ORIGINAL PAPER

Invariance of equilibrium to the strategy method II: experimental evidence

Daniel L. Chen[1] · Martin Schonger[2]

Abstract

The strategy method (SM) is, in practice, subject to a possibly severe economic-theoretical bias. Although many studies utilize SM to examine responses to rare or off-equilibrium behaviors unattainable through direct elicitation (DE), they ignore the fact that the strategic equivalence between SM and DE holds for the monetary payoff game but not the game participants actually play, which is in terms of utilities. We report three results. First, failing to account for estimation bias when decisions at one information set can influence utility at another may result in significant differences in decision-making. Second, the magnitude of this bias can be substantial, comparable to other measured treatment effects. Third, minor interventions targeting salience can amplify these differences similarly, causing treatment effects to differ significantly between SM and DE, even reversing in direction. These findings emphasize the need for reconsideration of the SM's reliability for economic research.

✉ Daniel L. Chen
daniel.chen@iast.fr

Martin Schonger
mschonger@ethz.ch

[1] Toulouse School of Economics, Institute for Advanced Study in Toulouse, University of Toulouse Capitole, Toulouse, France

[2] Center for Law and Economics, ETH Zurich, Zurich, Switzerland

Ⓐ Springer

1 Introduction

The study of human decision-making has long been a cornerstone of economics, but accurately measuring preferences has proven to be a complex challenge. One method gaining popularity in recent years is the strategy method (SM), which involves asking participants to indicate their choices at all information sets, enabling researchers to compare decisions at different points in a given scenario. However, while SM offers several advantages over traditional direct elicitation (DE), it also has its limitations. A fundamental distinction between the two methods is that the information sets for decision nodes differ, which can lead to different inferences. Nonetheless, SM remains a powerful tool for economists seeking to deepen their understanding of human behavior and the forces that shape it.

SM consists of asking participants to indicate their choices at all information sets rather than only those actually reached. One then compares the differences in decisions at different information sets. For example, to identify the effect of a low offer in an ultimatum game, one might compare the changes in decisions for the low-offer information set with the decisions for the high-offer information set. The appeal of SM comes from its simplicity as well as its potential to elucidate the equilibria that are actually played when theoretical models indicate there are multiple equilibria. SM also has the potential to circumvent many of the endogeneity problems that arise in estimating preferences when making comparisons between heterogeneous individuals.

Extant empirical research tends to rely on the behavioral validity of SM (Fischbacher et al., 2012). Brandts and Charness (2011, p. 376) write that, "according to the standard game-theoretic view, the strategy method *should* yield the same decisions as the procedure involving only observed actions" and provide empirical evidence against the claims in the literature. Chen and Schonger (2023) summarizes the theoretical views and presents a theorem (Moulin, 1986, pp. 84–86) arguing that SM is subject to a possibly *severe economic-theoretical bias*.

As evidence for the relevance of the theorem, we briefly revisit prior meta-analyses and conduct our own meta-analysis of ultimatum game experiments in the Online Appendix. We choose the ultimatum game because it is simple and one of the most employed games in experiments. But since the previous literature has highlighted that complexity is an important factor (Brandts & Charness, 2011), we also consider the three-player prisoners' dilemma.

In the meta-analysis, acceptance rates are 20 percentage points higher in the DE setting than in the SM setting. In the remaining analyses, we conduct our own experiments. First, we randomize whether the respondent, but not the poposer, is in SM or DE to ensure the proposal is the same in both treatments. The DE setting increases acceptances and is equivalent to an offer increase of 34% of endowment. Subsequent experiments allow the proposer to also know if the responder is in the DE or SM setting. Next, we manipulate the salience of off-equilibrium motivations. DE increases acceptance rates in the ultimatum game by 18 percentage points. When off-equilibrium motivations are made salient, the difference increases to 27 percentage points. In total, we report the results of five

🙫 Springer

analyses that all demonstrate the relevance of the theorem. As already mentioned, we do so in the context of simple games, like the ultimatum game and trust game, as well as more complex games, like the three-player prisoners' dilemma. In the trust game, DE respondents return three times the amount SM respondents return. In the three-player prisoners' dilemma, DE affects deductions of defectors.

The last two of our five analyses highlight how treatment effects can significantly differ between SM and DE, while also flipping in sign. When we interpret salience as the treatment effect of interest, we see evidence that salience has a weakly positive treatment effect under DE but is negative under SM. The difference in treatment effects is statistically significant at the 5% or 10% level.

The remainder of the paper is outlined as follows: Section 2 presents an experiment where the ultimatum game respondent is randomized to DE or SM. The appendix presents the experiment where DE versus SM differences extend to the trust game. Section 3 presents an experiment where the ultimatum game is randomized to DE or SM and where off-equilibrium considerations are randomly made salient. Section 4 presents a complex game, the three-player prisoners' dilemma. Section 5 concludes.

2 Ultimatum game: DE versus SM for respondent

2.1 Design

This study used MTurk. We first asked MTurk subjects to transcribe three paragraphs of text[1] to reduce the likelihood of their dropping from the study after seeing treatment—a technique to minimize differential attrition that may affect causal inference when using MTurk subjects (Chen & Reinhart, 2022; Chen & Horton, 2016; Chen & Yeh, 2010; Chen et al., 2017).[2] After the lock-in task, subjects have an opportunity to split with the recipient a 50 cent bonus (separate from the payment they received for data entry), up to 23 times the expected wage.[3] We had 156

---

[1] A sample paragraph of data entry was a Tagalog translation of Adam Smith's *The Wealth of Nations*: Kaya sa isip o diwa na tayo ay sa mga ito, excites ilang mga antas ng parehong damdamin, sa proporsyon ng kasiglahan o dulness ng kuru-kuro. Ang labis na kung saan sila magbuntis sa kahirapan ng mga wretches nakakaapekto sa partikular na bahagi sa kanilang mga sarili ng higit pa sa anumang iba pang; dahil sa takot na arises mula sa kathang isip nila kung ano ang kani-kanilang mga sarili ay magtiis, kung sila ay talagang ang wretches kanino sila ay naghahanap sa, at kung sa partikular na bahagi sa kanilang mga sarili ay talagang apektado sa parehong miserable paraan. Ang tunay na puwersa ng mga kuru-kuro na ito ay sapat na, sa kanilang mga masasaktin frame, upang gumawa ng na galis o hindi mapalagay damdam complained ng.

[2] This task was sufficiently tedious that no one was likely to do it "for fun," and sufficiently simple that all participants could do the task. The source text was machine-translated to prevent subjects from finding the text elsewhere on the Internet.

[3] A paragraph takes about 100 s to enter so a payment of 10 cents per paragraph is equivalent to $86.40 per day. The current federal minimum wage in the Unites States is $58/day. In India, payment rate depends on the type of work done, although the "floor" for data entry positions appears to be about $6.38/day (Payscale, Salary Snapshot for Data Entry Operator Jobs, http://www.payscale.com/research/IN/Job=Data_Entry-_Operator/Salary? Accessed June 17, 2011). In one data entry study, one worker emailed saying that $0.10 was too high and that the typical payment for this sort of data entry was $0.03

subjects split evenly between the role of proposer and respondent and between SM and DE (2×2 design). Instructions are in Online Appendix B.

In the ultimatum game (Fig. B.1), the proposer offers a split of $0.50 between herself and the responder, in increments of $0.05. In the DE treatment, the responder was informed about the amount offered and asked whether she accepts or rejects the offer (Fig. B.2). If accepted, both players received the payoff according to the split proposed by the proposer. If rejected, both players received zero payoff. In the SM treatment, the responder indicated whether she would accept or reject each possible offer without knowing the actual offer. If the responder rejected the offer actually made by the proposer, neither player received any bonus. The responder's behavior can be characterized by a rejection *threshold*, the minimum offer the responder is willing to accept (Fig. B.3). The proposer did not know the method of elicitation for the responder in order to hold proposer's decisions constant. We are interested in the average treatment effect of DE versus SM on the responder.

2.2 Results

Table 1 regresses an indicator for whether or not the ultimatum game offer was accepted on the treatment indicator, SM, using a linear probability model. Results are robust to using a probit specification. While there were 20 percentage points fewer acceptances in SM ($p < 0.1$) (Column 1), the effect becomes 22 percentage points and more significant ($p < 0.05$) when controlling for the amount offered (Column 2).[4] For each additional $0.01 offered, the acceptance rate increases by 2 percentage points ($p < 0.001$). In terms of magnitude, DE is equivalent to an additional 17 cents offer in a 0–50 ultimatum game, or roughly 34% of endowment. Including an interaction between offer and SM yields a significantly greater association of 1.7 percentage points acceptance rate per $0.01 offer amount ($p < 0.1$) (Column 3), which is analogous to what was found in the survey of prior literature in Online Appendix A.

3 Ultimatum game: DE versus SM and low versus high salience

3.1 Design

We chose to run our remaining studies in the lab, which may be a more controlled setting than MTurk. In this study, we ran the lab experiment at the MaXLab following their standard procedures in Magdeburg and using oTree (Chen et al., 2016). We collected data on 418 subjects across 16 experimental sessions (instructions

---

Footnote 3 (continued)

cents per paragraph. Our study involves $0.20 for a comparable task: reading essentially a single paragraph and making 1 decision, with an additional $0.50 possible.

[4] To put this in perspective, Oosterbeek et al. (2004) conducted a meta-analysis of 66 studies and found that SM reduced acceptance rates by 13%.

Table 1 Ultimatum game offer acceptance

	(1)	(2)	(3)
(Intercept)	0.917***	0.543***	0.784***
	(0.0467)	(0.126)	(0.214)
Strategy method	−0.202*	−0.223**	−0.629*
	(0.0846)	(0.0817)	(0.268)
Offer level		0.0155***	0.00552
		(0.00453)	(0.00814)
Strategy × offer level			0.0165[+]
			(0.00960)
Mean of Y	0.808	0.808	0.808
N	78	78	78

This table examines the determinants of whether the ultimatum game offer is accepted by the second player. Column (1) shows the raw correlation between acceptance and the treatment indicator (SM decision-making). Column (2) also controls for amount offered by the first player. Column (3) examines whether treatment affects the relationship between acceptance and amount offered

Robust standard errors in parentheses

$^{+}p<0.1$, $^{*}p<0.05$, $^{**}p<0.01$, $^{***}p<0.001$

are in Online Appendix B). In this study, the proposer knows the method of elicitation for the responder, so we examine and control for the offer. The endowment was €1.00. Roughly 70 participants were in each of six treatments (3×2 design), listed as follows with abbreviations in parentheses: Direct elicitation (DE)/strategy method (SM)/threshold method (SM-Th)×neutral (neu)/emotional (emo).

We introduce two variants of SM. In one variant, subjects report the *threshold* (where the responder had to state the minimum level of the offer that she would accept), and in another, they report their *strategy* (where the responder had to decide whether she would accept every theoretical offer that could be made by the proposer before the actual offer was revealed).[5] We also introduced a cross-cutting treatment to increase the salience of off-equilibrium payoffs (for a total of six possible groups, two emotional settings×three game variants). In the high salience treatment, the experiment changed two words: *proposer → dictator* and *respondent → subject*. The intervention involves only these two words to heighten emotional salience with terms like dictator and subject. If SM versus DE invariance is affected by a few words, the basis for using SM instead of DE would seem fragile. The intervention involves only these two words to heighten emotional salience with terms like dictator and subject.

---

[5] Some may argue that the threshold method is sufficient to capture SM, but many experimental studies document that subjects may have multiple switches when presented with the full strategy method.
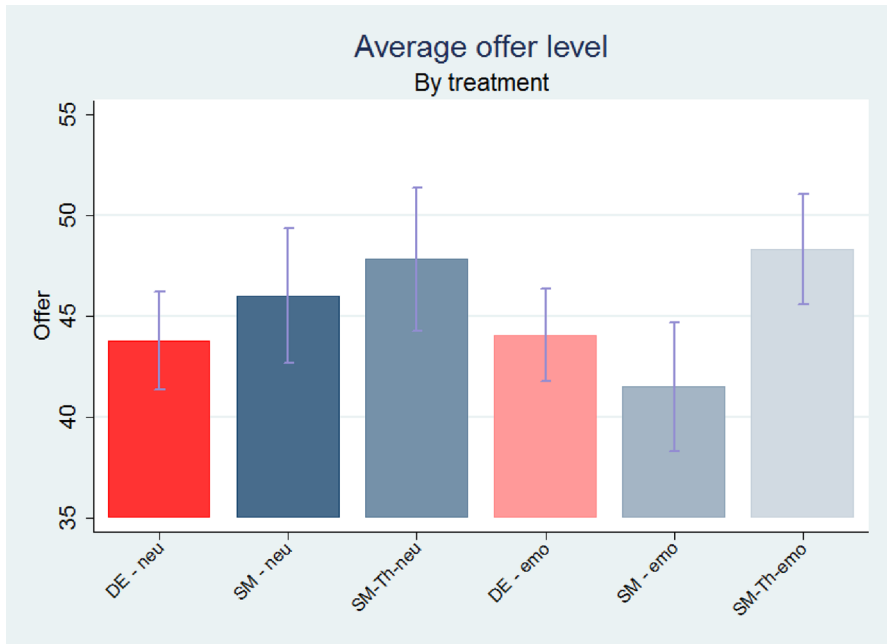
Fig. 1 Ultimatum game: average offer levels for different treatments with 95% confidence intervals

3.2 Results

We cannot reject the null that the proposer's offer is the same across treatments (see Fig. 1). Offers are slightly lower in DE than in SM, which is consistent with proposers being aware that responders are more likely to accept in DE. In Oosterbeek et al. (2004)'s meta-analysis of 66 studies, offered shares were significantly lower with DE by 2% ($p < 0.1$).

Figure 2 reports the natural pattern in ultimatum games: Acceptances are positively associated with the offered amount regardless of treatment. In Column 1, DE shows one observation per subject-pair. In Columns 2 and 3, SM and SM-Th show all possible observations per subject-pair. For the threshold method, we generate an acceptance or rejection for every possible offer. The display is intentionally saturated to illustrate the standard data analysis with SM.

Figure 3 shows that DE results in more acceptances, similar to the survey of prior literature and to our other experiment. In particular, the increase in acceptance is visible in both the low salience (neu) and high salience settings (emo). Increases in acceptance rates under DE are somewhat larger in the high salience setting, which suggests that salience of off-equilibrium considerations may drive some of the differences between DE and SM.[6] Notably, equilibrium behavior does not diverge between SM and SM-Th methods.

---

[6] Regression analyses indicate statistical significance level just shy of 10%.
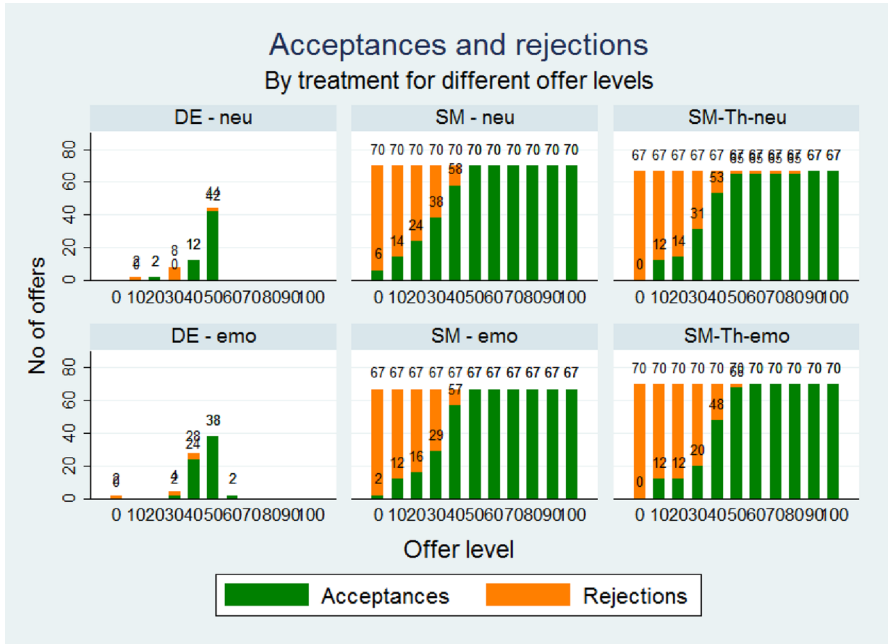
🖄 Springer

Fig. 2 Ultimatum game in laboratory: acceptances and rejections for different offer levels and different treatments
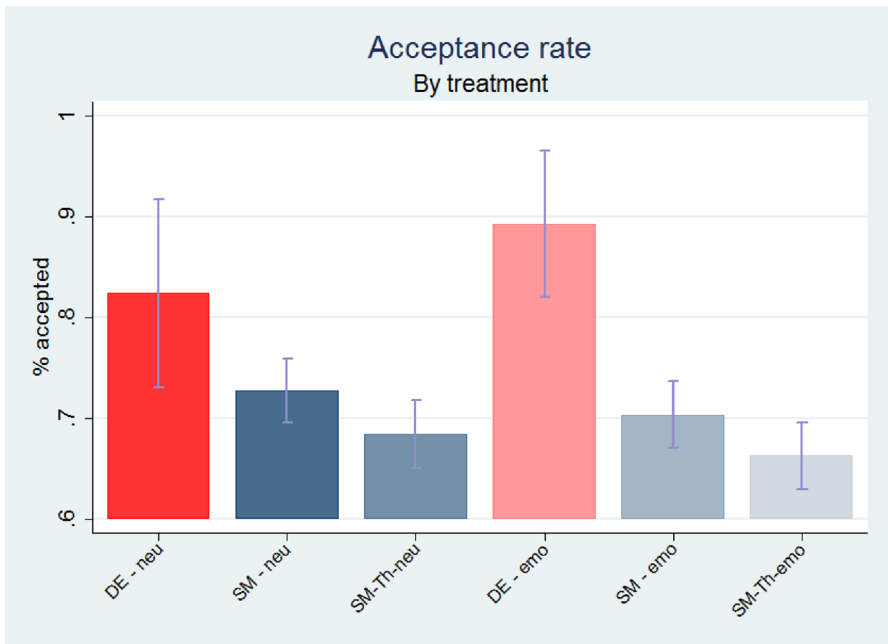

Fig. 3 Ultimatum game in laboratory: acceptance ratio for different treatments with 95% confidence intervals

Table 2 Acceptance rates in laboratory ultimatum game

	(1)	(2)	(3)
(Intercept)	0.824***	0.248***	−0.488+
	(0.0463)	(0.0406)	(0.284)
Strategy method	−0.0963*	−0.162***	0.625*
	(0.0490)	(0.0410)	(0.286)
Threshold method	−0.0434+	−0.0434**	−0.0943*
	(0.0235)	(0.0165)	(0.0399)
Emotions	0.0684	0.0659	0.355
	(0.0587)	(0.0498)	(0.325)
Strategy × emotions	−0.0928	−0.0903+	−0.427
	(0.0632)	(0.0523)	(0.327)
Threshold × emotions	0.00291	0.00291	0.00789
	(0.0336)	(0.0229)	(0.0545)
Offer level		0.107***	0.244***
		(0.00159)	(0.0488)
Strategy × offer level			−0.145**
			(0.0489)
Threshold × offer level			0.00848+
			(0.00465)
Emotions × offer level			−0.0541
			(0.0547)
Strategy × emo × offer			0.0620
			(0.0549)
Threshold × emo × offer			−0.000830
			(0.00634)
Mean of Y	0.702	0.702	0.702
N	3156	3156	3156

This table reports regression results for acceptance rate. SM-Th is treated as a subset of SM (i.e., the strategy dummy is set to 1 also for threshold method observations)

Robust standard errors in parentheses

$^+p<0.1$, $^*p<0.05$, $^{**}p<0.01$, $^{***}p<0.001$

We next examine these relationships in regression analysis. We create indicator variables for every treatment and their interaction (Table 2, Column 1). We include a control for offer level in Column 2 and interactions of offer level and treatment indicators in Column 3.

We begin with a large sample size for illustrative purposes, but later restrict to one observation per subject-pair. The fact that the proposer makes slightly lower offers in DE means that restricting to one outcome would lead to the erroneous conclusion of higher acceptances in SM.[7] Indeed, comparing Columns 1 and 2 shows that the

---

[7] Note that this was not necessary in the online experiment since the proposer did not know whether the responder was in DE or SM.

$\underline{\textcircled{2}}$ Springer

Table 3 Acceptance rates in laboratory ultimatum game

	(1)	(2)	(3)
(Intercept)	0.824***	0.248***	−0.488
	(0.0463)	(0.0406)	(0.284)
Strategy method	−0.117*	−0.184***	0.579*
	(0.0477)	(0.0402)	(0.285)
Emotions	0.0684	0.0659	0.355
	(0.0587)	(0.0497)	(0.325)
Strategy × emotions	−0.0923	−0.0898	−0.425
	(0.0610)	(0.0510)	(0.326)
Offer level		0.107***	0.244***
		(0.00159)	(0.0488)
Strategy × offer level			−0.141**
			(0.0488)
Emotions × offer level			−0.0541
			(0.0547)
Strategy × emo × offer			0.0618
			(0.0548)
Mean of Y	0.702	0.702	0.702
N	3156	3156	3156

This table reports regression results for acceptance rate. SM and SM-Th are pooled together

Robust standard errors in parentheses

$^{+}p<0.1$, $*p<0.05$, $**p<0.01$, $***p<0.001$

difference between SM and DE almost doubles from 9.6 percentage points higher acceptance rate in DE ($p < 0.05$) to 16.2 percentage points ($p < 0.001$) once the offer level is controlled for. This doubling did not occur in the experiment reported in Section B when the offer was added as a control, as the offerer in this experiment was unaware of the respondent's method of elicitation. Note that the high salience treatment further increases the difference in acceptance rates by 9 percentage points ($p < 0.1$) (Column 2). Here, we see that the "Emotions" treatment has significant interaction with SM rather than with SM-Th. If we interpret salience as the treatment effect of interest, we see evidence that salience has no significant treatment effect but is weakly positive under DE but appears weakly negative under SM, and the difference in treatment effects is statistically significant at the 10% level.

Since SM and SM-Th both involve off-equilibrium considerations and render similar results,[8] we pool these treatments in Table 3. Columns 1 and 2 confirm the lower acceptance rate in SM of 12 percentage points ($p < 0.05$) and 18 percentage points ($p < 0.001$) respectively. When we control for offer level (Column 2), this difference is

---

[8] The coefficient on Threshold in Columns 1 and 2 of Table II suggest that SM-Th renders 4.3 percentage points lower acceptance rate than SM ($p<0.1$, $p<0.01$).
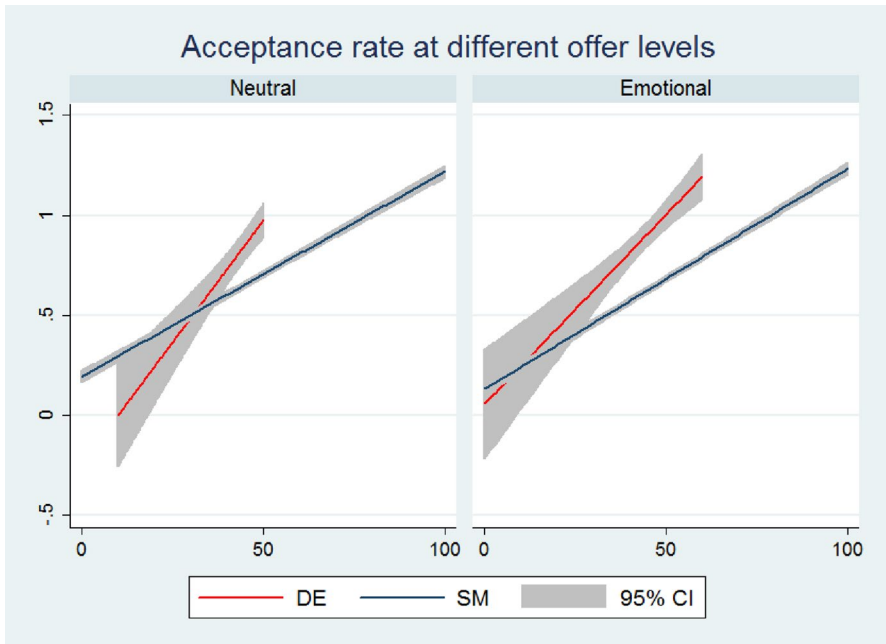
Fig. 4 Ultimatum game in laboratory: acceptance at different offer levels for DE and SM (pooled with threshold method)

highly significant. In Column 3, fully interacting offer with the treatments shows that while 1% of offer is associated with 24 percentage points higher acceptance rates ($p < 0.001$), SM reduces this association by 14 percentage points ($p < 0.01$) in the low salience setting. This interaction differs from the previous experiment and literature. The main result remains that behavior in DE and SM diverges rather than stays invariant.

We can visualize the different correspondence between acceptance rates and offer level for DE and SM in Fig. 4. DE responders are more than twice as sensitive to offers (the regression line for the raw data is red) than SM responders. This is true for both the low and high salience settings.

In sum, DE responders are 18 percentage points more likely to accept than SM responders in the low salience setting and are 27 percentage points more likely to accept in the high salience setting (Table 3, Column 2). Column 3 echoes Figure 4 as the coefficient on the interaction term of Strategy and Offer level suggests that differences between DE and SM responders grows with the offer level.

4 Three-player prisoners' dilemma: DE versus SM and low versus high salience

4.1 Design

In this study, we ran the lab experiment at the WiSo-Experimentallabor[9] following their standard procedures in Hamburg and used oTree (Chen et al., 2016). We collected data from 585 participants across 24 sessions. Subjects play the three-player prisoners' dilemma. We again implement a cross-cutting randomization of high versus low salience for a total of four treatments (SM vs. DE×emo vs. neu). As in the previous study, we designed the salience treatment to avoid framing effects. To manipulate salience, the experiment changed one word (*group* → *team*), and changed the background color (purple → red), when describing the game. The setting with *group* and purple is coded as Emotions = 0 and the setting with *team* and red is coded as Emotions = 1 in the data analysis. In color psychology, red tends to lead to feelings of excitement, while purple tends to calm (Valdez & Mehrabian, 1994; Elliot & Maier, 2014). A team is typically perceived as a group with a common purpose. Again, if invariance between SM and DE is affected by a few words or background color, the basis for using SM instead of DE would seem fragile. Instructions are in the Online Appendix B.

Participants are assigned to matches with three players each. In brief, as in the ultimatum game, DE responders were more cooperative than SM responders. They were less willing to punish non-cooperative first-stage behavior. Differences between DE and SM were affected by salience. We find similar results when we control for the first-stage outcome, restrict the sample to specific first-stage outcomes, or restrict to one observation per subject-first stage outcome. The complete experiment and results are reported in Online Appendix A.

5 Conclusion

Our study suggests that conventional SM estimates may be biased, leading to misleading treatment effects relative to DE. If DE is the gold standard for causal estimates, one possible solution for experiment methods is to collect pilot data that first tests whether SM and DE diverge before collecting additional data using SM. We leave empirical exploration of positive and negative bias for future work.

---

[9] We used a different lab because of the number of subjects we needed.

Data availability Data is available on request.

References

Anderson, L. R., Rodgers, Y. V., & Rodriguez, R. R. (2000). Cultural differences in attitudes toward bargaining. *Economics Letters, 69*(1), 45–54.

Andreoni, J., Castillo, M., & Petrie, R. (2003). What do bargainers' preferences look like? Experiments with a convex ultimatum game. *American Economic Review, 2003*, 672–685.

Angrist, J., & Pischke, S. (2008). *Mostly harmless econometrics: An empiricist's companion*. Princeton: Princeton University Press.

Blount, S. (1995). When social outcomes aren't fair: The effect of causal attributions on preferences. *Organizational Behavior and Human Decision Processes, 63*(2), 131–144.

Bornstein, G., & Yaniv, I. (1998). Individual and group behavior in the ultimatum game: Are groups more "rational" players? *Experimental Economics, 1*(1), 101–108.

Brañas-Garza, P., Cobo-Reyes, R., & Domínguez, A. (2006). "Si él lo necesita": Gypsy fairness in vallecas. *Experimental Economics, 9*(3), 253–264.

Brandts, J., & Charness, G. (2011). The strategy versus the direct-response method: A first survey of experimental comparisons. *Experimental Economics, 14*, 375–398.

Cameron, L. A. (1999). Raising the stakes in the ultimatum game: Experimental evidence from Indonesia. *Economic Inquiry, 37*(1), 47–59.

Carter, J. R., & Irons, M. D. (1991). Are economists different, and if so, why? *The Journal of Economic Perspectives, 5*(2), 171–177.

Chen, D. L., & Horton, J. J. (2016). Are online labor markets spot markets for tasks? A field experiment on the behavioral response to wage cuts. *Information Systems Research, 27*(2), 403–423. TSE working paper no. 16-675.

Chen, D.L., & Reinhart, E. (2022). Markets and morality: Do free markets corrode moral values? *Journal of Law, Economics, and Organization*, conditionally accepted.

Chen, D.L., & Schonger, M. (2023). Invariance of equilibrium to the strategy method: Theory. *Journal of the Economic Science Association*, forthcoming.

Chen, D. L., Yeh, S. (2010). *Does obscenity law corrode moral values and does it matter? Evidence from 1958–2008*. Technical report. Duke University mimeo.

Chen, D. L., Yeh, S., Levonyan, V., & Yeh, S. (2017). Do policies affect preferences? Evidence from random variation in abortion jurisprudence. Working paper.

Chen, D. L., Yeh, S., Schonger, M., & Wickens, C. (2016). oTree—An open-source platform for laboratory, online, and field experiments. *Journal of Behavioral and Experimental Finance, 9*, 88–97.

Croson, R. T. A. (1996). Information in ultimatum games: An experimental study. *Journal of Economic Behavior and Organization, 30*(2), 197–212.

Elliot, A. J., & Maier, M. A. (2014). Color psychology: Effects of perceiving color on psychological functioning in humans. *Annual Review of Psychology, 65*, 95–120.

Fershtman, C., & Gneezy, U. (2001). Strategic delegation: An experiment. *RAND Journal of Economics, 2001*, 352–368.

Fischbacher, U., Gächter, S., & Quercia, S. (2012). The behavioral validity of the strategy method in public good experiments. *Journal of Economic Psychology, 33*(4), 897–913.

Forsythe, R., Horowitz, J. L., Savin, N. E., & Sefton, M. (1994). Fairness in simple bargaining experiments. *Games and Economic Behavior, 6*(3), 347–369.

Gneezy, U., & Guth, W. (2003). On competing rewards standards: An experimental study of ultimatum bargaining. *The Journal of Socio-Economics, 31*(6), 599–607.

Güth, W., Marchand, N., & Rulliere, J.-L. (1997). *On the reliability of reciprocal fairness: An experimental study*. Berlin: Wirtschaftswissenschaftliche Fakultät.

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

Harrison, G. W., & McCabe, K. A. (1996). Expectations and fairness in a simple bargaining experiment. *International Journal of Game Theory, 25*(3), 303–327.

Henrich, J. (2000). Does culture matter in economic behavior? Ultimatum game bargaining among the machiguenga of the peruvian Amazon. *The American Economic Review, 90*(4), 973–979.

Henrich, J., Boyd, R., Bowles, S., Camerer, C., Fehr, E., Gintis, H., & McElreath, R. (2001). In search of homo economicus: Behavioral experiments in 15 small-scale societies. *The American Economic Review, 91*(2), 73–78.

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.

Hoffman, E., McCabe, K. A., & Smith, V. L. (1996). On expectations and the monetary stakes in ultimatum games. *International Journal of Game Theory, 25*(3), 289–301.

Moulin, H. (1986). *Game theory for the social sciences studies in game theory and mathematical economics* (2nd ed.). New York: NYU Press.

Munier, B., & Zaharia, C. (2002). High stakes and acceptance behavior in ultimatum bargaining. *Theory and Decision, 53*(3), 187–207.

Ong, Q., Riyanto, Y. E., & Sheffrin, S. M. (2012). How does voice matter? Evidence from the ultimatum game. *Experimental Economics, 15*(4), 604–621.

Oosterbeek, H., Sloof, R., & Van De Kuilen, G. (2004). Cultural differences in ultimatum game experiments: Evidence from a meta-analysis. *Experimental Economics, 7*(2), 171–188.

Oxoby, R. J., & McLeish, K. N. (2004). Sequential decision and strategy vector methods in ultimatum bargaining: Evidence on the strength of other-regarding behavior. *Economics Letters, 84*(3), 399–405.

Poulsen, A. U., & Tan, J. H. W. (2007). Information acquisition in the ultimatum game: An experimental study. *Experimental Economics, 10*(4), 391–409.

Roth, A. E., Prasnikar, V., Okuno-Fujiwara, M., & Zamir, S. (1991). Bargaining and market behavior in Jerusalem, Ljubljana, Pittsburgh, and Tokyo: An experimental study. *The American Economic Review, 1991*, 1068–1095.

Ruffle, B. J. (1998). More is better, but fair is fair: Tipping in dictator and ultimatum games. *Games and Economic Behavior, 23*(2), 247–265.

Schmitt, P., Shupp, R., Swope, K., & Mayer, J. (2008). Pre-commitment and personality: Behavioral explanations in ultimatum games. *Journal of Economic Behavior and Organization, 66*(3), 597–605.

Slembeck, T. (1999). Reputations and fairness in bargaining-experimental evidence from a repeated ultimatum game with fixed opponents (No. 9905002). University Library of Munich, Germany.

Slonim, R., & Roth, A. E. (1998). Learning in high stakes ultimatum games: An experiment in the Slovak Republic. *Econometrica, 1998*, 569–596.

Solnick, S. J. (2001). Gender differences in the ultimatum game. *Economic Inquiry, 39*(2), 189.

Solnick, S. J., & Schweitzer, M. E. (1999). The influence of physical attractiveness and gender on ultimatum game decisions. *Organizational Behavior and Human Decision Processes, 79*(3), 199–215.

Suleiman, R. (1996). Expectations and fairness in a modified ultimatum game. *Journal of Economic Psychology, 17*(5), 531–554.

Valdez, P., & Mehrabian, A. (1994). Effects of color on emotions. *Journal of Experimental Psychology: General, 123*(4), 394.

Weg, E., & Smith, V. (1993). On the failure to induce meager offers in ultimatum game. *Journal of Economic Psychology, 14*(1), 17–32.

Ⓓ Springer