utility from any sum than the giver, even if they value their own utility equally to others'. Similarly, if the decision maker is predisposed to maximize total wealth, there is nothing to be gained by giving to the other party. Both cases could violate assumptions of selfishness, but in the dictator game framing they result in identical behavior in comparison to a perfectly selfish decision maker.

Additionally, the total anonymity of the standard form dictator game experiment is not necessarily applicable to most real-world decision-making situations. In situations involving charity, most prospective givers are aware of not only some characteristics of the prospective recipient, but in street fundraising they meet representatives of the prospective recipient organization and may be observed by third parties. This even further applies to situations such as dividing funds among a group of colleagues or friends, where pressures and expectations would be assumed to be higher (and cultural factors may govern expected behavior, such as the related situation of who pays for a shared meal). In situations that are less comparable to the dictator game, such as 1 on 1 deal negotiations, similar pressures caused by interacting with another, non-anonymous, party apply. On the other hand, in some situations the barrier to give is higher, because people are not necessarily presented with the distinct decision to give or not. A person living in a world with bank accounts, instant transfers and the internet has the option to donate at almost any moment.

A question raised by comparison between behavior in experiments and in real life situations, by for example Levitt & List (2007), is: Why don't those dictators that are willing to share in an experimental setting then share in real life settings? As presented before by Levitt & List, there are several ways in which experiments may

differ from real life situations. In addition to possibly being “reluctant givers” as described by Lazear et al. (2011), some willing givers may simply not run into situations where donating is presented to them in a natural setting. Having the general awareness that charity is a possibility that exists is rather different than being reminded it exists and being presented the opportunity to do so. Individuals may not be aware of or willing to expend the effort to undertake an action that they would find beneficial. As an example, in the world of Internet video, Youtube users can “subscribe” so they are shown more videos from the creator and being reminded of newly released videos for no cost. Viewers by definition have chosen to watch the video and can switch at any time. Yet even though viewers have a desire to watch a certain type of content, and know the subscribe button exists, reminding them to press the button (even without emotional pleading) leads to a much larger increase in subscribers than not doing so. The users are effectively completely anonymous and have no special incentive or cost to press the button. Advertising works partially on the same principle: Customer awareness of a product is a prerequisite for customer purchases and making purchasing as easy as possible increases sales.

4.1 The limitations of the use of students

As is usually noted in most work in the field of experimental economics (or in experiments involving human behavior in general), most participants are usually students, often of the field being studied. This can give a distorted result if students do not represent the behavior in the general population. Students are by definition more educated (or at least in the process of getting more educated!) and younger than the average person. With population in developed countries growing older and younger age groups a smaller percentage of the population, one could argue that students have become less representative of the economically relevant populace (that is, excluding young children) since the origination of dictator game in the 1980s. On quite the opposite perspective, higher education rates have increased, and the lowest educated age groups have died off (and continue to do so) since the 1980s, perhaps making students education levels closer to the average over time. Furthermore, as noted by Levitt & List (2007), those students are often self-selected and maybe more likely to cooperate with the research and chase social approval. They also posit that market participants may have less pro-social characteristics than average, as they may have helped them achieve their market position.

As noted previously by Engel (2011), students average lower donation in dictator game experiments than others. Carter and Irons (1991) also performed experiments with the dictator game that provide support to the idea that economics students behave more selfishly (and are more willing to accept lower offers). This effect was existent from the beginning of their studies, thus providing evidence for the view that economics students are self-selected to more resemble the classic homo economicus. Other factors include economics students perhaps having a stronger conception of “performing well” where maximizing own income is seen as the winning play or following the models they know. The fact that they are more likely to know the experiment as the dictator game may introduce a psychological element of winning and losing. Economics students may also be less prone to misunderstanding the game and making choices that are mistaken based on this misunderstanding.

4.2 Dictators in daily life

Examples of behaviors that are like dictator game behavior can also be found in daily life. An example where generosity is shown towards strangers are so called “pay it forward” chains at drive-throughs at restaurants or coffee shops. A customer drives in, and after ordering for themselves, tells the shop that they will pay for the next order as well. The next person in line then receives their order for free and are given the option to pay it forward by paying the next order. Generally, those in line have very limited connection to each other, with social pressure coming from the staff of the establishment. These chains sometimes reach impressive lengths in the hundreds, even though each person in line cannot be sure whether they will end up in the positive or negative and by how much, as well as always having the option to not pay the order forward. Similarly to the dictator game, participants have very limited knowledge of the recipient, but have received a “donation” from a prior customer. This could be replicated in a dictator game setting with a paying it forward model or another model, where the dictator receives charity prior to their decision.

4.3 Game avoidance and explanations

The study of exits is constrained by the lack of reasonable comparison point, which makes it more difficult to assess the reasons for exit behavior. The large number of dictators willing to exit for sums smaller than the endowment raises the question, whether the exit is connected to the dictator game at all, or simply having to play any

game at all, or at least one that takes some effort, consideration and time. If some percentage of exits are due to these other factors, the willingness to exit may not signal that the dictator is specifically avoiding getting into altruistic situations in the real world either. A possible experiment is outlined below.

To test whether exit provoked by effort required to achieve a higher payoff, a situation as similar to a dictator game exit experiment as possible should be constructed. Several treatments could prove useful, if the specifics for the willingness to exit are of interest: A treatment where prospective dictators are provided the opportunity to exit quietly for a lower payout, another where the participant is simply offered the opportunity to exit for a lower payout or wait for the higher (dictator game endowment equivalent) payout, another where to obtain the higher payout the participant must complete a simple, but somewhat effort-requiring task (ideally something that all of the participants can complete, perhaps counting the number of objects in pictures where there are a significant number of them?) and finally one where the participants must participate in another, non-dictator game experiment that does not present a similar altruism dilemma to the dictator game to have the potential to obtain the higher payout. Then, the percentage of participants exiting can be used to test a hypothesis/multiple hypotheses.

If the hypothesis that dictators avoid the game to avoid having to give is true, we would expect the percentage exiting in the final two treatments to be substantially lower than in the first. On the other hand, if time is the primary concern, we would expect all treatments to exit at a similar rate. If the hypothesis is that effort required from the participant causes them to exit, those asked to wait should stay at a higher percentage. To obtain comparable results, the exit payouts and the highest possible payouts should be the same. The menial task could perhaps be something as simple as filling a survey, to ensure success if attempted, but this might be further from the nature of a game than some sort of easy task where success, not merely participation, is rewarded.

If time proves to be a significant factor, one will conclude that any exit experiment should not allow those choosing the exit option to leave earlier and should perhaps emphasize that those choosing to exit cannot leave earlier. An outcome that contradicts the “avoiding entering the situation to give”-hypothesis would imply that the propensity to exit is perhaps connected to some factor common to participating in

an experiment in general. Regardless, this hypothetical experiment should provide a baseline for any experiment concerning an early exit for a lower payout, which might be applicable in a more general sense to many different games in economics. The larger point is that changing how an experiment is conducted, how complex it is and how long it lasts all may have impact on the results, and thus the pursuit of very specific explanations may lead to contradictory results.

4.4 Complexity in explanatory variables

Many papers that try to connect dictator game behavior and motivations have a rather ambitious interpretation of the explanatory power of any single factor, while also assuming that decision makers are more competent than they might be. Humans are well-known to be susceptible to a wide variety of cognitive biases and miscalculations. Thus, it is imperative that any experiment be as clear and simple as possible, and the procedures of differing treatments be as similar as possible. Due to these factors, deriving complex explanations of behavior, or alternatively, strict and universal rules of behavior is fraught with danger. This is especially true when many of the more interesting experiments have very limited samples. Even when results appear to be significant, coincidences or mistakes in sampling can give skewed results. One could even say “Give me 20 dictators, and I can give you any result you want” and with such small samples, doctoring is not necessary. The fact that there are countless factors that influence economic behavior (business and marketing research has countless examples, such as aroma and music in stores, ex. Morrison et al. (2011)) makes compiling and comparing different studies more hazardous. Marketing uses seemingly unconnected factors to sell products and they clearly work, so economics would do well to learn from this.

Regardless, it is imminently clear that human behavior in the dictator game, in aggregate, does not obey the “standard economic theory”, as presented by Croson and Gächter (2010):

- (i) Economic decision makers are cognitively sophisticated – they are assumed to be rational; their preferences are complete, transitive and obey the other axioms of expected utility theory. Individuals also have rational expectations about relevant states of the world and the behavior of others. In strategic situations rationality is common knowledge, that is, everybody is rational, knows that

everybody is rational, knows that others know that everybody is rational and so on.

- (ii) Risk preferences obey the expected utility principle. Time preferences exhibit exponential discounting. Social preferences are assumed to be zero (or already captured in the utility function). In applications, functional forms for the utility function are specified, e.g., it is assumed that people have Cobb-Douglas utility functions or that they have constant relative or absolute risk aversion.
- (iii) Solution concepts are typically equilibrium concepts, like competitive equilibrium, Nash equilibrium, or subgame perfect Nash equilibrium.

Any theory must then surpass this flawed, but useful framework. The challenge to producing a superior framework is in the great heterogeneity in dictator game behavior, and thus, as mentioned before, producing a theory to replace standard economic theory in certain limited situations with a better one should be possible. By the very limitations of the subject of study, general laws should be looked at with skepticism. Additionally, in certain theoretical applications replacing the standard model is not necessary, possible or useful (in for example game theory). The most relevant component added is as mentioned previously utilizing other-regarding preferences in the rational framework, like in Andreoni and Miller (2002).

A component of experimental design that has not received much scrutiny so far is how participants react to the uncertainty of payoffs. They seem to have been used interchangeably with non-uncertain protocols, which may prove to be a source of bias. With research on how other protocols affect behavior (and not completely rationally), it is not unreasonable to scrutinize all possible sources of variation between experiments. Engel (2011) found that if a lottery determines whether the payoff might be replaced by a random draw, there was no significant effect in the meta-regression, but a significant negative effect on giving in individual data. On the other hand, Umer (2023) found in a meta-analysis of 94 studies that the two often used random payment mechanisms, random problem selection procedure (RPSP) and between-subject random incentivized system (BRIS) had no significant effect on behavior. Regardless, there is a significant amount of literature on human risk evaluation, with the forementioned expected utility principle resulting in different utility from receiving a 10% chance for a sum compared to certainly receiving a sum 1/10th the size, depending on the shape of the utility function.

A rather unsatisfying result in terms of economic impact might then be that yes, human beings are susceptible to similar patterns in the lab as have been discovered in psychological and behavioral experiments on different areas of research. Dictators are susceptible to their environment, to anchoring, to bias and may make errors. And yet dictators have selfish tendencies: many do not give at all (often some 30% of dictators), those that give typically give 50% or less, averaging out to approximately 30% or slightly less of the endowment. When given a more limited decision space, they often choose the more selfish option (As seen in for example the delegation experiments, or in Bolton et al. (1998) when presenting both higher and lower options than those freely chosen).

An apt comparison for the challenges of conducting meaningful dictator game experiments that attempt to grasp at some consistent rules or behaviors could be found in the words of the ancient Heraclitus: “It is not possible to step into the same river twice”. While the challenges of comparing different treatments using different samples of dictators (especially from vastly different countries, ages, times etc.), or especially different experimental framings is clearer, using the same dictators for subsequent experiments presents challenges with comparability of results as well. This is conferred in the additional quote “for it is not the same river and the same man”, where playing the dictator game (or having knowledge of it) may affect subsequent behavior, thus presenting treatments in random orders preferable in the case of repeated dictator games. This could be seen in the possible conditioning in Hamman et al. (2010), where despite returning to the standard game in the later rounds, the low amounts given in the earlier, delegated, rounds persisted. Similarly, repeating the game risks strategic considerations of fitting some parameter of the experimenter in order to get to participate in more rounds. Thus, most results may be relegated to simply stating that a factor increases or decreases giving as the total number of factors is too great to account for or model.

5 Conclusion

The very beginning period of Wild West exploration in dictator game behavior has been over for some time, and even though some uncharted territory remains, a basic lay of the land has been achieved. A relevant percentage of giving (often an average between 20% and 40%) with usually the majority giving a positive amount across a variety of demographically different dictator samples provides strong evidence that giving in the dictator game is a replicable result. The other main finding is that giving strongly depends on the framing, increasing or decreasing with changes, even to zero or near it. Additionally, there is also a collection of evidence that points towards considerable heterogeneity among dictators, even among groups that are more homogenous in some manner (such as economics students). As an example of this, a significant fraction, often over one third, give nothing and behave rationally in the classical or selfish manner.

On the other hand, any result arising from meta-analysis or comparison of different experiments must be approached with caution. Small changes in framing, situation or protocols and sampling may change behavior substantially, even without intent to change anything. Thus, comparing the giving of different genders or countries (among other examples) across different experiments, where those factors are not the ones under study, may lead to bias. The other factors in the sample of experiments completed in the US and the UK for example are not necessarily equivalent. If one sample contains more experiments on working adults, and the other more on students, these would have to be accounted for, but even more challenging are factors that are difficult to correct for such as differences in the wording of instructions or if in one set of experiments, newer studies are more common. Additionally, the bar to overturn the null hypothesis should be set quite a bit higher: For the first studies, simply the fact that giving occurred in some scale consistently was enough to overturn the null hypothesis of the orthodox selfish preferences, while rejecting the hypothesis that two groups have the same preferences requires a significant difference in behavior, and the elimination of confounding factors.

Dictator game studies, even if outwardly similar, are best used to compare treatments within the same experiment. An experiment whose focus is on the differences between giving for different age groups or genders is much more likely to have

treated both groups equivalently and also focused on possible distortions relating to the topic of study. Thus, instead of relying on data coming from other dictator studies, differences between groups should be approached from the perspective of setting a hypothesis and then testing them.

Related to the previous, dictator games are now at the phase, where more sophisticated hypothesis and theories are being, can be and should be tested. As put by Levitt and List (2007), "...theory helps us to understand the observed pattern of results and facilitates extrapolation of lab results to other environments." Earlier experiments were sufficient to demonstrate behavior deviating from game theoretic assumptions. Now is the time to experiment on how dictator game behavior relates to life outside the lab, what reasons cause certain behavior, and how those factors might be different in a variety of real-world situations. Additionally, testing psychological principles found in other behavioral economics should be worthwhile: The likely finding should be that those affect dictator game decision-making, but being aware of for example the effect of loss aversion (hypothetically affecting results if the split is done before vs after the endowment is presented to the dictator) can help avoid attributing differences in behavior mistakenly to differences in preferences.

A more comprehensive and standardized set of instructions and protocols for dictator game experiments should be formulated to avoid comparing experiments of different parameters by accident and to reduce the probability of accidentally presenting the game differently. To keep the games comparable and to prevent straying to an entirely different strategic game, dictators should not be recipients. If dictators should be asked to make multiple decisions, treatments should be presented randomly presented in different orders to mitigate order effects and previous decisions having a framing effect. It is worth considering whether multiple decisions or the possibility of them introduces strategic implications for dictators. Double-blind experiments are preferable, as they reduce the risk to introduce additional strategic considerations in the form of interacting with the experimenter. These are baselines, and if the specific treatment involves changing these instructions those modifications should be made, but any should be purposeful and as specific as possible. If experimental economics is to bring the methods of the natural sciences into the study of human behavior, it is imperative to bring the concern for possible disturbing factors as well.

Additionally, sample sizes have tended to be rather low, and because combining data from separate experiments presents complications, sample sizes should be robust, rather in the high tens (>50) than in the low tens per treatment. This goes along with care in sampling. This along with setting specific hypotheses to be tested prior to collecting data should help avoid dredging the data for any positives from many different factors measured. With the vast quantity of treatments, it bears remembering that at a 5%-significance level (or any level) a “significant” result could be coincidental, and some number of significant results likely are. One would expect 5% of significant results at that level to be false positives, and there to be some false negatives, where differences in preferences are not visible enough at a certain significance level. Thus testing for multiple factors should be avoided, in addition to the risk it poses in making instructions and protocols too complex and thus confusing.

Regardless, the dictator game has proven useful in certain areas. Because even small changes can cause significant changes in behavior, dictator game experiments could be used to distinguish significant changes in framing from insignificant ones. For example, the willingness to donate to a charity versus another, or versus a relative or some other recipient can provide a look into what dictators consider more or less worthy or important. This can then be used to more precisely formulate behavior in the real world through theories of what causes those differences. This also expands to using the dictator game as a comparison point to other experiments such as the ultimatum game which can help expand what factors matter in which type of situation. For this purpose, variant of the dictator game that fundamentally change it can be useful, as long as that difference is purposeful and clearly acknowledged. The experiments on children can also be of use when assessing the development of economic behavior. An interesting potential could be found in experimenting with the same people in identical fashion multiple times over years as has been done in medical and happiness studies; This could yield data on development of behavior over time at an individual level. Dictator game experiments have been used to test how behavior fits different utility functions (and to formulate fitting utility functions for behavior), formulating testable hypotheses of the “moral cost”- or “rules of fairness restricting selfishness”-types and testing these hypotheses for fit could help identify what is the more precise factor behind giving in the dictator game, and also shed light on behaviors we could then expect in related scenarios outside the lab.

Finally, as mentioned by Engel (2011) the standard stakes in the dictator game are \$10, with a mean of \$21.77 and a range of \$0-\$130. In developed countries, these are not sums that one would consider “life-changing”, and one could also argue that for an average person who lives in the developed world and earns over \$1000 a month, the change from \$10 to \$100 is rather limited in significance. Thus, to test the robustness of results on giving, it would be optimal to experiment with clearly higher stakes, closer to the monthly income of a person. This could be perhaps accomplished by conducting experiments in a society with much lower purchasing power, but that also allows other differences to affect the results. Regardless, higher stakes would “steelman” (or present in the most charitable way) the argument for selfishness and for giving as caused by mistakes or insignificance of the sum, and if giving persists, strengthen the finding of anonymous giving. In the same way, testing the utility functions of preferences at both higher and lower “costs of giving” or multipliers of the amount given could be used to test whether individual behavior truly is set based on one set of preferences. So far, multipliers have often ranged from 0.1 to 4, but multipliers from 0.01 to 100 would test consistency more strongly. The bottom line is that dictator games should explore a wide variety of settings to test behavior and theories in situations of varying gravity.

References

- Andreoni, J., & Miller, J. (2002). Giving according to GARP: An experimental test of the consistency of preferences for altruism. *Econometrica*, 70(2), 737-753.
- Bardsley, N. (2008). Dictator game giving: altruism or artefact?. *Experimental economics*, 11, 122-133.
- Benenson, J. F., Pascoe, J., & Radmore, N. (2007). Children's altruistic behavior in the dictator game. *Evolution and human Behavior*, 28(3), 168-175.
- Bilén, D., Dreber, A., & Johannesson, M. (2021). Are women more generous than men? A meta-analysis. *Journal of the Economic Science Association*, 7(1), 1-18.
- Bolton, G. E., Katok, E., & Zwick, R. (1998). Dictator game giving: Rules of fairness versus acts of kindness. *International journal of game theory*, 27, 269-299.
- Brañas-Garza, P. (2007). Promoting helping behavior with framing in dictator games. *Journal of Economic Psychology*, 28(4), 477-486.
- Broberg, T., Ellingsen, T., & Johannesson, M. (2007). Is generosity involuntary?. *Economics Letters*, 94(1), 32-37.
- Carter, J. R., & Irons, M. D. (1991). Are economists different, and if so, why?. *Journal of Economic Perspectives*, 5(2), 171-177.
- Cason, T. N., & Mui, V. L. (1998). Social influence in the sequential dictator game. *Journal of mathematical psychology*, 42(2-3), 248-265.
- Carpenter, J. P., Burks, S., & Verhoogen, E. (2005). Comparing students to workers: The effects of social framing on behavior in distribution games. In *Field experiments in economics* (Vol. 10, pp. 261-289). Emerald Group Publishing Limited.
- Cherry, T. L., Frykblom, P., & Shogren, J. F. (2002). Hardnose the dictator. *American economic review*, 92(4), 1218-1221.
- Chowdhury, S. M., Jeon, J. Y., & Saha, B. (2017). Gender differences in the giving and taking variants of the dictator game. *Southern Economic Journal*, 84(2), 474-483.
- Cochard, F., Le Gallo, J., Georgantzis, N., & Tisserand, J. C. (2021). Social preferences across different populations: Meta-analyses on the ultimatum game and dictator game. *Journal of Behavioral and Experimental Economics*, 90, 101613.

- 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.
- Dana, J., Weber, R. A., & Kuang, J. X. (2007). Exploiting moral wiggle room: experiments demonstrating an illusory preference for fairness. *Economic Theory*, *33*, 67-80.
- Dreber, A., Ellingsen, T., Johannesson, M., & Rand, D. G. (2013). Do people care about social context? Framing effects in dictator games. *Experimental Economics*, *16*, 349-371.
- Engel, C. (2011). Dictator games: A meta study. *Experimental economics*, *14*, 583-610.
- Ezquerro, L., & Kujal, P. (2020). Self-selecting into being a dictator: Distributional consequences. *Journal of Behavioral and Experimental Economics*, *87*, 101559.
- Fehr, E., Bernhard, H., & Rockenbach, B. (2008). Egalitarianism in young children. *Nature*, *454*(7208), 1079-1083.
- Fisman, R., Jakiela, P., Kariv, S., & Markovits, D. (2015). The distributional preferences of an elite. *Science*, *349*(6254), aab0096.
- Forsythe, R., Horowitz, J. L., Savin, N. E., & Sefton, M. (1994). Fairness in simple bargaining experiments. *Games and Economic behavior*, *6*(3), 347-369.
- Gomes, C. M., & McCullough, M. E. (2015). The effects of implicit religious primes on dictator game allocations: A preregistered replication experiment. *Journal of experimental psychology: General*, *144*(6), e94.
- Grech, P. D., & Nax, H. H. (2020). Rational altruism? On preference estimation and dictator game experiments. *Games and Economic Behavior*, *119*, 309-338.
- Grech, P. D., Nax, H. H., & Soos, A. (2022). Incentivization matters: a meta-perspective on dictator games. *Journal of the Economic Science Association*, *8*(1), 34-44.
- Gummerum, M., Hanoch, Y., Keller, M., Parsons, K., & Hummel, A. (2010). Preschoolers' allocations in the dictator game: The role of moral emotions. *Journal of Economic Psychology*, *31*(1), 25-34.
- Hamman, J. R., Loewenstein, G., & Weber, R. A. (2010). Self-interest through delegation: An additional rationale for the principal-agent relationship. *American Economic Review*, *100*(4), 1826-1846.

- Hoffman, E., McCabe, K., Shachat, K., & Smith, V. (1994). Preferences, property rights, and anonymity in bargaining games. *Games and Economic Behavior*, 7(3), 346-380.
- Houser, D., & Schunk, D. (2009). Social environments with competitive pressure: Gender effects in the decisions of German schoolchildren. *Journal of Economic Psychology*, 30(4), 634-641.
- Jakiela, P. (2013). Equity vs. efficiency vs. self-interest: on the use of dictator games to measure distributional preferences. *Experimental Economics*, 16, 208-221.
- Kahneman, D., Knetsch, J. L., & Thaler, R. H. (1986). Fairness and the assumptions of economics. *Journal of business*, S285-S300.
- Korenok, O., Millner, E. L., & Razzolini, L. (2014). Taking, giving, and impure altruism in dictator games. *Experimental Economics*, 17, 488-500.
- Larney, A., Rotella, A., & Barclay, P. (2019). Stake size effects in ultimatum game and dictator game offers: A meta-analysis. *Organizational Behavior and Human Decision Processes*, 151, 61-72.
- Lazear, E., Malmendier, U., & Weber, R. (2011). *Sorting, prices, and social preferences*. Cambridge, Mass., USA: National Bureau of Economic Research.
- Leibbrandt, A., Maitra, P., & Neelim, A. (2015). On the redistribution of wealth in a developing country: Experimental evidence on stake and framing effects. *Journal of Economic Behavior & Organization*, 118, 360-371.
- Luhan, W. J., Kocher, M. G., & Sutter, M. (2009). Group polarization in the team dictator game reconsidered. *Experimental Economics*, 12, 26-41.
- Morrison, M., Gan, S., Dubelaar, C., & Oppewal, H. (2011). In-store music and aroma influences on shopper behavior and satisfaction. *Journal of business research*, 64(6), 558-564.
- O'Garra, T., & Sisco, M. R. (2020). The effect of anchors and social information on behaviour. *Plos one*, 15(4), e0231203.
- Schurter, K., & Wilson, B. J. (2009). Justice and fairness in the dictator game. *Southern Economic Journal*, 76(1), 130-145.
- Sutter, M., Zoller, C., & Glätzle-Rützler, D. (2019). Economic behavior of children and adolescents—A first survey of experimental economics results. *European Economic Review*, 111, 98-121.

- Umer, H., Kurosaki, T., & Iwasaki, I. (2022). Unearned endowment and charity recipient lead to higher donations: A meta-analysis of the dictator game lab experiments. *Journal of Behavioral and Experimental Economics*, 97, 101827.
- Umer, H. (2023). Effectiveness of random payment in Experiments: A meta-Analysis of dictator games. *Journal of Economic Psychology*, 96, 102608.

Working papers

- Brookins, P., & DeBacker, J. M. (2023). Playing games with GPT: What can we learn about a large language model from canonical strategic games?. *Available at SSRN 4493398*.
- Fan, C., Chen, J., Jin, Y., & He, H. (2024, March). Can large language models serve as rational players in game theory? a systematic analysis. In *Proceedings of the AAAI Conference on Artificial Intelligence* (Vol. 38, No. 16, pp. 17960-17967).
- Xue, J., & Ohtsubo, Y. (2020). Altruistic Preferences in the Dictator Game: Replication of Andreoni and Miller (2002) in Japan. *Letters on Evolutionary Behavioral Science*, 11(1), 17-21.

Econometrics

- Afriat, S. (1967): "The Construction of a Utility Function From Expenditure Data," *International Economic Review*, 8, 67-77.
- Croson, R., & Gächter, S. (2010). The science of experimental economics. *Journal of Economic Behavior & Organization*, 73(1), 122-131.
- Houthakker, Henrick (1950): "Revealed Preference and the Utility Function," *Economica*, 17,159-174.
- Levitt, S. D., & List, J. A. (2007). What do laboratory experiments measuring social preferences reveal about the real world?. *Journal of Economic perspectives*, 21(2), 153-174.
- Varian, Hal R. (1982): "The Nonparametric Approach to Demand Analysis," *Econometrica*, 50,945-97
- Varian, Hal R. (1993): *Microeconomic Analysis*, Third Edition. New York: Norton.

