

An experimental analysis of the effect of recordkeeping over direct reciprocity*

Luis Paulo Guimarães dos Santos¹

 <https://orcid.org/0000-0001-9986-8237>

Email: lupaufba@gmail.com

Anderson José Freitas de Cerqueira^{2,3}

 <https://orcid.org/0000-0002-5063-9967>

Email: anderson.cerqueira@estacio.br

César Valentim de Oliveira Carvalho Júnior¹

 <http://orcid.org/0000-0003-0387-0872>

Email: cesarvalentim@ufba.br

¹ Universidade Federal da Bahia, Faculdade de Ciências Contábeis, Salvador, BA, Brazil

² Centro Universitário Estácio da Bahia, Departamento de Ciências Contábeis, Salvador, BA, Brazil

³ Universidade de Salvador, Departamento de Ciências Contábeis, Salvador, BA, Brazil

Received on 10.19.2019 – Desk acceptance on 11.29.2019 – 4th version approved on 05.25.2020 – Ahead of print on 11.09.2020

Editor-in-Chief: Fábio Frezatti

Associate Editor: Jacqueline Veneroso Alves da Cunha

ABSTRACT

This article experimentally investigates the impact of recordkeeping over the level of direct reciprocity. The study of reciprocity has been barely explored in accounting. This article helps to fill that gap by presenting the first investigation that provides experimental evidence of the causal relationship between recordkeeping and direct reciprocity. Reciprocity is a key aspect in the human cooperation process, based on its implications for the evolution of economic and social systems. However, understanding the mechanisms that promote it has been an important scientific challenge in various areas, such as biology, anthropology, sociology, psychology, and economics. The results of this study have an impact on accounting teaching as they provide a scientific basis that helps to improve the understanding of the role of accounting, through its most basic function, and its consequences for human cooperation. Moreover, they have implications for accounting research by showing the viability of using economic experiments to investigate emerging themes in accounting. Finally, from a practical viewpoint, the results of the research signal to the formulators of accounting control mechanisms the importance of considering the crowding out effect of those instruments over motivation. The research adopted a single-factor between-subjects experimental design with a pre-test and post-test and a control group. The causal inference was made using difference-in-differences regression models for panel data together with a variety of additional tests, aiming to give robustness to the results. The research presents evidence of the crowding out effect of recordkeeping over direct reciprocity in a trust game. This finding is important because it provides an explanation of how accounting, through its most basic function, influences human cooperation.

Keywords: recordkeeping, direct reciprocity, crowding out effect, trust game.

Correspondence address

Luis Paulo Guimarães dos Santos

Universidade Federal da Bahia, Faculdade de Ciências Contábeis

Avenida Reitor Miguel Calmon, s/n – CEP: 40110-903

Vale do Canela – Salvador – BA – Brazil

*Paper presented at the XVIII USP International Conference in Accounting, São Paulo, SP, Brazil, July of 2018.



1. INTRODUCTION

Human cooperation is a prominent question in behavioral sciences, given that many collective problems are characterized by a conflict of interests (Gächter & Herrmann, 2009). Within that context, reciprocity performs an important role, as it promotes trust and cooperation in social interactions, increasing the chances of a relationship of mutual trust developing (Malhotra, 2004). Moreover, reciprocity enables economic exchange relationships, helping to reduce transaction costs and the undesired consequences of incomplete contracts (Abraham et al., 2016; Barney & Hansen, 1995; Duffy et al., 2011; Gambetta, 1988).

Basu et al. (2009) were the first to study the role of formal recordkeeping in the evolution of cooperation, via its effect on reputation and reciprocity. The authors postulate that accounting, as an evolved economic institution, through its recordkeeping function, helps to promote the reciprocity needed for large scale exchanges and coordination in the interactions between strangers over time. Basu et al. (2009) presented experimental evidence suggesting that in the trust game with repeated interactions, recordkeeping alters the economic narrative by promoting reputation building and, as a result, reciprocity. According to the authors, their findings provide empirical support to the idea that institutions such as accounting can influence the propensity to act reciprocally and the cooperation between individuals. Basu et al. (2009) argue that this result occurs because recordkeeping, by complementing the resources of the brain, improves the memory of past interactions, meaning that the decisions regarding future interactions are conditioned by the partners' history of cooperation.

The findings of Basu et al. (2009) are compatible with those of other experimental studies that use the trust game with repeated interactions to address topics related to accounting [e.g., Lunawat (2013a, 2013b)]. However, due to the fact that they are configured for the participants to interact with different partners throughout the experiment and all the participants are informed of that condition at the start of the game, those studies create an incentive for strategic reputation building and emphasize indirect reciprocity, a situation in which the acts of one person toward another are rewarded/punished by third parties (Engelmann & Fischbacher, 2009; Nowak & Sigmund, 2005).

There is substantial literature suggesting that strategic reputation building induces increased competition [e.g.,

Abraham et al. (2016), Arnold and Schreiber (2013), Bohnet and Huck (2004), Brandts and Figueiras (2003), Camerer and Weigelt (1988), Engelmann and Fischbacher (2009), and Lunawat (2013a, 2013b)] and increases the efficiency of control mechanisms (Arnold & Schreiber, 2013; Rockenbach & Milinski, 2006). For that reason, it is probable that, in the experiment of Basu et al. (2009), being able to use recordkeeping to differentiate those who do not cooperate, or cooperate little, will mean that there is strategic concern about reputation, since, in the scenario in which one individual interacts with various players at the same time, the evaluation that is made about the reputation of one particular partner is affected by the behavior of the rest, so that the appreciation of reputation will always be relative.

The study of Basu et al. (2009) derived its results from an environment strongly anchored in strategic reputation building, in which indirect reciprocity performs a fundamental role. However, the evolution of cooperation can change in scenarios in which the interactions predominantly occur in the form of direct reciprocity and strategic reputation building is not possible. Direct reciprocity is an interaction involving only two people, in which one directly responds to the actions of the other, in a process that necessarily repeats over time (Engelmann & Fischbacher, 2009; Hizak et al. 2018). It is based on the idea that the behavior of one partner directly depends on the behavior of their counterpart (Baek et al., 2016; Hilbe et al., 2017; Rand et al. 2009; Van Veelen et al., 2012) and it fundamentally relies on the individuals' memory capacity (Milinski & Wedekind, 1998; Stevens & Hauser, 2004). In smaller groups, in which repeated interaction is frequent, cooperation may be primarily explained by direct reciprocity (Engelmann & Fischbacher, 2009).

On the other hand, recordkeeping, besides serving as a device for complementing memory, as argued by Basu et al. (2009), can work as a form of monitoring, acting as a control mechanism. Some studies have suggested that monitoring (like any other control mechanism) can reduce people's intrinsic motivation to cooperate, depending on the context in which it is applied (Arnold & Schreiber, 2013; Calabuig et al., 2016; Camera & Casari, 2017; Dickinson & Villeval, 2008; Enzle & Anderson, 1993; Falk & Kosfeld, 2006; Fehr & List, 2004; Fehr & Rockenbach, 2003; Houser et al., 2008; Rietz et al., 2017), according to the phenomenon known as the "crowding out" effect.

From that perspective, monitoring mechanisms can be interpreted as a hostile action (Falk & Kosfeld, 2006) or as a breach of the implicit contract based on mutual trust (Frey, 1993). When that occurs, the intrinsic motivation is eroded and the person monitored sees no impediment to behaving opportunistically and acting in their own self-interest (Frey, 1993). From the viewpoint of direct reciprocity, it is important to consider that aspect, given that, according to Frey (1993), the suppression of intrinsic motivation by external interventions, or the crowding out effect, will be stronger with direct and personal interactions.

Based on that scenario, the main aim of the present research was to investigate the effect of recordkeeping on direct reciprocity, considering the hypothesis that, in accordance with the crowding out effect, its use will negatively affect reciprocal exchange relationships. For that, a single-factor between-subjects experimental design was used, with a pre-test and post-test and a control group involving 64 undergraduate students, in a trust game environment with repeated interactions.

The main result documented presents evidence that recordkeeping reduces the propensity to form trust and, as a result, negatively affects the direct reciprocity when only that type of mechanism for promoting cooperation is possible. In the study, it was observed that, after the experimental manipulation, the level of reciprocity decreased, suggesting that the introduction of information records influenced the instrumental behavior of the trustees (player-B), making them reduce the return on the investments, probably in response to the perceived actions of the investors (player-A). The results show that, in the post-test, the mean of the values sent by the investors in the treatment group decreased and in the control group it increased.

The results are robust for a variety of statistical tests, but they differ from the one originally presented by Basu et al. (2009), since, while the latter identified that recordkeeping increased reciprocity, by promoting reputation building, this study recorded a reduction in the level of direct reciprocity. Taken together, however, these findings are complementary and suggest that recordkeeping, depending on the circumstances of

the interactions between the agents, can have diverse consequences in the cooperation in accordance with the crowding out effect.

In general, the findings of this article are consistent with the hypothesis that formal recordkeeping negatively affects direct reciprocity and it presents the contributions highlighted below. As far as we are aware, it is the first study to directly offer an experimental test of the impact of recordkeeping over direct reciprocity, helping to broaden the understanding of the role of accounting, based on its most basic function in the development of economic exchanges. Second, by presenting the most robust causal inference of the relationship between recordkeeping and direct reciprocity, it increases the empirical evidence about the role of accounting, as an evolved institution, in influencing the economic and social behavior of the agents in exchange relationships. Third, studying the relationship between direct reciprocity and recordkeeping may have important implications for the research into management accounting, as a relevant body of studies has investigated how control mechanisms, particularly incentive contracts [e.g., Christ (2013), Christ et al. (2008), Hales and Williamson (2010), Kelly and Tan (2010), and Marinich (2019)], can influence cooperation in firms. Considering that, in certain circumstances, recordkeeping itself can serve as a monitoring mechanism, studying its impact on direct reciprocity helps to broaden the scope of those studies.

Finally, from a practical viewpoint, the results of the research indicate that the accounting control mechanisms may impede the evolution of cooperation if they are designed without considering the crowding out effect over people's motivation.

Besides the Introduction, this article is constituted of four other parts, namely: section 2 summarizes the literature about the topic on which the research problem is based and develops the hypothesis; section 3 presents the methodological description, as well as the experimental design and statistical procedures used to explore the data from the experiment; and section 4 presents the results, together with the discussion of them. Finally, section 5 presents the conclusions of the study.

2. RELATED LITERATURE AND HYPOTHESIS

2.1 Definition of Direct Reciprocity

Reciprocity represents the responses that individuals give to the friendly or hostile actions of other people and it can be direct or indirect [for a more comprehensive

discussion about the concept of reciprocity, consult Fehr and Schmidt (2006) and Sobel (2005)]. According to Nowak and Sigmund (2005), in the case of direct reciprocity, one person responds directly to the actions of another, rewarding or punishing them, in a process

of interactions that necessarily repeats various times. In the case of indirect reciprocity, the acts of one person toward another are responded to by third parties and are intrinsically linked to reputation building (Engelmann & Fischbacher, 2009; Nowak & Sigmund, 2005; Rankin & Eggimann, 2009; Roberts, 2008).

According to Baek et al. (2016), direct reciprocity is a mechanism for the evolution of cooperation based on repeated interactions. When individuals repeatedly come together, they may use conditional strategies to impose cooperative results that would not be viable in situations with a single interaction.

Nowak and Sigmund (2005) explain that in cases of direct reciprocity cooperation will be sustained through exchanges of altruistic acts. An act is considered altruistic if it has a cost, but confers a benefit to another individual. In that type of interaction, the reciprocal relationship is direct between the parties and does not involve third party participation, as in indirect reciprocity. For that reason, as Stanca et al. (2009) underline, how people evaluate underlying intentions in the actions of their counterpart is a fundamental question for understanding how direct reciprocity works.

2.2 Recordkeeping, Reciprocity, and the Crowding Out Effect

Basu and Waymire (2006) argue that when people start to get involved in numerous and complex reciprocal exchanges, formal recordkeeping emerges to supplement the imperfect human memory, which cannot accurately keep track of the past behavior of many partners at the same time. For Basu et al. (2009), the simple ability to externally record exchanges can complement the memory resources of the brain, enabling the storage of information and a more effective memory of the past. At the same time, Mullins et al. (2013) emphasize that recordkeeping enables the data about transactions to be easily stored and recovered, serving as a guide for future reciprocal behaviors.

As Basu and Waymire (2006) clarify, accounting records institutionalize the memory of past transactions and, combined with norms of honesty incorporated into laws and other institutions, they serve to sustain the trust that enables reputation building and, consequently, cooperation between economic agents. For that reason, Basu et al. (2009) hypothesize that the possibility of recording exchanges externally increases the memory resources of the brain, widening the capacity to store and recover information from the past, thus promoting reciprocity.

Basu et al. (2009) examined the role of recordkeeping in the reputation building and reciprocal exchanges of an economy using the trust game with repeated interactions and they found support for the hypothesis that voluntary recordkeeping enables reputation building. According to Basu et al. (2009), the findings of the research are consistent with the archeological documents that suggest that records of pre-historic transactions and the invention of writing for recordkeeping were linked to an increase in the complexity of human interaction.

On the other hand, according to the motivation crowding theory approach (Frey & Jegen, 2001), recordkeeping, understood as a monitoring system, may have an adverse psychological result and reduce intrinsic motivation, eliminating the positive effect of the control mechanism. That process of eroding intrinsic motivation, depending on an external intervention, is known as the crowding out effect (Frey 1993; Frey & Jegen, 2001).

Osterloh and Frey (2002) argue that contractual relationships include an aspect of (traditional) extrinsic motivation and a relational aspect, focused on a reciprocal appreciation of the intrinsic motivation. If the relational part of the contract is breached, the reciprocal good faith is put into question, so that if a particular act is interpreted as having a merely instrumental purpose, the intrinsic motivation will be weakened because it will be perceived as a form of control. However, Frey (1993) formulates that, in impersonal and purely abstract interactions, the intrinsic motivation will not be negatively affected by monitoring. The crowding out effect will be more probable in personal interactions. Dickinson and Villeval (2008) directly tested that hypothesis and found its empirical support.

2.3 Development of the Hypothesis

Arnold and Schreiber (2013) elaborated an experiment to analyze the efficiency of auditing in controlling the costs reported by subordinate managers in the budgeting process. The authors documented that, in a context in which the participants faced new partners in each round, aspects of reputation strongly affected the behavior of the parties, meaning the auditing favored a reduction in budgetary slack and an increase in the managers' payoff. In contrast, in a scenario in which direct reciprocity was the standard, the control mechanism (auditing) had a different effect, as the benefits found in the previous scenario disappeared. According to the authors, the capacity of the subordinates to retaliate to their superiors' previous punishments has negative effects and reduces the efficiency of the auditing. These results suggest that

in cases of direct reciprocity the social norms perform an important role and their non-observance can undermine the efficiency of the control mechanisms.

Moreover, the findings of Arnold and Schreiber (2013) are compatible with the arguments that control mechanisms (such as monetary incentives, supervision/monitoring, the establishment of deadlines, competition, performance assessment, among others), under specific circumstances, can negatively influence people's intrinsic motivation, altering their social behavior in accordance with the crowding out effect [for a more detailed study regarding that subject, see Deci and Ryan (1985), Deci et al. (1999), Frey (1993), and Frey and Jegen (2001)].

Along those lines, Calabuig et al. (2016), Fehr and List (2004), Fehr and Rockenbach (2003), and Houser et al. (2008) discovered that control mechanisms based on punishments can impede reciprocity. Dickinson and Villeval (2008), Enzle and Anderson (1993), and Falk and Kosfeld (2006) found evidence of the crowding out

effect of monitoring over direct reciprocity. Conversely, Camera and Casari (2017) discovered that monitoring institutions lead to a reduction in cooperation.

Taken together, those studies indicate that monitoring, or any form of external intervention seen with the purpose of control, depending on the circumstances, can negatively affect motivation. Because in cases of direct reciprocity the interactions constitute relationships based on reciprocal altruistic exchanges, without involving the subjective evaluation of reputation, the presence of recordkeeping may be perceived as an external control intervention, causing the crowding out effect on intrinsic motivation. In the context of the trust game, the investor's intrinsic motivation is trust and that of the trustee is the willingness to reciprocate the trust received (Calabuig et al., 2016). With that, the following prediction can be made:

H₁: in environments with repeated interactions, the use of recordkeeping will reduce direct reciprocity.

3. METHODOLOGY

3.1 Design and Experimental Protocol

To test the previously mentioned hypothesis, this study uses an adaptation of the trust game with repeated interactions in a single-factor, pre-test/post-test experimental design, with a control group in which there is no other mechanism for promoting cooperation besides direct reciprocity. Moreover, unlike in the experiment of Basu et al. (2009), the players do not receive feedback about their partner's decisions, so that the only way to accompany the evolution of the interactions is with human memory. However, that scenario changes in the recordkeeping condition, when the participants can keep notes about the interactions and know that their counterpart can also do so.

Another important aspect of the experimental design adopted is the fact that the monitoring is set exogenously and not as a decision of the investor. That characteristic is important because the investor may not indicate their intentions through any other means other than sending resources. That configuration avoids the investor strategically using "non-monitoring" to indicate trust and induce reciprocity in the trustee.

As underlined by Campbell and Stanley (1963), the single-factor, pre-test/post-test and control group design adequately addresses the main threats that can traditionally compromise the internal validity of the experiment.

As a result of the pre-test, the design used increases the experimental control because it eliminates or sensitively reduces the confounding variables, providing a high degree of internal validity, as well as making the causality analysis robust (Cozby & Bates, 2011; Salkind, 2010). Moreover, the pre-test/post-test design enables a more sensitive analysis of the effect of the treatment, because each participant serves as their own control, making the effect of the experimental manipulation more prominent. According to Libby et al. (2002), this type of design is more effective when the salience of the treatment is desirable from the viewpoint of the objectives of the experiment. In the case of the experimental procedure carried out, it was crucial for the participants to carefully understand the task proposed, especially the use of recordkeeping, to guarantee their adequate performance. In that sense, the pre-test served as a learning process for the participants.

However, biases derived from the "demand characteristics" and "experimenter's expectations" represent threats that need to be considered, even though the solutions available to mitigate them are partial, as it is impossible to impede the participants from generating their own hypotheses related to the research (Shadish et al., 2002).

Thus, the following actions were taken to try to mitigate the effects of such threats: the experimental section was externally monitored by the coordinators of the research,

but they did not have contact with the participants before the end of the section; someone was designated as responsible who did not have knowledge about what exactly was being researched to monitor the experimental section; that person was instructed to maintain the least contact possible with the participants (basically, their function was to read the instructions and note down any doubts to discuss with the coordinators of the research who were outside the room).

On the other hand, due to the characteristics of the participants and of the sample selection process, the external validity of the research is compromised, so that its results cannot be generalized beyond the group that participated in the experiment or to other circumstances and environments.

To execute the experiment, software was developed based on a variant of the trust game presented by Berg et al. (1995) with the inclusion of repeated interactions (rounds). In total, the experiment processed 20 rounds: 10 in the pre-test and 10 in the post-test. The player pairs are the same until the end of the experiment. That characteristic of the game eliminates any possibility of reputation building via other mechanisms, other than the history of interactions between the same partners.

In the trust game used, the participants assume the role of investor (player-A) or trustee (player-B). Each player knows their own function, but does not know the function of any other participant. The game has two phases. In the first, the investor receives 10 currency units (lira), with the option of deciding the amount of the allowance they will keep and how much will be sent to the trustee. The investor keeps the entire resource that was not sent to the trustee. The allowance sent to the trustee is multiplied by 3. In the second stage of the game, the trustee decides how much of the tripled amount received will be sent (from 0 to the tripled quantity received) to the investor. At the end of each round, the investor's return is the value retained, that is, not sent to the trustee, plus the quantity received from the trustee. The trustee's profit, in turn, is the value received minus the quantity sent to the investor.

The participants were recruited among the students of the undergraduate course in accounting sciences of a public university of the state of Bahia. The promotion occurred through social media, banners, and email lists of the institution itself. During the recruitment phase, the participants were informed that the experimental section would consist of two activities (one presentation and one task) and would last between 90 and 150 minutes. At the time, it was also clarified that the details about the task would be given on the day of the experimental session

and that they would receive credits in complementary activities for their participation.

It is important to highlight that, at the faculty in which the research was carried out, there is no human research ethics committee, but studies conducted within the scope of its post-graduate program (as is the case here) undergo deliberation by the board. However, precisely because of that, during the act of recruiting, the students were duly informed about all the procedures that would be adopted in the experimental section, the objectives of the research, and the risks involved, among other relevant information. Moreover, as a rule, participation in the research would only be possible with the expressed consent of the participants.

Initially, the experimental session was formed of 64 students, each one randomly allocated as an investor or trustee, constituting 32 pairs. No student knew who their partner was. Next, the pairs were randomly separated into two groups (one for each experimental condition): treatment and control. In the treatment condition (recordkeeping), as of the 11th round, the participants had access to a text box on the screen of their computer in which they were supposed to make the alphanumeric records referring to the transactions with their partner. It was informed that making the records was a necessary condition for performing the task and that the information would be kept and could be consulted at any time.

Thus, it was possible to investigate the qualitative behavior (recordkeeping strategy) and quantitative behavior (number of characters) of the participants in the experiment. It was noted that the participants generally used recordkeeping to register information about the values sent and, primarily, those received by their partner in the game.

Before starting the experiment, the participants signed an attendance list. After distribution to the groups, the participants were duly seated at the computer terminals and received the initial instructions of the game. Next, the coordinator of the session made the presentation using a video projector to explain the experimental system.

The software was operated on the investors' computers with the purpose of storing the data generated throughout the experiment. Each pair's computers were linked up using an IP (internet protocol) after creating the game. The machines were numbered to enable the control and composition of the pairs. When starting the experimental section, the participants did not know the number of rounds the game would have. After the first round, the participants were informed that the game would have nine rounds. That aspect is important because it has an

implication in the measurement of the proxy used to operationalize the direct reciprocity construct.

At the start of the experiment, the participants accessed the initial screen of the game, related to registration of personal information, which enabled identification at the moment of searching the data in the game server, as well as reconciliation with the enrollment information. At the end of registration, the screen directed them to the start of the game. At that point, the participants could no longer communicate with each other or consult any material, being conditioned to only communicate with the moderator of the game.

In the first 10 rounds, both groups operated in the same way. From the 11th to the 20th round, the treatment group used recordkeeping. When they concluded the 20th round, the participants received a communication from the system itself about the end of the game, revealing the final balance obtained.

3.2 Participants and Incentive Structure

When working with students, the incentive often used in experiments in accounting, above all with the trust game and investment game, is monetary [e.g., Basu et al. (2009), Kanagaretnam et al. (2014), Lunawat (2016), and Maas et al. (2012)]. However, in the experiment developed, credits in complementary activities were offered to the participants. According to Slonim et al. (2013), incentives in the form of academic credits are just as powerful as monetary resources, for the purpose of attracting students to participate in experiments.

$$Recip = \beta_{0it} + \beta_{1it}(Dp) + \beta_{2it}(Rk) + \beta_{3it}(Dp \times Rk) + \varepsilon_{it}$$

in which *Recip* is the dependent variable of the model, a proxy for direct reciprocity, the value sent by the trustees in each round; (i) with relation to the parameters of the model, β_0 represents the expected value of the *Recip* variable related to the control group before the experimental manipulation, β_1 indicates how *Recip* behaves after the experimental manipulation and measures the difference (before and after) within the control group, β_2 measures the marginal effect of belonging to the treatment group before the experimental manipulation, and β_3 is the difference in differences (the difference within the control

At the university where the experiment was developed, students who participate in courses, activities, and institutional programs involving research, teaching, and extension can have their participation converted into curricular hours, at the discretion of the course panel. The time of participation in the experiment was divided as follows: hours of participation in presentations and hours of participation in the production of the extension activity.

3.3 Variables, Measurement, and Empirical-Statistical Model

The two main variables of the empirical model are direct reciprocity (abbreviated as *Recip*) and recordkeeping (abbreviated as *Rk*). As a construct is concerned, reciprocity will be operationalized using the values sent by the trustees throughout the 20 rounds. This form of operationalization of the construct is compatible with previous studies [e.g., Berg et al. (1995), Charness and Shmidoz (2014), Cox (2004), Engle-Warnick and Slonim (2004), and Glaeser (2000)].

In turn, the treatment was operationalized using a dummy variable where 0 and 1 indicate the absence and the presence of the treatment, respectively. With that, the empirical model of the research will be formed of two main variables: reciprocity (*Recip*) and recordkeeping (*Rk*).

The main analysis will be carried out based on the following difference in differences (*Diff-in-Diff*) regression model described in equation 1:

1

group minus the difference within the treatment group) and measures the effect of the treatment on the variable being studied; (ii) ε_i is the error term of the stochastic model; (iii) *Dp* is the dummy variable, which will be 1 when the data refer to the post-treatment and 0 when the data refer to the pre-treatment period; (iv) *Rk* is a dummy variable that will equal 1 when the participant is in the treatment group and 0 in the other situation.

The interest lies in the β_3 coefficient. According to the formulated hypothesis, it is expected to be significantly different from 0.

4. RESULTS

4.1 Evaluation of the Trust Game

To better understand the effect of recordkeeping on reciprocity, it is important to analyze how the trust and reciprocity relationships developed in the experiment conducted. The descriptive data reported below serve that purpose.

All in all, 32 pairs participated in the experiment, formed of 64 undergraduate students of the Accounting Sciences Faculty of the Federal University of Bahia. Of that total, 16 pairs were randomly allocated to each

experimental group. However, in the treatment group, four pairs were excluded from the analyses because they did not comply with the rules of the experiment task, as instructed, meaning the final sample was formed of 28 pairs (16 in the control group and 12 in the treatment group). With that configuration, 28 students performed the role of trustee and 28 performed the role of investor. Table 1 shows the summary of the demographic profile of the participants who performed the role of trustee and investor throughout the experiment.

Table 1

Demographic profile of the participants

Profile	Trustees			Investors		
	Control group	Treatment group	Total	Control group	Treatment group	Total
Women	8	6	14	8	5	13
Men	8	6	14	8	7	15
Mean age	26.6	28.1	28	26.5	27.4	28

Source: *Elaborated by the authors.*

A total of 560 decisions were taken for sending resources from the investors to the trustees and 560 were taken for sending them from the trustees to the investors. In only 11 investor decisions and 20 trustee decisions was the value sent 0 (1.96 and 3.57%, respectively). These percentages indicate that there was a high frequency of expression of trust and reciprocity.

On average, at the end of 20 rounds, each investor accumulated 265.18 liras and each trustee accumulated 171.53 liras. This difference between accumulated values is due to the investors starting with a 10 liras endowment in each round. The lowest and the highest values observed among the trustees were 57 and 311 liras. Among the investors, those values were 181 and 477. In equilibrium, the maximum value that each trustee or investor could earn would be 300 liras and for the duo, adding together the gains, it would be 600 liras. The maximum one pair managed to earn was 570 (311 for the trustee and 359 for the investor). Of the 28 pairs, the trustee accumulated a higher gain than the investor in only five.

The mean value sent by the investors, per round, was approximately 5.88 liras. Conversely, the mean value returned by the trustees was 9.21 liras. On average, the investors sent 58.8% of the value available in each round and the trustees returned 50% of the values received. Altogether, these findings are compatible with those

recorded in many experiments involving the trust game [see Johnson and Mislin (2010)].

With relation to the use of recordkeeping, it was possible to note that up to the 16th round, the participants added a large volume of information in each round. However, in the following rounds, there was a reduction in the inclusion of new characters. Therefore, it is possible to conjecture that the participants, as of the 17th round, were satisfied with the information accumulated, which enabled an understanding of the profile of their game partner. Among the trustees (player B), 12,700 characters were recorded, including the spaces between numbers and words [mean of 60.47 per round, standard deviation (SD) of 71, and median of 25.5]. The highest (lowest) extension record contained 206 (0) characters.

The main analysis of the research focused on the trustees' decisions and, therefore, took into account the amount of resources returned to the investors (return from the trustees) because it serves as a proxy for the direct reciprocity measure. However, the value sent by the investors is also important for understanding the dynamics of the experiment because it serves as a proxy for trust. Table 2 records the main descriptive statistics by experimental group for each one of the variables. Table 3 reveals the mean of the values sent per round before and after the experimental manipulation.

Table 2*Descriptive statistics of the values sent and returned throughout the sections*

Statistics	Values returned by the trustees (reciprocity)				Values sent by the investors (trust)			
	Control group before	Control group after	Treatment group before	Treatment group after	Control group before	Control group after	Treatment group before	Treatment group after
Mean	8.29	9.92	9.63	8.94	5.55	6.21	5.99	5.94
Standard error	0.48	1.23	0.51	0.57	0.22	0.24	0.29	0.30
Median	8.72	10.31	9.46	8.92	5.00	7.00	6.00	6.00
Standard deviation	1.52	3.89	1.61	1.79	2.82	3.00	3.16	3.28
Variance	2.30	15.16	2.59	3.20	7.97	9.01	9.96	10.78
Kurtosis	-1.36	-0.54	0.08	-0.77	-1.14	-1.19	-1.26	-1.22
Asymmetry	-0.39	0.27	-0.01	0.34	0.01	-0.34	-0.18	-0.26
Minimum	5.75	4.63	6.75	6.50	0.00	0.00	0.00	0.00
Maximum	10.19	16.88	12.33	12.08	10.00	10.00	10.00	10.00

Source: *Elaborated by the authors.***Table 3***Mean of the values returned and sent by the trustees and investors, respectively, by round and groups*

Round	Values returned by the trustees (reciprocity)				Values sent by the investors (trust)			
	Control group before	Control group after	Treatment group before	Treatment group after	Control group before	Control group after	Treatment group before	Treatment group after
1	5.75	4.63	6.75	10.17	4.13	6.25	4.33	6.08
2	6.81	10.69	10.08	7.75	4.56	6.88	5.67	5.67
3	7.06	7.81	8.83	6.50	4.13	6.38	5.33	5.17
4	9.19	11.63	8.67	10.83	5.31	6.13	5.17	4.83
5	7.06	6.63	12.33	9.75	6.00	7.19	7.00	4.58
6	8.25	9.94	8.50	8.75	5.63	6.06	5.58	6.17
7	10.19	5.25	9.25	7.17	6.38	6.00	6.50	4.33
8	9.69	11.88	9.67	7.33	6.50	5.63	5.92	6.08
9	9.56	13.88	11.25	9.08	6.25	5.44	7.25	9.17
10	9.31	16.88	11.00	12.08	6.63	6.19	7.17	7.33
Mean	8.29	9.92	9.63	8.94	5.55	6.21	5.99	5.94

Source: *Elaborated by the authors.*

Tables 2 and 3 show that, between the pre-test and the post-test, in the control group, the amount returned by the trustees increased approximately 19.93% (statistically significant, two-tailed Wilcoxon test p-value of 0.04) and in the treatment group there was a reduction of approximately 7.16% (statistically non-significant). Observing the SD in Table 2, it is noted that the dispersion before the experimental manipulation is exactly equal between the groups (SD = 2.1). However, after the manipulation, the SD increased approximately 16 and 15% in the control and treatment groups, respectively.

The data description also reveals that in the control group the amount of resources sent by the investors rose

approximately 11.94% between the pre-test and the post-test. That difference is statistically significant (p-value of 0.006 for the two-tailed Wilcoxon non-parametric test). Conversely, in the treatment group, there was a slight alteration (not significant, from a statistical viewpoint), as a decline of approximately 0.83% was observed. The alteration in the variance was also accentuated. However, the difference in differences is statistically significant (Mann-Whitney U test with a p-value of 0.03), suggesting a treatment effect on trust. The difference in differences is calculated as follows: [(mean of the control group in the post-test) – (mean of the control group in the pre-test)] – [(mean of the treatment group in the post-test) – (mean of the treatment group in the pre-test)].

As the trustees' returns may be explained by aversion to iniquity or pure altruism (Engelmann & Strobel, 2010; Fehr & Schmidt, 2006; Rabin, 1993; Smith, 2013), it is necessary to make an inference about the presence of reciprocity motivated by other factors conditioned by the investor's behavior in the experiment. According to Coricelli et al. (2006), if the trustee's return is simply motivated by pure altruism, then it should not depend on the investor's level of trust. However, Berg et al.

(1995) argue that reciprocity can be explained by the trust that one agent receives from an anonymous counterpart, that is, the trustee's reciprocity depends on the investor's trust. So, in the reciprocity hypothesis, the proportion returned by the trustee will be positively related with the quantity sent by the investor (Cochard et al., 2004). To evaluate these conjectures, Table 4 shows the contemporaneous correlation between trust and reciprocity.

Table 4

Correlation between trust and reciprocity in the experimental conditions

	Control group before	Control group after	Treatment group before	Treatment group after
Correlation between reciprocity and trust	0.70***	0.70***	0.78***	0.68***

Note: Reciprocity is represented by the value sent from the trustee to the investor and trust is represented by the value sent from the investor to the trustee. The results reported refer to the Pearson correlations; however, Spearman's correlations were run, which were also significant at 1%, and whose results were omitted.

*** = significant at 1%.

Source: Elaborated by the authors.

The data in Table 4 reveal the high contemporaneous correlation between trust and reciprocity in all the periods, but with a reduction of approximately 12.8% between the pre-test and the post-test in the treatment group and no alteration in the control group.

Taken together, these findings confirm the presence of reciprocity in the experiment and present the following situation: in the control group, the trustees returned an average of 49.72 and 53.13% of the resources received before and after the experimental manipulation, respectively. In the treatment group, in turn, those percentages were 53.59 and 50.16%. Those percentages are calculated as follows: (mean value returned by the trustee) ÷ (mean value sent by the investor × 3).

These results indicate that, in the control group, there was a practically equitable distribution of the gains for the increase in investor return. Conversely, in the treatment group, the scenario did not change, since in the post-test the division of gains was practically the same.

In short, the pattern of reciprocity and trust documented in this article is compatible with the findings of other studies involving the trust game [e.g., Basu et al. (2009), Berg et al. (1995), Bourgeois-Gironde and Corcos (2011), Cameron (2003), Cochard et al. (2004), Coricelli et al. (2006), Engle-Warnick and Slonim (2004, 2006a, 2006b), Johnsen and Kvaløy (2016), and Johnson and Mislin (2010)] and it indicates the adequacy of the experiment.

4.2 Effect of Recordkeeping over Direct Reciprocity

Due to the experimental design used, one preceding question to be observed in the data from the experiment is whether the control and experimental groups are equivalent in the pre-test. That equivalence should be evaluated in relation to the dependent variable. In this case, due to the random attribution to the groups, these are expected to be statistically equal in relation to the mean of the values observed for the *Recip* variable. The two-tailed Mann-Whitney *U* test indicates a p-value of 0.16, suggesting that in the pre-test the difference between the experimental groups is not significant from a statistical viewpoint.

In addition, the equivalence of the groups was also evaluated in relation to the trustees' returns (proxy for reciprocity) and to the values sent by the investors (proxy for trust). In both cases, in the pre-test, the control and treatment groups do not present statistically significant differences (p-values in the two-tailed Mann-Whitney *U* test of 0.17 and 0.38, respectively). Hence, equivalence between the groups is assumed.

To test the research hypothesis, equation 1, as already detailed in section 3, was used. Considering the longitudinal data structure, two panel data regression models were run (pooled and random effect). In addition, due to the relaxation of the assumption of normality of the residuals of the regression, an additional test was carried out using the bootstrapping resampling

procedure, simulating normality of the residuals based on the pooled regression. Table 5 presents the results of the regressions that test the effect of the recordkeeping over direct reciprocity.

Table 5
Results of the panel data regression for the Recip independent variable

Variables	Statistics	Pooled (OLS)	Random effect (GLS)	Bootstrapping simulating normal errors
Constant	Coefficient	8.27***	8.27***	8.27***
	Standard error	0.39	1.12	-
	Z-score	21.38	7.39	-
	p-value	0.00	0.00	0.00
Rk	Coefficient	1.36**	1.36	1.36**
	Standard error	0.62	1.80	-
	Z-score	2.19	0.76	-
	p-value	0.03	0.45	0.03
Dp	Coefficient	1.61***	1.61***	1.61***
	Standard error	0.55	0.55	-
	Z-score	2.95	2.95	-
	p-value	0.00	0.00	0.00
Dp x Rk	Coefficient	-2.30***	-2.30***	-2.30***
	Standard error	0.88	0.88	-
	Z-score	-2.61	-2.61	-
	p-value	0.00	0.00	0.00
N		560	560	5,000
Adjusted R ²		0.009	-	-
Durbin-Watson		0.87	1.545	-
Variance "between"		-	20.22	-
Variance "within"		-	27.40	-
F statistic(p-value)		3.39/(0.03)	-	-

Note: Due to there being repeated observations for the same individual in the sample, the regressions used Beck-Katz robust standard errors (panel-corrected standard errors – PCSE) to correct any possible intra-individual correlations, heteroscedasticity between the groups, and contemporaneous correlations. The intercept (constant) reflects the measurement of the variable being studied for the control group before the treatment. Rk is a dummy variable that takes the value 1 when referring to the treatment group and 0 in other cases, Dp is a dummy variable that takes the value 1 when the data refer to the post-treatment period and 0 for the pre-treatment period, Dp x Rk indicates the effect of the treatment, and the dependent variable Recip represents the mean value sent by the trustees in the environment of repeated interactions. The pooled regressions and those with random effects relaxed the assumption of normality of the residuals. So, the decision was made to estimate the p-value of the coefficients simulating normality of the residuals via the bootstrapping technique using 5,000 replications.

*, **, *** = significant at 10, 5, and 1%, respectively.

GLS = generalized least squares; OLS = ordinary least squares.

Source: Elaborated by the authors.

In the regression of Table 5, the coefficient of the constant (8.27) represents the mean of the values sent by the trustees of the control group before the experimental manipulation, serving as the basis for comparison. The coefficient of the Rk variable (1.36) is the mean marginal value of the amounts sent by the trustees of the treatment group before the experimental manipulation. The coefficient of the Dp variable (1.61) is the mean marginal value of the amounts sent by the trustees of the control group after the experimental manipulation. Finally, the coefficient of the Dp x Rk variable (-2.30) is the difference

in differences between the control and treatment groups and measures the treatment effect.

As is observed, the coefficient of Dp x Rk is negative and significant in the two main models and in the bootstrapping simulation of normality of the residuals. Also due to the relaxation of the normality assumption, a non-parametric test of the difference in differences was run, which indicated a statistically significant difference between the groups (p-value of 0.002 for the Mann-Whitney U test).

The experiment indicates that the recordkeeping caused a crowding out effect, reducing the level of direct

reciprocity, probably in response to the investor's trust. However, the sensitivity of the variation was greater in the treatment group, given that in the control group the 11.94% increase in the level of trust (recorded in Table 2) caused a 71.94% increase in the level of reciprocity (ratio of 6.03). Conversely, in the control group, the 0.83% reduction in trust led to a 21.26% reduction in reciprocity (ratio of 25.48). Clearly, the leverage was greater in the treatment group.

The fact that the investors did not alter the mean value sent to the trustees may explain the decline in reciprocity in the post-test of the treatment group. From the trustees' viewpoint, the variation in trust represents a variation in their gains. If, over the course of the experiment, the level of trust remains the same, by reducing reciprocity, the trustees can increase their wealth. In fact, in the experiment, in the treatment group, the mean gain per round for the trustees was 83.42 and 88.83 in the pre-test and post-test, respectively. Among the investors, that gain was 136.42 and 130.00 before and after the treatment, respectively. This suggests that the trustees acted strategically to increase their gains and punish the investors for breaching their trust.

Taken together, these results indicate the treatment effect, as the direct reciprocity decreased in the post-test, suggesting the crowding out effect of recordkeeping. This finding is compatible with other studies that show the crowding out effect of the monitoring mechanisms (Dickinson & Villeval, 2008; Enzle & Anderson, 1993; Frey, 1993; Rietz et al., 2017), but differs from the finding of Basu et al. (2009), who documented an increase in reputation.

In this aspect, it is important to highlight that, despite being complementary, the findings of this research cannot be directly compared with those of Basu et al. (2009), given that the latter emphasized reputation building. As previously discussed, it is probable that the mechanism for strategic reputation building improves the efficiency of recordkeeping as a control mechanism. In the experimental configuration of Basu et al. (2009), the trustees could offer higher returns to those investors who made greater investments to strategically build their

reputation. In that scenario, there would be a natural tendency to maintain high returns for the investors, as reputation increases every time return increases. Similarly, the investors could prioritize those trustees who offered the highest returns. Moreover, as argued by Frey (1993), the crowding out effect will probably not occur in impersonal or purely abstract interactions (as is the case of interactions based on subjective evaluations of reputation).

Also in relation to the results, with the GPower 3.1 software, a *post hoc* analysis was made of the statistical power relative to the non-parametric difference in differences test, as previously reported (two-tailed Mann-Whitney *U* p-value of 0.002). To facilitate the understanding, we chose to analyze only the non-parametric test because it involves a direct test for comparing groups.

Considering the p-value of 0.002 with the means of 1.61 and -0.69 (corresponding to the difference between the pre-test and the post-test) of the control and treatment groups, respectively, the grouped SD of 7.72, and using the minimum asymptotic relative efficiency (ARE) method, an effect size of approximately 0.30 and power of 0.74 were computed. According to the standards usually used, considering the Cohen's *d*, the size found is small.

With that, based on the sizes of the groups that served as the basis for the study (160 and 120), to achieve a power of 80%, for example, an effect size of 0.32 would be necessary. Conversely, considering the effect size actually found, a power of 80% would be obtained with the following minimum group sizes: 190 and 142 for the treatment and control, respectively.

4.3 Additional Analysis

Additionally, due to trust also explaining reciprocal behavior (as was revealed in the experiment) and the fact that some studies have suggested that gender can influence reciprocity [e.g., Buchan et al. (2008), Chaudhuri and Sbai (2011), Croson et al. (2008), and Dittrich (2015)], a new regression was run based on equation 1, in which these controls are included. The results of the new regression are revealed in Table 6.

Table 6*Results of the panel data regression for the Recip independent variable*

Variables	Statistics	Pooled (OLS)	Random effect (GLS)	VIF
Constant	Coefficient	-0.31	0.59	
	Standard error	0.42	0.77	-
	Z-score	-0.75	0.76	
	p-value	0.45	0.44	
Rk	Coefficient	0.63	0.71	
	Standard error	0.43	1.06	2.01
	Z-score	1.48	0.66	
	p-value	0.14	0.51	
Dp	Coefficient	0.56	0.67*	
	Standard error	0.42	0.40	1.76
	Z-score	1.35	1.67	
	p-value	0.17	0.09	
Dp x Rk	Coefficient	-1.18**	-1.29**	
	Standard error	0.61	0.63	
	Z-score	-1.93	-2.06	2.76
	p-value	0.05	0.04	
Gender	Coefficient	-0.39	-0.39	
	Standard error	0.35	1.16	
	Z-score	-1.12	-0.34	1.00
	p-value	0.26	0.74	
Trust	Coefficient	1.58***	1.42***	
	Standard error	0.06	0.06	1.00
	Z-score	24.64	22.83	
	p-value	0.00	0.00	
Adjusted R ²		0.50	-	-
Durbin-Watson		1.35	2.05	-
Variance "between"		-	8.07	-
Variance "within"		-	15.58	-
F statistic/(p-value)		132.49/(0.00)	-	-

Note: Due to there being repeated observations for the same individual in the sample, the regressions used Beck-Katz robust standard errors (panel-corrected standard errors – PCSE) to correct any possible intra-individual correlations, heteroscedasticity between the groups, and contemporaneous correlations. The intercept (constant) reflects the measurement of the variable being studied for the control group before the treatment. Rk is a dummy variable that takes the value 1 when referring to the treatment group and 0 in other cases, Dp is a dummy variable that takes the value 1 when the data refer to the post-treatment period and 0 for the pre-treatment period, Dp x Rk indicates the effect of the treatment, the dependent variable Recip represents the mean value sent by the trustees in the environment of repeated interactions, the Trust variable is measured by the amount of resources that the investor sends to the trustee, and Gender is a dummy variable where 1 represents the male gender and 0 represents the female gender.

*, **, *** = significant at 10, 5, and 1%, respectively.

GLS = generalized least squares; OLS = ordinary least squares; VIF = variance inflation factor.

Source: Elaborated by the authors.

The results in Table 6 show that the inclusion of the control variables improved the model, giving more robustness to the estimates with an expressive increase in the explanatory power (adjusted R² of 0.50). Moreover, the variable of interest continues to be significant at 5%. The gender variable did not present

significance. Furthermore, confirming the prediction of the literature, trust is positively correlated with reciprocity in both models. This result serves to confirm that, even controlling for trust, recordkeeping had an effect on direct reciprocity, suggesting the presence of the crowding out effect.

5. CONCLUSION

As highlighted by Mullins et al. (2013), recent studies derived from behavioral economics, paleography, evolutionary psychology, and anthropology suggest that recordkeeping helps to resolve the problem of cooperation in large groups, transcending the serious limitations of our psychology evolved through the promotion of reciprocal behaviors and reputation building and maintenance. However, that is only true in specific circumstances, given the crowding out effect of recordkeeping over direct reciprocity documented in this study.

In general, the results presented reinforce the idea that, contemporarily, reciprocity works as a response to trust and serves as a predictor of future trust, as widely documented in previous experimental studies. Specifically, the experimental manipulation had an effect on the instrumental behavior of the participants that performed the role of trustee, reducing the efficiency of cooperation, given that the mean marginal return of the trustees for the investors decreased in the post-test. These findings give support to the hypothesis raised in the research and are important because they show that recordkeeping influences the instrumental behavior of individuals based on self-interest.

As conjectured by Basu et al. (2009), it is probable that, by improving the memory of past interactions, the recording of information about the transactions influenced the subjective evaluation that the trustees made of the investors' real intentions, leading them to strongly reduce their propensity for reciprocity in the search to increase their gains and punish their counterpart for breaching the social contract (reciprocity norm). Based on what occurred in the control group, it was expected that in the treatment group there would be an increase in the resources sent to the trustees. However, a reduction occurred, signaling a possible crowding out effect of

recordkeeping in the environment in which only direct reciprocity is possible.

Taken together, the findings documented in this study serve to show the evolutionary role of accounting, even in its most basic function, and they are consistent with the idea defended by Waymire and Basu (2008) that this involves an evolved economic institution capable of influencing the pattern of human cooperation. For that reason, they have important practical implications, because they indicate that accounting control mechanisms can have adverse effects on human behavior if the social context in which they are applied is not considered.

However, the results of this research need to be considered in light of some restrictions. For example, extracurricular credits were used as an incentive mechanism. Even with the support of the literature for their use, compensating the participants with money could work as a more powerful incentive, affecting the individuals' level of commitment, as various studies involving the trust game have shown. Another limitation was not having evaluated the effect of the pre-test on the behavior of the treatment group. Future research could use an experimental design, such as the Solomon four-group design, to try to identify if that problem occurs.

Also within the field of suggestions for future research, some studies have documented that in environments with repeated interactions instrumental reciprocity is one of the main factors that explain the cooperative behavior of individuals (Cabral et al., 2014; Dreber et al. 2014). For that reason, it would be interesting for new studies to try to evaluate the effect of recordkeeping over instrumental reciprocity. Moreover, it is important for new studies to evaluate the role of altruism and aversion to inequality in the relationship between recordkeeping and reciprocity.

REFERENCES

- Abraham, M., Grimm, V., Neeß, C., & Seebauer, M. (2016). Reputation formation in economic transactions. *Journal of Economic Behavior & Organization*, 121, 1-14. <https://doi.org/10.1016/j.jebo.2015.10.010>
- Arnold, M. C., & Schreiber, D. (2013). Audits, reputation, and repeated interaction in a capital budgeting setting. *European Accounting Review*, 22(1), 185-213. <https://doi.org/10.1080/09638180.2011.631734>
- Baek, S. K., Jeong, H.-C., Hilbe, C., & Nowak, M. A. (2016). Comparing reactive and memory-one strategies of direct reciprocity. *Scientific Reports*, 6(1), 25676. <https://doi.org/10.1038/srep25676>
- Barney, J. B., & Hanson, M. H. (1995). Trustworthiness as a source of competitive advantage. *Strategic Management Journal*, 15(supplement), 175-190.
- Basu, S., & Waymire, G. B. (2006). Recordkeeping and human evolution. *Accounting Horizons*, 20(3), 201-229.
- Basu, S., Dickhaut, J., Hecht, G., Towry, K., & Waymire, G. (2009). Recordkeeping alters economic history by promoting reciprocity. *Proceedings of the National Academy*

- of Sciences, 106(4), 1009-1014. <https://doi.org/10.1073/pnas.0811967106>
- Berg, J., Dickhaut, J., & McCabe, K. (1995). Trust, reciprocity, and social history. *Games and Economic Behavior*, 10(1), 122-142. <https://doi.org/10.1006/game.1995.1027>
- Bohnet, I., & Huck, S. (2004). Repetition and reputation: Implications for trust and trustworthiness when institutions change. *American Economic Review*, 94(2), 362-366. <https://doi.org/10.1257/0002828041301506>
- Bourgeois-Gironde S., & Corcos, A. (2011). Discriminating strategic reciprocity and acquired trust in the repeated trust-game. *Economics Bulletin*, 31(1), 177-188.
- Brandts, J., & Figueras, N. (2003). An exploration of reputation formation in experimental games. *Journal of Economic Behavior & Organization*, 50(1), 89-115. [https://doi.org/10.1016/s0167-2681\(02\)00042-2](https://doi.org/10.1016/s0167-2681(02)00042-2)
- Buchan, N. R., Croson, R. T. A., & Solnick, S. (2008). Trust and gender: An examination of behavior and beliefs in the investment game. *Journal of Economic Behavior & Organization*, 68(3-4), 466-476. <https://doi.org/10.1016/j.jebo.2007.10.006>
- Cabral, L., Ozbay, E. Y., & Schotter, A. (2014). Intrinsic and instrumental reciprocity: An experimental study. *Games and Economic Behavior*, 87, 100-121. <https://doi.org/10.1016/j.geb.2014.05.001>
- Calabuig, V., Fatas, E., Olcina, G., & Rodriguez-Lara, I. (2016). Carry a big stick, or no stick at all: Punishment and endowment heterogeneity in the trust game. *Journal of Economic Psychology*, 57, 153-171. <https://doi.org/10.1016/j.joep.2016.09.006>
- Camera, G., & Casari, M. (2017). Monitoring institutions in indefinitely repeated games. *Experimental Economics*, 21(3), 673-691. <https://doi.org/10.1007/s10683-017-9532-5>
- Camerer, C., & Weigelt, K. (1988). Experimental tests of a sequential equilibrium reputation model. *Econometrica*, 56(1), 1. <https://doi.org/10.2307/1911840>
- Camerer, C. (2003). *Behavioral game theory: Experiments in strategic interaction*. Princeton University Press.
- Campbell D. T., & Stanley J. C. (1963). *Experimental and quasi-experimental designs for research*. Rand McNally & Company.
- Charness, G., & Shmido, V. (2014). Trust and reciprocity. *Foundations and Trends in Microeconomics*, 10(3), 167-207. <http://dx.doi.org/10.1561/07000000065>
- Chaudhuri, A., & Sbai, E. (2011). Gender differences in trust and reciprocity in repeated gift exchange games. *New Zealand Economic Papers*, 45(1-2), 81-95. <https://doi.org/10.1080/00779954.2011.556072>
- Christ, M. H. (2013). An experimental investigation of the interactions among intentions, reciprocity, and control. *Journal of Management Accounting Research*, 25(1), 169-197. <https://doi.org/10.2308/jmar-50443>
- Christ, M. H., Sedatole K. L., Towry K. L., & Thomas M. A. (2008). When formal controls undermine trust and cooperation. *Strategic Finance*, 2008(January), 38-44.
- Cochard, F., Nguyen Van, P., & Willinger, M. (2004). Trusting behavior in a repeated investment game. *Journal of Economic Behavior & Organization*, 55(1), 31-44. <https://doi.org/10.1016/j.jebo.2003.07.004>
- Coricelli, G., Morales, L. G., & Mahlstedt, A. (2006). The investment game with asymmetric information. *Metroeconomica*, 57(1), 13-30. <https://doi.org/10.1111/j.1467-999x.2006.00231.x>
- Cox, J. (2004). How to identify trust and reciprocity. *Games and Economic Behavior*, 46(2), 260-281. [https://doi.org/10.1016/S0899-8256\(03\)00119-2](https://doi.org/10.1016/S0899-8256(03)00119-2)
- Cozby, P. C., & Bates, S. C. (2011). *Methods in behavioral research* (11th ed.). McGraw-Hill.
- Croson, R., Buchan, N., & Solnick, S. (2008). Trust and gender: An examination of behaviour, biases, and beliefs in the investment game. *Journal of Economic Behaviour and Organization*, 68(3-4), 466-476.
- Deci, E. L., & Ryan, R. M. (1985). *Intrinsic motivation and self-determination in human behavior*. Plenum Press.
- Deci, E. L., Koestner, R., & Ryan, R. M. (1999). A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation. *Psychological Bulletin*, 125(6), 627-668. <https://doi.org/10.1037/0033-2909.125.6.627>
- Dickinson, D., & Villeval, M.-C. (2008). Does monitoring decrease work effort? The complementarity between agency and crowding-out theories. *Games and Economic Behavior*, 63(1), 56-76. <https://doi.org/10.1016/j.geb.2007.08.004>
- Dittrich, M. (2015). Gender differences in trust and reciprocity: Evidence from a large-scale experiment with heterogeneous subjects. *Applied Economics*, 47(36), 3825-3838. <https://doi.org/10.1080/00036846.2015.1019036>
- Dreber, A., Fudenberg, D., & Rand, D. G. (2014). Who cooperates in repeated games: The role of altruism, inequity aversion, and demographics. *Journal of Economic Behavior & Organization*, 98, 41-55. <https://doi.org/10.1016/j.jebo.2013.12.007>
- Duffy, J., Xie, H., & Lee, Y.-J. (2011). Social norms, information, and trust among strangers: Theory and evidence. *Economic Theory*, 52(2), 669-708. <https://doi.org/10.1007/s00199-011-0659-x>
- Engelmann, D., & Fischbacher, U. (2009). Indirect reciprocity and strategic reputation building in an experimental helping game. *Games and Economic Behavior*, 67(2), 399-407. <https://doi.org/10.1016/j.geb.2008.12.006>
- Engelmann, D., & Strobel, M. (2010). Inequality aversion and reciprocity in moonlighting games. *Games*, 1(4), 459-477. <https://doi.org/10.3390/g1040459>
- Engle-Warnick, J., & Slonim, R. L. (2004). The evolution of strategies in a repeated trust game. *Journal of Economic Behavior & Organization*, 55(4), 553-573. <https://doi.org/10.1016/j.jebo.2003.11.008>
- Engle-Warnick, J., & Slonim, R. L. (2006a). Inferring repeated-game strategies from actions: Evidence from trust game experiments. *Economic Theory*, 28(3), 603-632. <https://doi.org/10.1007/s00199-005-0633-6>
- Engle-Warnick, J., & Slonim, R. L. (2006b). Learning to trust in indefinitely repeated games. *Games and Economic Behavior*, 54(1), 95-114. <https://doi.org/10.1016/j.geb.2004.10.009>

- Enzle, M. E., & Anderson, S. C. (1993). Surveillant intentions and intrinsic motivation. *Journal of Personality and Social Psychology*, 64(2), 257-266. <https://doi.org/10.1037/0022-3514.64.2.257>
- Falk, A., & Kosfeld, M. (2006). The hidden costs of control. *American Economic Review*, 96(5), 1611-1630. <https://doi.org/10.1257/aer.96.5.1611>
- Fehr, E., & List, J. A. (2004). The hidden costs and returns of incentives – Trust and trustworthiness among CEOs. *Journal of the European Economic Association*, 2(5), 743-771. <https://doi.org/10.1162/1542476042782297>
- Fehr, E., & Rockenbach, B. (2003). Detrimental effects of sanctions on human altruism. *Nature*, 422(6928), 137-140. <https://doi.org/10.1038/nature01474>
- Fehr, E., & Schmidt, K. M. (2006). The economics of fairness, reciprocity and altruism – Experimental evidence and new theories. *Foundations*, 1, 615-691. [https://doi.org/10.1016/s1574-0714\(06\)01008-6](https://doi.org/10.1016/s1574-0714(06)01008-6)
- Frey, B. S. (1993). Does monitoring increase work effort? The rivalry with trust and loyalty. *Economic Inquiry*, 31(4), 663-670. <https://doi.org/10.1111/j.1465-7295.1993.tb00897.x>
- Frey, B. S., & Jegen, R. (2001). Motivation crowding theory. *Journal of Economic Surveys*, 15(5), 589-611. <https://doi.org/10.1111/1467-6419.00150>
- Gächter, S., & Herrmann, B. (2009). Reciprocity, culture, and human cooperation: Previous insights and a new cross-cultural experiment. *Philosophical Transactions of the Royal Society: Biological Sciences*, 364(1518), 791-806. <https://doi.org/10.1098/rstb.2008.0275>
- Gambetta, D. (1988). Can we trust? In D. Gambetta (Ed.), *Trust: Making and breaking cooperative relationships* (pp. 213-237). Blackwell.
- Glaeser, E., Laibson, D., Scheinkman, J., & Soutter, C. (2000). *The Quarterly Journal of Economics*, 115(3), 811-846. <https://doi.org/10.1162/003355300554926>
- Hales, J., & Williamson, M. (2010). Implicit employment contracts: The limits of management reputation for promoting firm productivity. *Journal of Accounting Research*, 48(1), 51-80. <https://doi.org/10.1111/j.1475-679X.2009.00352.x>
- Hilbe, C., Martinez-Vaquero, L. A., Chatterjee, K., & Nowak, M. A. (2017). Memory-n strategies of direct reciprocity. *Proceedings of the National Academy of Sciences*, 114(18), 4715-4720. <https://doi.org/10.1073/pnas.1621239114>
- Hizak, J., DMITROVIC, L. G., & CUBRILLO, M. (2018). *The role of memory in transition from direct to indirect reciprocity*. <https://arxiv.org/pdf/1810.02536v1.pdf>
- Houser, D., Xiao, E., McCabe, K., & Smith, V. (2008). When punishment fails: Research on sanctions, intentions and non-cooperation. *Games and Economic Behavior*, 62(2), 509-532. <https://doi.org/10.1016/j.geb.2007.05.001>
- Johnsen, Å. A., & Kvaløy, O. (2016). Does strategic kindness crowd out prosocial behavior? *Journal of Economic Behavior & Organization*, 132A, 1-11. <https://doi.org/10.1016/j.jebo.2016.09.016>
- Johnson, N. D., & Mislin, A. (2010). Trust games: A meta-analysis. *Journal of Economic Psychology*, 32(5). <https://doi.org/10.2139/ssrn.1702678>
- Kanagaretnam, K., Mestelman, S., Nainar, S. M. K., & Shehata, M. (2010). Trust and reciprocity with transparency and repeated interactions. *Journal of Business Research*, 63(3), 241-247. <https://doi.org/10.1016/j.jbusres.2009.03.007>
- Kanagaretnam, K., Mestelman, S., Nainar, S. M. K., & Shehata, M. (2014). Transparency and empowerment in an investment environment. *Journal of Business Research*, 67, 2030-2038. <https://doi.org/10.1016/j.jbusres.2013.10.007>
- Kelly, K., & Tan, P. M.-S. (2010). The effects of profit-sharing contract and feedback on the sustainability of cooperation. *Journal of Management Accounting Research*, 22(1), 251-269. <https://doi.org/10.2308/jmar.2010.22.1.251>
- Libby, R., Bloomfield, R., & Nelson, M. W. (2002). Experimental research in financial accounting. *Accounting, Organizations and Society*, 27(8), 775-810. [https://doi.org/10.1016/s0361-3682\(01\)00011-3](https://doi.org/10.1016/s0361-3682(01)00011-3)
- Lunawat, R. (2013a). The role of information in building reputation in an investment/trust game. *European Accounting Review*, 22(3), 513-532. <https://doi.org/10.1080/09638180.2012.748256>
- Lunawat, R. (2013b). An experimental investigation of reputation effects of disclosure in an investment/trust game. *Journal of Economic Behavior & Organization*, 94, 130-144. <https://doi.org/10.1016/j.jebo.2013.07.018>
- Lunawat, R. (2016). Reputation effects of information sharing. *Journal of Economic Behavior & Organization*, 131, 75-91. <https://doi.org/10.1016/j.jebo.2016.08.006>
- Maas, V., Van Rinsum, M., & Towry, K. (2012). In search of informed discretion: An experimental investigation of fairness and trust reciprocity. *The Accounting Review*, 87(2), 617-644. <https://doi.org/10.2308/accr-10205>
- Malhotra, D. (2004). Trust and reciprocity decisions: The differing perspectives of trustors and trusted parties. *Organizational Behavior and Human Decision Processes*, 94(2), 61-73. <http://dx.doi.org/10.1016/j.obhdp.2004.03.001>
- Marinich, E. (2019). Accounting information aggregation and managerial cooperation. *Journal of Management Accounting Research*. <https://doi.org/10.2308/jmar-17-033>
- Milinski, M., & Wedekind, C. (1998). Working memory constrains human cooperation in the prisoner's dilemma. *Proceedings of the National Academy of Sciences*, 95(23), 13755-13758. <https://doi.org/10.1073/pnas.95.23.13755>
- Mullins, D. A., Whitehouse, H., & Atkinson, Q. D. (2013). The role of writing and recordkeeping in the cultural evolution of human cooperation. *Journal of Economic Behavior & Organization*, 90(supplement), S141-S151. <https://doi.org/10.1016/j.jebo.2012.12.017>
- Nowak, M. A., & Sigmund, K. (2005). Evolution of indirect reciprocity. *Nature*, 437, 1291-1298. <https://doi.org/10.1038/nature04131>
- Osterloh, M., & Frey, B. (2002). Does pay for performance really motivate employees? In A. Neely (Ed.), *Business performance measurement: Theory and practice* (pp. 107-122). Cambridge University Press. <https://doi.org/10.1017/CBO9780511753695.008>
- Rabin, M. (1993). Incorporating fairness into game theory and economics. *The American Economic Review*, 83(5), 1281-1302.

- Rand, D. G., Ohtsuki, H., & Nowak, M. A. (2009). Direct reciprocity with costly punishment: Generous tit-for-tat prevails. *Journal of Theoretical Biology*, 256(1), 45-57. <https://doi.org/10.1016/j.jtbi.2008.09.015>
- Rankin, D. J., & Eggimann, F. (2009). The evolution of judgement bias in indirect reciprocity. *Proceedings of the Royal Society: Biological Sciences*, 276(1660), 1339-1345. <https://doi.org/10.1098/rspb.2008.1715>
- Rietz, T. A., Schniter, E., Sheremeta, R. M., & Shields, T. W. (2017). *Trust, reciprocity, and rules*. *Economic Inquiry*, 56(3), 1526-1542. <https://doi.org/10.1111/ecin.12512>
- Roberts, G. (2008). Evolution of direct and indirect reciprocity. *Proceedings of the Royal Society: Biological Sciences*, 275(1631), 173-179. <https://doi.org/10.1098/rspb.2007.1134>
- Rockenbach, B., & Milinski, M. (2006). The efficient interaction of indirect reciprocity and costly punishment. *Nature*, 444(7120), 718-723. <https://doi.org/10.1038/nature05229>
- Salkind N. J. (2002). *Encyclopedia of research design*, Vol. 10. Independence.
- Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Mifflin and Company.
- Slonim, R., Wang, C., Garbarino, E., & Merrett, D. (2013). Opting-in: Participation bias in economic experiments. *Journal of Economic Behavior & Organization*, 90, 43-70. <https://doi.org/10.1016/j.jebo.2013.03.013>
- Smith, A. (2013). *Reciprocity effects in the trust game*. *Games*, 4(3), 367-374. <https://doi.org/10.3390/g4030367>
- Sobel, J. (2005). Interdependent preferences and reciprocity. *Journal of Economic Literature*, 43(2), 392-436. <https://doi.org/10.1257/0022051054661530>
- Stanca, L., Bruni, L., & Corazzini, L. (2009). Testing theories of reciprocity: Do motivations matter? *Journal of Economic Behavior & Organization*, 71(2), 233-245. <https://doi.org/10.1016/j.jebo.2009.04.009>
- Stevens, J. R., & Hauser, M. D. (2004). Why be nice? Psychological constraints on the evolution of cooperation. *Trends in Cognitive Sciences*, 8(2), 60-65. <https://doi.org/10.1016/j.tics.2003.12.003>
- Van Veelen, M., Garcia, J., Rand, D. G., & Nowak, M. A. (2012). Direct reciprocity in structured populations. *Proceedings of the National Academy of Sciences*, 109(25), 9929-9934. <https://doi.org/10.1073/pnas.1206694109>
- Waymire, G. B., & Basu, S. (2008). Accounting is an evolved economic institution. *Foundations and Trends in Accounting*, 2(1-2), 1-174. <https://doi.org/10.1561/1400000011>