Do private incentives crowd out public good donations? Evidence from a lab-in-the-field experiment

Kira Lancker* Lakshmi Iyer† Faraz Usmani‡
James Allen IV§ Molly J. Doruska¶

Samba Mbaye‖ Aminata Sylla** Bineta Sow††
Jason R. Rohr‡‡ Christopher B. Barrett§§

November 5, 2025

Acknowledgements: We thank the field data collection team from the Centre de Recherche Pour Le Développement Économique et Social for their invaluable professionalism in primary data collection. We thank Kateri Mouawad for research assistance, and Krister Andersson, Florian Diekert, and Frikk Nesje, as well as seminar participants at the NAERE 2024 conference and Notre Dame for helpful comments and suggestions.
Declaration of funding sources: This research was supported by National Science Foundation grants DEB-2109293 and BCS-2307944, a

*Corresponding author. Department of Food and Resource Economics, University of Copenhagen, Rolighedsvej 23, 1958 Frederiksberg C, Denmark. Email: kl@ifro.ku.dk.

†Department of Economics and Keough School of Global Affairs, University of Notre Dame. Email: liyer@nd.edu.

‡Global Energy Alliance for People and Planet. Email: faraz.usmani@energyalliance.org.

§International Food Policy Research Institute. Email: j.allen@cgiar.org

¶Charles H. Dyson School of Applied Economics and Management, Cornell University. Email: mjd438@cornell.edu.

‖Université Gaston Berger, Saint Louis, Senegal. Email: sambambayeug-bcrdes@gmail.com

**Centre de Recherche Pour Le Développement Économique et Social, Dakar, Senegal. Email: miasylla@gmail.com

††Université Gaston Berger, Saint Louis, Senegal. Email: binetou452@gmail.com

‡‡Department of Biological Sciences, Eck Institute of Global Health, Environmental Change Initiative, 178 Galvin Life Science Center, University of Notre Dame, Notre Dame, IN 46556. Email: jrohr2@nd.edu.

§§Charles H. Dyson School of Applied Economics and Management and Jeb E. Brooks School of Public Policy, Cornell University. Email: cbb2@cornell.edu.

Frontiers Planet Prize, and a grant from the University of Notre Dame Poverty Initiative.

Data availability statement: All data and code can be made available to reviewers. Upon publication, an anonymized version of the data will be made publicly available.

Declaration of generative AI in scientific writing: Generative AI (ChatGPT) was used to improve language and writing. The authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

IRB and pre-registration: The experimental design and survey data collection were approved by the Cornell University Institutional Review Board for Human Participants (protocol 2109010544) and Senegal's Comité National d'Ethique pour la Recherche en Santé (protocol SEN23/119). This manuscript is part of the work related to pre-analysis plan AEARCTR-0013370 in the AEA registry (`https://www.socialscienceregistry.org/trials/13370`, case 2.1.3, p. 11 and 25). We note three deviations from the pre-analysis plan (PAP). First, we employ causal forests for heterogeneity analysis, which was not pre-specified. Second, for ease of presentation, we analyze the difference in donations between games as our outcome variable, rather than using the mathematically equivalent fixed effects specification outlined in the PAP. Third, while the PAP specified that game order would be determined by coin flip, it contained an error in describing the mapping between coin outcomes and game order. The actual implementation was: "tails" resulted in the incentivized game being played first, while "heads" resulted in the non-incentivized game being played first—the reverse of what was erroneously specified in the PAP.

Do private incentives crowd out public good donations? Evidence from a lab-in-the-field experiment

Abstract

We provide novel evidence on motivational crowding out —– the reduction in intrinsic motivation to donate due to private incentives —– in the context of public goods provision. In a lab-in-the-field experiment, we compare donation behavior with and without private incentives up to a fixed contribution threshold. We find that providing incentives reduces donations by 13% among individuals whose baseline contributions exceeded the threshold. This decline is consistent with a theoretical model where private incentives alter not only the level of giving, but also reduce the marginal utility of giving — crowding out intrinsic motivation beyond the incentive range. We show that motivational crowding out is strongest among more generous donors and individuals with higher socio-economic status.

JEL codes: H41; I18; C93; Q56
Keywords: motivational crowding out, impure public goods, warm glow, donation game, causal forests

1 Introduction

Many aspects of economic development and environmental sustainability depend on the willingness of individuals to contribute towards the provision of public goods. Such contributions typically reflect prosocial motivations, arising either from a "warm glow" effect of the act of contributing or from altruism. In some instances, contributions to public goods generate additional private benefits, transforming them into what economists call "impure" public goods. Standard economic theory predicts that such private incentives should increase public goods contributions, since they increase the marginal private material benefit of such contributions. However, offering private incentives might reduce the warm glow, or intrinsic satisfaction, of altruistic actions (Brekke et al., 2003; Cardenas et al., 2000; Frey and Oberholzer-Gee, 1997; Wollbrant et al., 2022). This behavioral "motivational crowding out" effect could reduce overall public good contributions.

We present new evidence on this behavioral crowding out hypothesis from a within-subject, lab-in-the-field experiment.We concentrate on the effect of positive, individual incentives, as are popular in payment for ecosystem services (PES) policies, for example. We study effects through reduced internal motivation, or reduced warm glow. Our study focuses not only on whether motivational crowding out occurs in the presence of private benefits, but also on isolating a behavioral channel to study how it interacts with established reference points and individual heterogeneity. This yields policy-relevant insights into the complexity of incentivizing public good contributions. The lab-in-the field setting allows simultaneously for estimating highly controllable within-individual effects, while allowing for a realistic setup in terms of donation purpose and study population. We set our experiment in rural Senegal, an environment where donations to public goods could be of particular importance, in contrast to settings with stronger tax-financed provision.

Our theoretical framework introduces a model in which individuals derive utility from the public goods that are funded by donations, as well as a "warm glow" utility based on the amount they donate. They also receive a private benefit from contributing up to a certain threshold level, with no additional benefit beyond that threshold. Whereas donations below the threshold may be affected by both altered levels of private benefits and "warm glow" experienced by the individual, donations above the threshold can only be affected by changes to the *marginal* "warm glow" utility (i.e., the utility function's slope, which determines the contribution level). This simple yet innovative model allows us to show that motivational crowding above the threshold only occurs if incentives change marginal warm glow, and, therefore allows us to design an experiment that exclusively tests for non-separable motivational crowding, independent of other mechanisms.

The threshold structure generates a directly testable hypothesis: if the private incentive decreases donations made above the threshold, then the contribution level is subject to social preferences that cannot be determined independently of the private incentive, what Bowles and Polania-Reyes (2012) term "non-separability". To give a real world example, if blood donors were paid an incentive for just one annual blood donation, then under non-separability, any (non-incentivized) donation beyond the first in that same year might be affected by a potential crowding effect. This would occur if the incentive changes the marginal utility of each donation that arises due to warm glow.

To isolate this behavioral channel, we play two donation games with 2,058 participants. Both played with a real monetary endowment provided by the study, but one with a private incentive to donate and one without.[1] Mimicking the theoretical model, the incentive applies only to donations up to a specified maximum threshold. Every individual in our study played both versions of the donations game in random order, enabling us to construct a within-subject estimate of how private incentives affect public goods contributions. Under rational utility maximization and separability, donations above the threshold should not change with the introduction of private benefits. Any disparity in donations would instead indicate a non-separable, behavioral effect on donations due to the private incentive.

We find that private benefits lead to motivational crowding out. Participants who were donating above the incentive threshold in the non-incentivized game donated 13% less, on average, in the incentivized game,

[1] We use the term "incentivized" to describe the donation game with the private incentive and "non-incentivized" for the donation game without private incentive. In both games participants receive an initial, real monetary endowment to ensure incentive compatibility.

indicating that the private incentives indeed change the marginal effect of "warm glow" on utility (i.e., social preferences are non-separable). Two pieces of evidence further underline that this decrease is due to diminished warm glow. First, motivational crowding out was greater among those who contributed more in the non-incentivized version of the game. Causal forest analysis of individual heterogeneity reveals that higher socioeconomic status—reflected in infrastructure (such as grid electricity or nearby healthcare), household wealth, and individual education—predicts larger motivational crowding out. This difference is consistent with our theory, suggesting that individuals who donate more overall react more negatively to private incentives. Second, this effect is particularly pronounced when the incentivized donation game was played second, ruling out anchoring or mean reversion explanations, which should be invariant to game order.

Our study advances both the theoretical and empirical understanding of public goods contributions. Theoretically, motivational crowding out can arise when private incentives interact with social preferences to reduce warm glow benefits (Andreoni, 1990). Psychologically, this can arise if private incentives diminish an individual's sense of self-esteem and autonomy by interfering with their feelings of self-determination or appreciation (Bénabou and Tirole, 2006; Frey and Jegen, 2001). Our lab-in-the-field experimental design explicitly isolates a behavioral effect consistent with the hypothesis that incentives interact with social preferences to change the *marginal* warm glow from donations, indicating non-separability.[2] We distinguish this behavioral channel from other mechanisms of how incentives could affect donations, such as reduced marginal opportunity costs or simple shifts in the warm glow function.

Distinguishing between changes in the marginal (slope) versus level (intercept) shifts of the warm glow function has important policy implications. If incentives merely displace the warm glow function, crowding out occurs only while donations actively increase the incentive.

Previous empirical evidence on motivational crowding out is mixed. Supportive findings come mainly from surveys (Agrawal et al., 2015; Chervier et al., 2019; Frey and Oberholzer-Gee, 1997; Oniki et al., 2023), student lab

[2] This differentiates our contribution from Bénabou and Tirole (2006), who study crowding due to (individually heterogeneous) reputational concerns.

experiments (Müller and Rau, 2020; Gneezy and Rustichini, 2000b; Xiao and Houser, 2011), or settings with coupled private–public good purchases (Engelmann et al., 2017; Guo et al., 2021; Munro and Valente, 2016). In contrast, we conduct a field study, where we study behavior when the public good contribution itself leads to private gains, instead of donations linked to another purchase.

Field experiments show varied results (Rode et al., 2015): While Cardenas et al. (2000) find motivational crowding out, others find mixed or null effects of private incentives (Handberg and Angelsen, 2019; Kerr et al., 2012; Vorlaufer et al., 2023) or even crowding in (Narloch et al., 2012). Reviews of PES schemes suggest effects depend on design features (Akers and Yasué, 2019; Huang et al., 2024). Moreover, Rode et al. (2015) and Van Hecken and Bastiaensen (2010) stress the importance of understanding the baseline of intrinsic motivations prior to the implementation of private incentives. For example, a private benefit could increase the marginal warm glow from donations if respondents read the incentive as a sign that their contributions are highly valued by society (Van Hecken and Bastiaensen, 2010). In contrast to most field studies, we conduct a within-individual analysis with randomized order, i.e., we compare how the same person contributes under two different scenarios, presented in random order. This obviates potential inaccuracy arising from unobserved heterogeneity among study participants, which complicates many prior studies. It also allows us to study outcome heterogeneity from changing baselines through game order randomization.

Our finding that motivational crowding out is higher among those who contribute more in the non-incentivized game—whether from stronger altruism or greater wealth— constitutes a substantial contribution to both academic and policy discussions. Adena et al. (2023) highlight the need to study differences in giving between the rich and the poor, and between developing and developed country contexts. They find stronger crowding in from donation matching incentives among poorer donors, theorizing that this results from a higher price sensitivity. Our work complements this finding, showing that the wealthy are more susceptible to crowding out even after price effects are ruled out. Prior work shows mixed evidence on heterogeneous effects across subgroups defined by gender, age or ability.[3] In contrast, we identify socioeconomic status—proxied by household wealth and education—as a key determinant of heterogeneity. Of course, this may be purely due to the fact that such individuals are likely to contribute more in the non-incentivized game, which we find strengthens motivational crowding out.

These results have practical policy implications. The heterogeneity we observe suggests that the distribution of donations across individuals within a community can influence the expected strength of motivational crowding out effects. Communities with a higher proportion of large donors may be more susceptible to motivational crowding out effects. Tying this back to our PES example, this suggests a need to carefully assess socioeconomic conditions when designing incentive programs like PES schemes. The order effects we observe—greater motivational crowding out when incentives follow voluntary contributions—caution against introducing payments in settings where voluntary provision already exists, especially if contributions exceed the proposed compensation threshold for many participants. If crowding out is indeed linked to reference points based on prior experience or perceived social norms, then policies that alter established systems that rely on social preferences may induce stronger motivational crowding out than policies that try to elicit entirely new public good contributions from private individuals.[4] In contrast, we find no strong declines when the incentivized game is played first, suggesting that behavioral effects of the incentive largely dissipate when the incentive is subsequently removed. In terms of prior literature, our results are more consistent with Vorlaufer et al. (2023), who find no effect six years after the incentive scheme had ended, rather than studies showing persistent crowding out after the private incentive was removed (Gneezy and Rustichini, 2000a; Guo et al., 2021; Oniki et al., 2023).

[3] Angrist and Lavy (2009) found that monetary rewards were more effective among women students while Mellström and Johannesson (2008) observed motivational crowding out only among women. Leuven et al. (2010) reported stronger effects for high-ability students compared to low-ability students. Deci et al. (1999) found that young children experienced more motivational crowding out than college students.

[4] Alternatively, the order effect may reflect stronger erosion of altruistic motivations in incentivized settings over time. We cannot test this hypothesis directly without having participants play incentivized games both first and second, though we are not aware of behavioral explanations that would support such a second-game crowding out mechanism.

2 Why might motivational crowding out arise?

We begin with a theoretical model to provide intuition on non-separable motivational crowding and its implications for game design. We model motivational crowding following the notion of non-separability, whereby the contribution level subject to social preferences cannot be determined independently of the private incentive. First, we write a simple individual-level model to articulate the channels through which a private incentive can affect donations. Second, we use the model to show under what conditions a private incentive affects non-separable motivational crowding and to make predictions on whether we would observe motivational crowding out or motivational crowding in.

We model utility-maximizing individuals deciding how much to contribute to a public good donation that offers a private monetary benefit as an incentive for donating. Individuals must divide a wealth endowment between donating, which gives a positive "warm glow" and adds to the provision of the public good itself, and keeping it for private consumption. The model articulates three channels through which the individual's optimal public good donation is affected by the size of the private incentive: (1) a positive "income effect" as the private incentive increases the individual's budget, (2) an ambiguous "level effect on warm glow" as the private incentive may increase or decrease the real value of the donation, shifting the "warm glow" experienced by the individual, and (3) an ambiguous "functional effect on warm glow" as the private incentive might change the marginal effect of donations on the warm glow experienced by the individual.

In this third channel, if the private incentive diminishes the marginal effect of donations on warm glow, then the private incentive will decrease the individual's optimal donation. Alternatively, if the private incentive intensifies the marginal effect of donations on warm glow—e.g., if individuals interpret private incentives as a signal that donations are urgent and highly valued—then the private incentive will increase the individual's optimal donation. Following Bowles and Polania-Reyes (2012), we call this third channel "non-separable motivational crowding". This channel violates a standard implicit assumption in economics that incentives only affect behavior by altering the economic costs and benefits of certain activities (as in channels 1 and 2) but are *separable* from social preferences.[5]

We then show the model's predictions for how donations are affected by a threshold private incentive—that is, an incentive that remunerates donations up to a certain threshold, but which is fixed above that threshold. We note that the private incentive must be increasing in the size of the donation to activate the first two channels, but this is not required to activate the third channel. Below the threshold, where the private incentive is still increasing in donations, all three channels are potentially active, yielding an ambiguous prediction. Above the threshold at which the private incentive is capped, only the third channel is potentially active, still yielding an ambiguous prediction. By empirically testing the effect of introducing a threshold private incentive on donations to individuals who would have made donations above the threshold otherwise, we can isolate the effect of "non-separable motivational crowding" of public goods contributions by private incentives (channel 3).

2.1 Theoretical Model

Consider a setting where N individuals share a commonly provided public good. Assume an individual i can divide endowment $\bar{w^i}$ into a public donation g^i and a private account x^i, where the public donation potentially yields associated private monetary benefit b^i. The full endowment $\bar{w^i} = w^i + b^i$ is comprised of an initial endowment w^i and the private incentive linked to donating with benefit $b^i = b(g^i)$. Benefits are only given for non-zero donations and are increasing and concave in g^i (i.e., $b^i(0) = 0$, $b^i_{g^i}(g^i) \geq 0$ and $b_{g^i\,g^i}(g^i) \leq 0$). Aggregate donations across individuals yield G, the supply of the public good. Individual i receives utility $u^i(x^i, G, I^i)$ from the private account, the public good, and warm glow $I^i = I^i(g^i, b^i)$, with I^i increasing and concave in g^i (i.e., $I^i_{g^i}(g^i, b^i) > 0$ and $I^i_{g^i\,g^i}(g^i, b^i) < 0$). We further assume that I^i is either decreasing in b^i (i.e., $I^i_{b^i}(g^i, b^i) < 0$), or that I^i is increasing in b^i (i.e., $I^i_{b^i}(g^i, b^i) > 0$) while I^i is still concave

[5] The mathematical concept of separability has different applications in economics. Our application in this paper differs from its alternative applications, e.g., in agricultural household models in development economics.

overall in g^i, that is, $I^i_{g^i}(g^i, b^i) + I^i_{b^i}(g^i, b^i) b_{g^i}(g^i)$ decreases in g^i.[6] The former case could represent that the incentive directly decreases the "warm glow" from donating purely altruistically, potentially because the individual feels that their autonomy or appreciation is compromised. The latter case could represent that the incentive signals to the individual that the contribution is really valuable, thereby increasing the "warm glow" from donating.

The individual thus maximizes utility subject to the budget constraint and aggregate donations accounting:

$$max_{x^i, g^i} \left\{ u^i(x^i, G, I^i(g^i, b^i(g^i))) \right\} \text{ s.t.} \tag{1}$$

$$x^i + g^i = w^i + b^i(g^i) \tag{2}$$

$$G = \sum_j g^j. \tag{3}$$

The individually optimal donation level g^i can be found by solving (see Appendix A.1):

$$u^i_G = u^i_{x^i} - u^i_{x^i} b^i_{g^i}(g^i) - u^i_{I^i} \left(I^i_{g^i}(g^i, b^i) + I^i_{b^i}(g^i, b^i) b^i_{g^i}(g^i) \right), \tag{4}$$

where subscripts denote first derivatives. The socially exemplary level g^* is defined by the Lindahl-Samuelson condition for the optimal provision of public goods, which can be found as the Pareto optimum using Lagrange optimization (see Appendix A.1 for details), which in the case of symmetric individuals is:

$$N u^i_G = u^i_{x^i} - u^i_{x^i} b^i_{g^*}(g^*) - u^i_{I^i} \left(I^i_{g^*}(g^*, b^i) + I^i_{b^i}(g^*, b^i) b^i_{g^*}(g^*) \right). \tag{5}$$

In words, the sum over all individuals' public-private marginal rates of substitution should equal the marginal rate of transforming the private into the public good. The individual will under-provide the public good, i.e $g^i < g^*$: While the right hand sides of equations (4) and (5) coincide whenever the individually and socially optimal donation levels are equal ($g^i = g^*$), the left hand sides still differ, since the individual only takes into account their own marginal benefit from the public good, while the

[6] This latter assumption serves to ensure that the problem has a well defined maximum.

social planner accounts for the marginal benefit for all in society. The above assumption of symmetry is invoked for ease of notation but does not affect results. In particular, the public good would always be under-provided, which the reader can verify given the derivation in Appendix A.1.

The effect on both g^i and g^* of introducing a private benefit is ambiguous and operates through three potential channels. First, if $b^i_g(g^i) > 0$ in the relevant part of the optimality condition, then both g^* and g^i increase with the introduction of the private benefit, because the private incentive reduces the cost of the public good relative to the private good. If this term is large enough, donations can be higher than without incentive (Gneezy and Rustichini, 2000b). Second, under the same condition, a direct effect $-u_{I^i} I^i_{b^i}(g^i, b^i) b^i_{g^i}(g^i)$ increases the right hand side if $I^i_{b^i}(g^i, b^i) < 0$, and vice versa. The intuition is that the warm glow function is shifted due to b^i. For example, if a voluntary worker suddenly got paid for half of their hours, they might feel that only the other half of the work is actually a voluntary contribution. Then, one might expect that the overall warm glow function simply shifts downwards, accounting now only for the voluntary contribution that is considered unpaid. Note that both channels thus far depend on $b^i_{g^i}(g^i) > 0$.

Third, the private benefit might act through $I^i_{g^i}(g^i, b^i)$, which we call the "non-separable motivational crowding effect".[7] The private incentive may change the slope of the warm glow function in donation g^i. This change in slope depends on the sign of the cross-derivative $I^i_{g^i\, b^i}(g^i, b^i)$. Non-separable motivational crowding does not depend on $b^i_{g^i}(g^i) > 0$ as the effect persists even when a private benefit is present but in the relevant region, $b^i_{g^i}(g^i) = 0$. The negative sign of the cross-derivative would suggest that the incentive and social preferences are substitutes: the effect of g^i on warm glow is smaller, the larger the incentive. In the above example about voluntary work, this would mean that the subject would also experience a change in their warm-glow feeling even for the hours of the work that remained unpaid, due to getting payment for the other half of hours worked.

In summary, the overall effect of a private benefit is ambiguous and

[7] Modeling non-separability as the incentive changing the marginal utility of public goods contributions was inspired by Bowles and Polania-Reyes (2012)'s model for state-dependent preferences.

consists of a mix of an "income effect" from cost reduction (channel 1), a level effect of the benefit on warm glow (channel 2), and the behavioral non-separable motivational crowding effect (channel 3).

2.2 Isolating non-separable motivational crowding

Of these three possible channels, we are mainly interested in better understanding non-separable motivational crowding. To isolate this effect, we use the fact that non-separable motivational crowding does not depend on the assumption that $b^i_{g^i}(g^i) > 0$.

Now consider a situation where donations trigger a private benefit, but only up until a certain threshold. That is, $b^i_{g^i}(g^i) > 0$ only up to a threshold level, after which $b^i_{g^i}(g^i) = 0$ and $b^i(g^i) = \bar{b}$. Then, the individual's condition in the domain above the threshold simplifies to:

$$u^i_G = u^i_{x^i} - u_{I^i} I^i_{g^i}(g^i, \bar{b}) \tag{6}$$

$$u^i_{x^i} = u^i_G + u_{I^i} I^i_{g^i}(g^i, \bar{b}) \tag{7}$$

Following Brekke et al. (2003), let utility be additively separable $U(x^i, G, I^i) = u(x^i) + v(G) + I^i(g^i, \bar{b})$ where $u_{x^i}(x^i) > 0$, $u_{x^i x^i}(x^i) < 0$, $v_G(G) > 0$, $v_{GG}(G) < 0$, $I_{g^i}(g^i) > 0$, and $I_{g^i g^i}(g^i) < 0$. We use this simplifying assumption to demonstrate characteristic pathways, in order to clarify our understanding of non-separable motivational crowding out, which motivates our experimental setup. Then, the total derivative of equation (6), with respect to $\bar{b}$ becomes:

$$v_{GG} \frac{dg^i}{d\bar{b}} = u_{xx} \frac{dx}{d\bar{b}} - I^i_{g^i g^i} \frac{dg^i}{d\bar{b}} - I^i_{g^i \bar{b}} \tag{8}$$

The cross-derivative $I^i_{g^i \bar{b}}$ plays a decisive role in the donation game. Keep the overall endowment, $\bar{w}^i = w^i + \bar{b}$, constant across different realizations of $\bar{b}$ in the different games, such that the budget constraint (as long as $\bar{b}$ is binding) is now

$$x^i + g^i = \bar{w}^i = w^i + \bar{b}. \tag{9}$$

As the right hand side is fixed, this entails that $\frac{dx^i}{d\bar{b}} = -\frac{dg^i}{d\bar{b}}$ and inserting into (8) we obtain:

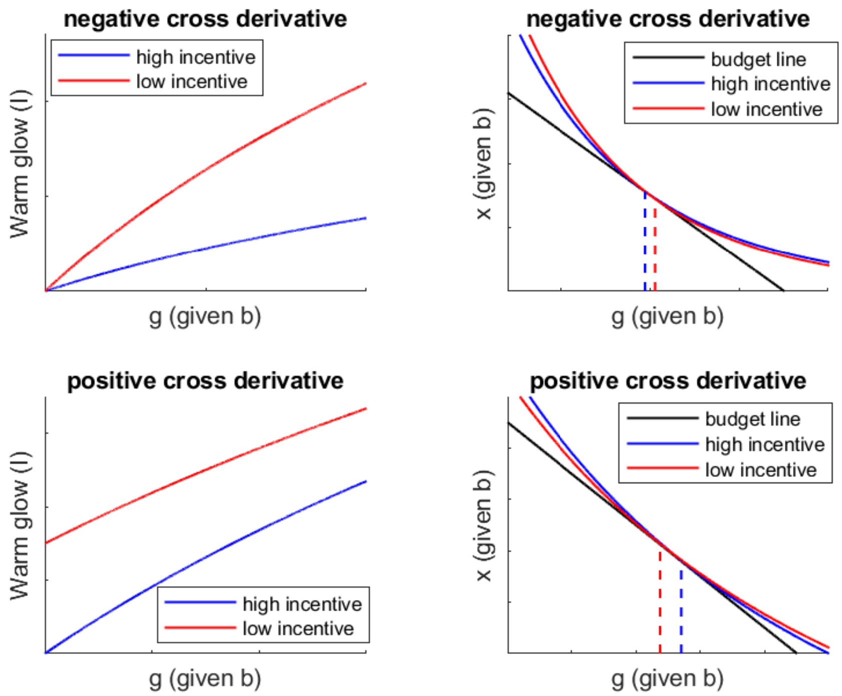
$$\frac{dg^i}{d\bar{b}} = \frac{-I^i_{g^i\bar{b}}}{v_{GG} + u_{x^ix^i} + I^i_{g^ig^i}} \tag{10}$$

An increase in the private incentive $\bar{b}$ has an ambiguous effect on the individual's response in g^i as the denominator is negative given our concavity assumptions, and the numerator may be negative or positive depending on the sign of the cross-derivative $I_{g\bar{b}}$. Therefore, the model makes the following predictions:

- If $I^i_{g^i\bar{b}} = 0$, then $\frac{dg^i}{d\bar{b}} = 0$. That is, there is no "non-separable motivational crowding" effect.
- If $I^i_{g^i\bar{b}} < 0$, then $\frac{dg^i}{d\bar{b}} < 0$. That is, the private benefit diminishes the marginal effect of donations on warm glow, which decreases the individually optimal donation.
- If $I^i_{g^i\bar{b}} > 0$, then $\frac{dg^i}{d\bar{b}} > 0$. That is, the private benefit intensifies the marginal effect of donations on warm glow, which increases the individually optimal donation.

We illustrate the intuition for this prediction in Figure 1, which shows how these conditions would affect the warm glow function I^i in the left column and the optimal choice of the donation level, visualized as the tangential point of budget constraint and indifference curve, in the right column given some fixed private incentive $\bar{b}$. The figure compares the case with a negative cross-derivative in the upper panels and a positive cross-derivative in the lower panels for an arbitrary functional form of the warm glow function. With a negative cross-derivative $I^i_{g^i\bar{b}} < 0$, we expect the warm glow function I^i to diverge as the cross-derivative further widens the gap between different levels of the private incentive as g^i increases. Widening this gap shifts the right-hand side (RHS) $u^i_G + u_{I^i}I^i_{g^i}(g^i, \bar{b})$ of the optimality condition (7) downward for higher levels of the private incentive $\bar{b}$, which, in turn, results in a lower level of donations g^i. Conversely, with a positive cross-derivative $I^i_{g^i\bar{b}} > 0$, the warm glow function I^i converges as the cross-derivative shrinks the gap between different levels of the private

Figure 1: Non-separable motivational crowding depends on the sign of the cross derivative



Notes: The upper two panels show results for a negative cross derivative $I^i_{g^i\bar{b}}$ (non-separable motivational crowding out). The lower two panels show results for a positive cross-derivative (non-separable motivational crowding in). The right side panels show utility tangential to the budget line for different levels of the incentive, while keeping the overall endowment $\bar{w}_i$ constant. With a negative cross-derivative, the individual's chosen donation level g^i decreases in the incentive, while the opposite holds for a positive cross-derivative.

incentive as g^i increases. The RHS of the optimality condition shifts upward for higher levels of the private incentive $\bar{b}$, which, in turn, results in a higher level of optimal donations g^i.

Furthermore, the size of the effect may depend on the level of wealth w^i. An increase in wealth shifts the left hand side of equation (7) downwards, intersecting with the RHS at overall higher donation levels g^i (see Figure A.1 in Appendix A.2). The RHS curve flattens over g^i due to diminishing returns, resulting in a larger absolute crowding-out effect from the incentive at overall higher donations. However, it is conceivable that wealth also alters functional forms in other ways, e.g. when wealthier individuals inherently experience a stronger or weaker motivational crowding response. Therefore, it is an empirical question whether this increase in crowding due to a higher level of wealth holds.

2.3 Hypotheses

The expected sign of the cross-derivative $I^i_{g^i\bar{b}}$ is mainly a behavioral question. The literature on motivational crowding suggests a negative cross-derivative when private incentives are introduced to systems previously supported by pro-social norms alone. We formulate the following hypothesis: (H1) on average, participants will display non-separable motivational crowding out behavior when a private incentive is introduced into a system where pro-social norms constitute the reference point. The effect of private incentives without such a reference point of pro-social norms remains unclear, as there is no literature investigating this case to the best of our knowledge. Furthermore, we hypothesize that (H2) people with higher overall donation amounts experience stronger motivational crowding out and that (H3) wealthier people experience stronger motivational crowding out. This is supported by the theoretical model, expecting decreasing marginal returns in the donation level, and also seems sensible assuming that those deriving more intrinsic utility from giving have "more to lose".

3 Data and empirical strategy

3.1 Data collection and sample

We investigate these hypotheses using a pre-registered lab-in-the-field experiment that collected data from 2,080 households in 104 villages in the Saint-Louis and Louga regions of northern Senegal, between January and April 2024. Villages were selected as part of a broader project, and were required to be located near water sources.[8] Survey households were

[8] For more details, see the pre-analysis plan AEARCTR-0013370, (case 2.1.3, p. 11 and 25). All data used here is part of the baseline data collection prior to any further treatments in connection to the broader project, which therefore do not confound the results in this paper.

randomly selected from village residents, stratified by relative wealth; half the households were selected from the poorer half of village residents.

We interviewed one member of each selected household, and collected data on household composition, income and living standards, agricultural practices, and agreement with statements on community versus individual resource ownership.[9] At the end of the household survey, we invited participants to take part in two donation games. Of the 2,070 participants who played the games, our final analytical sample comprises 2,058 participants (4,116 total observations), after dropping any observations with donation amounts exceeding the maximum level.[10]

Descriptive statistics for our sample are presented in Table 1: 48% were women, 71% did not complete primary school, and the average respondent was 48 years old. The average sample household had 8.3 total members and 3.3 children. As is common in rural Senegal, most households engaged in agricultural production—62.7% cultivated crops like rice, cassava, and onions, and 84% owned cattle—though many households also identified commerce and formal employment as their primary income source. Communities varied in their distance to cities, with 92.3% having primary schools, 82.6% having electricity connection and 62.4% having a healthcare facility.[11]

3.2 Experimental design

We embedded a field experiment at the end of the survey to estimate the effect of a threshold private incentive on donations. The experiment involved two donation games (also called dictator games): a standard or "non-incentivized" donation game, and an incentivized donation game. These games were presented in random order and respondents' donation decisions were unobserved by outsiders (the script appears in Appendix E). In both games, individuals were given a wealth endowment to ensure that the games

[9] A separate community survey gathered data on village characteristics.

[10] We exclude these observations given that they must be mis-recorded as per the experiment design. Results remain qualitatively robust both to including the observations unchanged and to including them while replacing the mis-measured value with the sample mean for the affected game. Results are available from the Stata scripts.

[11] Additional sample summary statistics can be found in Table B.2 for individual- and household-level variables and Table B.3 for village-level variables.

were incentive-compatible and thus that every participant's behavior aligned with their true preferences.

In the non-incentivized game, individuals allocated the CFA 1,200 wealth endowment (about half a daily agricultural wage)[12] between a private portion they kept and a public contribution to a village-serving organization previously chosen by the village chief.[13] We doubled aggregate contributions from all participants in the village and publicly handed over the aggregate donation to the chosen organization on the same day, ensuring individual-level anonymity to minimize social pressure effects. [14]

The incentivized game altered these donation incentives: respondents received an initial endowment of CFA 1,000 to be divided between private and public portions, but earned an individual benefit of CFA 200 if they contributed at least CFA 200 to the public good (the "threshold"). Thus, donations of CFA 200 or more resulted in participants receiving an additional CFA 200 on top of their initial CFA 1,000 endowment, keeping their overall wealth endowment consistent with the non-incentivized game while limiting the maximum donation to CFA 1,000.[15] All other aspects remained unchanged, and participants knew from the start that they would play two games.

By using a threshold private incentive, we are able to identify the non-separable motivational crowding effect of private incentives on donations and empirically test the ambiguous predictions of the theoretical model, where $b^i_{g^i}(g^i) = 0$ beyond a certain point—i.e., where additional donations beyond g^i do not increase the private benefit in the relevant domain.

To ensure comprehension, enumerators demonstrated the division of endowments and explained that total donations would be topped up and publicly donated to the community cause. We encouraged participants to

[12] At the time of implementation, CFA 1,200 was equivalent to about USD 2.

[13] A third of village chiefs chose the local mosque as the organization targeted for donation, 39% chose the health facility, and 28% chose the local school.

[14] Doubling of aggregate contributions could itself be seen as crowding-in inducing incentive (Adena and Huck, 2022; Diederich et al., 2022; Huck and Rasul, 2011; Karlan et al., 2011). However, given that this detail does not vary across games, it should not change relative results, and it also differs from varying private benefits.

[15] The CFA 200 threshold was set based on the lowest contribution level observed during pilot testing of the non-incentivized donation game in August 2022.

ask questions and only commenced the game after answering all questions. Only 212 participants (10%) donated below the CFA 200 threshold in the incentivized game, forfeiting the bonus payment. Enumerator assessments indicated that 93% of participants understood the games well. Notably, most participants who donated below the threshold level in the incentivized game were assessed as having full comprehension, suggesting that their choices reflected preferences rather than confusion.

The order of games was randomized via coin flip, in which interviewers were trained to flip a physical coin in respondent's view and then record the result into the tablet-based questionnaire. Yet the order assignment in the data is clearly skewed with 31% (647/2,458) of respondents playing the incentivized game first and 69% (1,411) playing the non-incentivized game first. We can rule out mis-recording of implemented game order because the tablet-based questionnaire automatically implemented game order based on the recorded coin toss result. We suspect that some interviewers may have foregone the coin toss and answered the coin toss question independently, preferring to play the non-incentivized game first since its associated coin toss result was listed first and the game itself was less difficult to explain. This irregularity leads us to implement a battery of alternative specifications and robustness checks, including showing that our results are robust to including enumerator fixed effects and among the sub-sample of interviewers with more balanced ordering. Additionally, balance tests reveal no imbalances among socio-economic variables, such as respondent age, gender, household size, wealth indicators, or spatial village characteristics (see Table 1 for a subset, and Appendix Table B.2 and Table B.3 for the broad set of all variables used).[16] We conclude that the order of games remains plausibly exogenous to respondent characteristics.

Based on the structure of these games, we anticipate that those contributing less than CFA 200 in the non-incentivized donation game should increase their contributions to CFA 200 in the incentivized version, and

[16] The only notable differences were a 4.6 percentage point higher probability of having no education among those playing the incentivized game first (small relative to the 71% baseline), village-level rice prices (CFA 4 difference), and some household characteristics (participation in a women's group, participation in a savings group, whether the household treats water, whether the household had commerce as main income source, and Lockean beliefs). These unbalanced variables represent 9% of all variables tested.

Table 1: Balance table of a subset of household and village level variables

Variable	Descriptives					Balance Test	
	Obs.	Mean	Std. dev.	Min.	Max.	coeff	pval
Respondent's gender (0=male)	2058	0.478	0.478	0	1	-0.027	0.252
Respondent's age	2052	48.241	15.299	16	91	-0.304	0.680
Education, <primary (0)	2058	0.426	0.772	0	3	0.068	0.068
Household size	2058	8.27	3.758	2	55	-0.168	0.347
No. of children in household	2058	3.304	2.302	0	26	-0.123	0.261
Agric. land cultivation	2045	0.627	0.484	0	1	-0.003	0.890
Hh wealth index	2058	0.367	0.143	0.063	0.875	-0.004	0.508
Village Population size	2038	2354.26	3309.965	90	30000	31.552	0.842
Presence: school	2058	0.923	0.267	0	1	0.007	0.590
Presence: health facility	2058	0.624	0.485	0	1	0.001	0.974
Presence: electr. conn.	2058	0.826	0.38	0	1	-0.009	0.608

Notes. Summary statistics and balance test of coin toss for game order. We conduct a t-test with robust standard errors for categorical variables. This table shows a subset of all variables considered, see Appendix B for full list of variables. Agric. land cultivation is an indicator variable that takes the value of one if at least one member of the household cultivated land in the last growing season. The balance coefficient represents the difference between the average characteristics of respondents who played the incentivized game first, and those who played it second. "Test pval" represents the p-value of a t-test for whether this difference in means is statistically significant.

those contributing more than CFA 1,000 in the non-incentivized game should mechanically reduce their contributions as the maximum allowable donation in the incentivized game was CFA 1,000 (due to the altered initial endowment). Importantly, if there is no motivational crowding out (or in), respondents who contribute between CFA 200 and 1,000 in the non-incentivized donation game should not alter their contributions in the incentivized game, a direct test of our hypothesis H1.

3.3 Estimation approach

Our analysis proceeds in three stages. First, since each respondent participated in both games, we can compare individual-level contributions between game variants using two-sided t-tests to examine whether including private benefits induces behavioral changes.

Second, we isolate the motivational crowding effect through regression analysis using between-game differences in donation levels as our outcome variable. In order to account for the game design, our main specification controls for whether baseline contributions were below CFA 200 or above

CFA 1,000, as follows:

$$\Delta g_i = \beta_1 + \beta_2 Below200_i + \beta_3 Above1000_i + \epsilon_i \qquad \text{(Main)} \qquad (11)$$

where Δg_i is individual i's contribution in the incentivized game minus their contribution in the non-incentivized game; $Below200_i$ is a binary variable that equals 1 if the individual contributed less than the threshold value (CFA 200) in the non-incentivized game, and zero otherwise; $Above1000_i$ is similarly a binary variable that equals 1 if the individual contributed more than CFA 1,000 in the non-incentivized game, and zero otherwise; and ϵ_{ikv} is a heteroskedasticity-robust error term.

In this most parsimonious specification, our parameter of interest is β_1, the estimate of motivational crowding, after controlling for monetary incentive effects and the mechanical effects of the game design. Note that β_1 captures only the behavioral part of the response due to non-separability, as respondents who contribute ≥ 200 and $\leq 1,000$ in the non-incentivized public good donation game had no monetary incentive to change their contributions. The separability hypothesis (H1) implies $\beta_1 = 0$ while $\beta_1 < 0$ (> 0) if there is motivational crowding out (crowding in). We expect $\beta_2 > 0$, reflecting increased donations from monetary incentives for below-threshold contributions, and $\beta_3 < 0$, reflecting mechanical reductions due to the lower donation ceiling.

Third, to test for heterogeneity in motivational crowding out, we estimate a series of models that add additional variables to the main equation above. We focus on three important characteristics. First, we check that pure "order effects" that could arise from playing two games consecutively are not driving the results by disentangling order and treatment effect (equation (12)). Since our main outcome Δg_i is the incentivized game contribution *minus* the non-incentivized game contribution, we set $order_i = 1$ if the incentivized game is played second and $order_i = -1$ if the non-incentivized game is played second. That way, these "pure order effects" enter the outcome in the same direction. Then, for example, if participants tend to donate less in the second game regardless of type (e.g., due to game learning effects), $\gamma < 0$.

$$\Delta g_i = \beta_1 + \beta_2 Below200_i + \beta_3 Above1000_i + \gamma order_i + \epsilon_i \tag{12}$$

Second, we examine whether the reference point matters, namely whether playing the incentivized game after the non-incentivized game increases motivational crowding out (equation (13)) (i.e., $incsecond_i$ takes value one if the incentivized game was played second, zero otherwise). If individuals who played the non-incentivized game first perceive the introduction of the private incentive as a stronger crowding-out of their warm-glow motivation than those who played the incentivized game first, we would expect $\gamma < 0$. Note that due to playing one game of each order with each respondent, we can only control for either the effect of playing any game second (pure order effect, see above) or the effect of playing the incentivized game second as described here.

$$\Delta g_i = \beta_1 + \beta_2 Below200_i + \beta_3 Above1000_i + \gamma incsecond_i + \epsilon_i \tag{13}$$

Third, we additionally analyze the effect of the individual's donation level in the non-incentivized game (equation (14)). If it is the case that motivational crowding out is largest among those who had the highest prosocial motivations, then we expect $\delta < 0$.

$$\Delta g_i = \beta_1 + \beta_2 Below200_i + \beta_3 Above1000_i + \gamma incsecond_i + \delta donationlevel + \epsilon_i \tag{14}$$

The order of games was decided based on a coin toss. However, as described in prior section, order assignment is clearly skewed with 69% playing the non-incentivized game first, likely due to interviewer implementation error. We showed in section 3.2 that this randomization imbalance is independent of major socio-economic variables. To further address doubts that this randomization imbalance could have systematically affected results, we implement two key robustness checks alongside our main results. First, noting that some interviewers had less balanced ordering than others, we identify a sub-sample of five (out of 20) interviewers with each game order used in at least 40% of the subsample. We re-estimate on that subsample of 417 participants observed by these five enumerators whose game order appears reasonably random. In this sample, 56% of participants played the incentivized game last and 44% played the non-incentivized game last. Second, we control for enumerator fixed effects, which restricts identification of the order effect to within-enumerator variation, such that enumerators with greater variation in game order contribute more information to this estimate than enumerators with little variation, replicating the key feature of randomization.

Finally, we examine how motivational crowding out varies with socioeconomic characteristics. Since our survey collected information on several different socioeconomic characteristics, we conduct a machine learning (causal forest) analysis to systematically identify how individual and characteristics are associated with heterogeneous responses to incentives. Implementation details are available in Appendix C.

Robustness checks (see Appendix D) include respondent comprehension checks, checks regarding game order dependence on interviewer, a Poisson distribution model, a model that is nonlinear in non-incentivized donation level using the inverse hyperbolic sine transformation, a subset of observations by order to account for endogeneity concerns, as well as a fully interacted specification including interactions e.g. with respect to the "Below 200" and "Above 1000" thresholds with donation level and game order. Results remain robust to these refinements.

4 Estimating motivational crowding out

4.1 Do private benefits cause motivational crowding?

Comparing donations across the two game versions with both summary statistics and regression analysis suggests offering private benefits leads to motivational crowding out. Average donations were CFA 475 (40% of the maximum possible donation) in the non-incentivized donation game and CFA 410 (41% of the maximum) in the incentivized game (Table 2).[17] The CFA 65 average difference is meaningfully large (14% of the average non-incentivized game donations) and statistically significant at the 1% level in a two-sided *t*-test.

This unconditional difference may underestimate the true motivational crowding out effect by including increased contributions from those who initially contributed less than CFA 200 without the incentive, or overestimate it by including mechanical reductions from those who contributed more than CFA 1,000 in the non-incentivized game. Table 2 breaks down average donations by these thresholds and still finds a statistically significant and meaningfully large decrease in incentivized donations among those participants who donated between CFA 200 and CFA 1000 in the non-incentivized game. We also control for these factors in the regression analysis below.

The distribution of donations shows clear patterns. The mode was CFA 500, with most respondents donating at or above the threshold of CFA 200 (see Figure B.1 in Appendix B). Only 2% (41 respondents) donated nothing in the non-incentivized game versus 3% (73) in the incentivized game, while 7% (147) and 9% (192) gave the maximum possible amount in the non-incentivized game (CFA 1,200) and the incentivized game (CFA 1,000), respectively. Additional summary statistics can be found in Appendix B.

Table 2: Average donation levels in incentivized and non-incentivized games

		Incentivized (A)		Non-incentivized (B)		Difference (A - B)	
	Obs	Mean	SD	Mean	SD	Mean	*p*-value
Full Sample	2,058	410.0	252.0	475.3	286.1	-65.4	<0.001
Non-inc. $\in$ [0,200)	116	245.7	275.5	64.6	51.0	181.1	<0.001
Non-inc. $\in$ [200,1000]	1,792	385.9	208.9	441.4	195.1	-55.6	<0.001
Non-inc. $\in$ (1000,1200]	150	824.7	303.9	1197.7	16.7	-373.0	<0.001

Notes: This table presents summary statistics for donation amounts by participants across different games and thresholds. We report two-sided p-values of t-tests in the final column.

Next, we test for motivational crowding out using regression analysis. Controlling for donations below CFA 200 and above CFA 1,000 in the pure donation game (equation (11)), we find that participants in the intermediate

[17] In neither game were donations statistically significantly different between donation causes (school, mosque or clinic).

range reduce their contributions by a statistically significant CFA 56 in the incentivized game, approximately 13% of the mean non-incentivized donations for that cohort (Table 3, model 1). This suggests that the private incentive has reduced the marginal warm glow from donations for the average participant, as in the two upper panels of 1. This share is comparable to findings in Guo et al. (2021) (8-22% reduction in voluntary donations upon introduction of a private incentive) and Gneezy and Rustichini (2000b) (8-36% reduction compared with non-incentivized donations). Results in the subsample- and enumerator fixed effects-models (Table 3, model (1) subs. and (1) FE, respectively) are consistent with the main result, showing a negative and significant treatment effect.

The other coefficient estimates are consistent with predicted patterns. Specifically, the estimated β_2 coefficient for donating less than CFA 200 in the non-incentivized game is positive and the estimated β_3 coefficient for donating more than CFA 1000 in the non-incentivized game is negative. Participants who donated less than the threshold of CFA 200 in the non-incentivized game significantly increased their contributions in the incentivized game, with the combined effect $(\beta_1 + \beta_2)$ estimated at CFA 181 (se= 27.47, $p < 0.01$), statistically indistinguishable from the threshold of 200. For participants who donate above CFA 1,000 in the non-incentivized game, the combined effect $\beta_1 + \beta_3$ shows a CFA 373 reduction (se= 24.81, $p < 0.01$). Interestingly, participants who donate above 1,000 CFA in the non-incentivized game decrease their donations by more than the combined mechanistic and motivational crowding out effects. This suggests that crowding out effects may vary with individuals' baseline (non-incentivized) generosity (hypothesis H2), which we explicitly test below.

Game order effects (Table 3, model (2)) are substantial. But the decline in donations does not arise purely because the incentivized game is more often played second. The treatment effect, as represented by the constant, remains negative and highly significant even after controlling for a pure order effect (Table 3, model 2, last three columns). The order effect is also negative and significant, which could arise due to participants' learning about the game.

Table 3: Main effects estimation results

	(1), main	(1), subs.	(1), FE	(2)	(2), subs.	(2), FE
Constant (β_1)	-55.565***	-59.664***	-55.158***	-35.815***	-50.427***	-35.316***
	(4.76)	(12.15)	(2.40)	(5.14)	(12.04)	(3.10)
Below 200 (β_2)	236.651***	535.980***	235.799***	228.671***	530.285***	226.791***
	(27.88)	(102.22)	(42.59)	(26.96)	(98.90)	(42.94)
Above 1000 (β_3)	-317.435***	-326.922***	-322.351***	-314.781***	-331.234***	-318.722***
	(25.26)	(46.84)	(29.21)	(24.82)	(45.42)	(28.06)
Effect of playing a game second				-52.510***	-67.300***	-52.794***
				(5.14)	(12.10)	(6.30)
Enum. FE	no	no	yes	no	no	yes
Subsample	no	yes	no	no	yes	no
R^2	0.182	0.276	0.182	0.223	0.328	0.222
N	2058	417	2058	2058	417	2058
RMSE	216.7	249.6	215.6	211.2	240.8	210.3
AIC	27982.0	5790.0	27957.6	27877.2	5760.9	27856.8
BIC	27998.9	5802.1	27968.9	27899.7	5777.0	27873.7
	mean donations (all models)					
non-incentivized game	475.3					
incentivized game	410.0					

Notes. The outcome variable is the difference in donations, i.e. the incentivized donations minus the non-incentivized donations. The table shows point estimates and robust standard errors in parentheses. Parameters as in equations (11) and (12). All models show the treatment effect condensed in the Constant (β_1). subs.: Model run for a subsample of observations from the five enumerators with the most balanced order of games. FE: Model with enumerator fixed effects (and standard errors clustered at enumerator level). N: number of observations. RMSE: root mean squared error. AIC: Akaike information criterion. BIC: Bayesian information criterion. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

4.2 Does the reference point matter?

Next, we analyze the interaction of private incentives and game ordering by controlling for playing the incentivized game second (in equation (13)). The incentivized game reduces donations mainly when the incentivized game is played after the non-incentivized game (Table 4, model (3))—that is, for participants who experienced the private incentive as a change to an initial state that offered no material incentive. The negative coefficient estimate of CFA 105 (se= 10.28, $p < 0.01$) can be attributed to two mechanisms: a negative interaction effect between the treatment and game order (playing the incentivized game second), and a pure order effect. Those two are confounded in this estimate. Meanwhile, the constant is estimated at CFA 17 (se= 8.72, $p < 0.10$), showing that relative to the non-incentivized game donation average across game order, the incentivized game being played first has a positive and marginally significant markup. This measures donations in the incentivized game if played first, against all donations in the non-incentivized game, whether played first or second. Therefore, again, this is a mix between the pure order and the treatment mechanism, and the

Table 4: Order effects estimation results

	(3), main	(3), subs.	(3), FE
Play incentivized game second	-105.019***	-134.599***	-105.588***
	(10.28)	(24.20)	(12.61)
Constant (β_1)	16.695*	16.873	17.477*
	(8.72)	(19.23)	(8.71)
Below 200 (β_2)	228.671***	530.285***	226.791***
	(26.96)	(98.90)	(42.94)
Above 1000 (β_3)	-314.781***	-331.234***	-318.722***
	(24.82)	(45.42)	(28.06)
Enum. FE	no	no	yes
Subsample	no	yes	no
R^2	0.223	0.328	0.222
N	2058	417	2058
RMSE	211.2	240.8	210.3
AIC	27877.2	5760.9	27856.8
BIC	27899.7	5777.0	27873.7

Notes. The outcome variable is the difference in donations, i.e. the incentivized donations minus the non-incentivized donations. The table shows point estimates and robust standard errors in parentheses. Parameters as in equations (11) and (13). All models show a negative treatment effect in "Play incentivized game second" when the incentivized game was played second. The Constant (β_1) shows a mix of a pure order (positive due to the specification of the outcome variable) and a negative treatment effect when the non-incentivized game was played second, where the combination is positive when the order effect is dominant. subs.: Model run for a subsample of observations from the five enumerators with the most balanced order of games. FE: Model with enumerator fixed effects (and standard errors clustered at enumerator level). N: number of observations. RMSE: root mean squared error. AIC: Akaike information criterion. BIC: Bayesian information criterion. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

significant positive coefficient here shows that the order effect outweighs the treatment effect (essentially, CFA 17 represents the difference of the treatment effect (CFA -36) and pure order effect (CFA -53) in Table 3, model 2). We conclude that a treatment effect is present when the incentivized game is played first, but it is relatively weak. The linear combination of the effect of playing the incentivized game second (CFA -105) with the constant (CFA 17) is estimated at CFA -88 (se= 5.45, $p < 0.01$), representing 19% of average non-incentivized game donations, and proving a strong net motivational crowding out effect when the incentivized game is played second. Results in the subsample- and enumerator fixed effects-models are again consistent with the main result.

Overall, we find motivational crowding out no matter the order of game play, although the estimated effect is substantially larger when the incentivized game is played second, after a reference point has been set by playing the non-incentivized game.

These order-related findings help rule out alternative interpretations. If the lower donations in the incentivized game stemmed from anchoring or signaling effects around the CFA 200 threshold, we would expect to observe a similar decrease for participants who played the incentivized game first. Similarly, mental accounting effects from the lower initial endowment (CFA 1,000 versus 1,200) should manifest consistently across order conditions. Neither pattern appears in our data. Regression-to-the-mean effects similarly cannot explain the patterns we observe.

An alternative way to control for order effects is to run a random effects specification, incorporating an order effect and a treatment-order-interaction (Appendix D.2). Order by itself, other than in interaction with treatment, does not have a significant impact on donations, although diagnostic tests favor the fixed effects specification.[18]

[18] Further robustness checks include respondent comprehension checks and a Poisson distribution model (see Appendix D).

4.3 How does motivational crowding vary by donation level?

Next, we test for heterogeneity in responses according to a participants' level of non-incentivized donation. To do this, we include the non-incentivized contribution level as an additional explanatory variable (equation (14)), the associated coefficient estimate is negative and statistically significant (Table 5, model 4). Model fit improves substantially in terms of R^2 and information criteria.

An increase in non-incentivized game donations by CFA 100, all else equal and controlling for non-incentivized game donations not between CFA 200-1,000, is on average associated with CFA 47 lower donations in the incentivized game. Motivational crowding out increases both in absolute and relative terms in non-incentivized game donations.

This result holds in the subsample- and enumerator fixed effects-models as well as in a fully interacted specification (Appendix D.5), and a model that is nonlinear in donation level (Appendix D.4).[19] To account for a concern that including the donation level from the non-incentivized game may lead to endogeneity bias, Appendix D.6 shows highly robust results.

Figure 2 illustrates this relationship by plotting the difference in donations across games against non-incentivized game donations for model 5, after controlling for the rational and mechanistic effects of donating below CFA 200 or above CFA 1,000 in the non-incentivized game. The relationship is plotted separately by game order. The positive and significant β_1 coefficient estimate in this specification means that those with the lowest levels of non-incentivized donations increase their donations when incentives are provided. We see that when the incentivized game is played first, the treatment effect becomes zero for non-incentivized donations above the modal contribution of CFA 500, with reductions reaching up to 28% of non-incentivized donations among the largest contributors. However, when the incentivized game is played second, significant negative effects emerge for non-incentivized game donation levels above CFA 290, reaching reduc-

[19] Causal forest results support including the non-incentivized game donation amount as covariate, it is the most important predictor of treatment effect heterogeneity (Appendix D.7, Table D.5).

Table 5: Donation size effects estimation results

	(4), main	(4), subs.	(4), FE	(5)	(5), subs.	(5), FE
Non-incentivized donation level	-0.469***	-0.531***	-0.490***	-0.463***	-0.555***	-0.492***
	(0.02)	(0.05)	(0.03)	(0.04)	(0.07)	(0.05)
Order × Donation-level				-0.009	0.046	0.003
				(0.04)	(0.09)	(0.04)
Play incentivized game second	-104.196***	-136.079***	-101.070***	-99.830***	-160.898***	-102.262***
	(9.44)	(21.85)	(13.02)	(20.96)	(52.37)	(19.21)
Constant (β_1)	223.130***	281.205***	230.632***	220.289***	294.186***	231.417***
	(13.15)	(30.15)	(18.51)	(18.44)	(42.49)	(23.41)
Below 200 (β_2)	52.026**	283.505***	49.626	52.168**	283.610***	49.584
	(26.45)	(97.53)	(35.28)	(26.35)	(97.27)	(35.20)
Above 1000 (β_3)	39.813	39.365	46.922	40.265	38.592	46.812
	(31.66)	(56.67)	(36.88)	(31.38)	(56.60)	(35.96)
Enum. FE	no	no	yes	no	no	yes
Subsample	no	yes	no	no	yes	no
R^2	0.351	0.452	0.357	0.351	0.452	0.357
RMSE	193.1	217.7	191.3	193.2	217.9	191.4
AIC	27509.7	5678.0	27468.4	27511.6	5679.6	27470.4
BIC	27537.8	5698.2	27490.9	27545.3	5703.8	27498.5

Notes. The outcome variable is the difference in donations, i.e. the incentivized donations minus the non-incentivized donations. The table shows point estimates and robust standard errors in parentheses. Parameters as in equations (11) and (14). All models show a negative treatment effect in "Play incentivized game second" when the incentivized game was played second. All models further show a stronger treatment effect for a higher "Non-incentivized donation level". The Constant (β_1) shows a mix of a pure order, donation size, and treatment effect. subs.: Model run for a subsample of observations from the five enumerators with the most balanced order of games. FE: Model with enumerator fixed effects (and standard errors clustered at enumerator level). N: number of observations. RMSE: root mean squared error. AIC: Akaike information criterion. BIC: Bayesian information criterion. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

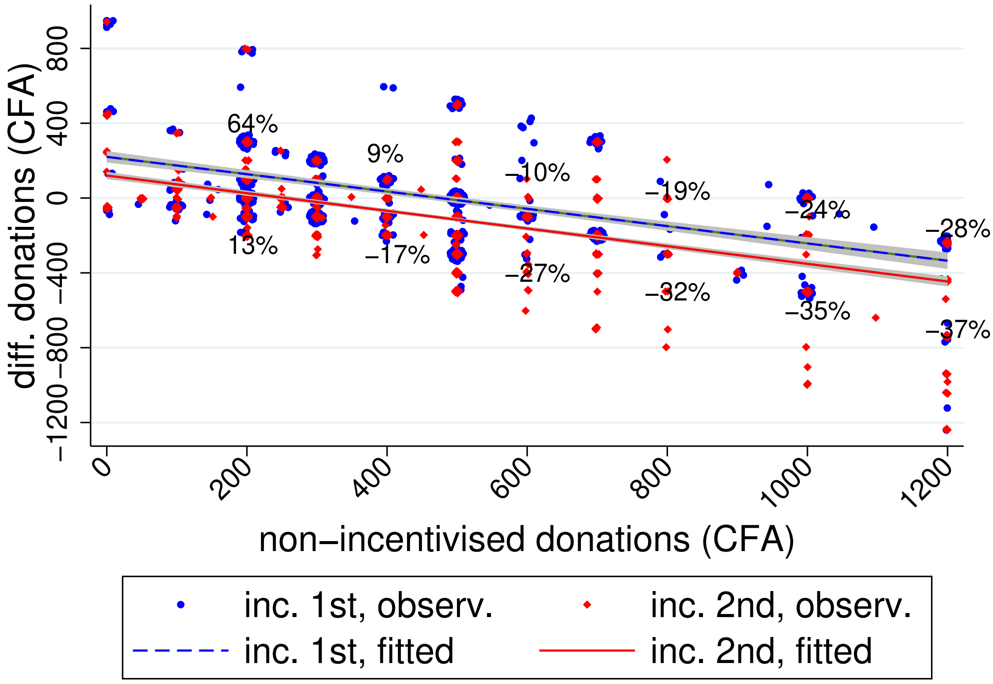
tions of up to 37% of non-incentivized game donations. Participants with stronger pro-social motivation to donate appear particularly susceptible to motivational crowding out, especially when they have already donated out of warm glow motivations without any material incentives.

We also estimate model 5 (Table 5), adding an interaction effect between order and donation level. That coefficient estimate is not significant and does not improve model fit, indicating that the effect of the donation level is independent of game order. Coefficient estimates remain unchanged (model 5). We rely on model 4 in the subsequent analysis of heterogeneity.

4.4 Heterogeneity analysis

Having established that motivational crowding out varies with game order and baseline donation levels, we now examine whether certain types of households or communities are more susceptible to this effect. To systematically identify such patterns across our rich set of household and community characteristics, we employ causal forests, a machine learning

Figure 2: Effect of game order and non-incentivized donation size on motivational crowding out



Notes: Results shown after controlling for lower and upper threshold indicators given estimates from model 5. Red circles and blue diamonds show observations when the incentivized game was played first and second, respectively (jittered slightly, for better readability). Red solid and blue dashed lines show linear fit lines with added 95% confidence bands in grey. Slopes of the blue and red line are not statistically significantly different. Percentages indicate the fitted average percentage reduction relative to non-incentivized donations for the respective size of non-incentivized donations.

method developed by Athey and Imbens (2016). This approach builds many decision trees, where each tree recursively partitions the data to identify subgroups with different treatment effects. By averaging across thousands of such trees, each built using random subsamples of the data and covariates, the method can detect complex patterns of treatment effect heterogeneity while avoiding spurious findings that could arise from multiple hypotheses testing. The methodology is particularly well-suited to our setting given the large number of household and community characteristics we observe.[20]

Table 6: Variable importance measures for top drivers of heterogeneity

Covariate	VIM (scaled)
Household wealth index	100.0
Respondent has no education	92.2
Number of households in community	72.8
Distance to nearest healthcare facility	68.3
Access to grid electricity in community	62.4

Notes. This table lists the name and variable importance measure (VIM) for the top five variables identified by the causal forest as drivers of heterogeneity in the impact of the impure donation game on donation amount. Mean (SD) of VIMs across all potential splitting variables listed in Table C.1 is 15.0 (20.3).

Table 6 presents the top five characteristics most strongly associated with heterogeneous responses to the incentivized donation game, other than the non-incentivized donation level.[21] The analysis reveals that socioeconomic factors—particularly household wealth and education level—are the strongest predictors of heterogeneous treatment effects. Specifically, the household wealth index emerges as the most influential factor, followed closely by whether the respondent has no formal education. Community-level factors (such as the number of households, remoteness, and access to grid electricity) are also associated with treatment effect heterogeneity.

Regression analysis confirms these patterns (Table 7). Higher household wealth is associated with significantly larger reductions in donations under the incentivized game. A one-unit increase in the household wealth index (ranging from 0 to 1) corresponds to an additional CFA 36 reduction in donations (column 1). Conversely, individuals with no formal education exhibit less motivational crowding out, with their reduction in donations being CFA 11 smaller than those with some education (column 2).

[20] We describe our implementation of the causal forest approach in detail in Appendix C.

[21] Causal forest results find that the non-incentivized donation amount is the single most important predictor of treatment effect heterogeneity (Appendix D.7, Table D.5). In this section, we examine the role of other characteristics.

Table 7: Household and community characteristics associated with motivational crowding out

	(1)	(2)	(3)	(4)	(5)
Constant (β_1)	-59.5***	-80.7***	-72.1***	-73.3***	-62.9***
	(0.45)	(0.31)	(0.24)	(0.24)	(0.34)
Below 200 (β_2)	1.58**	2.07**	2.46**	2.82***	2.19**
	(0.73)	(0.84)	(0.96)	(0.94)	(0.91)
Above 1000 (β_3)	-5.54***	-6.35***	-8.37***	-8.49***	-7.49***
	(0.71)	(0.64)	(0.83)	(0.80)	(0.74)
Wealth index	-36.4***				
	(1.16)				
No formal education		10.9***			
		(0.36)			
Number of households in community			0.000062		
			(0.00027)		
Distance to nearest healthcare facility				0.36***	
				(0.036)	
Access to grid electricity in community					-11.0***
					(0.40)
R^2	0.36	0.30	0.051	0.11	0.26
RMSE	7.11	7.40	8.66	8.37	7.65
			mean donations (all models)		
non-incentivized game			475.3		
incentivized game			410.0		

Notes. This table examines how the effect of the incentivized game on donations varies across different household and community characteristics. The outcome variable represents the predicted effect of playing the incentivized game for each household in our sample, estimated using the causal forest approach described in Appendix C. Negative values indicate larger reductions in donations when playing the incentivized game. The regression is weighted by the inverse of the variance of the predicted treatment effects. Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.
In an alternative fully interacted specification, where we allow for interactions also between the socio-economic variable and the threshold controls, main effects remain robust. We only find one significant interaction, namely a small negative interaction of wealth with the "Above 1000" control.

Community-level socioeconomic indicators also predict heterogeneous responses. While the raw number of households in a community shows no significant relationship with motivational crowding out (column 3)[22], access to infrastructure appears important: households in communities with grid electricity access reduce their donations by an additional CFA 11 (column 5), while those further from healthcare facilities show less motivational

[22] When we discretize this variable into terciles, we find that households in communities in the second and third terciles of size reduce their donations by an additional CFA 7.7 and CFA 3.3, respectively, compared to those in the smallest tercile (see Table C.2)

crowding out, with each additional kilometer associated with CFA 0.36 smaller reductions in donations (column 4).

These results paint a consistent picture: motivational crowding out is strongest among wealthier, better educated households in more developed communities—precisely those groups more likely to make larger baseline donations.[23] This finding has important implications for targeting incentive-based public goods policies.

5 Discussion and conclusions

We use a lab-in-the-field experiment in rural Senegal to test for motivational crowding out, namely, whether incentivizing contributions to public goods can paradoxically result in decreased contributions. We posit the existence of a behavioral channel, whereby paying people for their contributions could lead to a decline in the marginal "warm glow" of giving. This model implies that donations should increase for below-threshold givers, but decline for those who contributed larger amounts without the incentive.

We confirm these predictions using data from rural Senegal, with more than 2,000 respondents who played both the incentivized and non-incentivized versions of the game. While we observe the predicted positive monetary incentive effect among those whose non-incentivized game donations were below the CFA 200 threshold, we find that donations are 13% lower among those above the incentive threshold, confirming hypothesis H1. These motivational crowding out effects are stronger for respondents who played the incentivized game after the non-incentivized one, suggesting that monetary rewards diminish the warm glow they had previously experienced from purely voluntary giving. Our results are consistent with a theoretical model that features non-separability, where incentives alter not just the level but the slope of warm glow in donations.

Our heterogeneity analysis reveals three key patterns. First, motivational crowding out increases with baseline generosity. Participants who contributed the most in the non-incentivized game experienced the biggest

[23] The wealth index and grid electricity access correlate positively with non-incentivized donations, while lack of education correlates negatively (all $p < 0.01$). Community size and healthcare facility distance show no significant correlations.

motivational crowding out effects, with up to 37% reduction, confirming hypothesis H2. The negative treatment effect emerges at different thresholds depending on game order: above CFA 500 (affecting 22% of participants) when the incentivized game was played first, and above CFA 290 when the incentivized game was played second (affecting 74% of participants). Second, socioeconomic status strongly predicts susceptibility to crowding out, with wealthier and better educated households showing larger reductions in donations. Third, community characteristics also matter: households in villages with better infrastructure are more susceptible to motivational crowding out than those in more remote communities. These two results support hypothesis H3 — which posited that wealthier individuals experience stronger motivational crowding out — by showing that socio-economic status across multiple dimensions amplifies crowding out.

Our results have important implications for the design of policies that rely on individual contributions towards public goods, such as donation tax incentives or environmental conservation policies based on PES. The non-separable motivational crowding out suggests that incentive effects can persist beyond inbuilt thresholds. Tax incentives typically include a ceiling, such that effects from non-separable motivational crowding out may be realistically relevant. For example, tax incentives in Denmark are capped at about 19,000 DKK (about 3,000 USD) per tax-payer, above which the effect of the tax incentive according to our study will be negative, since non-separable motivational crowding out is no longer (partially) offset by a positive incentive effect.

The strong socioeconomic gradient in crowding out suggests that monetary incentives for public goods may be more effective in communities with lower socioeconomic status, less infrastructure, and greater geographic isolation. This pattern admits two alternative, potentially complementary interpretations. First, socioeconomic status may affect motivational crowding out mainly through its association with higher non-incentivized game donations, which face stronger motivational crowding out due to diminishing returns. Alternatively, wealth, education, and infrastructure might somehow directly influence susceptibility to motivational crowding out. The latter interpretation adds nuance to Henrich et al.'s (2001; 2005) findings that greater community-level market integration—a positive correlate of greater monetization of the economy and of higher average well-being—is associated with more prosocial behavior across different societies. Our within-society analysis reveals that those who are most prosocial—as measured by non-incentivized game donations—are also most susceptible to motivational crowding out when material incentives are introduced. While Henrich et al. (2005) note that "individual-level economic and demographic variables do not consistently explain game behavior," our within-group study clearly isolates individual socioeconomic factors that matter for motivational crowding out, if not for baseline pro-sociality. Thus, the mechanisms influencing prosocial behavior may differ from those affecting susceptibility to motivational crowding out. In that regard, our paper also differs from theoretical findings in Bénabou and Tirole (2006). In their model, the private incentive attracts is relatively more effective for donor who are more money-oriented, and the change in the composition of the donor group leads to crowding-out due to reputational concerns. However, in contrast to our results, crowding-out susceptibility at the individual level is independent of intrinsic pro-social motivations in Bénabou and Tirole (2006).[24]

Alternatively, it has been previously hypothesized (Andersson et al., 2018; Huck et al., 2015) that motivational crowding out is smaller for individuals with a stronger dependence on the funded good (here school, mosque or health clinic), because these individuals are already driven by economic considerations for contributions, rather than purely pro-social ones. In our case, it may be that lower socio-economic status is associated with a higher dependence on the goods, and therefore lower motivational crowding out. This would be in line with Müller and Rau (2020), who find that crowding out increases with intrinsic motivation.

Non-separability is also relevant for future theoretical work. Modeling should focus on preferences that demonstrate non-separability between private benefits and social motivations.

Several important questions remain for future research. While we focused on monetary incentives, it remains unclear whether other forms of private benefits—such as improved immunity through vaccination or enhanced ecosystem services from conservation—would similarly undermine prosocial

[24] This can be seen from the cross derivative of individual contribution level a in their equation (12) with respect to incentive y and the intrinsic propensity to contribute v_a.

motivations. There is also promise in further research on whether collective, group-level incentives might be less prone to motivational crowding out than individual level incentives (Andersson et al., 2018; Hayes et al., 2019). Understanding how different framings of private benefits affect crowding out, and testing the generalizability of our results beyond rural Senegal, are important next steps. The ongoing need to induce private contributions to public goods and policymakers' increasing reliance on incentives that deliver private benefits to donors—as distinct from those that change the effective price of donations, such as contribution-matching incentives—make these questions highly salient.

References

ADENA, M., R. HAKIMOV, AND S. HUCK (2023): "Charitable Giving by the Poor: A Field Experiment in Kyrgyzstan," *Management Science*, 70, 633–646.

ADENA, M. AND S. HUCK (2022): "Personalized fundraising: A field experiment on threshold matching of donations," *Journal of Economic Behavior & Organization*, 200, 1–20.

AGRAWAL, A., A. CHHATRE, AND E. GERBER (2015): "Motivational Crowding in Sustainable Development Interventions," *American Political Science Review*, 109, 470–487.

AKERS, J. F. AND M. YASUÉ (2019): "Motivational Crowding in Payments for Ecosystem Service Schemes: a Global Systematic Review," *Conservation & Society*, 17, 377–389.

ANDERSSON, K., N. COOK, R. GRILLOS, M. LOPEZ, C. SALK, G. WRIGHT, AND E. MWANGI (2018): "Experimental evidence on payments for forest commons conservation," *Nature Sustainability*, 1, 128–135.

ANDREONI, J. (1990): "Impure Altruism and Donations to Public Goods: A Theory of Warm-Glow Giving," *The Economic Journal*, 100, 464–477.

ANGRIST, J. AND V. LAVY (2009): "The Effects of High Stakes High School Achievement Awards: Evidence from a Randomized Trial," *American Economic Review*, 99, 1384–1414.

ATHEY, S. AND G. IMBENS (2016): "Recursive partitioning for heterogeneous causal effects," *Proceedings of the National Academy of Sciences*, 113, 7353–7360.

ATHEY, S., J. TIBSHIRANI, AND S. WAGER (2019): "Generalized random forests," *The Annals of Statistics*, 47.

BOWLES, S. AND S. POLANIA-REYES (2012): "Economic Incentives and Social Preferences: Substitutes or Complements?" *Journal of Economic Literature*, 50, 368–425.

Brekke, K., S. Kverndokk, and K. Nyborg (2003): "An Economic Model of Moral Motivation," *Journal of Public Economics*, 87, 1967–1983.

Bénabou, R. and J. Tirole (2006): "Incentives and Prosocial Behavior," *American Economic Review*, 96, 1652–1678.

Cardenas, J., J. Stranlund, and C. Willis (2000): "Local Environmental Control and Institutional Crowding-Out," *World Development*, 29, 1719 – 1733.

Chervier, C., G. Le Velly, and D. E. de Blas (2019): "When the Implementation of Payments for Biodiversity Conservation Leads to Motivation Crowding-out: A Case Study From the Cardamoms Forests, Cambodia," *Ecological Economics*, 156, 499–510.

Deci, E. L., R. Koestner, and R. M. Ryan (1999): "A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation." *Psychological Bulletin*, 125, 627–700.

Diederich, J., C. C. Eckel, R. Epperson, T. Goeschl, and P. J. Grossman (2022): "Subsidizing unit donations: matches, rebates, and discounts compared," *Experimental Economics*, 25, 734–758.

Engelmann, D., A. Munro, and M. Valente (2017): "On the behavioural relevance of optional and mandatory impure public goods," *Journal of Economic Psychology*, 61, 134–144.

Frey, B. and F. Oberholzer-Gee (1997): "The Cost of Price Incentives: An Empirical Analysis of Motivation Crowding-Out," *The American Economic Review*, 87, 746–755.

Frey, B. S. and R. Jegen (2001): "Motivation Crowding Theory," *Journal of Economic Surveys*, 15, 589–611.

Gneezy, U. and A. Rustichini (2000a): "A Fine is a Price," *The Journal of Legal Studies*, 29, 1–17.

——— (2000b): "Pay Enough or Don't Pay at All," *The Quarterly Journal of Economics*, 115, 791–810.

Guo, Q., E. Wang, Y. Nie, and J. Shen (2021): "Revisiting the impact of impure public goods on consumers' prosocial behavior: A lab experiment in Shanghai," *Bulletin of Economic Research*, 73, 51–66.

Handberg, O. N. and A. Angelsen (2019): "Pay little, get little, pay more, get a little more: A framed forest experiment in Tanzania," *Ecological Economics*, 156, 454–467.

Hayes, T., T. Grillos, L. L. Bremer, F. Murtinho, and E. Shapiro (2019): "Collective PES: More than the sum of individual incentives," *Environmental Science & Policy*, 102, 1–8.

Henrich, J., R. Boyd, S. Bowles, C. Camerer, E. Fehr, H. Gintis, and R. McElreath (2001): "In Search of Homo Economicus: Behavioral Experiments in 15 Small-Scale Societies," *American Economic Review*, 91, 73–78.

Henrich, J., R. Boyd, S. Bowles, C. Camerer, E. Fehr, H. Gintis, R. McElreath, M. Alvard, A. Barr, J. Ensminger, N. S. Henrich, K. Hill, F. Gil-White, M. Gurven, F. W. Marlowe, J. Q. Patton, and D. Tracer (2005): "'Economic man' in cross-cultural perspective: Behavioral experiments in 15 small-scale societies," *Behavioral and Brain Sciences*, 28, 795–815.

Huang, Y., X. Zhang, X. Sheng, Y. Wang, and K. M. Y. Leung (2024): "The impact of payment for ecosystem service schemes on participants' motivation: A global assessment," *Ecosystem Services*, 65, 101595.

Huck, S. and I. Rasul (2011): "Matched fundraising: Evidence from a natural field experiment," *Journal of Public Economics*, 95, 351–362, charitable Giving and Fundraising Special Issue.

Huck, S., I. Rasul, and A. Shephard (2015): "Comparing Charitable Fundraising Schemes: Evidence from a Natural Field Experiment and a Structural Model," *American Economic Journal: Economic Policy*, 7, 326–69.

Karlan, D., J. A. List, and E. Shafir (2011): "Small matches and charitable giving: Evidence from a natural field experiment," *Journal of Public Economics*, 95, 344–350, charitable Giving and Fundraising Special Issue.

Kerr, J., M. Vardhan, and R. Jindal (2012): "Prosocial behavior and incentives: Evidence from field experiments in rural Mexico and Tanzania," *Ecological Economics*, 73, 220–227.

Leuven, E., H. Oosterbeek, and B. van der Klaauw (2010): "The Effect of Financial Rewards on Students' Achievement: Evidence from a Randomized Experiment," *Journal of the European Economic Association*, 8, 1243–1265.

Mellström, C. and M. Johannesson (2008): "Crowding out in Blood Donation: Was Titmuss Right?" *Journal of the European Economic Association*, 6, 845–863.

Munro, A. and M. Valente (2016): "Green goods: Are they good or bad news for the environment? Evidence from a laboratory experiment on impure public goods," *Environmental and Resource Economics*, 65, 317–335.

Müller, S. and H. A. Rau (2020): "Motivational crowding out effects in charitable giving: Experimental evidence," *Journal of Economic Psychology*, 76, 102210.

Narloch, U., U. Pascual, and A. G. Drucker (2012): "Collective Action Dynamics under External Rewards: Experimental Insights from Andean Farming Communities," *World Development*, 40, 2096–2107.

Oniki, S., M. Berhe, T. Negash, and H. Etsay (2023): "Do economic incentives crowd out motivation for communal land conservation in Ethiopia?" *Forest Policy and Economics*, 150, 102948.

Rode, J., E. Gómez-Baggethun, and T. Krause (2015): "Motivation crowding by economic incentives in conservation policy: A review of the empirical evidence," *Ecological Economics*, 117, 270–282.

Tibshirani, J., S. Athey, E. Sverdrup, and S. Wager (2024): *grf: Generalized Random Forests*, r package version 2.3.2.

Van Hecken, G. and J. Bastiaensen (2010): "Payments for Ecosystem Services in Nicaragua: Do Market-based Approaches Work?" *Development and Change*, 41, 421–444.

Vorlaufer, T., S. Engel, J. de Laat, and B. Vollan (2023): "Payments for ecosystem services did not crowd out pro-environmental behavior: Long-term experimental evidence from Uganda," *Proceedings of the National Academy of Sciences*, 120, e2215465120.

Wager, S. and S. Athey (2018): "Estimation and Inference of Heterogeneous Treatment Effects using Random Forests," *Journal of the American Statistical Association*, 113, 1228–1242.

Wollbrant, C., M. Knutsson, and P. Martinsson (2022): "Extrinsic rewards and crowding-out of prosocial behaviour," *Nature Human Behavior*, 6, 774–781.

Xiao, E. and D. Houser (2011): "Punish in public," *Journal of Public Economics*, 95, 1006–1017.

Supplemental Appendix

A Proof and illustration regarding the theoretical model

A.1 Derivation of the individual and socially desirable level of donations

An individual maximizes utility subject to the budget constraint:

$$\begin{aligned}\mathcal{L} =& u^i(x^i, G, I^i) + \lambda^i(x^i + g^i - w^i - b^i) + \varphi(\sum_j g^j - G) \\ &+ \zeta(I^i(g^i, b^i) - I^i) + \nu^i(b^i(g^i) - b^i)\end{aligned} \tag{15}$$

The related first order condition is:

$$u^i_G = u^i_{x^i} - u^i_{x^i} b^i_{g^i}(g^i) - u^i_{I^i}\left(I^i_{g^i}(g^i, b^i) + I^i_{b^i}(g^i, b^i) b^i_{g^i}(g^i)\right), \tag{16}$$

A social planner would aim for the Pareto optimal supply of a public good for N individuals, subject to budget constraints for all individuals and the accounting equation for the public good. To simplify notation, define $\psi^i = 1$ and $\bar{u}^i = 0$ for $i = 1$, $i \in N$, then set up the Lagrangian:

$$\begin{aligned}\mathcal{L} =& \sum_j \psi^j \left(u^j(x^j, G, I^j) - \bar{u}^j\right) + \lambda^j(x^j + g^j - w^j - b^j) + \varphi(\sum_j g^j - G) \\ &+ \zeta(I^j(g^j, b^j) - I^j) + \nu^j(b^j(g^j) - b^j)\end{aligned} \tag{17}$$

This results in the Lindahl-Samuelson condition, where the asterisk marks the socially optimal solution:

$$\frac{1}{\psi^i}\sum_i \psi^j u^j_G = u^i_{x^i} - u^i_{x^i} b^i_{g^*}(g^*) - u^i_{I^i}\left(I^i_{g^*}(g^*, b^i) + I^i_{b^i}(g^*, b^i) b^i_{g^*}(g^*)\right). \tag{18}$$

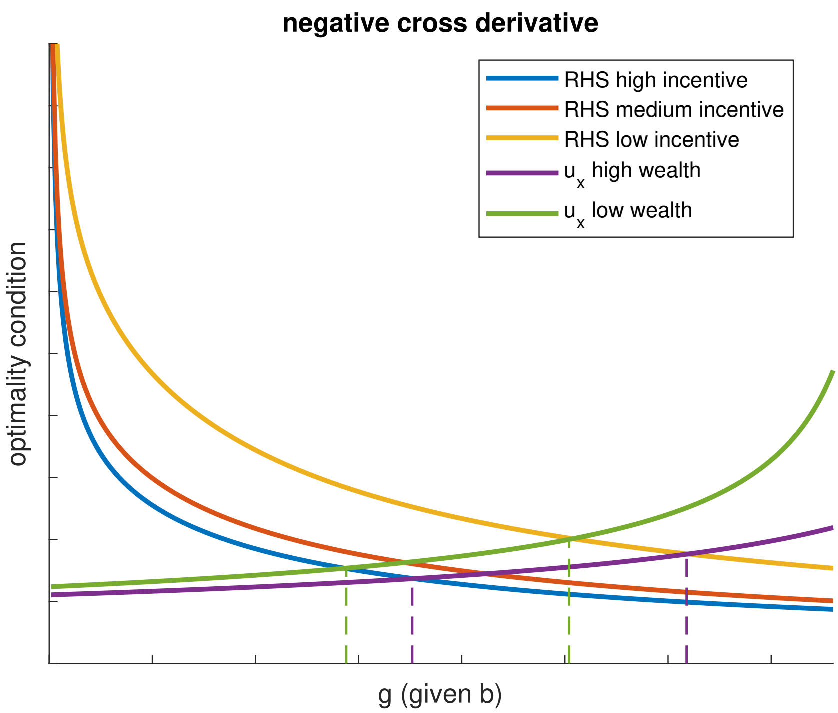
The Lindahl-Samuelson condition under symmetry is:

$$Nu^i_G = u^i_{x^i} - u^i_{x^i} b^i_{g^*}(g^*) - u^i_{I^i}\left(I^i_{g^*}(g^*, b^i) + I^i_{b^i}(g^*, b^i) b^i_{g^*}(g^*)\right). \quad (19)$$

Without private incentive, this simplifies to the standard Lindahl-Samuelson-condition. With private incentive, the effect of b on g^* is ambiguous.

A.2 Illustration of a potential wealth effect in the theoretical model

Figure A.1: Absolute motivational crowding depends on wealth



Notes: The figure shows the right and left and side of equation 7 for the case of a negative cross derivative $I^i_{g^i\bar{b}}$ (non-separable motivational crowding out). The optimal donation level lies at the intersection point. Higher wealth shifts the left hand side (marginal utility of the private good) to the lower right and leads to a higher absolute crowding out effect.

B Summary statistics and balance tables

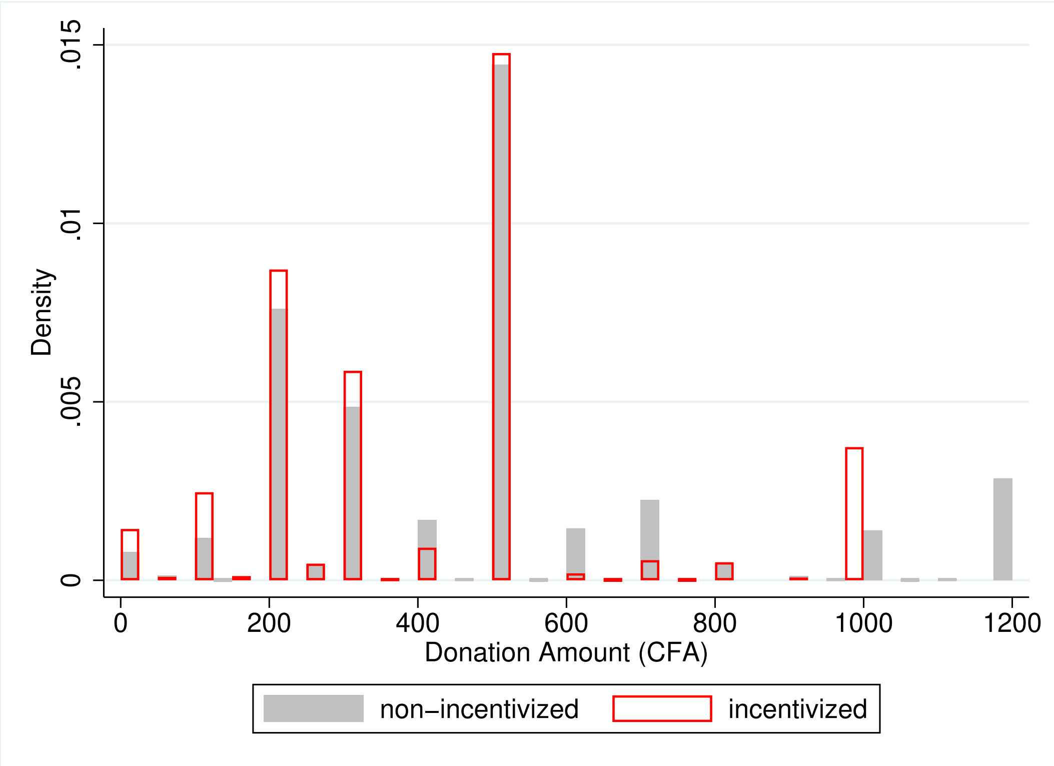


Figure B.1: Histogram to show distribution of donations for each game type

Table B.1: Donation amounts by donation destination

	Non-Incentivized Game	Incentivized Game
Donation Destination: School		
Count	587	587
Mean	463.46	403.01
St. Dev.	266.29	240.47
Min	0	0
Max	1200	1000
Donation Destination: Mosque		
Count	671	671
Mean	478.47	406.33
St. Dev.	290.27	245.82
Min	0	0
Max	1200	1000
Donation Destination: Clinic		
Count	800	800
Mean	481.38	418.09
St. Dev.	296.57	265.06
Min	0	0
Max	1200	1000
Total		
Count	2058	2058
Mean	475.32	409.95
St. Dev.	286.12	251.97
Min	0	0
Max	1200	1000

Notes. Summary statistics of the donation amount in the non-incentivized game and incentivized game by donation destination in CFA.

Table B.2: Balance table of household level variables

Variable	Descriptives					Balance Test	
	Obs.	Mean	Std. dev.	Min.	Max.	coeff	pval
Respondent's gender (0=male)	2058	0.478	0.478	0	1	-0.027	0.252
Respondent's age	2052	48.241	15.299	16	91	-0.304	0.680
Resp.'s education (categ., 0-4)	2058	0.426	0.772	0	3	0.068	0.068
Less than primary level (0)	2052	0.707	0.455	0	1	-0.046	0.035
Primary level (1)	2052	0.153	0.36	0	1	0.014	0.423
Secondary level (2)	2052	0.098	0.298	0	1	0.026	0.081
Higher level (3)	2052	0.026	0.159	0	1	0.001	0.925
Respondent's ethnicity:							
Wolof	2052	0.393	0.488	0	1	-0.004	0.865
Peulh	2052	0.484	0.5	0	1	0.014	0.564
Moor	2052	0.079	0.27	0	1	0.002	0.861
Other	2052	0.044	0.205	0	1	-0.012	0.193
Participate: women's group	2040	0.662	0.473	0	1	0.047	0.033
Participate: savings groups	2025	0.459	0.498	0	1	0.094	0.000
Owns mobile money acc.	2049	0.91	0.287	0	1	0.002	0.872
Community beliefs	2058	0.068	0.252	0	1	0.023	0.074
Lockean beliefs	2058	0.091	0.288	0	1	0.036	0.017
Household size	2058	8.27	3.758	2	55	-0.168	0.347
No. of children in hh	2058	3.304	2.302	0	26	-0.123	0.261
Agric. land cultivation	2045	0.627	0.484	0	1	-0.003	0.89
Own cattle	2058	0.84	0.367	0	1	-0.019	0.276
Livestock ownership	2058	10.051	17.123	0	308	-0.655	0.42
Hh wealth index	2058	0.367	0.143	0.063	0.875	-0.004	0.508
Access to electricity	2058	0.826	0.38	0	1	0.002	0.914
Toilet type	2058	3.272	1.226	0	6	0.034	0.563
Treats water	2057	0.284	0.451	0	1	-0.073	0.000
Water source:							
Own tap	2037	0.657	0.475	0	1	-0.031	0.173
Public tap	2037	0.168	0.374	0	1	-0.003	0.871
Neighboring tap	2037	0.091	0.287	0	1	0.008	0.556
Unprotected well	2037	0.03	0.172	0	1	0.017	0.06
Source/ stream	2037	0.04	0.195	0	1	0.003	0.731
Main income source:							
Fisheries	2058	0.096	0.295	0	1	0.017	0.226
Craft	2058	0.05	0.217	0	1	0.000	0.988
Commerce	2058	0.31	0.463	0	1	0.053	0.018
Service	2058	0.041	0.198	0	1	0.017	0.089
Employed	2058	0.136	0.342	0	1	-0.008	0.603
Transport	2058	0.044	0.205	0	1	0.015	0.142
Harvest	2058	0.129	0.336	0	1	-0.022	0.162
Main cooking fuel:							
Charcoal	2058	0.232	0.422	0	1	-0.03	0.129
Wood	2058	0.675	0.468	0	1	0.025	0.261
Gas	2058	0.091	0.288	0	1	0.004	0.757
Food insecurity (self-rep.)	2058	2.890	3.221	0	12	0.123	0.423
Illness in household	2058	0.961	0.194	0	1	-0.017	0.087

Notes. Summary statistics and balance test of coin toss for game order. We conduct a t-test with robust standard errors for categorical variables. Resp.'s education is a categorical variable for the respondent's education level. No. of children in the hh is the number of children in the household. Agric. land cultivation is an indicator variable that takes the value of one if at least one member of the household cultivated land in the last growing season. Toilet type is a categorical variable for the type of toilet in the household. Food insecurity (self-rep.) is the count of affirmative answers to a 12-question food security module. The balance coefficient represents the difference between the average characteristics of respondents who played the incentivized game first, and those who played it second. "Test pval" represents the p-value of a t-test for whether this difference in means is statistically significant.

Table B.3: Balance table of village level variables

Variable		Descriptives				Balance Test	
	Obs.	Mean	Std. dev.	Min	Max	coeff	pval
Population size	2038	2354.26	3309.965	90	30000	31.552	0.842
No. of households	2038	258.066	548.885	25	5000	6.007	0.819
Distance major city	2058	18.731	13.555	1	64	-1.098	0.088
Distance to market	2038	16.828	14.561	0	60	0.448	0.519
Distance water point	2058	0.300	1.132	0	11	-0.085	0.112
Distance health facility	2058	3.342	5.829	0	37	0.237	0.391
Department:							
St. Louis	2058	0.251	0.434	0	1	-0.032	0.113
Podor	2058	0.356	0.479	0	1	0.018	0.437
Dagana	2058	0.336	0.472	0	1	0.027	0.241
Louga	2058	0.058	0.233	0	1	-0.012	0.253
Donation recipient type:							
School	2058	0.285	0.452	0	1	0.001	0.962
Clinic	2058	0.326	0.469	0	1	0.011	0.61
Mosque	2058	0.389	0.488	0	1	-0.012	0.591
Presence: public transp.	2058	0.566	0.496	0	1	-0.014	0.55
Presence: school	2058	0.923	0.267	0	1	0.007	0.59
Presence: health facility	2058	0.624	0.485	0	1	0.001	0.974
Presence: money agent	2058	0.479	0.5	0	1	-0.023	0.343
Presence: electr. conn.	2058	0.826	0.38	0	1	-0.009	0.608
Daily wage ag	2058	2395.044	539.215	1300	4000	-20.386	0.426
Daily wage non-ag	2058	3142.857	791.913	2000	5000	14.814	0.694
Trade in key agricultural goods:							
Corn	2058	0.722	0.448	0	1	-0.009	0.684
Millet	2058	0.924	0.265	0	1	-0.011	0.386
Sorghum	2058	0.597	0.491	0	1	0.017	0.46
Cowpea	2058	0.981	0.136	0	1	-0.004	0.559
Tomatoes	2058	0.771	0.421	0	1	-0.01	0.605
Peanuts	2058	0.952	0.213	0	1	0.006	0.521
Market prices for rice	2058	365.525	40.059	275	500	-4.35	0.022
Market prices agr. goods	2058	687.658	195.866	150	1100	13.493	0.147
Presence of community organizations:							
Agriculture groups	2058	0.913	0.283	0	1	0.017	0.187
Credit groups	2058	0.182	0.386	0	1	0.023	0.222
Youth groups	2058	0.771	0.421	0	1	-0.017	0.395
Religious groups	2058	0.846	0.361	0	1	-0.005	0.76
Dev. projects ag	2058	0.24	0.427	0	1	-0.019	0.351
Dev. projects water	2058	0.29	0.454	0	1	-0.031	0.147

Notes. Summary statistics and balance test of coin toss for game order. We conduct a t-test with robust standard errors for categorical variables. Presence: electr. conn. is an indicator that takes the value of one if a village has network electricity. Dev. projects ag and Dev. projects water are indicator variables for a village having development projects in agriculture or water, respectively. The balance coefficient represents the difference between the average characteristics of respondents who played the incentivized game first, and those who played it second. "Test pval" represents the p-value of a t-test for whether this difference in means is statistically significant.

C Analyzing heterogeneity using causal forests

This appendix provides technical details of the causal forest methodology used to analyze heterogeneous treatment effects in Section 4.4 of the main text.

C.1 Overview

Causal forests are an extension of random forests, adapted specifically for estimating heterogeneous treatment effects. The method builds on causal trees (Athey and Imbens, 2016), which partition the covariate space into subgroups to identify regions where treatment effects are relatively homogeneous. Unlike standard decision trees used for prediction, causal trees are designed to estimate treatment effects by comparing outcomes between treated and control units within each partition.

While individual causal trees can be powerful for uncovering heterogeneity, they are prone to overfitting and can be sensitive to small changes in the data. Causal forests address these concerns by combining many causal trees, each constructed on a random subsample of the data and considering a random subset of available covariates at each split (Wager and Athey, 2018). This approach helps reduce overfitting and increase stability. Additionally, causal forests can produce valid confidence intervals for estimated treatment effects through techniques like "honest" trees, where separate subsamples are used for determining tree structure and estimating treatment effects (Athey et al., 2019).

C.2 Implementation details

Our analysis proceeded in five steps:

Step 1: Base model estimation

We first estimated the specification shown as specification (4) in Table 5, comprehensively controlling for mechanical aspects of the public goods game, including order effects, non-incentivized game donation amount, and threshold effects.

Step 2: Residual preparation

We predicted residuals from this regression and added back the effects of the incentivized game treatment (β_1), game order, and non-incentivized donation amount. These "adjusted" residuals served as our outcome variable for the causal forest analysis.

Step 3: Causal forest specification

We fit a causal forest using the `grf` package (Tibshirani et al., 2024) in `R`, with the following specifications:

- Treatment variable: an indicator for playing the incentivized game
- Number of trees: 15,000
- Covariates: Over 120 variables (see Table C.1)
- Random covariate selection: One-third of covariates considered as potential splitting variables for each tree
- Minimum leaf size: 25 observations
- Variable discretization: All continuous variables converted to categorical variables to address potential bias

Step 4: Variable importance calculation

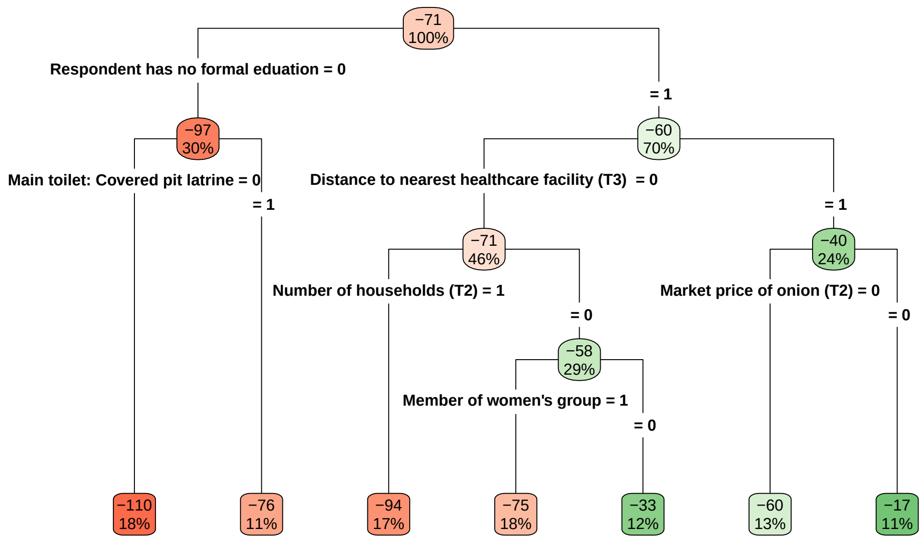
We computed variable importance measures (VIMs) across all trees to identify key drivers of heterogeneity. These measures reflect how frequently each variable is used for splitting, with greater weight given to splits closer to the tree root. For discretized variables, we summed VIMs across constituent dummy variables and rescaled all measures to range from 0 to 100. The top five variables based on VIMs used for our main analysis are shown in Table 6.

Step 5: Treatment effect estimation

For our main analysis (shown in Table 7), we predicted conditional average treatment effects (CATEs) for each household and conducted heterogeneity tests using the top five variables by VIM score following the empirical specification outlined in equation (14). As a robustness check, we also repeated this analysis using discretized versions of the three continuous variables (wealth index, number of households in community, and distance to nearest healthcare facility) that were used to generate the causal forest. The results, shown in Table C.2, are qualitatively similar to our main findings.

C.3 Illustrative causal tree

Figure C.1: Illustrative causal tree



Notes. This figure shows an illustrative causal tree showing heterogeneity in the effect of playing the incentivized game on donation amounts. Each node shows the average treatment effect (in CFA) at the top and the percentage of the sample in that node at the bottom. Positive values indicate an increase in donations in the incentivized game, while negative values indicate a decrease. Red and green shading indicate relatively more negative and positive values, respectively. This tree is constructed with a minimum node size of 225 observations and considers all available splitting variables shown in Table C.1, in contrast to the main causal forest analysis. "T2" and "T3" refer to the second and third tercile, respectively.

We generate a simple, illustrative causal tree to demonstrate the intuition behind our approach (Figure C.1). The tree shows how the effect of playing

Note that the approach we use to fit this tree differs from that we use for our main causal forest analysis in two key ways: (1) the illustrative tree uses a larger minimum the incentivized game on donation amounts varies across different subgroups of participants.

At the root, we see that the average effect across the full sample is a reduction of CFA 71 in donations when playing the incentivized game. The tree then splits the sample based on various characteristics.

The first split is based on education, separating those with no formal education (right branch, 70% of sample) from those with some education (left branch, 30% of sample). For those with some education, the effect is larger (a reduction of CFA 97) compared to those with no education (reduction of CFA 60).

Further splits reveal additional heterogeneity. Among those with some education, the presence of a covered pit latrine creates another meaningful split—those without covered pit latrines show a larger reduction in donations (CFA 110, 18% of sample) compared to those with them (CFA 76, 11% of sample).

For participants with no formal education, the tree branches based on distance to healthcare facilities, number of households in the village, women's group membership, and market prices of onions. For example, among those closer to healthcare facilities (46% of sample), villages with the number of households in the second tercile show a larger reduction in donations (CFA 94, 17% of sample) compared to those with the number of households in other terciles who are members of women's groups (CFA 75, 18% of sample). The smallest reduction in donations (CFA 17) is observed among those with no education, further from healthcare facilities, and in areas with onion market prices in the first or third terciles (11% of sample).

This single-tree illustration demonstrates how causal trees enable data-driven identification of subgroups with heterogeneous treatment effects by recursively partitioning the data into subgroups based on covariates. However, individual trees can be susceptible to overfitting and sensitive to idiosyncrasies in the data. Our main causal forest analysis, which aggregates information across 15,000 trees, addresses these concerns.

node size (at least 225 observations in each leaf) to create a less complex, more intuitive tree structure for illustration purposes; and (2) it considers all possible splitting variables shown in Table C.1 instead of a randomly selected subset.

Table C.1: Potential splitting variables for causal forest

Level	Variable	Discretization (if applicable)
Household	Gender of respondent	
	Participation in women's groups	
	Participation in informal savings groups	
	Mobile money account ownership	
	Illness in household in last 12 months	
	Agricultural land cultivation	
	Main lighting fuel is electricity	
	Treats water	
	Own cattle	
	Total cattle ownership	Tercile dummies
	Time spent collecting water	Tercile dummies
	Length of lean period in last 12 months	Tercile dummies
	Number of children in household	Tercile dummies
	Household wealth index	Tercile dummies
	Household size	Tercile dummies
	Age of respondent	Tercile dummies
	Main water source for household use	Separate dummy for each category
	Main toilet type	Separate dummy for each category
	Main income source	Separate dummy for each category
	Main cooking fuel	Separate dummy for each category
	Ethnicity of respondent	Separate dummy for each category
	Education level of respondent	Separate dummy for each category
	Agreement with statements on community versus individual resource ownership	Separate dummy for each category
Community	Presence of public transportation in village	
	Presence of school in village	
	Presence of health facility in village	
	Presence of mobile money agent in village	
	Presence of grid electricity in village	
	Typical daily agricultural wage (CFA)	Above-median dummy
	Typical daily non-agricultural wage (CFA)	Above-median dummy
	Village population	Tercile dummies
	Number of households in village	Tercile dummies
	Distance to nearest city (km)	Tercile dummies
	Distance to nearest weekly market (km)	Tercile dummies
	Distance to nearest water point (km)	Tercile dummies
	Distance to nearest health facility (km)	Tercile dummies
	Market price of rice	Tercile dummies
	Market prices of onions	Tercile dummies
	Department (administrative unit)	Separate dummy for each category
	Presence of community organizations	Separate dummy for each category
	Recent development projects in village	Separate dummy for each category
	Availability of key agricultural commodities in local market	Separate dummy for each category
	Donation recipient (i.e., school, mosque, or clinic)	Separate dummy for each category

Notes. This table lists the variables specified in the causal forest algorithm as potential splitting variables. Where relevant, it also indicates how continuous variables were discretized to mitigate the bias tree-based methods have towards continuous variables.

Table C.2: Household and community characteristics associated with motivational crowding out

	(1)	(2)	(3)
Constant (β_1)	-68.1***	-68.6***	-72.6***
	(0.24)	(0.33)	(0.29)
Below 200 (β_2)	1.78**	2.56***	3.60***
	(0.75)	(0.80)	(0.98)
Above 1000 (β_3)	-5.81***	-7.57***	-7.91***
	(0.72)	(0.76)	(0.77)
Wealth index - second tercile	-5.98***		
	(0.39)		
Wealth index - third tercile	-13.0***		
	(0.39)		
Number of households in community - second tercile		-7.74***	
		(0.47)	
Number of households in community - third tercile		-3.33***	
		(0.43)	
Distance to nearest healthcare facility - second tercile			-4.17***
			(0.49)
Distance to nearest healthcare facility - third tercile			5.04***
			(0.43)
R^2	0.36	0.18	0.21
RMSE	7.10	8.02	7.87

Notes. This table examines how the effect of the incentivized game on donations varies across different household and community characteristics, focusing on terciles of of each continuous variable (wealth index, number of households in community, and distance to nearest healthcare facility), with the first tercile serving as the reference group. The outcome variable represents the predicted effect of playing the incentivized game for each household in our sample. Negative values indicate larger reductions in donations when playing the incentivized game. The regression is weighted by the inverse of the variance of the predicted treatment effects. Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

D Robustness checks

All results are available upon request, if not printed below.

D.1 Comprehension checks

In terms of comprehension, we found that part of the sample exhibited remarkable behavior, such as donating below the threshold in the incentivized game (212 participants (10%)), forfeiting the bonus payment. Furthermore, for 29 participants (1%), there was a difference of more than CFA 800 between their donations in the two games.

Enumerator assessments indicated that 93% of participants understood the games well, while 7% understood only certain aspects or very little. Those with poor comprehension largely overlapped with participants showing significant discrepancies in their donation amounts.

The majority of participants who donated below threshold in the incentivized game were assessed as having full comprehension. Their low donations may not stem from comprehension issues, but rather from a deliberate choice to minimize contributions, possibly reflecting strong disagreements with the purpose of the donations, as 42% occurred when the donations were targeted to the local mosque, versus just 29% for both school and health clinic donations.

We run the two main models 2 and 4 (from Table 3 and Table 5), again while dropping these observations. For model 2, results are almost identical to those obtained earlier, with the order effect changing slightly to CFA -79 (se=9.61, $p < 0.01$) (from CFA -105). For model 4, we again observe no qualitative differences. The constant and the linear effect of non-incentivized donation size remain very robust. The order effect again changes to -77 (se=8.54, $p < 0.01$) (from 104).

A similar exercise eliminating 142 participants who according to enumerator assessment did not comprehend well leads to results that are very close to the results reported in the main part of the paper.

Detailed results are available in the accompanying Stata files.

D.2 Random effects model

Since each individual played exactly once the non-incentivized and once the incentivized donation game, direct order effect and the interaction with the incentivized treatment dummy are not separable in a fixed effects specification. We therefore include a random effects model, with the donation in the respective game as outcome variable, in order to be able to include both the direct effect of order and its interaction with the incentivized treatment dummy. We compare this random effects model with a similar fixed effects model, which is equivalent to our OLS specification with the difference of donations as outcome variable. As errors correlate with observations lying above 1000 or below 200, we dis-include these two variables. We estimate the model as a random effects model using Stata's *xtreg* command, using robust standard errors to account for heteroscedasticity.

Let k be a subscript that indexes the type of game played, let g_{ik} denote the donation by individual i in game k, α_i an intercept, I_{ik} is a binary variable that is 1 if the observation is from the impure donation game and zero otherwise, O_{ik} is a binary variable that is 1 if the game played is the second game played and zero otherwise, u_i is an unobserved individual-specific effect uncorrelated with the explanatory variables, and ϵ_{ik} is an error term. We estimate:

$$g_{ik} = \alpha_i + \beta_1 I_{ik} + \beta_2 O_{ik} + \beta_3 O_{ik} I_{ik} + u_i + \epsilon_{ik} \tag{20}$$

Results are summarized in the first column of Table D.1. As before, we observe a statistically significant motivational crowding out effect (interaction of the incentivized treatment with order β_3) similar to the effect observed in model 2 (see section 3), but no motivational crowding out when the incentivized game is played first.

The linear combination of incentivized treatment effect, order effect, and the interaction of the two is estimated as CFA -100.689 (6.049), i.e., crowding-out amounts to 21% of average non-incentivized game donations, and significant at the 1% level. However, a test of overidentifying restrictions using Stata's *xtoverid* rejects H0 that the additional random effects assumptions hold (test stat. 309.109, pval=0.000), such that we conclude to continue with the fixed effects model.

Table D.1: Robustness results

	Random effects	Poisson, as model 2	Poisson, as model 4
Constant (β_1)	-4.408	7.104***	7.286***
	(12.77)	(0.007)	(0.011)
Below 200 (β_2)		0.181***	0.025
		(0.020)	(0.019)
Above 1000 (β_3)		-0.323***	-0.001
		(0.030)	(0.035)
Game played second	-16.087		
	(13.70)		
Play incentivized game second	-80.194***	-0.09***	-0.090***
	(23.21)	(0.009)	(0.008)
Non-incentivized donation level			-0.000***
			(0.000)
within R^2	0.114		
RMSE	165.3	211.044	192.918

Notes. Results from a random effects model (original type is model 2) and poisson models of the original model 2 and model 5 types. The outcome variable is the donation in the respective game for the random effects model, and the difference between donations (incentivized minus non-incentivized donations) for the Poisson models. RMSE: root mean squared error. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

D.3 Model with a Poisson distribution

Given that our participants can only decide on their donations in steps of CFA 50, the data could be interpreted as count data. For robustness, we therefore run a regression with a Poisson specification. Since our outcome variable, the difference in donations sizes across games, may be negative, we transform this outcome variable by adding a constant equal to the minimum observed difference (-1,200). Table D.1 summarizes results in the two last columns, for model types mirroring original models 2 and 4, respectively. Treatment effects remain robust, we find significant motivational crowding out in particular when the incentivized game is played second (model 2 type), and for high non-incentivized game donation levels (model 4 type). RMSE values are similar to those obtained in our linear model specifications. An alternative negative binomial specification produces similar results.

D.4 Donation size dependence using a non-linear specification of non-incentivized donations

When controlling for non-incentivized game donation size (column 4 in Table 3), results showed that the motivational crowding out effect depends on non-incentivized game donations. We test whether this relationship is actually linear using an inverse hyperbolic sine transformation of non-incentivized game donations. The treatment effects are significant and signs remain robust. The linear combination of treatment effects for non-incentivized donations at CFA500 amounts to -110.816 (se=5.60) if the incentivized game was played second and CFA -8.357 (se= 8.36) if the incentivized game was played first, which is comparable with our main results findings (see Figure 2). Model statistics are slightly worse than for the original model.

Table D.2: Nonlinear specification results

	IHS
Constant (β_1)	706.657***
	(60.05)
Below 200 (β_2)	-110.107***
	(21.27)
Above 1000 (β_3)	-201.149***
	(26.82)
Play incentivized game second	-102.459***
	(9.47)
Non-incentivized donation level, IHS transformed	-103.509***
	(8.93)
R^2	0.324
RMSE	197.0

Notes. Model with nonlinear specification of the non-incentivized donation effect, using the inverse hyperbolic sine transformation (IHS). The outcome variable is the difference in donations, i.e. the incentivized minus the non-incentivized donations. RMSE: root mean squared error. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

D.5 Fully interacted design

In the main specification, for simplicity we do not include all conceivable interaction effects while still running the model on all observations, which may raise concerns about interpretation of the coefficient estimates. We therefore provide below results from a fully interacted design. Results remain robust both with respect to the donation level effect as well as with respect to the effect of playing the incentivized game second, for which the parameter is reduced in absolute terms but not significantly different from the parameter estimated in specification (4, Table 5.

D.6 Subset with treatment in second game

Endogeneity may arise in specification 4 (see Table 5), where we regress the difference in donations on the donation in the non-incentivized game. If the incentivized game was played first, lasting effects of the incentive could affect the baseline donation, violating the exogeneity assumption. To address this, we re-estimate specification 4 using only the subset of observations in which the incentivized game was played after the non-incentivized game, ensuring that the treatment could not have influenced the baseline donation. We find that (Table D.4) the effect of the non-incentivized donation level is very close to the effect found in the main specification (-0.469). The constant is very close to the main specification's linear combination of constant and order effect of CFA 118.934 (se= 10.71, $p<0.01$), which is the reference estimate since we can not include order as a covariate in this robustness specification. We conclude that the results remain robust, and it is therefore unlikely that endogeneity is a concern. The alternative specification with enumerator fixed effects likewise does not differ substantially from the corresponding results for the full sample.

D.7 Validating the importance of the standard game donation amount as a driver of heterogeneity in the causal forest approach

As an additional check, we expanded our causal forest analysis to include the standard (non-incentivized) game donation amount alongside the house-

Table D.3: Robustness check: Fully interacted design

	Fully interacted design
Constant (β_1)	201.106***
	(19.70)
Below 200 (β_2)	354.747***
	(100.61)
Above 1000 (β_3)	1590.373***
	(326.13)
Play incentivized game 2nd	-77.821***
	(23.38)
Non-incentivized donation level	-0.426***
	(0.04)
Below 200 x Play inc. game 2nd	-239.280**
	(116.69)
Above 1000 x Play inc. game 2nd	-4446.205***
	(474.46)
Below 200 x Non-inc. donation level	-3.873***
	(0.95)
Above 1000 x Non-inc. donation level	-1.313***
	(0.30)
Play inc. game 2nd x Non-inc. donation level	-0.049
	(0.05)
Below 200 x Play inc. game 2nd x Non-inc. donation level	2.524**
	(1.11)
Above 1000 x Play inc. game 2nd x Non-inc. donation level	3.727***
	(0.44)
R^2	0.371
RMSE	190.4
AIC	27457.3
BIC	27524.9

Notes. The outcome variable is the difference in donations, i.e. the incentivized donations minus the non-incentivized donations. The table shows point estimates and robust standard errors in parentheses. RMSE: root mean squared error. AIC: Akaike information criterion. BIC: Bayesian information criterion. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table D.4: Results on subset where incentivized game was played second

	(4), subset by order	(4), subset by order, FE
Constant (β_1)	125.345***	130.212***
	(12.57)	(18.25)
Below 200 (β_2)	27.252	26.655
	(26.14)	(29.02)
Above 1000 (β_3)	42.355	46.665
	(39.88)	(52.54)
Non-incentivized donation level	-0.480***	-0.491***
	(0.03)	(0.04)
Enum. FE	no	yes
R^2	0.331	0.331
N	1411	1411
RMSE	186.0	184.4
AIC	18755.8	18728.8
BIC	18776.9	18744.5

Notes. The outcome variable is the difference in donations, i.e. the incentivized donations minus the non-incentivized donations. The table shows point estimates and robust standard errors in parentheses. We estimate on the subset of observations where the incentivized game was played after the non-incentivized game. Parameters as in equation (14). FE: Model with enumerator fixed effects (and standard errors clustered at enumerator level). N: number of observations. RMSE: root mean squared error. AIC: Akaike information criterion. BIC: Bayesian information criterion. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table D.5: Variable importance measures for top drivers of heterogeneity using adjusted residuals from model 2 in Table 3

Covariate	VIM (scaled)
Standard game donation amount	100
Distance to nearest healthcare facility	8.57
Number of households in community	4.90
Market price of rice	3.80
Age of respondent	3.28

Notes. This table lists the name and variable importance measure (VIM) for the top five variables identified by the causal forest as drivers of heterogeneity in treatment effects, including standard game donation amount as a potential splitting variable alongside household and community characteristics. VIMs are scaled relative to the most important variable, with mean (SD) of VIMs across all potential splitting variables listed in Table C.1 plus standard game donation amount being 2.20 (10.9).

hold and community characteristics as potential drivers of heterogeneity. For this analysis, we first obtained residuals from the specification of Table 3 (model 2), then added back the effects of the incentivized game treatment (β_1) and game order to create adjusted residual donation amounts. Using these adjusted residuals as the outcome variable, we implemented the causal forest algorithm considering all variables listed in Table C.1 plus the standard game donation amount. As shown in Table D.5, the standard game donation amount emerges as the strongest predictor of heterogeneous treatment effects, with a scaled VIM substantially higher than other characteristics. This finding provides additional support for including the standard game donation amount as a key control variable in our main analysis of treatment effect heterogeneity.

E Scripts for the donation game

Before entering the household, toss a coin. Note the result here.
Coin toss result: Heads () Tails ()

If the coin shows tails, play game A (pure public good) first, and game B (impure public good) second. If the coin shows heads, play game B (impure public good) first and game A (pure public good) second. Follow the respective script carefully.

Tails: game A (pure public good) first, and game B (impure public good) second

We would now like to give you 2200 FCFA in appreciation of your hospitality and the time you have taken with us. You are free to keep those funds for yourself. We will, however, invite you to make confidential contributions to [INSERT GIFT PURPOSE]. We will distribute this gift as part of two activities.

In the first activity, you will receive 1200 CFA in an envelope (hold up an envelope). Once you receive the 1200 CFA, we will ask you to divide up your 1200 CFA in two parts. One part you will put in your pocket to keep. You and your family can decide what to do with it. The other part, you put back into the envelope as a contribution to [INSERT GIFT PURPOSE].

I will then record your decision and seal your envelope for the [INSERT GIFT PURPOSE]. Only I will know your decision; I will not share this information with anyone in the village. No one else will know what you decide. This is your decision and yours only. You can decide to put as much or as little as you want into the envelope. It can be 0 or 1200 CFA or any 100 CFA increment in between. There is no right or wrong decision. It is just a personal choice.

Once we have finished the survey in this village, we will meet openly at [INSERT TIME AND PLACE] to hand over the community gift. You and your household are cordially invited to join us there. There, one of my colleagues or I will sum up all the community gifts of all participants in this village from the sealed envelopes. The envelopes are not marked, so no one will be able to tell what any one individual contributed. Our research team will add another half times the same amount from our research team funds, such that the total sum available to the common purpose is one and a half times as high as the sum donated by all participants. So, if the total amount contributed by the group is 4000 FCFA, we will add 2000 FCFA and donate a total of 6000 FCFA. This total amount will then be donated to [INSERT GIFT PURPOSE].

Enumerator will proceed to do a demonstration of the game with a small amount of FCFA. After demonstration:

B1. Do you have any questions about the game?

[0] No

[1] Yes (If yes, answer any questions about the game and then ask again. Do not continue until the respondent says no.)

B2. Are you willing to participate in the game?

[0] No

[1] Yes

If No, continue to second game.

If Yes, and consent is given, proceed and hand out the envelope. Have the participant place their contribution in the envelope. Record the contributed amount and seal the envelope. Do not pressure people to make a decision quickly. Give them sufficient time.

Note: Sometimes participants might ask what they can do with the money they have. Emphasise that is up to them. They should treat this money as they would any other income they earned.

Take the sealed envelope and put it into a basin that is dedicated to this purpose. Then proceed:

Thank you. Now, for the second activity, you will receive 1000 FCFA in an envelope (hold up an envelope), and again divide up your 1200 FCFA in two parts. One part you will put in your pocket to keep. The other part, you put back into the envelope as a contribution to [INSERT GIFT PURPOSE]. I will then record your decision and seal your envelope for the [INSERT GIFT PURPOSE]. Again, only I will know your decision. At the donation ceremony at [INSERT TIME AND PLACE], we will again add your contribution in the sealed envelope to the community total. We will again add half the same amount from our research team funds, thus increasing the total sum available to the common purpose to one and a half times the sum donated by participants. This total amount will then be donated to [INSERT GIFT PURPOSE].

But there is one difference from the previous activity: If you donate at least 200 CFA to the community gift, I will directly here give you an additional 200 CFA back to keep for yourself and your family.

Enumerator will proceed to do a demonstration of the game with a small amount of FCFA. After demonstration:

B1. Do you have any questions about the game?

[0] No

[1] Yes (If yes, answer any questions about the game and then ask again. Do not continue until the respondent says no.)

B2. Are you willing to participate in the game?

[0] No

[1] Yes

If No, thank the participant for their time and depart.

If Yes, and consent is given, proceed and hand out the envelopes. Have the participant place their contribution in the envelope. Do not pressure people to make a decision quickly. Give them sufficient time. Record the contributed amount and seal the envelope. If at least 200 FCFA are in the envelope, hand the participant an additional 200 FCFA.

Note: Sometimes participants might ask what they can do with the money they have. Emphasize that is up to them. They should treat this money as they would any other income they earned. Take the sealed envelope and put it into a basin that is dedicated to this purpose. Then proceed:

Thank you for taking the time to respond to my questions. Please join us at [insert DAY/TIME] at [LOCATION] for the donation ceremony with the other villagers participating in this study.

Heads: game B (impure public good) first, and game A (pure public good) second

We would now like to give you 2200 FCFA in appreciation of your hospitality and the time you have taken with us. You are free to keep those funds for yourself. We will, however, invite you to make confidential contributions to [INSERT GIFT PURPOSE]. We will distribute this gift as part of two activities.

In the first activity, you will receive 1000 FCFA in an envelope (hold up an envelope). Once you receive the 1000 FCFA, we will ask you to divide up your 1000 FCFA in two parts. One part you will put in your pocket to keep. You and your family can decide what to do with it. The other part, you put back into the envelope as a contribution to [INSERT GIFT PURPOSE].

I will then record your decision and seal your envelope for the [INSERT GIFT PURPOSE]. Only I will know your decision; I will not share this information with anyone in the village. No one else will know what you decide. This is your decision and yours only. You can decide to put as much or as little as you want into the envelope. It can be 0 or 1000 FCFA or any 100 FCFA increment in between. There is no right or wrong decision. It is just a personal choice. However, if you donate at least 200 FCFA to the community gift, I will directly here give you an additional 200 FCFA back to keep for yourself and your family.

Once we have finished the survey in this village, we will meet openly at [INSERT TIME AND PLACE] to hand over the community gift. You and your household are cordially invited to join us there. There, one of my colleagues or I will sum up all the community gifts of all participants in this village from the sealed envelopes. The envelopes are not marked, so no one will be able to tell what any one individual contributed. Our research team will add one half times the same amount from our research team funds, thus we will increase the total sum available to the common purpose to one and a half times the amount donated by participants. So, if the total amount contributed by the group is 4000 FCFA, we will add 2000 FCFA and place a total of 6000 FCFA on the table. This total amount will then be donated to [INSERT GIFT PURPOSE].

Enumerator will proceed to do a demonstration of the game with a small amount of FCFA. After demonstration:

B1. Do you have any questions about the game?

[0] No

[1] Yes (If yes, answer any questions about the game and then ask again. Do not continue until the respondent says no.)

B2. Are you willing to participate in the game?

[0] No

[1] Yes

If No, continue to second game. If Yes, and consent is given, proceed and hand out the envelopes. Have the participant place their contribution in the envelope. Do not pressure people to make a decision quickly. Give them sufficient time. Record the contributed amount and seal the envelope. If at least 200 FCFA are in the envelope, hand the participant an additional 200 FCFA.

Note: Sometimes participants might ask what they can do with the money they have. Emphasise that is up to them. They should treat this money as they would any other income they earned. Take the sealed envelope and put it into a basin that is dedicated to this purpose. Then proceed:

Thank you. Now, for the second activity, you will receive 1200 CFA in an envelope (hold up an envelope), and again divide up your 1200 FCFA in two parts. One part you will put in your pocket to keep. The other part, you put back into the envelope as a contribution to [INSERT GIFT PURPOSE]. I will then record your decision and seal your envelope for the [INSERT GIFT PURPOSE]. Again, only I will know your decision. At the donation ceremony at [INSERT TIME AND PLACE], we will again add your contribution in the sealed envelope to the community total. We will again add one half the same amount from our research team funds, thus increasing the total sum available to the common purpose to one and a half times the total amount donated by participants. This total amount will then be donated to [INSERT GIFT PURPOSE].

But there is one difference from the previous activity: This time, no matter how much you donate, I will not hand out the additional FCFA to you. So there will be no additional payout to you, regardless of the amount that you decide to put into the envelope.

Enumerator will proceed to do a demonstration of the game with a small amount of FCFA. After demonstration:

B1. Do you have any questions about the game?

[0] No

[1] Yes (If yes, answer any questions about the game and then ask again. Do not continue until the respondent says no.)

B2. Are you willing to participate in the game?

[0] No

[1] Yes

If No, thank the participant for their time and depart.

If Yes, and consent is given, proceed and hand out the envelope. Have the participant place their contribution in the envelope. Record the contributed amount and seal the envelope. Do not pressure people to make a decision quickly. Give them sufficient time.

Have the participant place their contribution in the envelope. Do not pressure people to make a decision quickly. Give them sufficient time. Record the contributed amount and seal the envelope. Take the sealed envelope and put it into a basin that is dedicated to this purpose. Then proceed:

Thank you for taking the time to respond to my questions. Please join us at [insert DAY/TIME] at [LOCATION] for the donation ceremony with the other villagers participating in this study.

Donation ceremony (same for both versions)

At the donation ceremony, present a box with all the sealed envelopes. Then, open the envelopes and take out the funds from both games. Do this quickly and try not to show too much how much is in each envelope. Count the total and announce the total. Then, double the total and place the full amount on the table. Donate the full amount to a representative for the chosen community gift.