Dictator games: a meta study

Christoph Engel

Abstract Over the last 25 years, more than a hundred dictator game experiments
have been published. This meta study summarises the evidence. Exploiting the fact
that most experiments had to fix parameters they did not intend to test, in multiple
regression the meta study is able to assess the effect of single manipulations, con-
trolling for a host of alternative explanatory factors. The resulting rich dataset also
provides a testbed for comparing alternative specifications of the statistical model for
analysing dictator game data. It shows how Tobit models (assuming that dictators
would even want to take money) and hurdle models (assuming that the decision to
give a positive amount is separate from the choice of amount, conditional on giving)
provide additional insights.

Keywords Dictator game · Meta-study · Multiple regression · Tobit · Hurdle model

JEL Classification C24 · C91 · D03

1 Introduction

Some 25 years ago, Daniel Kahneman first had students in his classroom play an
ultimatum game (Güth et al. 1982), which was then followed by the question:

> "You will be matched at random with two other students, and you will get to
> share some money with one or both of them. If the two people made different

C. Engel (✉)
Max Planck Institute for Research on Collective Goods, Kurt-Schumacher-Straße 10, 53113 Bonn,
Germany
e-mail: engel@coll.mpg.de

decisions in the first stage (e.g. one of them took $10 and one took $18), then you must make a decision about how to allocate the money. Call the person who took $10 and gave the other one $10 student E (for even). Call the person who took $18 and gave the other one $2 student U (for uneven). Your choices are as follows: you may allocate $5 to yourself, $5 to student E, and nothing to student U; or you may allocate $6 to yourself, nothing to student E, and $6 to student U".

74% of their participants chose the first option although this cost them $1 (Kahneman et al. 1986:S290 f.). This is how a literature started to which in the meantime 129 contributions have been published, testing a total of 616 different treatments.[1] The experimental paradigm has proven so powerful precisely because it is so simple. Actually the successors of Kahneman have further simplified the game (starting with Forsythe et al. 1994). The game now is typically stripped of the third party punishment component and played in isolation. The dictator's action space is complete, so that she may distribute the pie at her will between the recipient and herself.

Kahneman invented the game as part of his programme that turned textbook assumptions into behavioural hypotheses. While normally a sizeable fraction of participants does indeed give nothing, as predicted by the payoff maximisation hypothesis, only very rarely this has been the majority choice. It by now is undisputed that human populations are systematically more benevolent than *homo oeconomicus*. Later experiments have explored this predisposition in two dimensions: situational and demographic. The former implicitly sticks to the claim that, at least at the population level, behavioural dispositions are human universals. It engages in refining the conditions under which benevolence is to be expected. The latter puts the research question upside down and uses the extremely simple design as a tool for quantifying systematic behavioural differences between populations.

The dictator game has become popular among experimentalists. In the one year of 2008, 30 new papers with this game have been published. It therefore is time to take stock, and to make the existing body of evidence accessible. Yet this paper aims beyond merely providing orientation. Implicitly, through their design choices, experimenters have generated data on independent variables they have not explicitly set out to test. They for instance have played a one-shot game with students, asking dictators to divide a pie of $10 given to them between themselves and an anonymous recipient from the same subject pool. What looks like a perfectly standard dictator game implicitly provides data on one shot versus repeated games; on games with students versus other populations; on manna from heaven versus earned money; on stakes; on a specified degree of social distance; on dictator-recipient anonymity versus dictator identification. This evidence is untapped as yet. It is useful for two purposes. The basis for testing the effect of isolated manipulations becomes much broader. More interestingly even, it becomes possible to control for alternative explanations to a degree that by far transcends what is feasible in individual experimental studies. That way one learns which effects are robust, and how big effects are once one controls for other factors that have been shown to matter for the willingness of dictators to give.

---

[1]For detail, see the list of papers used for the meta study in Appendix 1 in electronic supplementary material.

The remainder of this paper is organised as follows. Section 2 explains how the sample has been collected and addresses analytic methodology. Section 3 treats all experiments as contributions to one question: how much are dictators willing to give? Section 4 focuses on individual experimental manipulations. Section 5 simultaneously uses all these independent variables to explain generosity with multiple regressions. In this section I also discuss different options for specifying the statistical model. Section 6 concludes.

2 Data and methodology

In disciplines like medicine or psychology, meta-analysis is standard.[2] In econometrics it is also becoming increasingly frequent.[3] By contrast, the number of meta studies on economic experiments is still fairly limited (Druckman 1994; Harless and Camerer 1994; Croson and Marks 2000; Zelmer 2003; Huck et al. 2004; Oosterbeek et al. 2004; Blackwell 2007; Engel 2007; Prante et al. 2007; Jones 2008; Hopfensitz 2009; Percoco and Nijkamp 2009; Weizsäcker 2010). A word on the potential and the limitations of the method may therefore be in order.

Meta-analysis is best understood as quantitative literature review (Stanley 2001). It is much more objective, and usually also more informative, than narrative review. It exploits the fact that there is a whole body of studies that are sufficiently similar to make their joint analysis meaningful; with lab experiments, this condition is much easier to fulfil than with field data, given they are standardised by design (Thompson and Pocock 1991). Nonetheless, different experiments have related, but different research questions. They may be differently well executed. Non-results are difficult to publish, which may lead to publication bias. Heterogeneity between studies can be pronounced, which may be due to unobserved explanatory factors (all of this had already been discussed by Druckman 1994). These concerns can be mitigated by the techniques presented in this section. Still results from meta analysis should be read with caution. If a researcher remains sceptical about a finding from meta-analysis, she should design a new experiment that is specifically targeted to precisely this research question.

Two papers have engaged in a similar exercise. In 2003, Camerer has done a meta-study of 11 experiments (Camerer 2003:57 f.). In 2008 Cardenas and Carpenter have done the same for 10 experiments playing the dictator game in developing countries (Cárdenas and Carpenter 2008:317). The present meta-study covers all 129 papers published between 1992 and the end of 2009, including 4 papers to come out in 2010 but already available through advance access. 4 papers do not report sample size. The remaining papers cover a total of 41,433 observations.

The search for papers has started with the two predecessor studies. The keyword "dictator game" produced 89 hits in the EconLit database, and 240 hits in RePEc. Finally I have checked the references of the papers thus found. Unsurprisingly, these

---

[2]The database PsychInfo reports 4,685 entries with the word "meta analysis" in the title of the paper.

[3]The database EconLit reports 491 entries with the keyword "meta analysis", most of which refer to work in econometrics.

sources frequently pointed to the same publications. Moreover, in particular but not only in RePEc, not so rarely one and the same publication is listed repeatedly within the same database, usually since earlier and later versions are kept. Subtracting duplicates, the resulting gross sample includes 261 papers. I have excluded 76 papers since they test a different game. I in particular have left aside papers that give the recipient any kind of power (and thereby bring the experiment close to an ultimatum game); papers (like Kahneman et al. (1986) that started the literature) that mix the motive of benevolence with one of sanctioning the recipient; twelve papers that limit the dictator's action space such that it does not include the option to give nothing (again like Kahneman et al. 1986). Six experiments play the right game, but ask a different research question. They for instance want to know how much unendowed observers expect dictators to give (Aguiar et al. 2009). Another 15 papers do not report the data such that it can be used for meta-analysis. Seven papers have no original data. Another five are surveys or (the two) meta-studies. 20 are pure theory papers. One is a pure econometrics paper. The 131 papers included in the final sample comprise a total of 616 treatments. A complete list of these papers is made available online.[4] I deliberately include working papers, even if they have never been published, to guard against publication bias (as recommended by Thompson and Pocock 1991).

In dictator games, the dependent variable is (quasi)continuous. Dictators may give any fraction between 0 and 100% of their endowments, albeit frequently only in integers of $10. Classic meta-analysis works with one datapoint per experiment and condition. In my case, this is the mean fraction of the pie that dictators give recipients. For 498 treatments, this measure has been reported. For the remaining treatments, it can be calculated from the pie size and information on individual contributions.

While all experiments reported in this paper play a dictator game, they use this game to test a rich array of research questions. To do so, they manipulate a host of context factors. Given manipulations differ, one should expect differences in outcome. More importantly, heterogeneity across studies can be traced back to explicit differences in experimental design, including factors held constant in the individual experiment, but different across experiments. The appropriate technique for capturing this observed heterogeneity is meta regression (Harbord and Higgins 2008).

Meta-analysis exploits the fact that individual studies have different size and different variance.[5] So called fixed effects meta-regression weights each data point by the respective standard error.[6] The weight thus reflects the precision of the information from each study. Yet in my sample standard deviations or standard errors are only reported for 191 treatments. Fortunately, for another 254 treatments, the standard error can be reconstructed from information on the distribution of give rates. Frequently, this information is presented graphically, as a histogram or as a cumulative distribution, with no exact numbers reported. I then have checked whether my

---

[4]doi:10.1007/s10683-011-9283-7.

[5]While this is standard in the neighbouring disciplines, weighting by a measure for precision is still unusual in experimental economics, but see (Croson and Marks 2000; Zelmer 2003) who weight by the number of observations.

[6]Terminology slightly differs from the terminology that is standard in econometrics. The "fixed effects" model actually does not have an individual specific effect in the model, but exclusively captures differences between studies by the weight.

measurement of the height of bars is sufficiently precise. I have accepted the outcome if numbers add up to a result within [.95 1.05]. I thus have accepted a translation mistake of at most 5%. This procedure seemed more reliable than ex post readjusting the visual impression such that the numbers add up to exactly 1. Meta-regression works with the resulting 445 datapoints (and drops 171 datapoints for which no information about standard errors can be recovered).

Fixed effects meta-analysis assumes that all studies are taken from a sample with the same residual error. So called random effects meta-analysis relaxes this assumption. It takes the possibility into account that different experiments not only differ by observed design features, but also by unobserved factors. Following the procedure proposed by DerSimonian and Laird (1986), random effects meta-regression allows the standard error of each study to itself result from a random draw. It thus works with model

$$y_i = \mathbf{x'_i}\beta + u + e_i$$

where y_i is the fraction of the pie one participant gives the recipient, $\mathbf{x'_i}\beta$ is a vector of explanatory variables with its coefficient vector, $u \sim N(0, \tau^2)$ is the between studies error, with variance τ^2, and $e_i \sim N(0, \sigma_i^2)$ is the within study (residual) error. Consequently, each study is weighted with $1/(\sigma_i^2 + \tau^2)$ (Harbord and Higgins 2008). The weight thus also reflects how much the respective treatment contributes to explaining the between studies variance.[7] This procedure mitigates the loss in information resulting from unobserved heterogeneity across studies.

Normally, meta-analysis cannot go any further. Fortunately, the literature on dictator games frequently not only reports means and standard errors, but also distributions. The measure is slightly imprecise though. Even if dictators' action space was not limited to integers from 0 to 10, usually only deciles are reported. Moreover, I mostly had to translate graphs into numbers for the purpose, with a small risk of translation error. With these limitations, based on distribution information and on the N of the respective study, for those 83 papers with complete distribution information, I am able to reconstruct the original data, resulting in 20,813 observations.[8] The procedure to do so is straightforward. If a publication reports 100 observations per treatment, and if in the first treatment 30% have kept their entire endowment, this results in 30 reconstructed datapoints with the independent variables from this treatment and 0% giving as the normalised dependent variable. Note that I have full information of independent variables in that they are fix per treatment.

I use this second data set for complementary analysis. It in particular makes it possible to go beyond means per treatment as the exclusive dependent variable, and to also investigate distributions. If I work with this data set, I have to account for the fact that each study has generated multiple data points. I use two complementary approaches. Clustering at the level of treatments corrects standard errors for the potential lack of independence. Treatment dummies control for unobserved hetero-

---

[7]For the restricted maximum likelihood estimator of τ^2, see Harbord and Higgins (2008).

[8]Prante et al. (2007), Hopfensitz (2009), Weizsäcker (2010) also work with meta datasets.

geneity.[9] Along with the analysis based on individual data, classical meta-analysis is always reported since it is more conservative, and since it can cover more papers.

The reconstructed meta-dataset provides a further opportunity. As is patent from the graphs usually reported in dictator game manuscripts, giving is not distributed normally. A substantial fraction tends to give nothing. Frequently there is a second peak at the equal split. The rich dataset that reflects characteristic distributions makes it possible to discuss alternative statistical approaches for analysing this data.

As is well known, omitted variables may induce bias. Since potentially experiments not only differ by the factors the respective design expressly manipulated, it is fortunate that experimenters also had to report those features of their designs which they held constant. In the multiple regressions of Sect. 5, I use this information as control variables. Ultimately, readers should trust these findings more than the regressions just testing the effect of one exogenous variable in isolation. I nonetheless first present these findings in Sect. 4 for three reasons: for some variables, I do only have information for a limited number of experiments. The reader may be interested to learn which variables explain a considerable fraction of the variance (measured by the gain in adjusted R^2 from adding them), and which explain little. Finally, the analysis of single variables provides the opportunity to introduce these variables. Readers should see the reported coefficients more as a somewhat more informative form of descriptives, rather than the definite result.

3 Overall generosity

If one calculates the grand mean from all reported or constructed means per 616 treatments, dictators on average give 28.35% of the pie. As Fig. 1 demonstrates, the distribution of means is left skewed. Dictators are more likely to give little. In 6 of 616 treatments, they on average give zero. Only in one of these six cases, this results from the fact that dictators had a take option.[10]

Technically, meta-analysis is more reliable since it weights observations by their precision. It can be undertaken for those 445 treatments for which standard errors are reported or can be reconstructed. Interestingly, random effects meta-analysis almost perfectly matches the unweighted grand mean. It establishes a give rate of 28.3%.[11] In principle, this is a very reliable result. The null hypothesis that the give rate is 0 is rejected with $z = 35.44$, $p < .0001$. Yet between studies heterogeneity is pronounced. Between studies variance[12] explains 97.1% of the overall variance. Since we have

---

[9]This approach has also been used by Croson and Marks (2000), Hopfensitz (2009).

[10]In Brosig et al. (2007) giving nothing established the equal split between both players' endowments; in the pertinent treatment of Leider et al. (2009) social distance was high and efficiency of giving low; in Oxoby and Spraggon (2008) the dictator had earned the pie.

[11]The result from fixed-effects meta-analysis differs though. If weights are not corrected for the contribution of a treatment to explaining between studies variance, the estimated give rate goes down to 20.4%.
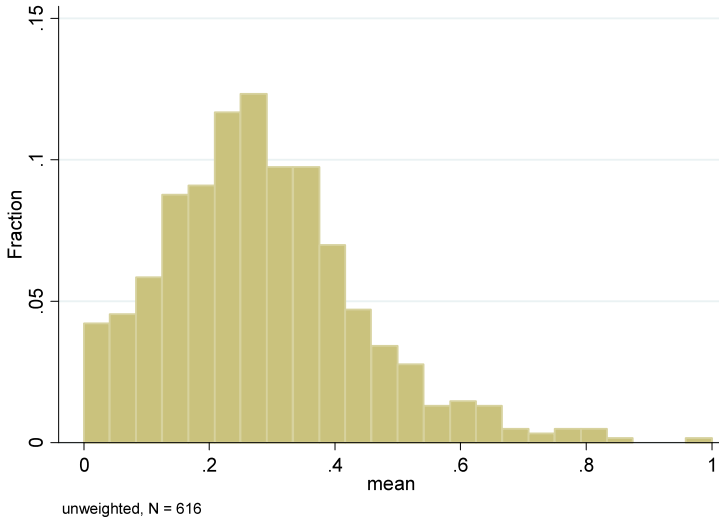
[12]In technical terms, parameter τ^2.

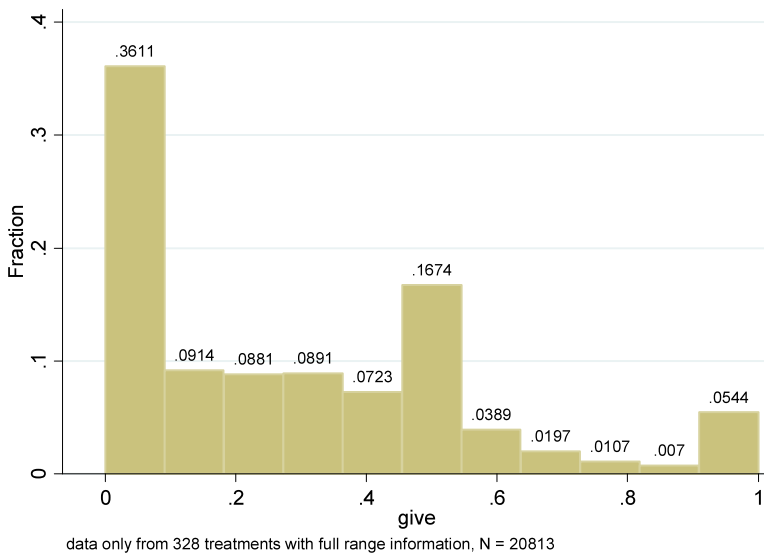**Fig. 1** Distribution of mean giving per treatment

Fig. 2 Distribution of individual give rates

information on a host of independent variables, this is a strong indicator that meta-regression is preferable over mere meta-analysis.

Finally, using the supplementary dataset with reconstructed individual observations, we learn that contributions are very unevenly distributed over the unit interval, see Fig. 2. 36.11% of all participants give nothing to the recipient. 16.74% choose the equal split. As many as 5.44% give the recipient everything.

4 Explanatory factors

Experiments are designed to test the effect of one, maybe two or three, manipulations. In the motivation of their papers, experimentalists link their tests to the related literature. Yet if experiments are not meant to test formal theory, which is normally not the case in dictator games, it is not always clear how one effect is related to another. In a fully satisfactory way, these links could only be built if the mediating factors, from the environment or from participants' attitudes, were isolated; this would require new experiments. Here, I must content myself with ex post generating order in this literature. In so doing, I start with those manipulations that are closest to standard economic theory. A number of experiments have slightly altered the incentive structure (4.1). Others have given the dictator all the power, but have exposed her to social control. This tests the power of social expectations and of the risk of social sanctions (4.2). Money maximising agents only care about their own payoff. Many experimental games point to the fact that participants are also sensitive to relative, not only to absolute payoffs. Many dictator games have explored the motivating force of payoff comparisons by making the dictator and the recipient differently deserving (4.3). Some experiments have introduced context by framing the situation. Since context defies standardisation, for this meta-study only one manipulation is analysed in detail: the effect of having participants handle real coins or notes (4.4). I may be willing to give to the victims of an earthquake at the opposite side of the globe, and my neighbours may have annoyed me so much in the past that they would be the least to whom I would want to be generous. Yet if such contextual factors are excluded by design, one would expect that people give the less, the higher the social distance between themselves and the recipient (4.5). Finally, demographics like age, gender, race, or the country of origin matter (4.6). Many of these explanatory variables are dummies, some are ordinally scaled, some are continuous. For detail, the reader is referred to the descriptive statistics in the online Appendix 2.[13]

As is typical for meta-studies, these independent variables are not distributed equally across the (meta) sample. Moreover some independent variables are not reported in all studies. For detail, the interested reader is again referred to Appendix 2. Since the dataset is rich, despite the imbalance many effects turn out significant. In this section, I introduce the independent variables one by one and present tests. Even if effects are significant, once one controls for study heterogeneity, the additional explanatory power due to a single explanatory variable is usually fairly low. I nonetheless report regressions throughout. Knowing the marginal effect and its significance level is more informative than just descriptive statistics, which would be the alternative. Also, low explanatory power is what one should expect. All models aim at explaining variance in giving in the entire dataset. Yet the variance of individual explanatory factors across studies is low for most individual explanatory factors. As laid out in Appendix 2, most individual explanatory factors are distributed very unevenly in this dataset. Therefore frequently a small number of datapoints have to provide all the explanation. In Sect. 5, models are presented that simultaneously control for multiple explanatory factors. They explain a much higher fraction of the variance.

---

[13]doi:10.1007/s10683-011-9283-7.

4.1 Incentives

In most dictator games, the action space is quasi continuous. For instance the stake is $10, and dictators can give any amount of integer dollars they deem fit, but they usually cannot split the amount by cents. Sometimes, experimenters have (further) constrained the action space. They for instance have excluded the equal split, or they have only given dictators a choice between keeping everything and contributing half of the pie (both is, e.g. used in Bolton et al. 1998). Meta-regression with this explanatory variable is insignificant. However if one uses individual data, limiting the action space has a significant positive effect: OLS, with fixed effects for each treatment (not reported), cons .342***, limited action space .121***, $N = 20,813$, additional adj.$R^2 < .001$.[14]

In the standard dictator game, the dictator may be sure that her decision determines her payoff. Some papers have introduced uncertainty about the dictator's own payoff. A lottery determines whether the dictator's choice is implemented, or whether it is replaced by a random draw (a key contribution is Andreoni and Bernheim 2009). Again meta-regression is insignificant. However using individual data, one establishes a significant negative effect. The less the dictator is sure that intended benevolence becomes effective, the less she gives in the first place: OLS, cons .342***, degree of uncertainty $-.265$***, $N = 20,813$, additional adj.$R^2 < .001$.

Normally, there is just one dictator. What if more than one dictator competes? The key contribution to this question is a study with school children. After the whole class had played, the payoff ranking was posted on the blackboard, and the dictator with the highest payoff got an additional treat (Houser and Schunk 2009). Once more, meta-regression is insignificant, as is a regression with individual data.[15]

It is standard in dictator games to pay out each individual choice. Yet if experimenters have repeated the game, or if they wanted to use high stakes, they sometimes have only paid some choices, or some dictators for that matter, at random. Occasionally, experimenters have not used pecuniary incentives at all, and have just asked the hypothetical question how much participants would be happy to give, were they to have the money. Again, the effect of this manipulation is insignificant in meta-regression and when using individual data.[16]

The standard stakes in a dictator game are $10. Especially when repeating the game, or when using the strategy method (Selten 1967), stakes can be fairly low.

---

[14]All OLS models have heteroskedasticity robust standard errors. Throughout the paper, significance levels are reported by *** $p < .001$, ** $p < .01$, * $p < .05$, + $p < .1$. Since the model contains a regressor for each treatment of an experiment, the overall adjusted R^2 (which is .230) is not informative. Instead the increase in the adjusted R^2 resulting from adding the regressor(s) in question is reported. In the concrete instance it is close to 0, despite the fact that the regressor is highly significant, $t = 4.59$, $p < .001$.

[15]If one does not add study dummies (but clusters standard errors for treatments), dictators give less if they are under competitive pressure, OLS, cons .273***, dictator competition $-.074$*, $N = 20,813$, adj.R^2 .0008.

[16]If one does not add study dummies (but clusters standard errors for treatments), one finds a significant negative effect for random pay, while hypothetical payment is still not significantly different from paying out every choice, OLS, cons .377***, random pay $-.117$***, hypothetical pay $-.045$, $N = 20,813$, adj.R^2 .010.

An additional challenge stems from the fact that many experiments have been run in countries with a national currency other than the dollar. If papers have indicated the exchange rate at the time of the experiment, I have used the dollar equivalent. Nonetheless, for this independent variable, I only have 440 (rather than 616) treatments. Stakes range from $0 to $130, with mean $21.77. This time, both meta-regression and a regression with the original data are insignificant. Meta-regression is significant, though, if one reduces the sample to those 158 treatments that have manipulated stakes. Higher stakes reduce the willingness to give. If there is more to gain, dictators keep more, not only in absolute, but also in relative terms (although the effect is very small): meta-regression, cons .326***, stake $-.002$**, $N = 158$, adj.R^2 .036.

The standard dictator game is one shot. Some experimenters have repeated the game, but changed recipients every round. Others have used the strategy method. I have coded both manipulations as multiple tests. Multiple tests have a clear negative effect on giving: meta-regression, cons .297***, multiple tests $-.071$**, $N = 445$, adj.R^2 .058.

What if the dictator is not an individual, but a group? Again, both meta-regression and regression with individual data are insignificant.[17]

4.2 Social control

Many experimenters have wondered whether dictator greed would disappear with increased social control. The most obvious way to increase social control is making recipients identify the person in whose hands they are. One for instance makes dictators stand up, so that recipients observe who decides (Frey and Bohnet 1995). Here, meta-regression yields a surprising result. The effect of identification is insignificant. But it becomes significant if one controls for stakes. Yet then the effect is negative: meta-regression, cons .283***, identification $-.060$*, stake .0002, $N = 346$, adj.R^2 .045. At the level of individual data, the effect reverses and thus is strongly positive, as one would expect: OLS, cons .342***, identification .658*, $N = 20,813$, additional adj.$R^2 < .001$. Comparing distributions, one sees the likely reason for the surprising result of meta-regression, Fig. 3. If dictators are identified, this has three effects: they are less likely to give nothing; the mode shifts from 0 to .5; dictators are less likely to give more than half of the pie. Seemingly, even in the dictator game benevolence may crowd out.

A less intrusive form of increasing social control is giving the dictator a social cue. For instance one study has shown part of the dictators three large dots that were arranged in a way that evokes a face (Rigdon et al. 2009). Since standard dictator games are devoid of any social cues, it is meaningful to compare them. While the effect is not significant in meta-regression, it is when using individual data (and controlling for study heterogeneity), OLS, cons .342***, social cue .121***, $N = 20,813$, additional adj. $R^2 < .001$.

---

[17]There is a highly significant effect when using individual data and not adding treatment dummies (but clustering standard errors at the treatment level). Groups then are more selfish than individuals: OLS, cons .274***, group decision $-.083$***, $N = 20,813$, adj.R^2 .002.
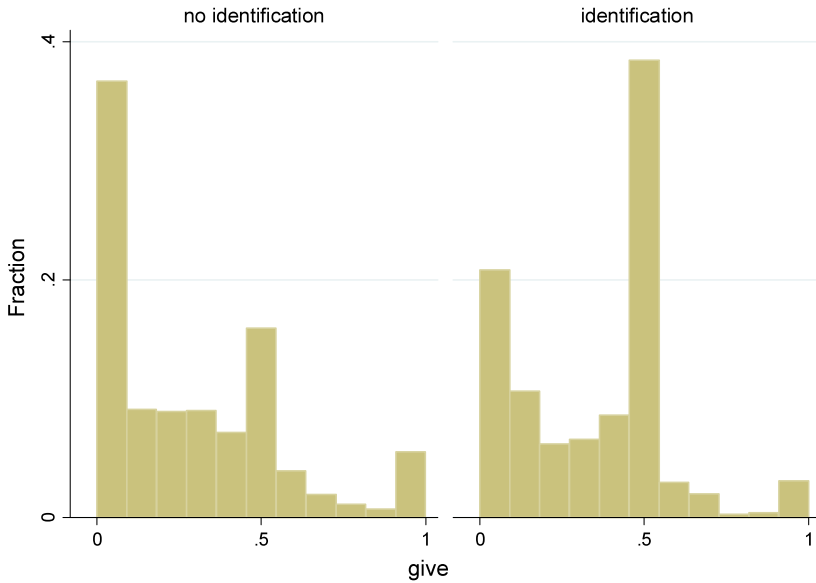
Fig. 3 Effect of dictator identification based on individual data

Other experimenters have not increased, but reduced the degree of social control, compared to the standard design of dictator games. To that end, they have given the dictator a chance to hide her decision, such that the recipient does not learn (for sure) that her payoff is due to a dictator's decision (see e.g. Dana et al. 2006). This reduces dictators' generosity: meta-regression, cons .288***, concealment option −.077*, $N = 445$, adj. R^2 .007.

Early on, experimenters were concerned that dictator-recipient anonymity would be insufficient to measure true benevolence. They were afraid that participants would give because they were aware the experimenter would learn their choices. Fairly elaborate procedures have been invented to also guarantee dictator-experimenter anonymity (starting with Hoffman, McCabe et al. 1994). Both in meta-regression and when using individual data, such doubleblind protocols do not have a significant effect. This changes if one controls for one shot versus repeated interaction. One then finds a weakly significant, small reduction of generosity: meta-regression, cons .309***, repeated interaction −.082***, doubleblind −.036[+], $N = 445$, adj. R^2 .066. This suggests that the choice between single and double blind designs matters only in more complicated designs.

4.3 Distributive concerns

Ultimately, a dictator game is about distribution. One should therefore expect that dictators give more if the recipient's claim to her fair share is made more legitimate. One way of manipulating legitimacy is property rights. In the standard dictator game, the dictator receives money from the experimenter and is free to give a fraction to the recipient. What if, instead, the money is allocated to the recipient, but the dictator
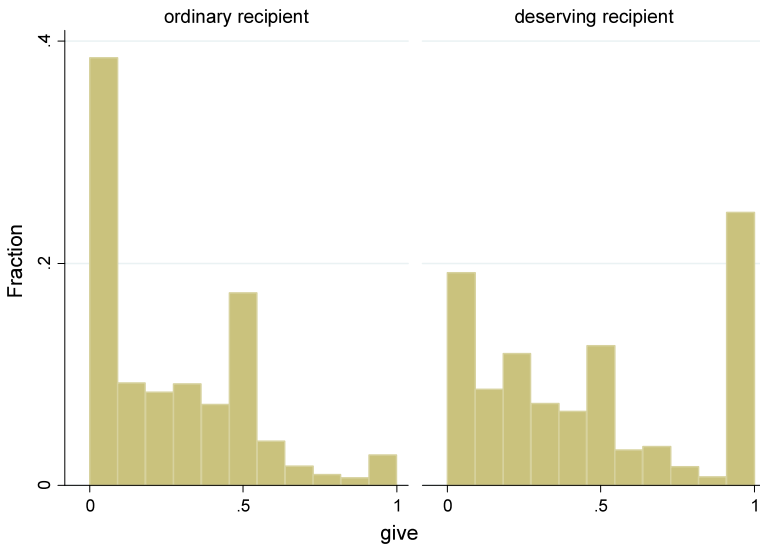
Fig. 4 Giving to a deserving recipient

is free to take a fraction (as first investigated by Eichenberger and Oberholzer-Gee 1998)? While this does not have a significant effect in meta-regression, it does when using individual data, and controlling for study heterogeneity, cons .342***, take option −.342***, additional adj.$R^2 < .001$. The constant and the regressor almost perfectly neutralise each other, indicating that, when they have the option to also take money, on average dictators leave the recipient with no gain.

In the standard dictator game, the recipient is not needy. Many experiments have checked how generosity increases if it is made clear that the recipient is deserving. They for instance have asked dictators whether they were happy to make a donation to a charity. This has a clear and strong effect: meta-regression cons .261***, deserving recipient .115***, $N = 445$, adj.R^2 .075. Figure 4 demonstrates that this can be decomposed into two effects. If the recipient is deserving, only half as many dictators keep everything. More than 20% give everything. The difference in distributions is significant: Epps Singleton, $p < .0001$.[18]

Dictators do take even less if the recipient had earned the money into which they are now free to tap: meta-regression, cons .286***, recipient earned money −.059, takeoption −.084, earned*takeoption .616+, $N = 445$, adj.R^2 .008.

The standard dictator game is a zero sum game. One dollar won for the recipient is one dollar lost for the dictator. Quite a few experiments have turned the game into a positive sum game. They for instance have stipulated that one dollar lost for the dictator increases the recipient's payoff by three dollars. Other experiments have reduced the efficiency of giving, so that one dollar lost for the dictator creates less than a dollar for the recipient. These manipulations have a significant effect in the

---

[18]On methodology see Epps and Singleton (1986), Goerg and Kaiser (2009). The test is preferred over Kolmogorov Smirnov since it has higher statistical power.
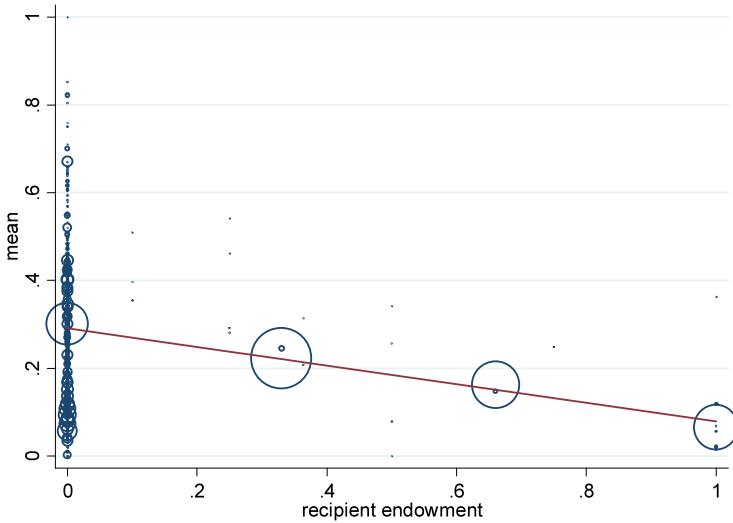
Fig. 5 Effect of upfront endowment to recipient

expected direction. The larger the multiplier, the more dictators give, and vice versa: meta-regression, cons .248***, multiplier (fraction or multiple of 1) .030*, $N = 445$, adj.R^2 .016.

What if there is more than one recipient? One might have thought that the presence of a second recipient serves as an excuse for giving little to both. Actually, the opposite is true. If there is another recipient, both receive substantially more: meta-regression, cons .280***, multiple recipients .112*, $N = 445$, adj.R^2 .022.

In the standard dictator game, the recipient is poor while the dictator is rich. If the recipient also receives an endowment upfront (as, for instance, in Eckel et al. 2005), this strongly reduces giving: meta-regression cons .291***, recipient endowment $-.213$***, $N = 445$, adj.R^2 .052. As Fig. 5 shows, if the recipient has received a positive endowment at the start of the interaction, the reduction is almost perfectly proportional to the size of the endowment; datapoints almost perfectly lie on the regression line.

Perceived distributive fairness is also altered if the pie is not manna from heaven, but if the dictator had to earn it. This strongly reduces her generosity: meta-regression, cons .295***, earned pie $-.169$***, $N = 445$, adj.R^2 .084. The mere fact that the dictator is ostentatiously entitled with the pie does not have a significant effect in meta-regression. With original data, and controlling for study heterogeneity, there is a small, weakly significant effect in the expected direction: cons .342***, entitlement $-.098^+$, additional adj.$R^2 < .001$.

4.4 Framing

The standard dictator game is presented as a context free opportunity structure, giving the dictator all the decision power, and the recipient none. Experimenters have added a variety of frames to the dictator game, like buying and selling a commodity
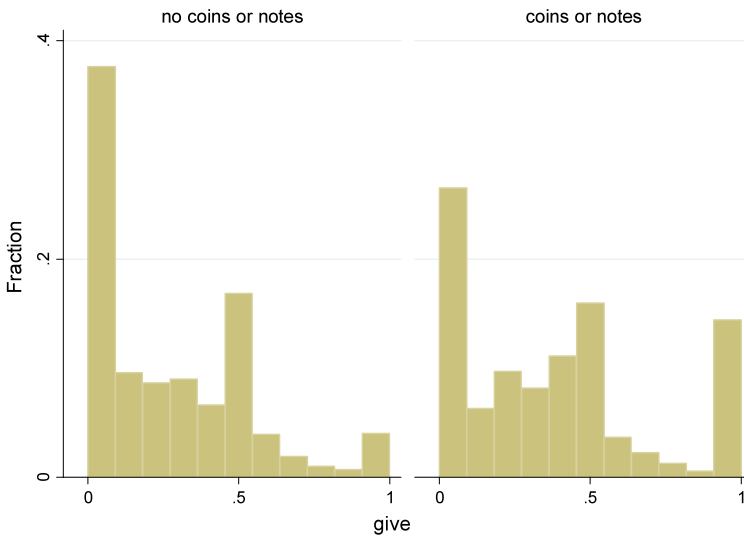
Fig. 6 Effect of handling real money

(Hoffman et al. 1994). Individual frames can be quite influential (see e.g. Brañas-Garza 2007). Yet the abstract fact that a frame has been added does not have a significant effect; each frame is different. Another manipulation that is related to framing has a remarkably strong effect though. If dictators handle coins or notes, they give substantially more: meta-regression, cons .268***, real money .075***, $N = 445$, adj.R^2 .023. As Fig. 6 demonstrates, the effect has two sources. If they handle real money, participants are less likely to give nothing. More remarkably even, they become much more likely to give everything. The difference in distributions is significant: Epps Singleton, $p < .0001$.

4.5 Social distance

In the standard dictator game, it is common knowledge that the dictator and the recipient are randomly selected members of the same student subject pool. Experimenters have manipulated the degree of social distance, ranging from a total stranger over another member of the subject pool to a member of a more closely defined group, to the friend of a friend of a friend, to the friend of a friend, to an immediate friend (this has been an active area of research recently, see Leider et al. 2009; Brañas-Garza et al. 2010; Goeree et al. 2010).[19] In the meta-regression, the degree of social distance has a surprising negative effect. The model predicts that dictators give the less the closer the recipient: meta-regression, cons .355***, degree of social distance −.063***, $N = 445$, adj.R^2 .098. Inspecting Fig. 7, one should become suspicious though. If dictators and recipients are any closer than being members of the

---

[19]Note that social distance measures proximity, not distributive concerns. If, for instance, dictators are invited to give to a charity, social distance is high, but there is a higher than ordinary social expectation of giving.
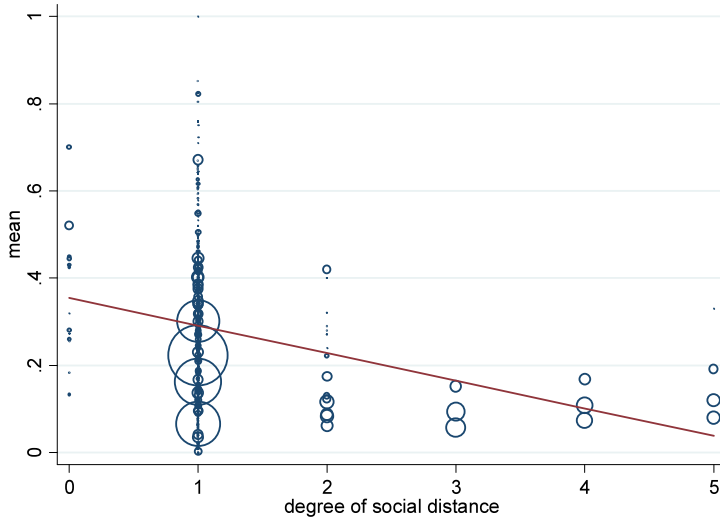
Fig. 7 Effect of social distance

same subject pool, their generosity is practically stable. The significant effect is a statistical artefact, driven by other explanatory factors inducing ordinary members of the subject pools to contribute substantially more. This of course strongly indicates that multiple regression is preferable over single regression.[20]

4.6 Demographics

The normal participant of a dictator game is a student. On average, non-students give much more: meta-regression, cons .398***, student $-.151$***, $N = 445$, adj.R^2 .153. As Fig. 8 shows, students are much more likely to give nothing, and they are much less likely to choose the equal split, or to even give everything. The difference in distributions is statistically significant: Epps Singleton, $p < .0001$. These are remarkable findings. Experimental economists have often been criticised for generating artificial findings with a subject pool that has little to do with the much harsher reality of economic relations (see e.g. Levitt and List 2007). At least in dictator games, the opposite turns out to be true. Student experiments *underestimate* the deviation from the textbook prediction.[21]

Starting with Eckel and Grossman (1998), many have tested gender effects. Since in ordinary papers on dictator games gender is not reported, meta-regression with all data would not be meaningful. If one confines the sample to those papers that have explicitly tested gender, it turns out that women give significantly more: meta-regression, cons .212***, female .058*, $N = 12$, adj.R^2 .106.

---

[20]If one controls for unobserved heterogeneity, in multiple regression one recovers the positive effect of social proximity found in those studies that explicitly manipulated this factor, see below Sect. 5.

[21]Note, however, that the effect reverses in model 3 of Table 1, i.e. when controlling for other explanatory factors and for unobserved heterogeneity by way of treatment dummies.
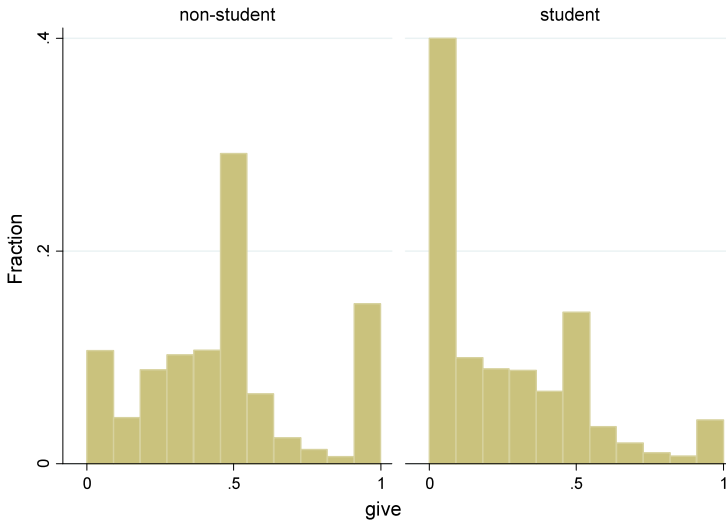
Fig. 8 Comparing students to non-students

If one does the same with dictator race, there is no significant effect. For this independent variable, meta-regression is the only option since no paper that has manipulated dictator race has also reported distributions.

Women do not only give more in dictator games, they also get more as recipients. In a meta-regression confined to those experiments that have explicitly tested recipient gender, this factor alone explains 73.2% of the observed variance, cons .052, female recipient .150***, $N = 39$. If one controls for recipient gender, dictator gender is insignificant, cons .041, female dictator .016, female recipient .143*, $N = 33$, adj.R^2 .7203. Recipient race does not have a significant effect.

Anthropologists have frequently used the dictator game as a technology for quantifying the sociability of developing or indigenous cultures (the classic study is Henrich and Boyd 2005). If one codes Western subject pools as 0, pools from developing countries as 1, and pools from indigenous societies as 2, one establishes a remarkable significant positive effect: meta-regression, cons .275***, society characteristic .053**, $N = 445$, adj.R^2 .024. The more a society is indigenous, the more dictators are willing to share. If one treats the two non-Western origins as categorical variables, seemingly the difference between Western and developing countries disappears, while the difference between developed and indigenous societies is all the stronger: meta-regression, cons .275***, developing country .050, indigenous society .108**, $N = 445$, adj.R^2 .020.

As Fig. 9 demonstrates, this impression is misleading. In Western societies, the typical picture from student populations is replicated. Close to 40% of participants give nothing, less than 20% choose the equal split, some 5% give everything. Both in developing countries and in indigenous societies, giving more than 50% is rare. However, in indigenous societies the equal split is the mode, and giving little is rare. Developing countries are in the middle. Giving nothing is much less frequent than in Western societies, but much more frequent than in indigenous societies. Contribu-
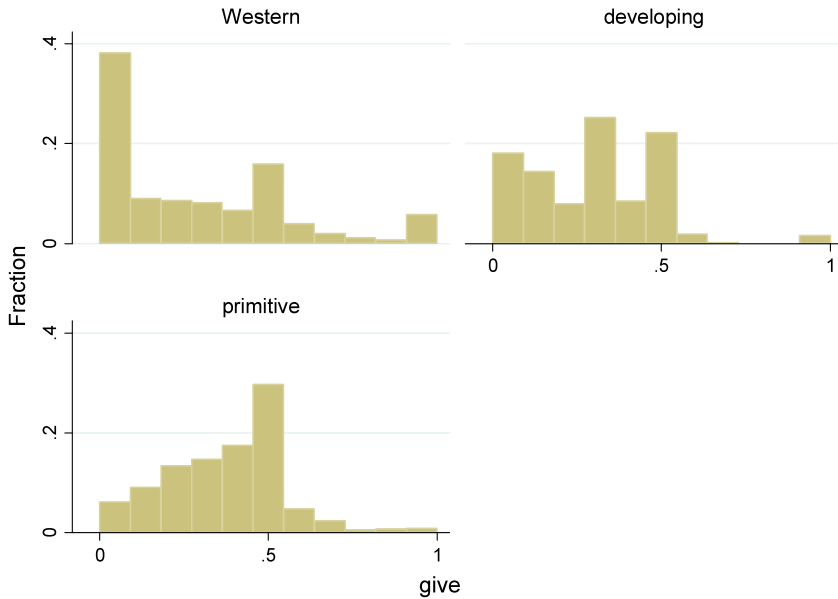
Fig. 9 Society of origin

tion patterns basically distribute over the range [0, .5]. All bilateral comparisons of distributions are statistically significant: Epps Singleton, $p < .0001$.

Age also has a strong effect. If one codes children with 0, students with 1, middle-aged adults with 2, and the elderly with 3, there is a highly significant, substantial effect: meta-regression, cons .187***, age .098***, $N = 445$, adj.R^2 .038. If one treats each age class as a categorical variable, in meta-regression, the behaviour of children is not significantly different from the behaviour of students, while there is a significant difference with respect to the remaining age classes: meta-regression, cons .269***, child .036, middle age .138**, elderly .443***, $N = 445$, adj.R^2 .122. Again distributions are more informative than means. Children are unlikely to give more than half of the pie, and many give less. This explains why there is no difference in means, compared to students. Yet children are much less likely to give nothing. Giving nothing is even rarer in participants of middle age, and it never happens in the elderly. For people of middle age, the equal split is the mode, while for the elderly this is giving everything. All bilateral comparisons of distributions are statistically significant: Epps Singleton, $p < .0001$.

5 Multiple regression

In single regression most effects turns out significant. Yet single regression has little explanatory power. This becomes patent through the measure for the adjusted R^2. In most regressions, it is below .1, implying that more than 90% of the variance remains unexplained. Compared to the meta-analysis of means, which left 97.1% of
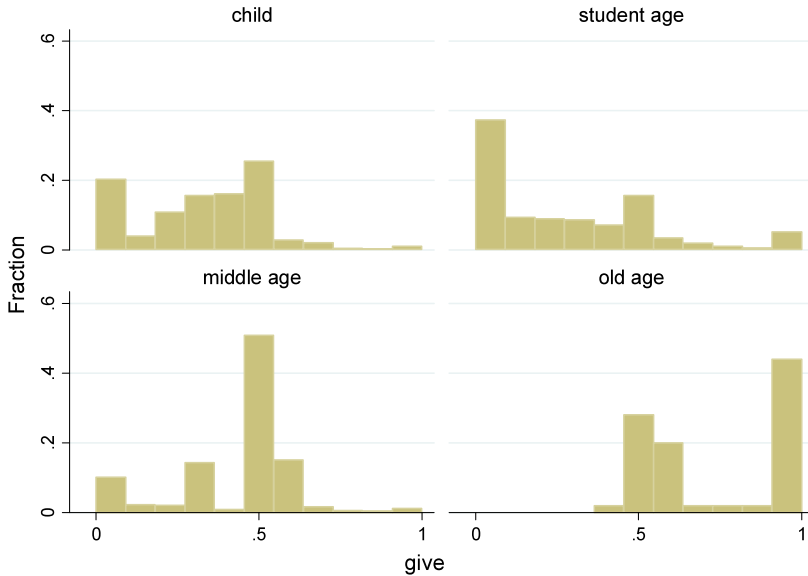
Fig. 10 Effect of age

the variance unexplained, this is only a marginal improvement. Multiple regression has a considerably better fit. Table 1 uses complementary analytic strategies. The complete meta-regression model, i.e. the regression with weighted study means as the dependent variable, explains almost half of the variance, even if one adjusts the R^2 for the fact that one uses 24 explanatory variables.[22]

In meta-regression, not all regressors turn out significant. The degree of uncertainty is insignificant, as is the kind of incentives, dictator identification, the presence of a social cue, dictator-experimenter anonymity, a take option, the use of real money, the fact that the dictator has middle age, and that she comes from a developing country or from a indigenous society. Table 2 collects those regressors that remain at least marginally significant if one controls for all other explanatory variables. Old age has by far the strongest positive effect. A fairly strong effect also stems from having more than one recipient. The remaining three positive variables concern different dimensions of recipient need and legitimacy. By contrast, if the recipient's desire is only weakly legitimate, this strongly reduces giving. This may result from the fact that the recipient had an endowment in the first place, or that the dictator had to earn the pie. In demographic terms, children and students give less. Dictators also exploit concealment options. Finally they give less if the game is repeated, if the action space

---

[22]In these regressions three explanatory variables are left out since including them would severely reduce the number of observations: gender, race, and stakes. To check for potential problems of multicollinearity, I have calculated all pairwise correlations. None is above .3711, and most are considerably below. I also have calculated variance inflation factors. The highest is 4.56. Both suggests that collinearity is not an issue.

Table 1 Multiple regressions. Codes are explained in Sect. 4

	Meta-regression	Ols no treat dummies	Ols treat dummies	Tobit	Logit 0	Truncated Ols	Logit 50	Logit 100
limited action space	−0.062+	0.038	−0.029	0.027	0.281	0.131**	−1.093**	0.05
degree of uncertainty	−0.036	−0.07	−0.654***	−0.068	0.303	−0.2	0.584	0
incentive	−0.01	−0.040**	−0.052**	−0.067***	0.418**	−0.004	−0.305*	−0.142
repeated game	−0.064**	−0.018	−0.130*	−0.024	0.037	−0.025	−0.566***	0.409
group decision	−0.054+	−0.108*	−0.015	−0.103+	−0.07	−0.210***	−0.097	0
identification	0.042	0.049*	0.243***	0.077*	−0.522*	0.035	0.575**	0.016
social cue	0.005	−0.031	0.225***	−0.033	−0.026	−0.06	0.062	−0.343
concealment	−0.065*	−0.028	−0.123***	−0.035	0.083	−0.032	−0.147	−0.14
double blind	−0.024	−0.021	−0.262***	−0.023	0	−0.028	0.185	−0.447
take option	0.067	−0.037	−0.038	−0.083	0.443	0.043	0.371	0.125
deserving recipient	0.086***	0.168***	0.534***	0.226***	−0.913***	0.117*	−0.520*	1.830***
recipient earned	0.128*	0.169***	−0.132*	0.275***	−0.922**	0.220***	1.006***	0.866
efficiency recipient	0.026+	0.007	−0.180***	0.022	−0.294**	−0.026	−0.394*	0.875**
multiple recipients	0.148***	0.038*	−0.065+	0.028	0.187*	0.125***	−0.611***	0.539+
recipient endowment	−0.173***	−0.058	0.204**	−0.147	0.886*	0.062	−0.795+	0.248
dictator earned	−0.174***	−0.191***	−0.126***	−0.374***	1.489***	−0.213***	−1.556***	0
real money	0.025	0.062+	0.216***	0.076+	−0.141	0.092+	−0.443*	2.058***
degree of social proximity	−0.053***	0.002	0.191***	0.005	0.036	0.017	−0.832**	2.675***
student	−0.104**	−0.220***	0.216***	−0.233***	0.456	−0.301***	−0.076	−2.185***
child	−0.117**	−0.195**	−0.007	−0.172*	−0.311	−0.385***	0.932+	−6.404***
middle age	0.001	−0.044	0.435***	0.031	−1.475+	−0.258***	1.888**	−2.434***
old age	0.336***	0.189**	0.181+	0.247**	0	0	1.293*	−1.384*
developing country	0.015	0.01	0.231***	0.042	−0.617*	−0.052+	0.211	−1.329*
indigenous society	−0.009	−0.098	−0.047	−0.027	−1.847*	−0.329***	0.492	−4.147***
constant	0.416***	0.518***	0.210*	0.460***	−1.377+	0.651***	0.556	−5.241***
N	603	20813	20813	20813	20663	13298	20813	19402
adj. R^2 / pseudo R^2	.485	0.149	(0.242)	0.105	0.074		0.077	0.268

is limited, if groups decide and, surprisingly, the smaller the social distance to the recipient.

If one replaces study means by the original data (column 2 of Table 1), one works with 20,813 instead of 445 datapoints. One should expect that this makes it easier to establish significance; standard errors are the standard deviation, divided by the square root of N, after all. This indeed holds for dictator identification and for the use of real coins and notes. Both now also have their positive effect if one controls for all the other explanatory variables. Interestingly, when controlling for alternative explanations the effect of incentivising individual choices turns out negative; dictators give less. The equally negative effect of group decision is now significant at conventional levels. Two more regressors become significant at a more demanding level. Yet there are also three regressors that were significant in meta-regression, and that become insignificant when using individual choices. Two more regressors are only significant at a lower level. Yet if one reduces the sample of the meta-regression to those studies that reported distribution information, one finds the same or even lower significance levels. The differences are thus due to sample attrition. As usual, working with individual rather than aggregate data leads to considerably more unexplained variance; means tend to exhibit greater regularity.

Model 3 also works with the original data, but adds treatment dummies to the model. This has an advantage and a disadvantage. Any unexplained heterogeneity across studies is "dummied out". This yields many more significant regressors. Participants give less the higher the uncertainty. They give more if there is a social cue. They give less when double blind procedures are used. They give more if they handle real money, when they have middle age or when they come from a developing country. All of this is as expected. It is interesting that, in this specification, students give substantially more. It is comforting that, in this model, a higher degree of social proximity has the expected positive sign (not a negative one, as in meta-regression). Yet other regressors are surprising. The negative effect of being a child is no longer significant, nor is the positive effect of old age. Higher efficiency for the recipient has a counterintuitive negative effect, and the fact that the recipient already holds an endowment has an equally counterintuitive positive effect. All these surprising effects result from the fact that almost all the variance with respect to these explanatory variables results from variance between, not within treatments. Therefore in this specification the effects picked up by both meta-regression and OLS without treatment fixed effects go into the treatment dummies. For the explanatory variables no variance is left, or it comes from a few outliers. Consequently eventually both specifications of the OLS model should be seen as complementary, not as mutually exclusive.

As Fig. 2 demonstrates, a fairly substantial fraction of dictators maximises individual money earnings and gives zero. One may therefore wonder whether OLS is the appropriate functional form. One may instead consider the individual data to be left censored. The standard treatment of censored data is a Tobit model. It assumes that there are dictators who would have given a negative amount, had they not been prevented from this by experimental design. As the experiments demonstrate that give participants a take option, this assumption is tenable. If they have a take option, quite a few participants seize it. This even holds if they can both give and take

(Bardsley 2008). In the Tobit model, significance levels look very similar to OLS (without treatment dummies). However, a number of coefficients become substantially larger.[23] This in particular holds for the fact that the recipient is deserving, and that she or the dictator earned the pie. Also the effect of old age becomes more pronounced.

An equally plausible analytic approach assumes that the decision to make a positive contribution, and the decision how much to give, conditional on the willingness to give at all, are two separate processes. Then, a hurdle model is appropriate. It first analyses the binary decision to give or not to give with a logit model. For the size of positive contributions, conditional on giving anything, it uses OLS, adjusting the distribution assumption to the fact that observations are taken from a truncated normal distribution (for background see McDowell 2003).[24]

From this exercise, one learns that the willingness to give anything is indeed not driven by the same forces as the choice of the size of the contribution to the recipient. Let us first compare the logit component with the least squares model of all data. One now finds a strongly significant effect of the factor by which dictator generosity translates into recipient benefit. The larger this factor, i.e. the higher the efficiency of a dollar given, the less likely a dictator is to give nothing. Conversely, the larger the upfront endowment of the recipient, the more a dictator is inclined to keep everything. In this model, one also reestablishes the difference between Western and developing countries or indigenous societies. In developing countries, dictators are less likely to give nothing. In indigenous societies this is extremely unlikely. Finally there is a fairly pronounced, although only weakly significant negative effect of middle age.

For the truncated least squares model, two comparisons are of interest: with unconditional least squares, and with the logit model. The first comparison reveals that dictator-experimenter anonymity significantly reduces the willingness to give, as does middle age and being a member of a indigenous society, while a limited action space increases conditional giving; none of these four effects was significant in the unconditional model. For another four regressors, the significance level goes up, compared to the unconditional model. This holds true for the negative effect of group decision and of being a child, and for the positive effect of the recipient having earned her endowment, and of the dictator being faced with multiple recipients.

Interestingly, in the very same respects truncated OLS also differs from the hurdle equation. The logit model is either insignificant in these respects, or it has a lower level of significance. This suggests that the hurdle model indeed outperforms OLS. Apparently, the decision to give nothing is indeed driven by forces that differ from the conditional choice of a contribution level. Three regressors only matter for the conditional choice, not for the willingness to give at all. Students and groups make smaller contributions. If dictators handle real coins and notes, this slightly increases their contributions, conditional on their willingness to give. By contrast, a number

---

[23]To allow for direct comparisons with the OLS model, the Tobit model reports marginal effects, not the coefficients of the latent variable.

[24]The number of observations is, of course, confined to those participants who have made a positive contribution.

of explanatory factors only matter for the willingness to give anything, not for the size of the contribution. This holds for the question whether each and every choice is paid out, for the factor by which a dollar given translates into a dollar received, for the question whether the recipient has her own endowment, and whether the dictator comes from a developing country. Actually, no more than three regressors significantly explain both the willingness to make a positive contribution and its size. If the recipient has earned the pie, she is more likely to get something, and the size of the contribution increases. Interestingly, the remaining two regressors that are significant in both models predict conflicting effects. Being of middle age and being a member of a indigenous society reduces the likelihood of giving nothing, but it also reduces the size of contributions. Such dictators think they should give something, but they give rather little.

While giving nothing is the mode, the distribution of dictator giving may also be described as tri-modal, with the equal split and giving everything as the two other peaks. This invites two more logit models. The first tests the hypothesis that a dictator goes for the equal split, against the nul that she makes another decision. This model too is revealing. The equal split is most popular in middle age and in indigenous societies. If the recipient identifies the dictator, this also pushes the dictator towards the equal split. By contrast, if the game is repeated, or if there are multiple recipients, the equal split becomes considerably less likely.

The final logit model tests the hypothesis that the dictator gives everything, against the nul that she makes a different choice.[25] The strongest force pulling dictators into this direction is social proximity. Handling real coins and notes and knowing that the recipient is deserving also make total generosity much more likely. If a dollar given is worth more to the recipient, this also makes it more likely that the dictator gives all. By contrast, the elderly, dictators from developing countries and students are quite unlikely to donate the entire pie. This is extremely unlikely in middle aged dictators, in members of indigenous societies, and in children. Note that this final model explains a larger fraction of the variance than any other model working with the original data.

Table 2 summarises the evidence. In this table, if a cell is empty, the respective regressor does not have a significant effect. + signs indicate a significant positive effect, − signs indicate a significant negative effect. If either meta-regression or least squares at the level of individual data demonstrates a positive effect, one should expect less dictators to give nothing. This expectation is borne out for deserving recipients and for recipients who have earned their endowment. By the same logic, dictators who have earned the pie give less and are more likely to give nothing. One has a similar picture if dictators handle real coins and notes and if each individual choice is paid, although here the overall effect is insignificant in meta-regression. The logic is violated by the effect of deciding for multiple recipients. The overall effect is positive. Nonetheless, dictators are more likely to give nothing. Yet they are also more likely to give everything. This manipulation thus polarises choices.

---

[25]The number of observations is smaller since with 3 independent variables participants never gave everything.

Table 2 Significant effects in multiple regression

	Meta-reg	Ols no treat dummies	Ols treat dummies	Tobit	Logit 0	Truncols	Logit 50	Logit 100
deserving recipient	+	+	+	+	−	+	−	+
recipient earned	+	+	−	+	−	+	+	
old age	+	+	+	+			+	−
multiple recipients	+	+	−		+	+	−	+
real money		+	+	+		+		+
identification		+	+	+	−		+	
efficiency recipient	+				+			+
dictator earned	−	−	−	−	−	−	−	
student	−	−	+	−		−	−	−
child	−	−		−		−	+	
group decision	−	−						
incentive	−	−	−	−	+	−		
limited action space	−					−	−	
repeated game	−		−				−	
concealment	−		−					
degree of social proximity	−		+				−	+
recipient endowment	−		+		+			
middle age			+		−		+	−
indigenous society					−	−		−
developing country			+		−	−		−
double blind			−					
degree of uncertainty								
social cue			+					
take option								
constant	+	+	+	+	−	+		−

Being a student, being a child or deciding as a group have a significant overall effect, but do not significantly influence the decision to make a positive contribution. If the action space is limited, the game is repeated, the dictator can conceal her action, social proximity is different from being a member of the same subject pool or if the recipient has a positive endowment, this explains differences in study means, but it does not explain individual choices (in the reduced sample). Having middle age, being a member of a indigenous society or of a developing country or providing dictator-experimenter anonymity does not explain overall effects. Yet these regressors become significant if one separately analyses the decision to make a positive contribution, and its size. Manipulating the degree of uncertainty, providing the dictator with a social cue, or giving her a take option does not have a significant effect once one controls for the other explanatory factors.

6 Conclusion

The main substantive results of this study can be read off Table 2 and Fig. 11. The more consistent effects are across specifications of the statistical model (Table 2), the more firmly an effect is established. The following effects are very robust: If the recipient is deserving, she gets more. If the dictator has old age, she gives more. If the dictator is identified, she is more generous. Children give less, as do groups. Other effects only show up when means are replaced by distributions. If one does, one finds that dictators give more when handling real money, and when they are identified. If choices are incentivised, this reduces generosity. Yet other effects only become visible if one further controls for unobserved heterogeneity. One then finds that dictators give more when they get a social cue, when they are middle aged (rather
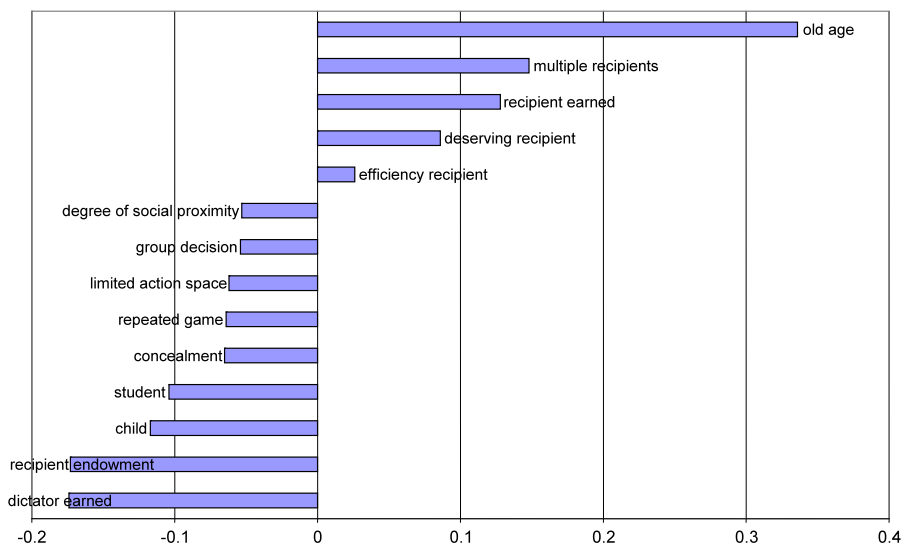

Fig. 11 Significant regressors in the complete model meta-regression

than being a student), or when they come from a developing country. Further effects were already visible in meta-regression, but become apparent in the original data only when controlling for unobserved heterogeneity. One then finds that dictators give less if the game is repeated, and if they have a concealment option.

For some effects, it is important to separately analyse the decision to give anything, and the amount given, conditional on passing the hurdle. One then for instance better understands the effect of identification. If dictators cannot hide behind anonymity, they feel urged to give something; they do not feel obliged to give more. The same holds if a dollar given is multiplied before it reaches the recipient. If dictators had to earn the pie, they are significantly more likely to keep everything. The same holds, once their choices are incentivised and if the recipient also has an endowment. Middle aged dictators, and those from developing countries or indigenous societies are significantly less likely to keep their entire endowments.

Effect sizes are visualised in Fig. 11. By far the biggest effect is demographic. Pensioneers give more than 30% more than average, children give more than 10% less. Next in line are distributional concerns. If the dictator had to earn the pie, or if the recipient has her own endowment, generosity drops by almost 20%. By contrast, the recipient keeps more than 10% more if she had to earn the pie, and she receives some 8% more if she is visibly deserving. Situational features have smaller effects. If there are multiple recipients, generosity goes up by more than 10%. If the game is repeated, if the action space is limited, or if groups decide, generosity goes down by less than 10%.

The picture is thus complicated. There is, of course, also a simple message. Originally, Daniel Kahneman and his co-authors had invented the dictator game to refute the income maximisation assumption of economics textbooks. If one translates the assumption into the categorical hypothesis that all humans maximise income, their first experiment was sufficient to falsify the hypothesis. After 25 years of experimentation, the original result has not been refuted. Actually, in the sample of those experiments that provide distribution information, the probability of giving a positive amount, and hence violating the income maximisation hypothesis, comes close to the original result. In Kahneman's experiment, 74% did not seize the option to gain an additional dollar. In later dictator game experiments, 63.89% of all participants, i.e. 13,298 of 20,813 dictators, made a positive contribution.

Yet if this was everything one wanted to test, 129 experiments would certainly have been an overkill. The research question becomes richer, and more meaningful for understanding social interaction, if one turns the finding upside down. While 63.89% violate the income maximisation hypothesis, 36.11% do not. Hence more than a third of a typical population does indeed consist of subjects who have no reticence to leave a recipient with nothing, although this recipient is at their mercy. Clearly, generosity is not a human universal. Moreover, those who are willing to make a donation do neither give everything, nor do they split the pie equally. On average, they give 42.64%. Hence even those who in principle are generous to a degree exploit the opportunity to their advantage. Even generous subjects thus tend to have a selfish side.

Consequently, the ample evidence on dictator games is better read as an exercise in exploring human heterogeneity. The radically simple design of the game makes

it a powerful tool for the systematic variation of conditions that moderate sociality. As differential psychology, using different experimental approaches, has established long ago (Ross and Nisbett 1991), (perceived) situation and personality interact. In and of themselves, neither situational nor personality variables suffice to predict behaviour. One must understand how a person with a certain personality trait reacts to the situation, as she perceives it. Dictator games are helpful for exploring human sociality, because the situation can be tightly controlled, because the simple game can be played with all classes of participants and, of course, because decisions are incentivised.

Ironically, the radically simple dictator game provides ample evidence that human sociality is far from simple. The existing evidence is best compared with a map. After 25 years of expeditions into human sociality, some major roads have been safely documented. Yet aside those roads, there are still many white spots. It is therefore safe to predict that dictator game experimentation will continue to thrive. To date, it is an open question though whether experimenters will ever be able to draw a near complete map of sociality. It is even less clear whether those relying on this knowledge in the interest of improving their individual well-being or of making better policy will be able to correctly infer where an interaction partner is located in this landscape. The better the driving forces of human behaviour are understood, the more it becomes obvious that institutions are not only needed for overcoming antisocial motives. They are at least as important for making behaviour predictable (Engel 2005).

References

Aguiar, F., Branas-Garza, P., Cobo-Reyes, R., Jimenez, N., & Miller, L. M. (2009). Are women expected to be more generous? *Experimental Economics*, *12*, 93–98.

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

Bardsley, N. (2008). Dictator game giving. Altruism or artefact? *Experimental Economics*, *11*, 122–133.

Blackwell, C. (2007). A meta-analysis of tax compliance experiments. http://aysps.gsu.edu/isp/files/ISP_CONFERENCES_TAX_COMPLIANCE_AND_EVASION_BLACKWELL.pdf.

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*, 477–486.

Brañas-Garza, P., Cobo-Reyes, R., Espinosa, M. P., Jiménez, N., Kovárík, J., & Ponti, G. (2010). Altruism and social integration. *Games and Economic Behavior*, *69*(2), 249–257.

Brosig, J., Riechmann, T., & Weimann, J. (2007). Selfish in the end? An investigation of consistency and stability of individual behavior. http://mpra.ub.uni-muenchen.de/2035/1/MPRA_paper_2035.pdf.

Camerer, C. F. (2003). *Behavioral game theory. Experiments in strategic interaction*. New York: Sage.

Cárdenas, J. C., & Carpenter, J. (2008). Behavioural development economics: lessons from field labs in the developing world. *Journal of Development Studies*, *44*, 311–338.

Croson, R. T. A., & Marks, M. (2000). Step returns in threshold public goods. A meta- and experimental analysis. *Experimental Economics*, *2*(3), 239–259.

Dana, J., Daylian, M. C., & 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*, 193–201.

DerSimonian, R., & Laird, N. (1986). Meta-analysis in clinical trials. *Controlled Clinical Trials*, *7*, 177–188.

Druckman, D. (1994). Determinants of compromising behavior in negotiation. A meta-analysis. *Journal of Conflict Resolution*, *38*(3), 507–556.

Eckel, C. C., & Grossman, P. J. (1998). Are women less selfish than men?: Evidence from dictator experiments. *Economic Journal*, *108*, 726–735.

Eckel, C. C., Grossman, P. J., & Johnston, R. M. (2005). An experimental test of the crowding out hypothesis. *Journal of Public Economics*, *89*, 1543–1560.

Eichenberger, R., & Oberholzer-Gee, F. (1998). Rational moralists: the role of fairness in democratic economic politics. *Public Choice*, *94*, 191–210.

Engel, C. (2005). *Generating predictability. Institutional analysis and institutional design*. Cambridge: Cambridge University Press.

Engel, C. (2007). How much collusion? A meta-analysis on oligopoly experiments. *Journal of Competition Law and Economics*, *3*, 491–549.

Epps, T. W., & Singleton, K. J. (1986). An omnibus test for the two-sample problem using the empirical characteristic function. *Journal of Statistical Computation and Simulation*, *26*, 177–203.

Forsythe, R., Horowitz, J. L., Savin, N. E., & Sefton, M. (1994). Fairness in simple bargaining experiments. *Games and Economic Behavior*, *6*, 347–369.

Frey, B., & Bohnet, I. (1995). Institutions affect fairness. *Journal of Institutional and Theoretical Economics*, *151*, 286–303.

Goeree, J. K., McConnell, M. A., Mitchell, T., Tromp, T., & Yariv, L. (2010). The 1/d law of giving. *American Economic Journal: Microeconomics*, *2*(1), 183–203.

Goerg, S. J., & Kaiser, J. (2009). Non-parametric testing of distributions—the epps-singleton two-sample test using the empirical characteristic function. *Stata Journal*, *9*, 454–465.

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

Harbord, R. M., & Higgins, J. P. T. (2008). Meta-regression in stata. *Stata Journal*, *8*, 493–519.

Harless, D. W., & Camerer, C. F. (1994). The predictive utility of generalized expected utility theories. *Econometrica*, *62*, 1251–1289.

Henrich, J., & Boyd, R. (2005). Economic man' in cross-cultural perspective. Behavioral experiments in 15 small-scale societies. *Behavioral and Brain Sciences*, *28*, 795–815.

Hoffman, E., McCabe, K., Shachat, K., & Smith, V. L. (1994). Preferences, property rights, and anonymity in bargaining games. *Games and Economic Behavior*, *7*, 346–380.

Hopfensitz, A. (2009). Previous outcomes and reference dependence. A meta study of repeated investment tasks with and without restricted feedback. http://mpra.ub.uni-muenchen.de/16096/1/Hopfensitz_2009.pdf.

Houser, D., & Schunk, D. (2009). Social environments with competitive pressure: gender effects in the decisions of German schoolchildren. *Journal of Economic Psychology*, *30*, 634–641.

Huck, S., Normann, H.-T., & Oechssler, J. (2004). Two are few and four are many. Number effects in experimental oligopolies. *Journal of Economic Behavior & Organization*, *53*, 435–446.

Jones, G. (2008). Are smarter groups more cooperative? Evidence from prisoner's dilemma experiments, 1959–2003. *Journal of Economic Behavior & Organization*, *68*(3-4), 489–497.

Kahneman, D., Knetsch, J. L., & Thaler, R. (1986). Fairness and the assumptions of economics. *Journal of Business*, *59*, S285–S300.

Leider, S., Möbius, M. M., Rosenblat, T., & Quoc-Anh, D. (2009). What do we expect from our friends? https://mercury.smu.edu.sg/rsrchpubupload/15345/ExpectFriends.pdf.

Levitt, S. D., & List, J. A. (2007). What do laboratory experiments measuring social preferences reveal about the real world? *Journal of Economic Perspectives*, *21*, 153–174.

McDowell, A. (2003). From the help desk. Hurdle models. *Stata Journal*, *3*, 178–184.

Oosterbeek, H., Sloof, R., & van de Kuilen, G. (2004). Cultural differences in ultimatum game experiments. Evidence from a meta-analysis. *Experimental Economics*, *7*, 171–188.

Oxoby, R. J., & Spraggon, J. (2008). Mine and yours. Property rights in dictator games. *Journal of Economic Behavior & Organization*, *65*, 703–713.

Percoco, M., & Nijkamp, P. (2009). Estimating individual rates of discount. A meta-analysis. *Applied Economics Letters*, *16*(12), 1235–1239.

Prante, T., Thacher, J. A., & Berrens, R. P. (2007). Evaluating Coasean bargaining experiments with meta-analysis. *Economics Bulletin*, *3*(68), 1–7.

Rigdon, M., Ishii, K., Watabe, M., & Kitayama, S. (2009). Minimal social cues in the dictator game. *Journal of Economic Psychology*, *30*, 358–367.

Ross, L., & Nisbett, R. E. (1991). *The person and the situation. Perspectives of social psychology*. New York: McGraw-Hill.

Selten, R. (1967). Die Strategiemethode zur Erforschung des eingeschränkt rationalen Verhaltens im Rahmen eines Oligopolexperiments. *Beiträge zur experimentellen Wirtschaftsforschung*. Ernst Sauermann. Tübingen: Mohr: 136–168.

Stanley, T. D. (2001). Wheat from chaff. Meta-analysis as quantitative literature review. *Journal of Economic Perspectives*, *15*(3), 131–150.

Thompson, S. G., & Pocock, S. J. (1991). Can meta-analyses be trusted? *Lancet*, *338*(8775), 1127–1130.

Weizsäcker, G. (2010). Do we follow others when we should? A simple test of rational expectations. *American Economic Review*, *100*, 2340–2360.

Zelmer, J. (2003). Linear public goods. A meta-analysis. *Experimental Economics*, *6*, 299–310.