



Shall we pay all? An experimental test of Random Incentivized Systems

Sophie Clot^{a,*}, Gilles Grolleau^b, Lisette Ibanez^c

^a University of Reading, Whiteknights street, Reading, United Kingdom

^b CEE-M, Montpellier SupAgro, INRA, CNRS, Univ. Montpellier, Univ. Bourgogne Franche-Comté, Burgundy School of Business-CEREN, France

^c CEE-M, INRA, Montpellier SupAgro, CNRS, Univ. Montpellier, France



ARTICLE INFO

Keywords:

Random Incentive Systems
Experimental methodology
Dictator game

ABSTRACT

This study seeks to investigate the impact of payment procedures on behaviour in a standard Dictator game. Although the Between-subjects Random Incentivized System (BRIS) is increasingly utilized in experimental economics, its use could be contested from a purely theoretical point of view. In the BRIS, all subjects play the Dictator game, but only a percentage of subjects (e.g., 1 out of 10) are paid. The aim of this article is to test whether the results obtained using this system are consistent with those obtained using a conventional full payment system in which every subject receives payment. An additional, hypothetical treatment is also run. Results show a very similar pattern between the BRIS and the full payment treatment, whereas the hypothetical payment mechanism leads to far fewer egoistic players and far more egalitarians. These findings lend support to the use of BRIS as a valid experimental methodology.

Experimental economics is a fast-growing field challenged by robustness and external validity concerns. Ever-larger samples are desired, which presents researchers with logistical and financial challenges (Charness et al., 2013). The Between-subject Random Incentivized System (BRIS) has emerged as a potential solution that could enable researchers to test ideas among large samples of subjects without the logistical constraints that arise with full payment systems. Additionally, it reduces the significant costs of paying all subjects and eases the organizational constraints associated with large-scale experiments. Using this mechanism, all subjects play the game, but only a percentage of subjects are paid. If experimenters maintain the endowment at the original level under this random pay procedure, the expected value of the stake (i.e. an individual's statistical expectation of a gamble's outcome) is lowered, and the effectiveness of the incentive becomes debatable from a methodological standpoint. Given that economic experiments are characterized by the use of monetary incentives, this is an important issue to address. In contrast to psychologists, economists argue that conditioning experimental earnings on individual decisions more accurately elicits 'true' preferences and therefore more closely corresponds to the 'real' decisions that individuals would actually make in the field (Kagel and Roth, 1995). From this perspective, incentives should therefore be designed so that rational individuals are able to calculate the costs and benefits of every choice (Madsen and Stenheim, 2015). This position is based on the assumptions that individuals are rational, self-interested and seek to maximize their net payoffs.

The expected value of a 10€ stake that is associated with a one-in-ten probability of receiving payment, for example, is 1€. In order to obtain a situation similar to that of the full pay experiment (in which every player receives a 10€ payment), a prescription of expected value theory is to increase the stake (i.e., to 100€ in this example with a winning probability of 1/10). Moreover, evidence has shown that people are particularly attracted to small probabilities of large rewards and generally tend to overweight small probabilities (Kahneman and Tversky, 1979; Prelec, 1998). Without such insights, how can one explain the fact that sales of the Euromillion 2€ lottery ticket reach up to 3 million purchases per hour at their peak, while the expected gain per ticket is only 0.08€? In a medical study by Volpp et al. (2008), using lotteries as incentives significantly and positively altered health behaviours, showing that even rewards with small expected value may have an incentive effect. Another variation of the standard payment system is the use of hypothetical payments. Also, the literature suggests that hypothetical payments may have a limited incentive effect. In social experiments where payments are hypothetical rather than real, participants act more generously and keep less money for themselves (Sefton, 1992; Forsythe et al., 1994).

In this study, we run a Dictator Game in the laboratory to test whether behaviour under the Between-subject Random Incentivized System (BRIS) is consistent with that observed under a standard individual payment system (Full Payment Scheme) and/or whether these behaviours differ substantially from that observed in a hypothetical setting. Against the standard predictions offered by Expected Value

* Corresponding author.

E-mail addresses: s.clot@reading.ac.uk (S. Clot), grolleau@supagro.inra.fr (G. Grolleau), lisette.ibanez@inra.fr (L. Ibanez).

theory, observed behaviours in a 10€ full pay vs. a 10€ random pay (i.e. one-in-ten probability) experiment demonstrate very similar characteristics. Conversely, when the stake is increased in order to maintain a constant expected value relative to the 10€ full pay procedure (i.e., the chance to win 100€ with a one-in-ten probability), small behavioural deviations from the standard prediction emerge. Finally, when financial incentives are removed entirely, even greater discrepancies emerge: we observe fewer strictly egoistic players and more egalitarians.

1. Background

Despite its growing use (Clot et al., 2014; Brañas-Garza et al., 2013; Exadaktylos et al., 2013; Fong and Luttmer 2009; Carpenter et al., 2008; Langer and Weber 2008; Bettinger and Slonim 2007), no clear-cut evidence from the literature demonstrates that BRIS is statistically reliable. More specifically, the use of the strategy method in dictator games¹ offers an interesting BRIS variant of the standard version. With the strategy method, each subject recruited generate data, while only half of the subjects recruited generate data in the standard version. This paper uses a laboratory experiment based on the standard dictator game (Forsythe et al., 1994) to address the following question: how does the Between-subject Random Incentivized System (BRIS) impact individual decision making compared to the standard incentive system? The dictator game is one of the most popular games among experimentalists. Using a very simple design, it is able to quantify systematic behavioural differences between populations. It has been run in more than one hundred published papers over the last three decades (Engel, 2011), consistently demonstrating violations of the selfishness axiom. Moreover, given that dictator games are susceptible to numerous types of framing effects (Brañas-Garza, 2007; Zizzo, 2010), we contend that any test of invariance to framing effects should include a dictator game, especially in the light of several recent contributions that suggest that people's choices in dictator games change when the probability of realizing the decision changes (e.g., Kellner et al., 2015; Brock et al., 2013). In this work, we analyse how certainty in regard to the realization of expected pay-offs impacts pro-social behaviour. The selfishness axiom of standard economic theory assumes that individuals seek to maximize their net payoffs and thus predicts null donations in a dictator game, regardless of the incentive formula. Numerous studies on pro-social preferences consistently show, however, that a large proportion of people donate a positive amount of money to an anonymous participant (Engel, 2011). According to this literature, we therefore expect to observe positive donations in our experiment. From the expected value theory viewpoint, if participants expect getting less money (e.g., 1/10 to get 10€ compared to 1/10 to get 100€), they are mechanically driven to give less. Indeed, the expected value of their endowment is respectively 1€ (respectively 10€) for the Random pay experiment – Standard stake (Random pay experiment – High stake). For the sake of argument, let us assume that all winning participants (drawn by lots) are absolute altruists and give all what they get. Even in this polar case, what they can give is bounded by what they can get at best, namely either 10€ in a low or standard stake treatment and 100€ in a high stake treatment. In concrete terms, we expect lower donations as the expected value of the amount at stake is lower. Assuming that individuals are equally sensitive to probabilities and stakes, we expect that a higher, probabilistic stake that is equivalent in expected value to a lower, non-probabilistic stake should not alter individual decisions in a systematic way. Finally, in the hypothetical scenario, in which there is no cost of giving, the social desirability bias leads us to expect to observe higher donations than in either of the incentivized scenarios.

Psychologists generally assume that, even in the presence of

monetary rewards, other factors that influence individual decisions (e.g. conformity, social norms) tend to override, or at least significantly affect, payoff-based concerns (Madsen and Stenheim, 2015). Results regarding the comparison of hypothetical vs. incentivized experiments in the literature are mixed, suggesting context dependence (i.e. type of game, sample selection). El Harbi et al. (2015; see also Rubinstein, 2001, 2013), for instance, compared hypothetical and incentivized experimental methods in the context of positional concerns. They argued that, although they may be likely to overestimate the proportion of socially desirable choices,² hypothetical surveys nevertheless constitute a powerful complement to incentivized experiments insofar as they have the capacity to generate interesting qualitative insights at a low cost. Falk et al. (2016) developed a survey tool aiming to measure risk, time and social preferences based on hypothetical questions previously validated by financially incentivised experiments. Ben-Ner et al. (2008) found similar behavioural patterns between a dictator game played with actual money and a hypothetical dictator game, suggesting that the “average dictator is not acting less generously when payments are real than hypothetical”. Nevertheless, the effect of monetary incentives depends on two individual traits of participants: extraversion and agreeableness.

Studies about BRIS are scarce. A few exceptions apply, namely for the ultimatum game (Bolle, 1990), as well as less conventional experiments such as the ‘Deal or no Deal’ game (Baltussen et al., 2011). BRIS has also been studied using the Hybrid Random Incentivized System (HRIS) mechanism (Armantier, 2006; Stahl and Haruvy, 2006), which is a combination of BRIS and a Within-subject Random Incentivized System (WRIS).³ However, these studies embed confounding factors (i.e., collateral results), which obscure the issue. Additional probabilities inherent to either the experimental design (i.e., strategy method, risky choice) or the incentive mechanism (i.e., HRIS) make the full pay version of these experiments unsuitable as representations of baseline behaviour because they distort the elicitation of true preferences.

Bolle (1990) compared the outcomes of \$10 ultimatum games under random pay and full pay procedures. In the random pay experiment, 2 out of 20 subjects were randomly selected to receive payment. In the full pay experiment, all 24 participants were paid. The hypothesis of identical offers in experiments with random pay and full pay was not rejected. However, the ultimatum game in this experiment was run in its strategic version (i.e., each pair of subjects played both roles and a lottery subsequently determined which player would be the proposer). Given the role uncertainty in this design, even the full pay experiment does not perfectly correspond to decision scenarios in which the potential gain is certain. The same criticism applies to Baltussen et al. (2011) study. Their experiment mimics the choice problems in the TV game show Deal or No Deal. In each round of the game, a subject must choose between a sure alternative and a risky lottery with known probabilities. Even in the full pay version of the game, however, participant's observed behaviour is presumably driven not only by their social preferences, but also by their risk preferences.

² For instance, Rubinstein (2001, p. 626) states that “paying subjects will change the distribution of responses, however, assuming that we only want to confirm the existence of a plausible pattern of reasoning it seems unlikely that whether or not we pay the subjects will effect results more than an infinite number of other factors (such as, gender, age, profession, time of day, mood, etc.)”.

³ Broadly defined, Random Incentivized Systems include two different categories: the Within-subject Random Incentivized System (WRIS) and the Between-subject Random Incentivized System (BRIS). The WRIS consists in randomly paying only one task over a set of multiple tasks. An oft-cited argument in favor of this system is that the WRIS avoids the endowment effect (i.e. when a player's choice is influenced by the size of the potential gains the player has accrued from previous games). The validity of the Within-subject RIS has been investigated in many studies. While past research has shown that decisions follow the rule of isolation (i.e., each decision is made independently) more than reduction (Beattie and Loomes 1997, Cubitt et al 1998, Camerer 1989), recent research from Cox et al., 2015 suggests that subjects do not in fact isolate each individual decision in multiple decision experiments.

¹ Each subject plays the game as the Dictator, but would potentially be either a Dictator or Recipient with one half probability (Brañas-Garza et al. 2013)

In order to correctly assess the impact of a random payment alone, the experimental design should not involve additional probabilities and/or strategic interactions. It is for this reason that the dictator game in its simplest version constitutes an appropriate design to test the Between-subject Random Incentivised System (Sefton, 1992). Sefton (1992) compared data from random pay dictator games with data from full pay dictator games from Forsythe et al. (1994) and concluded that subjects were more altruistic under uncertainty. However, this experiment has several limitations in terms of current methodological standards. First, the financial stake was moderate (\$5 vs. a usual minimum of \$10), and second, the sample size was relatively small (69 pooled pairs for the full pay experiment and only 24 pairs in the random pay experiment). In addition, the hypothesis concerning the impact of random pay under constant expected stake value was not tested.

Given that the existing literature offers no clear conclusions regarding the validity of BRIS, we test this payment procedure in the lab using a large sample and a standard stake size. To do so, we compare a full pay experiment with two versions of a random pay experiment, as well as with a hypothetical experiment. Our contribution adds empirical evidence to an on-going debate regarding experimental methods in the field of economics (see Charness et al., 2016).

Our experiment includes one control (full pay) treatment and three test treatments. In the first random pay treatment we consider the same stake as in the control treatment, but use a lottery procedure to determine which players actually receive payment (i.e. giving players the chance to win 10€ with a one-in-ten probability). The second random pay treatment consists in a lottery with the same probability as in treatment one. However, the stake in treatment two is increased so as to maintain the same expected value as the stake in the control treatment (i.e. giving players the chance to win 100€ with a one-in-ten probability). The third treatment is hypothetical, offering the same stake as in the control treatment, but does not provide participants with any payment at the end of the experiment (i.e. they win 10 virtual euros).

2. Experimental design and data

2.1. Treatments

The experiment consists in a standard Dictator Game as proposed by Forsythe et al. (1994). A dictator game is a two-player game in which a certain amount of money is to be divided and one of the players (the dictator) decides how to divide the money.

Our design (Table 1) includes one control treatment and three test treatments, altering two variables: the endowment level and the probability of actually receiving the payment (which determines its expected value). With respect to the endowment, all participants are asked to make the following decision: ‘How would you like to divide this sum of money?’. In terms of actual payment at the end of the experiment, all participants in the control treatment are paid according to the dictator's decision, whereas in the Random pay experiments, a computerized lottery selects a subset of subjects who will be paid at the end of the experiment according to the decision they make. In the

Hypothetical pay experiment, participants receive a participation fee only.

Thus, in the control session (i.e., ‘Full pay experiment’), dictators are endowed with 10€ and all payments are realized.⁴ In treatment 1, dictators also play with a 10€ stake but only have a one-in-ten chance of actually being paid (‘Random pay experiment - Standard stake’). In treatment 2, dictators play with a 100€ stake and face the same one-in-ten chance of receiving payment (‘Random pay experiment - High stake’). The expected value of the stake (10€) is identical in the control treatment and in treatment 2. In treatment 3, dictators play a hypothetical game with a standard 10€ stake (‘Hypothetical pay experiment’) but don't receive any payment based on the decision they make in the experiment.

2.2. Experimental procedure

We conducted eight laboratory sessions, each with 20 participant pairs, at the University of Montpellier⁵ in March 2014 and September 2017.⁶ Participants were randomly assigned to two groups, split into rooms A and B. The two groups did not have any contact before, during or after the session. Participants were only allowed to talk to administrators. Subjects in room A were randomly matched with subjects in room B. Only players in Room A (i.e., dictators) made a decision, allocating either nothing or some portion of their endowment to player B. Before the experiment began, subjects were given ample time to read a written version of the instructions and to ask questions. Instructions for the dictator game were then read aloud to both groups.⁷

Subjects were informed about the payment procedure before beginning the experiment. Specifically, participants in the random experiments were told, ‘The computer will randomly select one player A for every 10 A players. Only selected A players will have their decision applied. The decisions of other participants will not be applied and their earnings will be zero’.

In the third treatment, participants were informed about the hypothetical nature of their earnings as follows: ‘Suppose a hypothetical sum of 10 euros has been allocated to each pair of players A and B. Player A can share this hypothetical amount with Player B. (...) Player A makes a proposal. The proposal consists of a hypothetical amount that the player B will receive. (...) The hypothetical amount that player A will receive is the total amount, 10 euros, less the hypothetical amount that player B receives.’

2.3. Subjects

A total of 320 participants were recruited from a subject pool drawn from a database of people who had completed a form indicating their interest in participating in economic experiments. None of the subjects had previously participated in a dictator game. All subjects were paid a 6€ show-up fee.

As shown in Table 2, samples are well balanced across control and test treatments. A Kruskal–Wallis H test shows no significant differences between groups with respect to the main socio-demographic variables.

3. Results

First, our results indicate that average transfers (in terms of percentage of the initial endowment) in the Full pay experiment and the Random pay experiments are very similar, but lower than those in the Hypothetical pay treatment. Subjects gave on average 2.7€ (SD = 0.21) in the Full pay experiment, 2.6€ (SD = 0.20) in the Random pay

Table 1
Between subject Random Incentive System - Experimental Design.

	Endowment	Probability	Expected Value
Control – ‘Full pay experiment’	10€	1/1	10€
Treatment 1 – ‘Random pay experiment – Standard stake’	10€	1/10	1€
Treatment 2 – ‘Random pay experiment – High stake’	100€	1/10	10€
Treatment 3 – ‘Hypothetical pay experiment’	10€	0	0€

⁴ Standard stakes in laboratory Dictator Games are usually 10\$ (Engel 2011).

⁵ Laboratoire d'Economie Experimentale de Montpellier

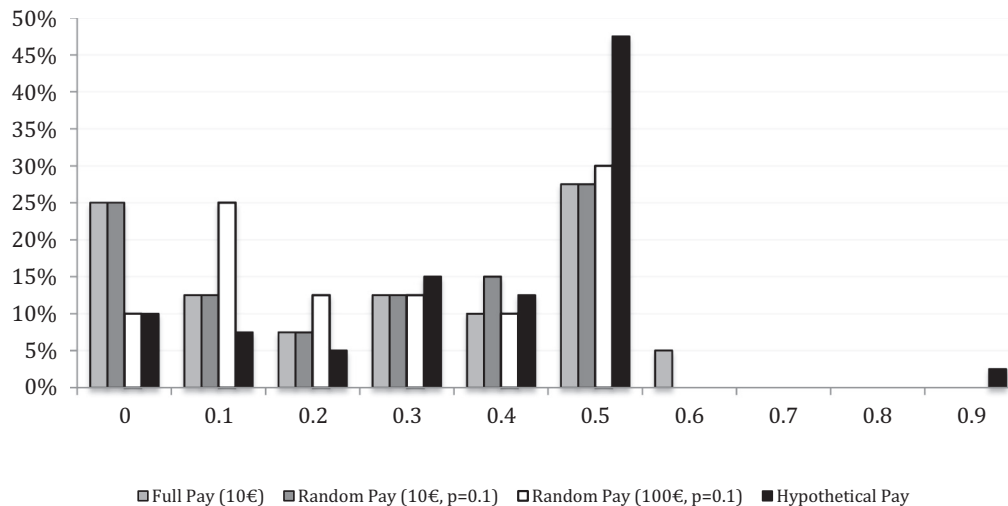
⁶ We thank an anonymous referee for suggesting an additional treatment, which explains the period that elapsed between the first treatments and the hypothetical one.

⁷ Full instructions are provided in the supplementary materials.

Table 2

Randomization table: Kruskal–Wallis H test of between group differences across subjects (dictators) from control and treatment groups.

(N = 120)	(1) Control	(2) Treatment 1	(3) Treatment 2	(4) Treatment 3	(5) H-statistics	(6) p-value
Socio demographic characteristics						
Gender ratio (% of male)	40%	50%	45%	45%	0.59	0.89
Age	26.12	23.17	27.72	23.97	2.2	0.53
Student ratio	85%	90%	90%	95%	0.59	0.89

**Fig. 1.** Histograms of dictator offers (fraction of endowment) among control (full pay experiment) and test treatments.

experiment – Standard stake (10€, $p = 0.1$), 26.6€ (SD = 0.19) in the Random pay experiment – High stake (100€, $p = 0.1$) and 3.7€ (SD = 0.19) in the Hypothetical pay experiment (10€, $p = 0$).

Second, we highlight two main findings that we discuss in more detail in the following sections.

Result 1: The three financially incentivised treatments (Full pay, Random pay - Standard stake and Random pay - High stake) yield similar transfer distributions (see Fig. 1 and Table 3).

Result 2: There are both significantly fewer selfish subjects (who transfer nothing) and significantly more egalitarian subjects (who transfer half of the endowment) in the hypothetical treatment.

3.1. Full pay experiment (10€) vs. Random pay experiment – Standard stake (10€, $p = 0.1$)

Dictator offers among the ‘Full pay experiment’ and the ‘Random pay experiment – Standard stake’ exhibit very similar distributions (see Fig. 1 and Table 3). A Mann–Whitney U test reports no significant difference between the distributions of dictator offers across the two treatments ($U = 0.275$; $p = 0.78$). A test of equivalence (TOST - two one sided test) corroborates this result, finding similarity at the 0.01 level.⁸

Purely selfish subjects, identified as those who make ‘zero offers’, appear in 25% of cases in the Full pay experiment, which is exactly the same proportion as in the Random pay version of the game. Egalitarian subjects, or those making fifty-fifty offers, comprise 28% of the sample in both versions, as well.

⁸ As pointed out by an anonymous referee, the absence of a significant difference in the statistical test (notably, testing the null hypothesis of equal distribution using the non-parametric Mann–Whitney U test) does not prove statistical equivalence. We therefore complemented the analysis with a test of equivalence (Lakens, 2017). More specifically, we used a two one sided test (TOST) of equivalence, which is designed to provide evidence regarding the absence of a difference large enough to be considered meaningful.

In Fig. 2, three quantile-quantile plots illustrate that the two distributions exhibit similar patterns. The majority of points in the plot for Treatment 1 fall around the 45° line, indicating that these two samples are characterized by very similar distributions.

3.2. Full pay experiment (10€) vs. Random pay experiment – High stake (100€, $p = 0.1$)

The ‘Random pay experiment - High stake’ generates a similar overall distribution as in the Full pay experiment, but with fewer purely selfish subjects. Although we find no statistical difference between the distribution of offers across these treatments ($U = -0.063$, $p = 0.94$) and a TOST is significant at the 0.01 level, the proportion of selfish subjects is significantly lower in the ‘Random pay experiment - High stake’ (10% vs. 25%; $U = 1.754$, $p = 0.07$). This results in a shift to low offers (ranging from 10% to 30% of the potential endowment), which occur more frequently than in the Full pay experiment (40% vs. 20%; $U = -1.94$, $p = 0.05$). The quantile-quantile plot of Treatment 2 in Fig. 2 shows small, but relatively frequent departures from the 45° line.

The dictator game results found by Cherry et al. (2002) corroborate the ‘zero offer’ pattern as being a decreasing function of the stake (‘zero offers’ decrease as the stakes increase). Possible explanations include a rise in culpability that is associated with purely selfish behaviour when the stake is high. Most importantly, this is an additional argument in favour of randomized incentive schemes using standard, rather than elevated, stakes as a valid preference elicitation mechanism.

3.3. Full pay experiment (10€) vs. Hypothetical pay experiment (10€, $p = 0$)

The ‘Hypothetical experiment’ generates a different distribution than the control treatment ($U = -1.949$, $p = 0.05$) that is driven by two main sources: 1) fewer selfish subjects (10% vs. 25%, $U = 1.754$,

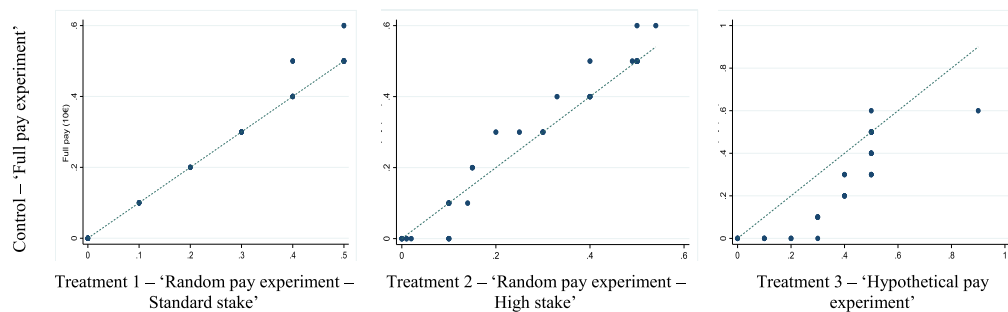


Fig. 2. Quantile-quantile plots of dictator offers between control and test treatments.

Table 3

Dictator offers (fraction of endowment) among control (Full pay experiment) and test treatments.

Dictator's offer (Fraction of endowment)	Control – 'Full pay experiment'	Treatment 1 – 'Random pay experiment – Standard stake'	Treatment 2 – 'Random pay experiment – High stake'	Treatment 3 – 'Hypothetical pay experiment'
0	0.25	0.25	0.10	0.10
0.1	0.13	0.13	0.25	0.08
0.2	0.08	0.08	0.13	0.05
0.3	0.13	0.13	0.13	0.15
0.4	0.10	0.15	0.10	0.13
0.5	0.28	0.28	0.30	0.48
0.6	0.05	0.00	0.00	0.00
0.7	0.00	0.00	0.00	0.00
0.8	0.00	0.00	0.00	0.00
0.9	0.00	0.00	0.00	0.03

$p = 0.07$) and 2) more egalitarian subjects (47.5% vs. 27.5%; $U = -1.836$, $p = 0.06$).

4. Discussion and conclusion

Our experiment demonstrates that subjects tend to be more influenced by the amount of the reward itself than by the probability of receiving it. For identical stake amounts, behaviour under random payment is strikingly similar to behaviour under a regular, non-probabilistic incentive system. For identical expected stake values, however, we observe significantly higher egoistic behaviour under a regular payment procedure than under a random payment procedure. Both findings challenge conventional predictions and may widen the scope of methods available to researchers seeking to test hypotheses under logistical and financial constraints. Further research should extend this work by testing the Between-subject Random Incentivized System with other games as well as with varying probabilities of payment. Another non-negligible issue concerns how different payment procedures could affect participation in future experiments. On the one hand, participants who are motivated by earning money at the end of an experiment may be discouraged by the use of BRIS. On the other hand, anecdotal evidence suggests that many individuals do not stop gambling when they do not earn (and indeed even when they lose) money repeatedly. A related issue that deserves more attention is the possible long term effects of the widespread use of BRIS, especially among risk-averse subjects. Indeed, a widespread use of BRIS can lead to biased samples and results because of an involuntary selection bias against risk-averse subjects. Rather than proposing a conclusive finding regarding the effect of probabilistic outcomes on prosociality in the dictator game, we prefer to invite further research on the issue.

Although further investigation into the external validity and robustness of these findings is required, they conceivably have interesting potential implications for policy makers. When designing innovative

policy measures in order to encourage adherence to better habits in a variety of domains (e.g., environment, health, transportation security, etc.), for example, the BRIS procedure could be used to significantly reduce the costs of implementing financial incentives. Random (dis) incentivized interventions could also be a cost-effective alternative to existing strategies for improving compliance with public policies insofar as these strategies are often either too complex or too expensive to implement on large scales. As such, not only could the BRIS represent a useful tool for researchers seeking to maximize the value of their funding for economic research, but it could also have promising potential applications for policymakers seeking to pursue public policy objectives in a context of budgetary constraints. This insight echoes the sizeable literature in psychology on intermittent reinforcement, originating from the works of Skinner in the 1950s. Possible applications are numerous and wide-ranging, such as the implementation of random rewards for proper waste management or random fines for disregarding mandatory health measures, and constitute interesting possible directions for future research.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at [doi:10.1016/j.socrec.2018.01.004](https://doi.org/10.1016/j.socrec.2018.01.004).

References

- Armantier, O., 2006. Do wealth differences affect fairness considerations? *Int. Econ. Rev.* 47 (2), 391–429. <http://dx.doi.org/10.1111/j.1468-2354.2006.00383.x>.
- Baltussen, G., Post, T.G., Van den Assem, M.J., Wakker, P.P., 2011. Random incentive systems in a dynamic choice experiment. *Exp. Econ.* 15 (3), 418–443. <http://dx.doi.org/10.1007/s10683-011-9306-4>.
- Beattie, J., Loomes, G., 1997. The impact of incentives upon risky choice experiments. *J. Risk Uncertainty* 14 (2), 155–161.
- Ben-Ner, A., Kramer, A., Levy, O., 2008. Economic and hypothetical dictator game

- experiments: incentive effects at the individual level. *J. Socio-Econ.* 37 (5), 1775–1784.
- Bettinger, E., Slonim, R., 2007. Patience among children. *J. Public Econ.* 91 (1–2), 343–363.
- Bolle, F., 1990. High reward experiments without high expenditure for the experimenter? *J. Econ. Psychol.* 11, 157–167.
- Brañas-Garza, P., Kovarik, J., Neyse, L., 2013. Second-to-fourth digit ratio has a non-monotonic impact on altruism. *PLoS ONE* 8 (4), e60419. <http://dx.doi.org/10.1371/journal.pone.0060419>. 2013.
- Brañas-Garza, P., 2007. Promoting helping behavior with framing in dictator games. *J. Econ. Psychol.* 28 (4), 477–486.
- Brock, J.M., Lange, A., Ozbay, E.Y., 2013. Dictating the risks. experimental evidence on norms of giving in risky environments. *Am. Econ. Rev.* 103, 415–4337.
- Camerer, C.F., 1989. An experimental test of several generalized utility theories. *J. Risk Uncertainty* 2 (1), 61–104.
- Carpenter, J., Connolly, C., Knowles-Myers, C., 2008. Altruistic behavior in a representative dictator experiment. *Exp. Econ.* 11, 282–298.
- Charness, G., Uri, G., Kuhn, M.A., 2013. Experimental methods: extra-laboratory experiments-extending the reach of experimental economics. *J. Econ. Behav. Organiz.* 91, 93–100.
- Charness, G., Gneezy, U., Halladay, B., 2016. Experimental methods: Pay one or pay all. *J. Econ. Behav. Organiz.* 131, 141–150.
- Cherry, T.L., Frykblom, P., Shogren, J.F., 2002. Hardnose the dictator. *Am. Econ. Rev.* 92 (4), 1218–1221.
- Clot, S., Grolleau, G., Ibanez, L., 2014. Do good deeds make bad people? A self-licensing experiment to revisit the norm enforcement debate. *Eur. J. Law Econ.* <http://dx.doi.org/10.1007/s10657-014-9441-4>. ISSN 0929-1261.
- Cox, J.C., Sadiraj, V., Schmidt, U., 2015. Paradoxes and mechanisms for choice under risk. *Exp. Econ.* 18, 215–250.
- Cubitt, R.P., Starmer, C., Sugden, R., 1998. On the validity of the random lottery incentive system. *Exp. Econ.* 1 (2), 115–131. <http://dx.doi.org/10.1007/BF01669298>.
- El Harbi, S., Bekir, I., Grolleau, G., Sutan, A., 2015. Efficiency, equality, positionality: What do people maximize? Experimental vs. hypothetical evidence from Tunisia. *J. Econ. Psychol.* 47, 77–84.
- Engel, C., 2011. Dictator games: a meta-study. *Exp. Econ.* 14, 583–610. <http://dx.doi.org/10.1007/s10683-011-9283-7>.
- Exadaktylos, F., Espín, A.M., Brañas-Garza, P., 2013. Experimental subjects are not different. *Nature Sci. Rep.* 3, 1213. <http://dx.doi.org/10.1038/srep01213>.
- Falk, A., Becker A., Dohmen, T., Huffman, D., Sunde, U., 2016. “The preference survey module: a validated instrument for measuring risk, time, and social preferences”. IZA Discussion Paper 9674. Bonn: Institute for the Study of Labor.
- Fong, C.M., Luttmer, E.F.P., 2009. What determines giving to hurricane katrina victims? Experimental evidence on racial group loyalty. *Am. Econ. J.* 1, 64–87.
- Forsythe, R., Horowitz, J.L., Savin, N.E., Sefton, M., 1994. Fairness in simple bargaining experiments. *Games Econ. Behav.* 6 (3), 347–369. <http://dx.doi.org/10.1006/game.1994.1021>.
- Kagel, J., Roth, A., 1995. *The Handbook of Experimental Economics*. Princeton University Press, Princeton.
- Kahneman, D., Tversky, A., 1979. Prospect theory: An analysis of decision under risk. *Econometrica* 47 (2), 263–292.
- Kellner, C., Reinstein, D., Riener, G., 2015. Stochastic income and conditional generosity, dice. , Discussion Paper, #197, Available at SSRN: <http://ssrn.com/abstract=2691027>.
- Lakens, D., 2017. Equivalence tests: a practical primer for *t*-Tests, correlations, and meta-analyses. *Social Psychol. Personality Sci.* 8 (4), 355–362.
- Langer, T., Weber, M., 2008. Does commitment or feedback influence myopic loss aversion? An experimental analysis. *J. Econ. Behav. Organiz.* 67 (3–4), 810–819.
- Madsen, D.O., Stenheim, T., 2015. Experimental methods in economics and psychology: A comparison. *Procedia - Social Behav. Sci.* 187, 113–117.
- Prelec, D., 1998. The probability weighting function. *Econometrica* 66, 497–527.
- Rubinstein, A., 2001. A theorist's view of experiments. *Eur. Econ. Rev.* 45, 615–628.
- Rubinstein, A., 2013. Response time and decision making. *Judgment Decis. Making* 8 (5), 540–551.
- Sefton, M., 1992. Incentives in simple bargaining games. *J. Econ. Psychol.* 13, 263–276.
- Stahl, D.O., Haruvy, E., 2006. Other-regarding preferences: Egalitarian warm glow, empathy, and group size. *J. Econ. Behav. Organiz.* 61 (1), 20–41. <http://dx.doi.org/10.1016/j.jebo.2004.10.008>.
- Volpp, K.G., Loewenstein, G., Troxel, A.B., Doshi, J., Price, M., Laskin, M., Kimme, S.E., 2008. A test of financial incentives to improve warfarin adherence. *BMC Health Serv. Res.* 8, 272–278.
- Zizzo, D.J., 2010. Experimenter demand effects in economic experiments. *Exp. Econ.* 13 (1), 75–98.