Journal of Accounting Research

DOI: 10.1111/1475-679X.12628
Journal of Accounting Research
Vol. No. June 2025
Printed in U.S.A.

CHICAGO BOOTH

Use and Design of Peer Evaluations for Bonus Allocations

MANUEL GRIEDER* AND KARL SCHUHMACHER†

Received 1 December 2022; accepted 4 March 2025

ABSTRACT

We conduct an experiment to investigate the use of peer evaluations for compensation purposes. Although organizations often rely on peer evaluations for incentive compensation, it is not well understood how peer feedback should be used and designed to ensure non-distorted evaluations and motivate effort provision. We study peer evaluations in form of bonus allocation proposals, thereby enabling a quantifiable test of our hypothesis. We distinguish between discretionary use (i.e., allocation by the manager) and formulaic use (i.e., allocation by the team via the average) of self-including and self-excluding proposals. We find that, relative to self-including proposals, self-excluding proposals are less distorted, irrespective of use, but lead to more effort provision only under formulaic use. Under discretionary use, the benefits of self-excluding proposals are offset, as managerial biases enter bonus allocations. In sum, our findings illustrate benefits of delegating bonus allocations to teams through formulaic use of self-excluding peer evaluations

*Faculty of Economics, UniDistance Suisse; School of Management & Law, Zurich University of Applied Sciences; †Goizueta Business School, Emory University

Accepted by Regina Wittenberg Moerman. We thank an anonymous reviewer for very helpful comments and suggestions. We also appreciate insightful feedback from Lorenz Goette, Deborah Kistler, Victor Maas, Serena Loftus, Michael Majerczyk, Hong Qu, Jan Schmitz, Kevin Smith, Doug Stevens, Matthias Sutter, Ivo Tafkov, Kristy Towry, Grazia Xiong, and Christian Zehnder, seminar participants at the Goizueta Business School experimental brownbag series, workshop participants at Georgia State University, and conference participants at the AAA Annual Meeting 2023, the MAS Midyear Meeting 2023, SESARC 2023, and IOEA 2024. Karl Schuhmacher is grateful to the Goizueta Business School for financial support. The authors have no potential conflicts of interest to declare. All errors are our own. An online appendix to this paper can be downloaded at https://www.chicagobooth.edu/jar-online-supplements.

and extend the understanding of how organizations can effectively incorporate peer evaluations into incentive compensation.

JEL codes: C92, D23, D86, D90, M40, M41, M52

Keywords: peer evaluations; team-based compensation; mutual monitoring; mechanism design

1. *Introduction*

Peer evaluations are common practice, because peers usually have better information about each other's actions than their superiors (via mutual monitoring). Organizations increasingly rely on peer evaluations for incentive compensation (Edwards and Ewen [1996], Appelo [2015], Deb, Li, and Mukherjee [2016]). As such, a burgeoning literature examines the use of peer evaluations in bonus pool arrangements (Arnold, Hannan, and Tafkov [2018, 2020], Dong, Falvey, Luckraz [2019], Abbink, Dong, and Huang [2022]). In these arrangements, an objective measure of group performance typically determines the size of the bonus pool, while subjective signals, such as peer evaluations, inform the allocation of the bonus pool among employees (Baiman and Rajan [1995], Rajan and Reichelstein [2006, 2009]).[1] Yet, it is not well understood how peer feedback should be used and designed to ensure non-distorted evaluations and motivate effort provision. Our study employs a bonus pool arrangement to compare the effectiveness of a discretionary use (i.e., managers keep discretion) and a formulaic use (i.e., managers commit to the average) of self-including and self-excluding peer evaluations in motivating employee effort.

Studying the use of peer evaluations for compensation purposes is important, as millennials and generation Zs, a growing share of the workforce, "embrace personal responsibility" and want their voices to be heard, especially as it pertains to their compensation (see Deloitte [2021]). The extent to which peer feedback is used for incentive compensation varies in practice and is subject to debate. In many organizations, managers elicit peer feedback but maintain discretion over the allocation of bonuses. For example, at Goldman Sachs and LivingSocial, employees provide peer evaluations that help managers set annual bonuses (e.g., Silverman and Kwoh [2012], Alexander [2022]). Similarly, at Crane, peer evaluations inform discretionary bonus pool allocations among subunit presidents. In other organizations, managers do not maintain discretion and, instead, commit to formulaically aggregating peer feedback to determine bonuses, effectively delegating incentive compensation to their employees. For instance, at Agility Scales, peer evaluations are directly entered into the bonus formula (Hannah [2019]). Similarly, Galbraith [1997] reports on

[1] Survey evidence indicates that most US firms use bonus plans that are based on both firm profits and subjective signals of performance (Murphy and Oyer [2003], WorldatWork and Compensation Advisory Partners [2021a]).

a formulaic aggregation via the average of peer bonus pool allocation proposals at National City Bank (now Citibank). Google, Meta, Microsoft, and Zappos use peer-to-peer bonuses, enabling their employees to reward each other through peer assessments (e.g., Emmonds [2019]). Given these trends toward utilizing peer feedback for incentive compensation, it is critical to understand how peer evaluations should be used and designed to best motivate employee effort.

In line with prior work, we study peer evaluations in form of bonus pool allocation proposals (e.g., Arnold, Hannan, and Tafkov [2018, 2020]). The bonus pool arrangement lends itself to our research question because it creates—like the public goods game [PGG]—opportunities for free riding (i.e., employees generate, with their efforts, a common pool that is shared among them). Yet, unlike the classic PGG, bonus pools are not necessarily evenly split. Rather, the allocation of the pool is based on subjective signals—here, peer allocation proposals—enabling a reward system that can limit free riding but also create incentives for employees to distort proposals in their own favor, similar to peer performance evaluations used for incentive pay.[2]

Within this setting, we compare a discretionary use (i.e., allocation by an impartial manager) and a formulaic use (i.e., allocation by the team via the average) of self-including proposals and self-excluding proposals. Self-including proposals allow employees to propose an allocation of the bonus pool among all team members including themselves, whereas self-excluding proposals prohibit self-allocations. Prior research does not compare discretionary use vs. formulaic use or self-including design vs. self-excluding design. In the accounting literature, Arnold, Hannan, and Tafkov [2018] focus on the combination of discretionary use and self-including design. They find that managers' bonus allocations lead to more effort provision when self-including proposals are made available to managers than when they are not. In the common pool literature in economics, Dong, Falvey, and Luckraz [2019] focus on the combination of formulaic use and self-excluding design. They find that bonus allocations based on a linear aggregation of self-excluding proposals result in more effort provision than even-split allocations. Yet, it remains unknown across the two literatures whether a discretionary use or a formulaic use of either self-including or self-excluding (peer) allocation proposals is better suited to motivate effort provision.

To address this gap, we first outline the monetary incentives underlying those different uses and designs within an analytical framework. The predictions of the analytical framework greatly vary with small changes in the assumptions, illustrating the need for an empirical investigation. Distinct

[2] A key feature of bonus pool arrangements, relative to other common pool arrangements, is that the reward system is self-funding, as reward and punishment solely materialize via the allocation of the pool (i.e., a zero-sum game).

from the analytical framework, we develop a hypothesis drawing on behavioral research.

Our reasoning builds on the notion that employees provide more effort the more they expect their bonuses to be fair. We define bonuses as fair—and, similarly, proposals as non-distorted—if they reflect employees' relative contributions to the bonus pool. Whether bonuses will be fair depends on (1) employees' allocation proposals and (2) the transformation of these proposals into bonuses. First, we argue that the self-excluding design leads to less distortion in proposals than the self-including design under both discretionary use and formulaic use. We expect employees to neglect incentives to distort self-excluding proposals, which arise under discretionary use but not under formulaic use (as will be described). Second, we argue that compared to formulaic use, managers' transformations of proposals into bonuses are more prone to bias when proposals are self-excluding than self-including. Specifically, we expect it to be more complex for managers to derive fair bonuses from self-excluding proposals than from self-including proposals. Therefore, we hypothesize that the positive effect of the self-excluding design, relative to the self-including design, on effort provision is less pronounced under discretionary use than under formulaic use.

To test this hypothesis, we conduct an incentivized laboratory experiment with business school students. We assign participants to groups of four—one manager and three employees—that stay together for a total of eight rounds, each consisting of three stages. Stage 1 is similar to a PGG in that employees choose costly effort contributions, which are linearly aggregated to form group output. A fixed percentage of group output determines the bonus pool. While the manager only observes group output, employees also observe individual contributions. In stage 2, each employee proposes an allocation of the bonus pool. In stage 3, the bonus pool is allocated among the three employees. Within this set-up, we manipulate—between-groups—the design in stage 2 (self-including vs. self-excluding) and use in stage 3 (formulaic vs. discretionary) of employees' allocation proposals. The self-including design allows employees to propose an allocation of the pool among all three employees, while the self-excluding design requires an allocation among the other two employees. Under formulaic use, the average of all proposals determines employees' final bonuses (i.e., allocation by the team), whereas, under discretionary use, the manager keeps full discretion over the final bonus allocation (i.e., allocation by the manager).

The results of the experiment support our hypothesis. We find that, under formulaic use, the self-excluding design leads to significantly more effort provision than the self-including design. Under discretionary use, this positive effect is significantly attenuated, such that we do not find a significant difference in effort provision between the self-including and self-excluding designs. Overall, the formulaic use of self-excluding proposals leads to more effort provision than each of the other three conditions. Perceptions of distributive fairness explain these differences in effort. In line with the behavioral reasoning underlying our hypothesis, the self-excluding design leads to less distortion in proposals than the self-including design, irrespective of use. However, under discretionary use, this effect is counteracted, as managerial biases enter the transformation of proposals into final bonuses more when proposals are self-excluding than self-including. That is, managers tend to allocate more (less) to low (high) contributors than suggested by the average of employees' proposals than when the design is self-excluding.

Our study makes several contributions. First, we contribute to a burgeoning literature that examines the use of peer evaluations for incentive compensation (e.g., Deb, Li, and Mukherjee [2016], Arnold, Hannan, and Tafkov [2018, 2020]). Specifically, we advance the understanding of how different uses and designs of peer evaluations affect employee behavior. We find that the combination of formulaic use and self-excluding design results in the most favorable outcomes. Thus, organizations may consider limiting managers' discretion in setting bonuses and separating employees' self-evaluations from evaluations of their peers. In this vein, we also contribute to research on calibration committees that hold managers accountable for their bonus allocations (e.g., Demeré, Sedatole, and Woods [2019], Grabner, Künneke, and Moers [2020]).

Second, our study adds to the literature on bonus pools in accounting and common pools in economics, which both examine different uses and designs of peer allocation proposals within bonus pool arrangements. While the literature in accounting focuses on a transfer of employees' allocation proposals to an impartial manager, who can use that information to allocate bonuses (e.g., Arnold, Hannan, and Tafkov [2018, 2020]), the literature in economics focuses on a self-governing use, thereby enabling employees to allocate bonuses among themselves (e.g., Dong, Falvey, and Luckraz [2019], Abbink, Dong, and Huang [2022]). We contribute to both literatures by showing that the relative effectiveness of the two uses depends on the design of the allocation proposals. Only when the design is self-excluding, formulaic use (i.e., allocation by employees) motivates more effort than discretionary use (i.e., allocation by manager). This result seems to be driven by fairness considerations, because, only when the design is self-excluding, formulaic use leads to fairer bonuses than discretionary use. As such, we also add to accounting research that examines how control choices, like choices about the use and design of peer evaluations, affect fairness perceptions and cooperation in organizations (e.g., Coletti, Sedatole, and Towry [2005], Tayler and Bloomfield [2011], Maas and van Rinsum [2013], Cardinaels and Yin [2015]).

Third, our study contributes to the broader management literature on delegation (Nagar [2002], Moers [2006], Graham, Harvey, and Puri [2015]) and answers calls for more research on the delegation of decision rights to teams rather than individuals (Fehrler and Janas [2021]). In that regard, we also add to analytical work on bonus pools in the accounting literature (Rajan and Reichelstein [2006, 2009]) by considering the delegation of the allocation decision to employees. Notably, the formulaic use of self-excluding proposals extends the solution of the moral hazard problem between managers and employees to the moral hazard problem among employees.

The remainder of this paper proceeds as follows. Section 2 reviews the literature and outlines the analytical framework and hypothesis development. Section 3 describes the experimental setting, design, and procedures. Section 4 summarizes the results, and section 5 concludes.

2. *Literature Review, Analytical Framework, and Hypothesis Development*

2.1 BONUS POOL ARRANGEMENTS IN ACCOUNTING LITERATURE

Subjective information is omnipresent in performance appraisals and compensation systems. In principal–agent relationships, such non-verifiable information creates a moral hazard problem, as the principal could claim ex post that the agents performed worse than they did, resulting in lower bonuses for agents and larger profits for the principal. Bonus pool arrangements eliminate this moral hazard problem (e.g., Baiman and Rajan [1995], Rajan and Reichelstein [2006, 2009]) in that the principal commits ex ante to a formulaic determination of the size of the bonus pool, using an objective measure of group performance (e.g., a fixed percentage of profits). Because the bonus pool is only shared among the agents, the principal can credibly use subjective signals of individual performance to allocate the pool among the agents, thereby providing them with incentives to contribute effort (Ederhof, Rajan, and Reichelstein [2011]).

Survey evidence confirms that discretionary bonus pool arrangements are widely used in practice (e.g., Murphy and Oyer [2003], WorldatWork and Compensation Advisory Partners [2021a, 2021b, 2021c]). Moreover, several studies examine these arrangements in laboratory experiments (e.g., Fisher et al. [2005], Bailey, Hecht, and Towry [2011], Maas, van Rinsum, and Towry [2012], Arnold, Hannan, and Tafkov [2018, 2020], Arnold and Tafkov [2019], Majerczyk and Thomas [2021]). For example, Fisher et al. [2005] find that it is indeed beneficial when the principal (1) fixes the size of the bonus pool ex ante and (2) uses subjective signals of individual performance to allocate the bonus pool among the agents.

Extending this work, Arnold, Hannan, and Tafkov [2018, 2020] consider peer evaluations as subjective signals of individual performance, assuming agents are better informed about each other's performance than the principal (i.e., mutual monitoring). Specifically, they focus on self-including peer evaluations, allowing agents to also submit assessments of their own performance. Such a self-including design creates a new moral hazard problem, because agents benefit from distorting peer evaluations in their own favor, thereby weakening the incentive effects of bonus pool arrangements. Arnold, Hannan, and Tafkov [2018] operationalize peer evaluations as bonus pool allocation proposals sent to an uninformed principal. They find that, when self-including proposals are available, the principal's allocation leads to more effort provision than when such proposals are not available to the principal, suggesting that—while susceptible to opportunism—self-including proposals contain valuable information. In addition to this work in accounting, research in economics also examines the use of peer evaluations in bonus pool arrangements to address the free-rider problem in the common pool and PGG literatures.

2.2 BONUS POOL ARRANGEMENTS IN ECONOMICS LITERATURE

In a classic PGG, the marginal benefit of contributing to the common pool is known ex ante and calibrated such that it is optimal for all players, collectively, to contribute (i.e., cooperate) but for each player, individually, to free ride on the contributions of others. Experimental studies provide evidence that, in repeated PGGs, some players initially cooperate, while others free ride; yet, over time, average contributions decline, as initial cooperators observe free riders and lower their contributions (e.g., Kim and Walker [1984], Andreoni [1988], Isaac and Walker [1988], Chaudhuri [2011]). This strategy of adapting one's own contributions based on the contributions of others is called *conditional cooperation* (Fischbacher, Gächter, and Fehr [2001]).

To address the free-rider problem in the PGG, prior studies consider punishment and reward systems that enable players—after the pool is evenly split—to punish and/or reward each other. Such systems reduce free riding and can sustain cooperation (e.g., Ostrom, Walker, and Gardner [1992], Fehr and Gächter [2000b], Sefton, Shupp, and Walker [2007], Gächter, Renner, and Sefton [2008], Rand et al. [2009]). Different from bonus pool arrangements, the punishment and reward systems in these studies create social losses or gains.[3]

More recent studies in the common pool literature consider *self-funding* reward systems, where punishment and reward materialize only via the allocation of the pool, absent an initial even-split allocation. These reward systems differ from discretionary bonus pool arrangements solely in that the allocation right is not given to an impartial principal but directly to the agents who observe each other's contributions.[4] We categorize this stream of research into studies with self-including and self-excluding designs of peer allocation proposals.

Studies with *self-including* design typically examine the delegation of the allocation right to a single agent who allocates the pool among all agents

---

[3] Punishment creates a cost for both the sender and the receiver of the punishment, leading to a social loss; similarly, social gains arise, as the cost incurred for rewarding is usually smaller than the reward received (Rand et al. [2009]).

[4] Stoddard, Walker, and Williams [2014] and Stoddard, Cox, and Walker [2021] consider a third-party allocator who benefits from contributions to the pool (i.e., principal). Their setting captures discretionary bonus pool arrangements but assumes that the principal perfectly monitors individual contributions.

including themself (e.g., van der Heijden, Potters, and Sefton [2009], Drouvelis, Nosenzo, and Sefton [2017], Karakostas et al. [2023]).[5] These studies find that bonus allocations by a single agent are not fully opportunistic as they lead to higher contributions than even-split allocations. Notably, one agent who allocates the entire pool has similar incentives (for distortion) as many agents who each allocate a share of the pool.

Studies with *self-excluding* design usually delegate the allocation right to all agents such that each agent allocates an equal, fixed share of the pool among all other agents excluding themself, rendering them indifferent about their allocation (Yang et al. [2018], Dong, Falvey, and Luckraz [2019], Abbink, Dong, and Huang [2022]). In an experiment, Dong, Falvey, and Luckraz [2019] find that such a formulaic use of self-excluding proposals outperforms even-split allocations. Galbraith [1997] describes a slightly different, but algebraically equivalent mechanism, which requires each agent to propose an allocation of the *entire* bonus pool among only other agents, where the average of all agents' proposals determines the final bonus allocation.

Dong, Falvey, and Luckraz [2019] argue that the formulaic use of self-excluding proposals outperforms even-split allocations because bonuses better reflect agents' relative contributions to the pool. Similarly, Yang et al. [2018, p. 9968] refer to "payoff-based conditional cooperation" to explain that agents will contribute more, the more they expect their payoffs—from the pool—to reflect their relative contributions. Arnold, Hannan, and Tafkov [2018, 2020] use a similar line of reasoning to explain the effectiveness of the discretionary use of self-including proposals.

While prior work examines the different uses and designs of allocation proposals, we are not aware of research that compares their effectiveness. As such, it is unknown how peer evaluations—here, bonus allocation proposals—should be used (formulaic vs. discretionary) and designed (self-including vs. self-excluding) to best motivate effort provision. Our study addresses this gap. In what follows, we outline the monetary incentives underlying those uses and designs with an analytical framework, before developing our hypothesis drawing on behavioral research.

2.3 ANALYTICAL FRAMEWORK

The analytical framework describes a bonus pool arrangement that reflects the settings in Arnold, Hannan, and Tafkov [2018, 2020], Dong, Falvey, and Luckraz [2019], Abbink, Dong, and Huang [2022], and our

[5] Interestingly, Karakostas et al. [2023] require the agent who allocates the bonus pool to contribute their entire endowment to the pool, which may create a sense of obligation by others to match that contribution (e.g., Schuhmacher, Towry, and Zureich [2022]). Furthermore, other design features (e.g., number of agents, aggregation function, etc.) also differ among those studies, making it challenging to compare them. Baranski [2016] extends the approach in those studies by allowing all agents to vote on the proposing agent's allocation, which is implemented if a majority accepts the proposal.

study. That is, we capture a three-stage game, outlining the problem of a manager (principal), who can use subjective peer evaluations, to motivate $n > 2$ employees (agents), each with an endowment of $E > 0$, to contribute effort to group output.

In stage 1, each employee i independently chooses how much costly effort, e_i, to contribute, where $0 \leq e_i \leq E$ with cost of effort of $c_i = e_i$ (i.e., $c_i'(e_i) = 1$). Employees' efforts are added and multiplied with productivity factor $M > 1$ to determine group output $G = M \sum_i (e_i)$, where $M\, e_i$ represents employee i's individual performance.[6] The manager keeps portion θ of group output (i.e., $0 < \theta \leq (M - 1)/M$). The remainder $(1 - \theta)$ forms the bonus pool $B = (1 - \theta)\ M \sum_i (e_i) = M_B \sum_i (e_i)$, such that $M_B \geq 1$ reflects the marginal increase of the bonus pool per unit of effort.

In stage 2, the manager observes only group output, whereas employees also observe each other's effort contributions (and, thus, individual performances).[7] To access employees' private information, the manager requests peer evaluations, asking each employee i to send an allocation proposal, p_i, with a proposed bonus, $p_{i,j} \geq 0$, for every employee j, where $\sum_j (p_{i,j}) = B$.[8]

Stage 3 determines the final bonus allocation, b, with a bonus of $b_i \geq 0$ for each employee i, such that the bonus pool, B, is allocated among all employees without residual, i.e., $\sum_i (b_i) = B$. Importantly, this requirement ($\sum_i (b_i) = B$) is a key feature of bonus pool arrangements, because it ensures (1) the manager's impartiality toward employees and (2) the self-funding nature of the reward system (i.e., the bonus allocation is a zero-sum game without social losses or gains).

Within this setting, we differentiate the self-including and self-excluding designs (in stage 2) and the formulaic and discretionary uses (in stage 3) of employees' bonus allocation proposals. Contrary to the self-including design, the self-excluding design requires employees to propose a bonus of zero to themselves, where $p_{i,j} = 0$ for $i = j$ (Dong, Falvey, and Luckraz

[6] In line with prior work, we operationalize individual performance as a perfect multiple of effort, abstracting away noise in the relation between effort and individual performance (thereby ensuring optimal experimental control).

[7] Consistent with Arnold, Hannan, and Tafkov [2018], Dong, Falvey, and Luckraz [2019], and Abbink, Dong, and Huang [2022], we assume mutual monitoring. This assumption also aligns with the team production literature (e.g., Che and Yoo [2001], Kvaløy and Olsen [2006], Alonso and Matouschek [2008]). Similarly, in the PGG literature, transparency about contributions is a standard feature in studies that (1) incorporate punishment or reward systems (see Fiala and Suetens [2017]) or (2) examine the benefits of peer leadership (see Eichenseer [2023]).

[8] Note that $\sum_j (p_{i,j}) = B$ is not a necessary requirement, as the manager can proportionally transform any proposal (e.g., peer evaluations elicited on Likert scales) to percentages such that the proposed bonuses add up to 100%. For example, consider a team of three employees and assume employee A evaluates as follows: A: 9 / B: 7 / C: 4. The corresponding percentages amount to: A: 45% (= 9/20) / B: 35% (= 7/20) / C: 20% (= 4/20). In line with prior research and without loss of generality, we require $\sum_j (p_{i,j}) = B$, as it ensures consistent scaling across all employees, which enables a formulaic use based on simple averaging and facilitates the manager's task.

[2019], Abbink, Dong, and Huang [2022]). Under formulaic use, the final bonus allocation is then determined via the average of all n employees' proposals (i.e., $b_i = \sum_k (p_{k,i})/n$). Thus, each employee's proposal allocates an ex ante fixed share, $1/n$, of the pool, effectively delegating the final bonus allocation to the employees. In contrast, under discretionary use, the manager keeps full discretion over the final bonus allocation. In sum, we differentiate among the following four conditions based on the use and design of employees' allocation proposals: [1] formulaic use and self-including design, [2] formulaic use and self-excluding design, [3] discretionary use and self-including design, and [4] discretionary use and self-excluding design.

Next, we outline, for each condition, the requirement for multiplier M_B to *induce* full effort as a *unique* Nash equilibrium, respectively, *support* full effort as a Nash equilibrium.[9] We extend the analysis in Dong, Falvey, and Luckraz [2019] and apply backward induction (from stage 3), assuming individuals are self-interested and risk-neutral.[10] We demonstrate that the requirement for multiplier M_B to *induce* (*support*) full effort depends on individuals' behavior when they are indifferent about their allocation decisions, as is the case in conditions [2], [3], and [4]. First, we assume indifference is resolved via random allocations, resulting in the same requirement for M_B in all four conditions: $M_B > n$ ($M_B \geq n$). Next, we consider other allocation behaviors to resolve indifference and illustrate how they affect the requirement for M_B in conditions [2], [3], and [4].

2.3.1. Conditions [1] and [2]: Formulaic Use (Allocation by the Team). Under formulaic use, the bonus allocation in stage 3 is formulaically derived and, therefore, deterministic—given employees' proposals in stage 2 and efforts in stage 1. Hence, the problem reduces to a two-stage game (i.e., stages 1 and 2).

In condition [1], each employee i receives a bonus of $b_i = \sum_k (p_{k,i})/n$ and strictly prefers to propose allocating the entire pool to themself ($p_{i,i} = B$), as b_i grows with i's self-allocation $p_{i,i}$, since $b_i{}'(p_{i,i}) = 1/n > 0$. Thus, each employee i will receive $b_i = B/n$ and contribute effort, if the marginal benefit of effort $b_i{}'(e_i) = M_B/n$ (weakly) exceeds the marginal cost $c_i{}'(e_i) = 1$, resulting in the following requirement for multiplier M_B to *induce* (*support*) full effort: $M_B > n$ ($M_B \geq n$).

In condition [2], each employee i receives a bonus of $b_i = \sum_{k \neq i} (p_{k,i})/n$, where $b_i{}'(p_i) = 0$. Because p_i does not affect b_i, each employee is indifferent about their stage 2 proposal, such that any proposal p_i prescribes a subgame-perfect, dominant-strategy Nash equilibrium in stage 2. Assuming

[9] In what follows, we use the phrase *induce full effort* to refer to a full-effort Nash equilibrium that is unique and the phrase *support full effort* to refer to any full-effort Nash equilibrium (i.e., not necessarily unique).

[10] While assuming self-interested and risk-neutral individuals may not be useful for making absolute predictions, as numerous PGG studies show, it may be useful for making relative predictions among the different conditions.

indifference is resolved via a random allocation, employees propose in expectation an even-split among all other $n - 1$ employees (where $p_{i,j} = B/(n - 1)$ for $i \neq j$ and $p_{i,j} = 0$ for $i = j$). Thus, each employee i expects to receive a bonus of $b_i = B/n$, resulting in the same requirement for multiplier M_B in condition [2] as in condition [1].

Dong, Falvey, and Luckraz [2019] extend this analysis by considering the following three *allocation behaviors* to resolve indifference about proposals: (A) egalitarian allocation behavior (i.e., even-split proposals), (B) proportional allocation behavior (i.e., proposals reflecting relative contributions to the pool), and (C) winner-takes-all allocation behavior (i.e., proposals allocating the entire pool to the highest contributor(s)—in equal shares if several exist). Following Dong, Falvey, and Luckraz [2019] and assuming ex ante a given allocation behavior under indifference, table 1, panel A, shows the requirement for multiplier M_B to *induce* (*support*) full effort.

We focus on proportional allocation behavior, as it mirrors the presumption of a benevolent manager in analytical work on bonus pools (e.g., Baiman and Rajan [1995]). Prior research also shows individuals have concerns for truth-telling (e.g., Abeler, Nosenzo, and Raymond [2019]) and equity (e.g., Adams [1963, 1965], Lawler [1968]), which can lead to proportional allocations in our setting. We define proposals to be *non-distorted* and, similarly, bonuses to be *fair*, when they represent employees' relative contributions to the pool.[11] Assuming non-distorted proposals (i.e., proportional allocations), the following requirement—see table 1, panel A—must hold for multiplier M_B to *induce* (*support*) full effort: $M_B > n/(n-1)$ ($M_B \geq n(n-1)/(n(n-1)-1)$).

This requirement reflects an efficiency loss, as values of M_B between 1 and $n/(n - 1)$ are not sufficient, despite non-distorted proposals, to induce full effort. To illustrate the efficiency loss, consider one employee who contributes some effort and $n - 1$ employees who do not contribute any effort. The efficiency loss arises, as the sole contributor must allocate an *ex ante* fixed share, $1/n$, of the pool among all other employees such that only $(n - 1)/n < 1$ of the pool remains for the contributor. Thus, even if all proposals are *non-distorted*, bonuses will not necessarily be *fair*.[12] Notably, the efficiency loss is transitive in that higher contributions result in larger bonuses, just not in perfect multiples of M_B (i.e., high contributors cross-subsidize low contributors).[13] Table 1, panel B, shows examples of the efficiency loss in condition [2] for $n = 3$ employees and $M_B = 1.5$.

[11] We distinguish *non-distorted proposals* from *fair bonuses*, since they differ for the self-excluding design, where a proportional transformation of proposals is needed, such that allocations to others reflect their relative contributions.

[12] As such, if employees want to resolve indifference in stage 2 based on some concern for equity (in bonuses), this desire cannot always be perfectly fulfilled with proportional allocation behavior (i.e., non-distorted proposals).

[13] This transitivity ensures that if all employees contribute equally, final bonuses will be fair (i.e., no efficiency loss). In general, the efficiency loss depends on employees' contributions (Dong, Falvey, and Luckraz [2019]). It is largest when only one employee contributes and de-

TABLE 1

Formulaic Use and Self-Excluding Design (Condition [2])

Panel A: Employees' allocation behaviors and corresponding full effort requirements				
Stage 2: Employees' Allocation Proposals	Employees are indifferent about their allocation proposals. To resolve indifference, we consider four different allocation behaviors (Dong, Falvey, and Luckraz [2019]):			
	(A) Egalitarian Even-split proposals	(B) Proportional Non-distorted proposals	(C) Winner-takes-all All-or-nothing proposals	(D) Random Arbitrary proposals
Stage 1: *Employees' Effort Contributions*	Assuming ex ante a given allocation behavior to resolve indifference in stage 2, the following requirement for multiplier M_B must hold to …			
… *induce* full effort.	$M_B > n$	$M_B > \frac{n}{n-1}$	$M_B > \frac{n}{n-1}$	$M_B > n$
… *support* full effort.	$M_B \geq n$	$M_B \geq \frac{n\ (n-1)}{n\ (n-1)-1}$	$M_B \geq 1$	$M_B \geq n$

(*Continued*)

TABLE 1—*(Continued)*

Panel B: Examples of efficiency loss under proportional allocation behavior ($n = 3$, $M_B = 1.5$)

	Example 1			Bonus		Example 2			Bonus
Employee	A	B	C	Pool	Employee	A	B	C	Pool
Effort	40	0	0	60	Effort	40	20	0	90
(%)	100%	0%	0%		(%)	67%	33%	0%	
Proposal	to A	to B	to C		Proposal	to A	to B	to C	
by A	–	30	30		by A	–	90	0	
by B	60	–	0		by B	90	–	0	
by C	60	0	–		by C	60	30	–	
Bonus	40.0	10.0	10.0		Bonus	50.0	40.0	0.0	
(%)	66.7%	16.7%	16.7%		(%)	55.6%	44.4%	0.0%	

	Example 3			Bonus		Example 4			Bonus
Employee	A	B	C	Pool	Employee	A	B	C	Pool
Effort	40	20	20	120	Effort	40	40	20	150
(%)	50%	25%	25%		(%)	40%	40%	20%	
Proposal	to A	to B	to C		Proposal	to A	to B	to C	
by A	–	60	60		by A	–	100	50	
by B	80	–	40		by B	100	–	50	
by C	80	40	–		by C	75	75	–	
Bonus	53.3	33.3	33.3		Bonus	58.3	58.3	33.3	
(%)	44.4%	27.8%	27.8%		(%)	38.9%	38.9%	22.2%	

Panel A illustrates, for the formulaic use and self-excluding design (condition [2]), four allocation behaviors employees may exhibit in stage 2 (i.e., egalitarian, proportional, winner-takes-all, and random) to resolve indifference about their allocation proposals and, assuming ex-ante a given allocation behavior, the requirement for multiplier M_B to *induce* full effort (in stage 1) as unique Nash equilibrium, respectively support full effort (in stage 1) as Nash equilibrium.

Panel B illustrates, for formulaic use and self-excluding design (condition [2]), examples of the efficiency loss under (B) proportional allocation behavior (i.e., non-distorted proposals). Notably, bonuses are transitive relative to efforts, and the bonus of employee A, who always contributes 40, increases, as the other two employees contribute more.

2.3.2. Conditions [3] and [4]: Discretionary Use (Allocation by the Manager). Under discretionary use, the manager is indifferent about the bonus allocation (in stage 3), such that any allocation prescribes a subgame-perfect, dominant-strategy Nash equilibrium in stage 3. Thus, the requirement for multiplier M_B to *induce* (*support*) full effort also depends on the manager's allocation behavior under indifference. If employees assume ex ante the manager resolves indifference via a random allocation, they can expect a final bonus of $b_i = B/n$, resulting in the same requirement for M_B in conditions [3] and [4] as in condition [1].

In contrast, if employees assume ex-ante that the manager takes the average of all proposals to derive final bonuses, discretionary use will perform equally well as formulaic use, resulting in the same requirement for multiplier M_B in condition [3] ([4]) as in condition [1] ([2]). In what follows, we illustrate that—relative to formulaic use—discretionary use has both upside and downside potential, depending on the manager's allocation behavior under indifference.

For example, condition [3] has upside potential relative to condition [1], since the manager can treat self-including proposals as if they were self-excluding. Specifically, the manager can ignore employees' self-allocations and proportionally transform allocations to others such that they add up to the entire pool. If employees assume that the manager takes the average of those transformed proposals to derive final bonuses, they will have no reason to distort self-including proposals, resulting in the same requirement for M_B in condition [3] as in condition [2].

In contrast, condition [4] has downside potential relative to condition [2] in that the manager can treat self-excluding proposals as if they were self-including (to eliminate the efficiency loss). Specifically, the manager can derive fair bonuses—i.e., non-distorted, self-including proposals—from any set of two non-distorted, self-excluding proposals, using the relative ratios of proposed bonuses in those proposals. Yet, in this case, employees have incentives to distort self-excluding proposals in the first place, as they no longer allocate an ex-ante fixed share of the bonus pool, resulting in the same requirement for M_B in condition [4] as in condition [1].[14]

creases with increasing contributions of others. (Note that, if at least two employees contribute some effort, the efficiency loss only materializes among those employees, whereas employees who do not contribute any effort receive a fair bonus of zero.) As such, the requirement for M_B to *support* full effort—i.e., prevent unilateral deviations from full effort—is weaker than the requirement to *induce* full effort. Moreover, the efficiency loss, $n/(n-1)$, decreases, as the number, n, of employees increases.

[14] In essence, employees want to avoid that their own proposals allocate large shares of the pool to others and, thus, distort their proposals to preserve more of the pool for themselves. The optimal distortion strategy depends on which set of two proposals the manager uses to derive final bonuses. Overall, there are $n\ (n-1)/2$ possible combinations that will result in the same solution of fair bonuses, if all n proposals are non-distorted. Yet, if at least one proposal is distorted, the $n(n-1)/2$ solutions may differ. If the manager takes all proposals

Therefore, the manager's allocation behavior under indifference determines—irrespective of the design—whether employees have incentives to distort their proposals. We are not aware of a stage 3 mechanism that induces truth-telling (i.e., non-distorted proposals) in stage 2, such that full effort is induced in stage 1 without efficiency loss, thereby outperforming condition [2].[15] Recall that a key feature of bonus pool arrangements is the self-funding nature, such that burning of the pool as a threat or use of the pool for costly verification are not feasible options. In sum, the predictions of the analytical framework greatly vary with small changes in the assumptions about individuals' allocation behavior under indifference, illustrating the need for a hypothesis drawing on behavioral research and an empirical examination via an experiment.

2.4 HYPOTHESIS DEVELOPMENT

We argue that employees are more likely to provide effort, the more they expect bonuses to be fair. Our line of reasoning builds on the idea of payoff-based conditional cooperation (Yang et al. [2018]), suggesting that the more employees' bonuses reflect their relative contributions to the pool, the greater the effort they will contribute. Prior research shows individuals are often willing to cooperate with others to pursue desirable group outcomes, even if they are better off defecting.[16] We expect the effectiveness of the different uses and designs of allocation proposals (in motivating effort) to depend on the extent to which employees believe those uses and designs prevent defectors from taking advantage of coop-

---

into consideration (e.g., by taking the average of all solutions), employees will benefit from making proposals uninformative (e.g., even splits).

[15] Notably, it is possible to construct a stage 3 mechanism that *supports* truth-telling in stage 2—assuming all other employees tell the truth—such that full effort is induced in stage 1 without efficiency loss (e.g., Moore and Repullo [1988], Battaglini [2002], Ambrus and Takahashi [2008]). For example, consider a bonus allocation in stage 3 based on self-including, "agreeing" proposals in stage 2. That is, if two or more employees propose the same allocation, the manager will allocate bonuses based on those "agreeing" proposals. If all employees tell the truth (i.e., proposals are non-distorted), no employee will have incentives to unilaterally deviate from truth-telling, because the manager will ignore the 'disagreeing' proposal. Thus, a mechanism exists that supports truth-telling in stage 2, such that full effort is induced in stage 1 without efficiency loss. However, under this mechanism, the stage 2 proposals of others matter and, thus, truth-telling is not induced—or part of a subgame-perfect, *dominant-strategy* Nash equilibrium in stage 2 (as is the case in condition [2]). In fact, under this stage 3 mechanism, which is based on 'agreement,' any proposal in stage 2 will be supported, if all employees coordinate (i.e., agree) on that proposal. Such coordination is complex, since each employee is interested in coordinating on a proposal that allocates more of the pool to themself.

[16] Thöni and Volk [2018] report that 61% of players in PGGs are conditional cooperators and only 19% are free riders, who always defect. Equity theory states that individuals care about the fair distribution of rewards (e.g., Adams [1963, 1965]). Similarly, a large body of work documents and formalizes non-standard preferences, like preferences for efficiency, fairness, and truth-telling (e.g., Rabin [1993], Fehr and Schmidt [1999], Bolton and Ockenfels [2000, 2006], Fehr and Gächter [2000a], Charness and Rabin [2002], Engelmann and Strobel [2004], Fehr, Naef, and Schmidt [2006], Abeler, Nosenzo, and Raymond [2019]).

erators, thereby ensuring fair bonuses. The fairness of bonuses depends on (1) employees' proposals and (2) the transformation of these proposals into final bonuses.

First, irrespective of specific use, we argue that employees expect the self-excluding design to result in less distorted proposals than the self-including design. We suggest that employees see opportunities to benefit from distorting self-including proposals but not self-excluding proposals, regardless of use. Drawing on bounded rationality (e.g., Rubinstein [1998]), we argue employees neglect possibilities—that arise under discretionary use but not under formulaic use—to benefit from either not distorting self-including proposals (in condition [3]) or distorting self-excluding proposals (in condition [4]). Thus, we predict that, irrespective of use, the self-excluding design leads to less distortion in proposals than the self-including design.

Second, we argue that—relative to formulaic use—managers' transformations of proposals into final bonuses are more prone to bias when proposals are self-excluding than self-including. We suggest that, under discretionary use, it is more complex for managers to derive fair bonuses from self-excluding proposals than from self-including proposals, even if they are non-distorted. Consider a pool of 180 that is allocated among employees A, B, and C, and assume that the following allocation is fair: 30/60/90. If all employees provide non-distorted, *self-including* proposals, each proposal reflects that allocation. Yet, if all employees provide non-distorted, *self-excluding* proposals, the proposals will differ: A: 0/72/108; B: 45/0/135; C: 60/120/0. While the manager can derive the fair allocation from any set of two self-excluding proposals via the ratios of proposed bonuses, it is more complex than with self-including proposals.[17] Bol [2011] finds that complexity exacerbates managerial biases like centrality bias—i.e., assessing low (high) performers better (worse) than justified. Thus, we predict that unintentional managerial biases in final bonuses are more pronounced when proposals are self-excluding than self-including.

In sum, we expect (1) that the self-excluding design results in less distorted proposals than the self-including design, irrespective of use, and (2) that, relative to formulaic use, discretionary use is more prone to managerial biases—in the transformation of proposals into final bonuses—when the design is self-excluding than self-including. Overall, we argue that the positive effect of the self-excluding design, relative to the self-including design, in establishing fair bonuses and, thus, motivating effort is weaker under discretionary use (i.e., allocation by the manager) than under for-

[17] In the example, employee A proposes a bonus for B that is two thirds of C's bonus, and employee B proposes a bonus for A that is one third of C's bonus. Given that the bonuses must add up to 180 ($= 1/3\ b_C + 2/3\ b_C + b_C$), the fair bonuses can be derived ($b_C = 90$). The complexity underlying the self-excluding design increases, relative to the self-including design, as the number of employees, n, increases. Notably, the efficiency loss that arises from taking the average of non-distorted, self-excluding proposals (35/64/81) is relatively small, given its transitivity.

mulaic use (i.e., allocation by the team).[18] We state the following hypothesis:

H1. *The positive effect of the self-excluding design—relative to the self-including design—on effort provision is weaker under discretionary use (i.e., allocation by the manager) than under formulaic use (i.e., allocation by the team) of allocation proposals.*

3. Method

To test our hypothesis, we conduct a laboratory experiment, which comprises eight rounds, each representing the three-stage game described in the analytical framework (see section 2.3).[19] We assign participants to groups consisting of one manager and three employees (i.e., $n = 3$), which remain grouped together throughout all eight rounds of the experiment.[20] We manipulate – between groups – the design in stage 2 (*self-including* vs. *self-excluding*) and use in stage 3 (*formulaic* vs. *discretionary*) of employees' bonus pool allocation proposals.

3.1 EXPERIMENTAL SETTING AND PARAMETER CHOICES

In stage 1 of every round, each employee receives an endowment of $E = 40$ points, where each point is worth US \$0.05. Employees simultaneously choose costly efforts $e_i \in \{0, 2, 4, \ldots, 40\}$, which must be even numbers to avoid rounding issues. Each point of effort increases group output by $M = 2$ points. The manager receives $\theta = 0.25$ of the output. The remaining 75% forms the bonus pool. That is, each point of effort increases the bonus pool by $M_B = 1.5$ points.[21]

[18] Note that when proposals are self-including, we expect formulaic use and discretionary use to result in similar effort provision. On the one hand, formulaic use may activate a sense of responsibility in employees and, thus, limit tendencies to distort self-including proposals. On the other hand, formulaic use may encourage more distortion in self-including proposals than discretionary use, as each proposal has an equal weight, such that employees may want to protect themselves more against defectors and, hence, distort proposals more heavily in their own favor.

[19] The experiment has been approved by the applicable university's Institutional Review Board. It was programmed using the software Lioness Lab (Giamattei et al. [2020]). We assume a finite time horizon, which is reflective of practice in that the composition of teams often changes over time. Predictions of a framework with one round translate to a framework with a finite number of rounds (based on backward induction).

[20] A team size of three employees is the smallest meaningful size for self-excluding proposals and provides the most stringent test of our hypothesis. Specifically, a smaller team size (1) increases the efficiency loss under formulaic use and self-excluding design, which biases against finding a positive effect of self-excluding proposals under formulaic use, and (2) reduces complexity in the manager's task under discretionary use and self-excluding design, which biases against the expected mitigating influence of discretionary use on the effect of the self-excluding design.

[21] We opt for clarity in the instructions with parameters that are cognitively easy to process. Moreover, a multiplier of $M_B = 1.5$ aligns with prior studies. Arnold, Hannan, and Tafkov

In stage 2 of every round, each employee proposes to the manager how to allocate the pool. Proposed bonuses, $p_{i,j} \geq 0$, must be multiples of three—to avoid rounding issues in stage 3—and add up to the pool, $\sum_j (p_{i,j}) = B$. While the self-including design enables an allocation among all three employees, the self-excluding design requires an allocation among the other two employees such that $p_{i,i} = 0$. We make proposals private in that they are not available to other employees.[22]

In stage 3 of every round, the bonus pool is allocated among the three employees such that individual bonuses, $b_i \geq 0$, add up to the pool, $\sum_i (b_i) = B$. Under formulaic use (i.e., allocation by the team), the final allocation is calculated as the average of the three employees' proposals, $b_i = \sum_k (p_{k,i})/3$. Under discretionary use (i.e., allocation by the manager), the manager has full discretion over the final bonuses and decides how to use the proposals. Thus, in any given round, employee i earns $\Pi_i = 40 - e_i + b_i$, while the manager earns $\Pi_M = 20 + 0.5 \sum_i (e_i)$.

At the end of every round, employees are informed of each other's bonuses and earnings in that round. This design choice aligns with Arnold, Hannan, and Tafkov [2018, 2020], as well as prior work on self-funding reward systems in the common pool literature (e.g., van der Heijden, Potters, and Sefton [2009], Dong, Falvey and Luckraz [2019], Karakostas et al. [2023]).[23]

3.2 PARTICIPANTS AND PROCEDURES

In April 2022, we ran our experimental sessions at a private US university in the Southeast. In total, 200 BBA students participated in the study.[24]

[2018] choose $M_B = 1.8$, while Dong, Falvey, and Luckraz [2019] choose $M_B = 1.2$ and $M_B = 1.8$. Further, it corresponds to a marginal per capita return [MPCR] of 0.5 in PGGs, which aligns with an average MPCR of 0.46 as reported in the meta-analysis by Fiala and Suetens [2017].

[22] This design choice aligns with Arnold, Hannan, and Tafkov [2018, 2020] and reflects common business practice, as peer evaluations are typically not public. Notably, prior work in economics (Dong, Falvey, and Luckraz [2019], Abbink, Dong, and Huang [2022]) makes employees' proposals transparent. Therefore, our study contributes to that literature also by examining the effectiveness of formulaic use and self-excluding design when proposals are private. The (non-)transparency of proposals among employees does not affect the predictions of the analytical framework.

[23] Fiala and Suetens [2017] report that 11% of PGG experiments make payoffs transparent. They find that transparency of payoffs decreases contributions, while transparency of contributions increases contributions, with the former effect outweighing the latter. Transparency of payoffs likely decreases contributions, as players observe free riding to be individually beneficial. It is unclear whether this effect translates to our setting, where free riding (i.e., not providing any effort) may not be beneficial.

[24] In appreciation of students' decision to participate, they received partial course credit. Participation was voluntary, and alternative options to obtain partial credit were offered. Participants have an average age of 20 years, 36% identify as female, and 89% indicate that English is their primary language. Testing for randomization success, we do not find significant differences in age ($F_{\text{Age}\,(3,196)} = 1.60$, $p = 0.191$), gender ($F_{\text{GenderID}\,(3,196)} = 1.16$, $p =$

Each session lasted about 75 minutes with an average payment per participant of US$20.46. We randomly assign participants to groups of four and, within each group, to the manager and employee roles. Further, we randomly assign each group to one of the four experimental conditions (and do so in a balanced way).

As participants enter the laboratory, they select a computer terminal, where all instructions are provided on the computer screen. Throughout the instruction phase, participants complete understanding questions that must be answered correctly (i.e., they have as many attempts as needed), ensuring their understanding. Once all participants within a group have completed the instruction phase, round 1 begins. After the final round, participants answer questions about the task, their decisions, and socio-demographic background.[25] We sum participants' earnings across all rounds, convert them to US dollars, and pay participants in cash, as they exit the laboratory.

4. Results

4.1 DESCRIPTIVE RESULTS AND HYPOTHESIS TEST

Our dependent variable of interest is *Effort*, which ranges from 0 to 40. Table 2 shows average *Effort*—per individual employee and round—for each condition. Under formulaic use (i.e., allocation by the team), average *Effort* is 19.646 for the self-including design and 27.444 for the self-excluding design, whereas, under discretionary use (i.e., allocation by the manager), average *Effort* is 19.744 for the self-including design and 19.667 for the self-excluding design. Figure 1, panel A illustrates the results graphically.

To test our hypothesis, we use regression analysis and cluster standard errors within teams. Specifically, we regress *Effort* on indicator variables for the design (i.e., *Self-excluding*) and use (i.e., *Discretionary*) of the allocation proposals. Table 3 depicts the results. Columns 1, 2, and 3 differ in the level of analysis: (1) decision level (i.e., one observation per employee and round), (2) employee level (i.e., one observation per employee—averaged across all eight rounds), and (3) team level (i.e., one observation per team—averaged across all three employees and all eight rounds). The results are consistent across the three levels.[26] Thus, we initially focus our analysis on the decision level (column 1).

0.326), or primary language ($F_{\text{Language }(3,196)} = 1.13$, $p = 0.336$, all untabulated) among the four experimental conditions.

[25] We ask employees to respond to questions relating to distributive fairness, procedural fairness, their contribution decisions, their allocation proposals, the bonus allocation procedure, their team, and general behaviors and attitudes. We include these questions to provide support for the predicted mechanism underlying our hypothesis. Section A of the online supplement provides more detail about those questions.

[26] Note that coefficients in columns 1, 2a, and 3 are identical, as average *Effort*—per round and employee—does not depend on the level of analysis (i.e., decision, employee, or team).

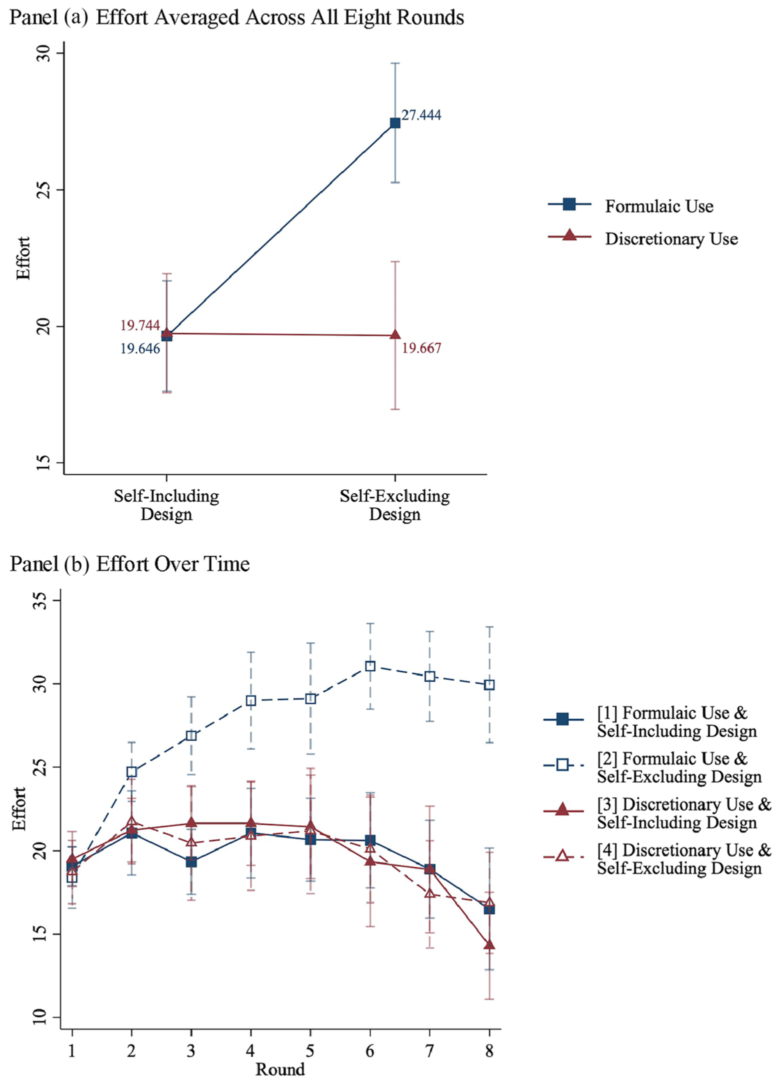

Fig. 1.—Effort. Panel A shows average *Effort*—per round and employee—under formulaic use (i.e., allocation by the team) in blue squares and under discretionary use (i.e., allocation by the manager) in red triangles for both self-including design and self-excluding design. Error bars represent plus/minus one standard error of the mean, clustered within teams. Panel B shows, for each round, average *Effort*—per employee—under formulaic use (i.e., allocation by the team) in blue squares and under discretionary use (i.e., allocation by the manager) in red triangles, separately for both self-including design (filled markers) and self-excluding design (unfilled markers). Error bars represent plus/minus one standard error of the mean, clustered within teams. Figure 2 provides the corresponding descriptive results.

TABLE 2
Effort: Descriptive Results

	Self-Including Design			Self-Excluding Design		
Formulaic use (allocation by the team)	19.646			27.444		
	(10.883) [$n = 288$]	(7.674) [$n = 36$]	(7.026) [$n = 12$]	(11.412) [$n = 288$]	(7.931) [$n = 36$]	(7.549) [$n = 12$]
Discretionary use (allocation by the manager)	19.744			19.667		
	(14.349) [$n = 312$]	(8.644) [$n = 39$]	(7.864) [$n = 13$]	(13.726) [$n = 312$]	(10.487) [$n = 39$]	(9.761) [$n = 13$]

This table shows average *Effort*—per employee and round—under formulaic use (i.e., allocation by the team) and discretionary use (i.e., allocation by the manager) for both the self-including design and the self-excluding design. See the appendix for variable definitions. For each condition, standard deviations are depicted in parentheses at three levels with the respective numbers of observations in brackets (left: decision level with one observation per round and employee; middle: employee level with one observation per employee averaged across all eight rounds; right: team level with one observation per team averaged across all three employees and all eight rounds).

We find that, under formulaic use, the self-excluding design leads to significantly more *Effort* than the self-including design ($\beta_1 = 7.799$, $p = 0.009$).[27] The negative interaction effect ($\beta_3 = -7.876$, $p = 0.041$, one-tailed) supports the hypothesis that the positive effect of the self-excluding design is significantly weaker under discretionary use than under formulaic use. Note that this interaction effect completely offsets the positive effect of the self-excluding design in our setting, such that, under discretionary use, *Effort* does not significantly differ between the self-including and self-excluding designs ($\beta_1 + \beta_3 = -0.077$, $p = 0.982$, untabulated).

The simple comparisons reported in table 3 illustrate that condition [2], formulaic use and self-excluding design, leads to significantly more *Effort* than each of the other three conditions ([1], [3], and [4]). Further, note that *Effort* does not significantly differ among those three conditions ($F_{(2, 49)} = 0.00$, $p = 0.999$, untabulated). In sum, our results support the hypothesis that the effect of the self-excluding design—relative to the self-including

Because we cluster standard errors within teams, *t*-statistics (in parentheses) barely differ among columns 1, 2a, and 3, despite the divergent numbers of observations. Clustering within teams—among the three employees per team across all rounds—accounts for interdependence of observations at the decision and employee levels (columns 1 and 2) and applies robust standard errors at all levels. Given the constraint that effort contributions must be multiples of two (between zero and 40), we also test if *Effort* is normally distributed, using the Shapiro–Wilk test. While the assumption of normality is rejected at the decision level ($p < 0.001$), it is not rejected at the employee level ($p = 0.139$) or team level ($p = 0.741$, all untabulated). Because the coefficients are identical across the three levels, we conclude that the choice to constrain effort contributions to multiples of two does not bias the coefficients. Similarly, the normality assumption for the error term is rejected at the decision level ($p < 0.001$) but not at the employee level ($p = 0.432$) or team level ($p = 0.516$, all untabulated).

[27] All tests are two-tailed, unless they relate to a directional prediction, in which case we report one-tailed tests.

TABLE 3
Effort: OLS Regression Results

	(1) *Effort*	(2a) *Effort*	(2b) *Distributive Fairness*	(2c) *Effort*	(3) *Effort*
Self-excluding (β_1)	7.799***	7.799**	0.780***	3.738	7.799***
	(2.71)	(2.68)	(3.03)	(1.60)	(2.62)
Discretionary (β_2)	0.098	0.098	−0.096	0.597	0.098
	(0.03)	(0.03)	(−0.32)	(0.26)	(0.03)
Self-excluding × *Discretionary* (β_3)	−7.876** †	−7.876** †	−0.902**	−3.175	−7.876** †
	(−1.77)	(−1.76)	(−2.50)	(−0.91)	(−1.72)
Distributive Fairness (β_4)				5.209***	
				(7.47)	
Constant (β_0)	19.646***	19.646***	−0.105	20.194***	19.646***
	(10.00)	(9.91)	(−0.47)	(12.18)	(9.70)
# Observations	1,200	150	150	150	50
# Teams	50	50	50	50	50
F_{Model}	3.32**	3.26**	10.60***	17.01***	3.12**
R^2	0.0636	0.1274	0.1602	0.3742	0.1523
Simple comparisons between conditions: [2] *Formulaic Use & Self-Excluding Design* versus …					
[1] *Formulaic Use & Self-Including Design*	7.799***	7.799**	0.780***	3.738	7.799**
	(2.71)	(2.68)	(3.03)	(1.60)	(2.62)
[3] *Discretionary Use & Self-Including Design*	7.701**	7.701**	0.875***	3.141	7.701**
	(2.57)	(2.55)	(3.80)	(1.33)	(2.50)
[4] *Discretionary Use & Self-Excluding Design*	7.778**	7.778**	0.998***	2.578	7.778**
	(2.31)	(2.29)	(4.96)	(0.95)	(2.24)

This table shows the results of regressing *Effort* (columns 1, 2a and c, and 3) and *Distributive Fairness* (column 2b) on indicator variables capturing our experimental manipulations (i.e., *Self-excluding* and *Discretionary*) and their interaction term. See the appendix for variable definitions. Columns 1, 2 (a, b, c), and 3 differ in the level of analysis, as illustrated by the number of observations (see # Observations): (1) decision level (i.e., one observation per round and employee), (2) employee level (i.e., one observation per employee—averaged across all eight rounds), and (3) team level (i.e., one observation per team—averaged across all three employees and all eight rounds). The table also illustrates simple comparisons between condition [2] (*Formulaic Use & Self-Excluding Design*) and, respectively, condition [1] (*Formulaic Use & Self-Including Design*), condition [3] (*Discretionary Use & Self-Including Design*), and condition [4] (*Discretionary Use & Self-Excluding Design*).

** and *** indicate statistical significance at the 5% and 1% levels, respectively, referring to two-tailed tests, unless marked with a †, in which case they refer to one-tailed tests based on a directional prediction. All *t*-statistics (reported in parentheses) are obtained using standard errors clustered within teams (see # Teams).

design—on effort provision is weaker under discretionary use than under formulaic use.

To better understand the mechanism underlying this finding, we examine responses to four post-experimental questions [PEQs] capturing participants' perceptions of distributive fairness in the allocation of bonuses.[28] Because employees provide one response to each PEQ, we examine *Distributive Fairness* at the employee level. Table 3, column 2b de-

[28] We ask participants to indicate the extent to which they agree, on a Likert-scale from 1 ("*to a small extent*") to 7 ("*to a large extent*"), with the following four statements: (1) "your individual bonuses reflect your contributions," (2) "your individual bonuses [are] appropriate for the

picts the results of regressing *Distributive Fairness* on the indicator variables capturing our manipulations. The pattern of the results mirrors the pattern in column 2a. Further, the simple comparisons show that participants perceive significantly higher levels of *Distributive Fairness* in condition [2], formulaic use and self-excluding design, than in each of the other three conditions ([1], [3], and [4]).

Column 2c illustrates the results of including *Distributive Fairness* as an explanatory variable in the main regression model. We find a significantly positive effect of *Distributive Fairness* on *Effort* ($\beta_4 = 5.209$, $p < 0.001$), while neither the effect of the self-excluding design ($\beta_1 = 3.738$, $p = 0.117$) nor the interaction effect ($\beta_3 = -3.175$, $p = 0.366$) is significant. Similarly, we do not find significant simple comparisons, suggesting that perceptions of *Distributive Fairness* explain the differences in *Effort* among the four conditions.[29] In sum, the results show that condition [2], formulaic use and self-excluding design, is perceived to lead to fairer bonuses than each of the other three conditions, potentially motivating employees to contribute more effort.

4.2 SUPPLEMENTAL ANALYSES

To gain additional insight into the mechanisms underlying our results, we examine in more detail (1) employees' effort contributions in stage 1 (section 4.2.1), (2) their allocation proposals in stage 2 (section 4.2.2), and (3) the final bonus allocations in stage 3 (section 4.2.3).

4.2.1. Employees' Effort Contributions (Stage 1). Figure 1, panel B, illustrates, for each condition, how *Effort* evolves over time. Starting from round 2, the pattern of our main results (see panel A) becomes discernible. In round 1, we do not find significant differences in *Effort* among the four conditions ($p > 0.900$, untabulated). Over time, however, *Effort* significantly increases only in condition [2] ($p = 0.007$, untabulated) but does not significantly change in condition [1], [3], or [4] (all $p > 0.240$, untabulated).[30]

---

outputs you have generated," (3) "your individual bonuses reflect your contributions to team output," and (4) "your individual bonuses [are] justified, given your performance." A factor analysis shows that the four PEQs load on one factor with eigenvalue of 3.210 (untabulated). Based on the loadings of the four PEQs (each > 0.800), we create the factor variable *Distributive Fairness.* Note that this variable can take on 2,401 potential values (= 7 Likert points $^{4\ \mathrm{PEQs}}$), such that we treat *Distributive Fairness* as non-ordinal.

[29] In section A of the online supplement, we report additional results from the post-experimental questionnaire. While several PEQs correlate with *Effort* and/or *Distributive Fairness,* they do not explain the differences in *Effort* among the four conditions. For example, in four additional PEQs, participants also indicate that they perceive more procedural fairness under formulaic use than under discretionary use ($F_{(1,49)} = 6.11$, $p = 0.017$), but these differences in procedural fairness do not seem to explain the differences in *Effort* among the four conditions in our setting.

[30] Note that the increase in *Effort* in condition [2] significantly differs from the change in *Effort* in each of the other three conditions (all $p < 0.020$, untabulated). These results are based on a regression model that includes regressors for our manipulations (i.e., *Self-excluding* and *Discretionary*), time (i.e., *Round*), and all four interaction terms.

Figure 2 shows histograms of employees' effort contributions for each condition and round. In round 1, *Effort* is in each condition concentrated at 20, the center of the possible range, with only a few observations at 0 and 40, the endpoints. Over time, effort contributions in condition [2] steadily shift toward 40, while effort contributions in conditions [1], [3], and [4] initially spread more equally but then shift toward 0.[31] Despite these shifts toward the endpoints, the majority of observations—in each condition and round—do not fall on either endpoint. In sum, our results show that employees are more inclined in condition [2], formulaic use and self-excluding design, to increase their effort contributions over time compared to conditions [1], [3], and [4].

4.2.2. Employees' Allocation Proposals (Stage 2). Our reasoning purports that employees are more motivated to contribute effort the more they expect final bonuses to be fair. Thus, we examine the extent to which final bonuses deviate from fair bonuses and investigate the sources of this deviation. We distinguish three potential sources: (1) distortion in employees' allocation proposals (in stage 2), (2) efficiency loss due to taking the average of self-excluding proposals (in stage 3), and (3) managerial bias under discretionary use (in stage 3). Table 4, panel A, shows the potential sources of deviation for each condition. While distortion is a potential source in all four conditions, efficiency loss is a potential source only in conditions [2] and [4], and managerial bias is a potential source only in conditions [3] and [4].

We first present evidence for our general reasoning (without distinguishing the three sources of deviation). Table 4, panel B, depicts, for each condition, the average absolute deviation of the final bonus from the fair bonus (per employee and round).[32] Consistent with our reasoning, this deviation is significantly smaller in condition [2], formulaic use and self-excluding design, than in each of the other three conditions (each $p <$ 0.050, one-tailed, untabulated).[33] Next, we isolate how much of this overall deviation is due to each potential source of deviation.

We focus first on distortion in employees' allocation proposals (stage 2) as a potential source. We predict that—irrespective of use—the self-excluding design leads to less distortion in the proposals than the self-including design. To test this prediction, we first create a measure that

[31] In the final round (i.e., round 8), 56% of effort contributions in condition [2] are larger than 30, compared with only 14%, 18%, and 18% in conditions [1], [3], and [4] respectively; similarly, only 8% of effort contributions in condition [2] are smaller than 10, compared with 33 (46, 38)% in condition [1] ([3], and [4]).

[32] For a given employee i, we derive the absolute deviation of the final bonus from the fair bonus (at the decision level) as $|b_{i,t,r} - 1.5\ e_{i,t,r}|$, where, in each team t and round r, (1) $b_{i,t,r}$ denotes employee i's *Final Bonus* ($i \in \{A, B, C\}$), (2) $e_{i,t,r}$ denotes i's *Effort*, and (3) 1.5 $e_{i,t,r}$ describes i's *Fair Bonus* (i.e., $M_B = 1.5$ points).

[33] Notably, overall deviation does not significantly differ among those three conditions ($p >$ 0.250, untabulated).

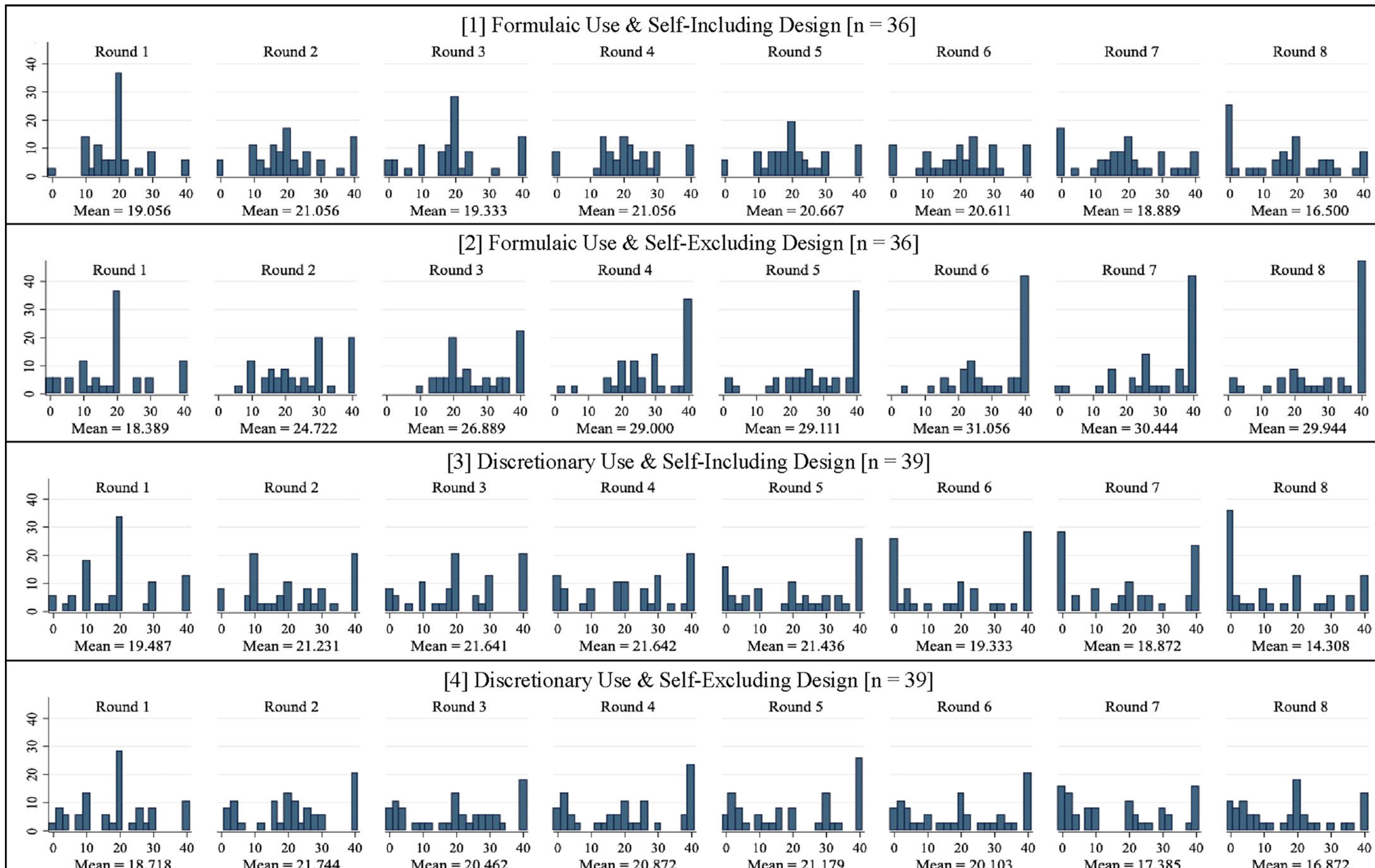

FIG. 2.—Effort Histograms. This figure illustrates, for each condition and round, histograms of *Effort*—i.e., the percentages with which each effort contribution (from 0 to 40) occurs.

TABLE 4
Absolute Deviation of Final Bonus from Fair Bonus

Panel A: Absolute deviation of final bonus from fair bonus—potential sources

	Self-Including Design	Self-Excluding Design
Formulaic use (allocation by the team)	Distortion in proposals (stage 2)	Distortion in proposals (stage 2) & Efficiency loss (stage 3)
Discretionary use (allocation by the manager)	Distortion in proposals (stage 2) & Managerial bias (stage 3)	Distortion in proposals (stage 2) & Efficiency loss (Stage 3) & Managerial bias (stage 3)

Panel B: Absolute deviation of final bonus from fair bonus—descriptive results

	Self-Including Design			Self-Excluding Design		
Formulaic use (allocation by the team)	7.514			4.674		
	(9.073)	(5.223)	(5.034)	(6.630)	(3.417)	(2.950)
	[$n = 288$]	[$n = 36$]	[$n = 12$]	[$n = 288$]	[n = 36]	[$n = 12$]
Discretionary use (allocation by the manager)	10.673			11.276		
	(12.807)	(7.149)	(6.701)	(16.196)	(9.481)	(8.959)
	[$n = 312$]	[$n = 39$]	[n = 13]	[$n = 312$]	[n = 39]	[$n = 13$]

Panel A shows for each of the four conditions the potential sources for the absolute deviation of final bonuses from fair bonuses: (1) distortion in proposals (stage 2), (2) efficiency loss (stage 3), and (3) managerial bias (stage 3).

Panel B shows the average *Absolute Deviation of Final Bonus from Fair Bonus*—per employee and round—under formulaic use (i.e., allocation by the team) and discretionary use (i.e., allocation by the manager) for both the self-including design and the self-excluding design. See the appendix for variable definitions. For each condition, standard deviations are depicted in parentheses at three levels with the respective numbers of observations in brackets (left: decision level with one observation per round and employee; middle: employee level with one observation per employee averaged across all eight rounds; right: team level with one observation per team averaged across all three employees and all eight rounds).

captures in each proposal the average distortion *per* proposed bonus: *Proposal Distortion*.[34] Table 5, panel A, depicts the descriptive results, while table 6, column 1, presents the results of regressing *Proposal Distortion* on

[34] For a given employee i, we define *Proposal Distortion* (at the decision level) as average distortion per proposed bonus in i's proposal and derive distortion in each proposed bonus as the absolute deviation from the corresponding fair bonus. For the self-excluding design, we consider only the two proposed bonuses for others and scale each by the percentage of the pool created by the other employees to ensure comparability with fair bonuses: $\mathit{Proposal\ Distortion}_{i,t,r} = \begin{cases} \sum_j |p_{i,j,t,r} - 1.5\ e_{j,t,r}|/3 & \text{for the self} - \text{including design} \\ \sum_{j \neq i} |p_{i,j,t,r} \cdot \frac{B_{t,r} - 1.5\ e_{i,t,r}}{B_{t,r}} - 1.5\ e_{j,t,r}|/2 & \text{for the self} - \text{excluding design} \end{cases}$, where, in each team t and round r, (1) $p_{i,j,t,r}$ denotes the proposed bonus by employee i for employee j ($i, j \in \{A, B, C\}$), (2) $e_{j,t,r}$ denotes j's *Effort*, (3) 1.5 $e_{j,t,r}$ describes j's *Fair Bonus*, and (4) $B_{t,r} = \sum_j (1.5\ e_{j,t,r})$ describes the bonus pool. For the self-excluding design, we scale each proposed

TABLE 5

Proposal Distortion, Distortion, and Distortion and Efficiency: Descriptive Results

Panel A: Proposal distortion: descriptive results

	Self-Including Design			Self-Excluding Design		
Formulaic use (allocation by the team)		14.208			4.795	
	(17.914) [$n = 288$]	(13.632) [$n = 36$]	(13.271) [$n = 12$]	(8.690) [$n = 288$]	(4.327) [$n = 36$]	(3.438) [$n = 12$]
Discretionary use (allocation by the manager)		11.186			5.824	
	(13.395) [$n = 312$]	(9.202) [$n = 39$]	(6.564) [$n = 13$]	(9.645) [$n = 312$]	(4.903) [$n = 39$]	(3.655) [$n = 13$]

Panel B: Distortion: Descriptive results

	Self-Including Design			Self-Excluding Design		
Formulaic use (allocation by the team)		7.514			3.890	
	(9.073) [$n = 288$]	(5.223) [$n = 36$]	(5.034) [$n = 12$]	(6.120) [$n = 288$]	(3.226) [$n = 36$]	(2.879) [$n = 12$]
Discretionary use (allocation by the manager)		9.147			4.763	
	(9.363) [$n = 312$]	(5.342) [$n = 39$]	(4.992) [$n = 13$]	(6.575) [n = 312]	(3.437) [$n = 39$]	(3.039) [$n = 13$]

Panel C: Distortion & Efficiency Loss: — descriptive results

	Self-Including Design			Self-Excluding Design		
Formulaic use (allocation by the team)		7.514			4.674	
	(9.073) [$n = 288$]	(5.223) [$n = 36$]	(5.034) [$n = 12$]	(6.630) [$n = 288$]	(3.417) [$n = 36$]	(2.950) [$n = 12$]
Discretionary use (allocation by the manager)		9.147			6.013	
	(9.363) [$n = 312$]	(5.342) [$n = 39$]	(4.992) [$n = 13$]	(6.914) [$n = 312$]	(3.184) [$n = 39$]	(2.660) [$n = 13$]

This table shows average *Proposal Distortion* (in panel A), average *Distortion* (in panel B), and average *Distortion & Efficiency Loss* (in panel C)—per employee and round—under formulaic use (i.e., allocation by the team) and discretionary use (i.e., allocation by the manager) for both the self-including design and the self-excluding design. See the appendix for variable definitions. For each condition, standard deviations are depicted in parentheses at three levels with the respective numbers of observations in brackets (left: decision level with one observation per round and employee; middle: employee level with one observation per employee averaged across all eight rounds; right: team level with one observation per team averaged across all three employees and all eight rounds).

the experimental variables.[35] In line with our prediction, we find that *Proposal Distortion* is significantly smaller for the self-excluding design than

bonus by the factor $\frac{B_{t,r}-1.5\ e_{i,t,r}}{B_{t,r}}$ such that, absent any distortion in the proposed bonus, we obtain the corresponding fair bonus.

[35] For ease of presentation, Table 6 focuses on the employee level. Again, the regression coefficients are identical to those at the decision and team levels, and t-statistics barely differ across levels due to clustering of standard errors within teams. Yet, because the distortion

TABLE 6

Proposal Distortion, Distortion, and Distortion and Efficiency: Regression Results

	(1) *Proposal Distortion*	(2) *Distortion*	(3) *Distortion & Efficiency Loss*
Self-excluding (β_1)	$-9.413^{***\,\dagger}$	$-3.624^{**\,\dagger}$	$-2.840^{**\,\dagger}$
	(−2.43)	(−2.22)	(−1.73)
Discretionary (β_2)	−3.022	1.634	1.634
	(−0.73)	(0.83)	(0.83)
Self-excluding × *Discretionary* (β_3)	4.051	−0.760	−0.294
	(0.93)	(−0.33)	(−0.13)
Constant (β_0)	14.208^{***}	7.514^{***}	7.514^{***}
	(3.80)	(5.29)	(5.29)
# Observations	150	150	150
# Teams	50	50	50
F_{Model}	4.86^{***}	4.63^{***}	2.99^{**}
R^2	0.1626	0.1902	0.1290
Simple Effects of …			
Self-excluding when *Discretionary* = 1	$-5.362^{***\,\dagger}$	$-4.384^{***\,\dagger}$	$-3.135^{**\,\dagger}$
($\beta_1 + \beta_3$)	(−2.62)	(−2.76)	(−2.04)
Discretionary when *Self-Excluding* = 1	1.029	0.873	1.339
($\beta_2 + \beta_3$)	(0.74)	(0.75)	(1.21)

This table shows the results of regressing *Proposal Distortion* (column 1), *Distortion* (column 2), and *Distortion & Efficiency Loss* (column 3)—at the employee level (i.e., one observation per employee)—on indicator variables capturing our experimental manipulations (i.e., *Self-excluding* and *Discretionary*) and their interaction term. See the appendix for variable definitions. The table also shows the simple effects of the indicator variables *Self-excluding* and *Discretionary* when the other variable equals 1.

** and *** indicate statistical significance at the 5% and 1% levels, respectively, referring to two-tailed tests, unless marked with a †, in which case they refer to one-tailed tests based on a directional prediction. All *t*-statistics (reported in parentheses) are obtained using standard errors clustered within teams (see # Teams).

for the self-including design (under both uses, $p < 0.010$, one-tailed). Similarly, proportional allocation behavior is significantly more frequent for the self-excluding design (53%) than for the self-including design (35%, $p = 0.001$, one-tailed, untabulated).[36] When the design is self-including, we also find evidence for opportunistic allocation behavior, as employees

measures in Table 6 are non-negative with several observations at zero, we must reject at all three levels the assumption of normality (untabulated). We confirm the results, directionally and inferentially, with a logarithmic transformation of the distortion measures, for which the normality assumption is not rejected at the employee or team levels (untabulated). Further, we confirm the results at the team level, directionally and inferentially, with Wilcoxon rank-sum tests (untabulated). We do the same for the results reported in Table 8.

[36] We consider allocation behavior to be proportional, if *Proposal Distortion* is weakly smaller than two, as (1) zero distortion is often not possible for the self-excluding design—given the requirement to propose bonuses in multiples of three—and (2) two being the smallest, nonzero *Proposal Distortion* for the self-including design.

propose to allocate significantly more to themselves than justified ($p < 0.001$, one-tailed, untabulated).[37]

To understand the extent to which distortion in the proposals contributes to the deviation of final bonuses from fair bonuses, we create a second measure: *Distortion*. This measure captures the deviation of the average proposed bonus for a given employee—with adjustment via scaling for the self-excluding design—from that employee's fair bonus.[38] For the self-excluding design, we make an explicit adjustment to isolate deviation due to distortion in the proposals (in stage 2) from deviation due to efficiency loss (in stage 3). Table 5, panel B shows the descriptive results. Note that in condition [1], formulaic use and self-including design, *Distortion* is identical to the deviation of final bonuses from fair bonuses (in table 4, panel B), as distortion in the proposals is the only potential source of deviation (as illustrated in table 4, panel A).

Table 6, column 2 shows the results of regressing *Distortion* on indicator variables capturing our manipulations. Consistent with our prediction, the self-excluding design leads to significantly less *Distortion* (in bonuses) than the self-including design both under formulaic use ($\beta_1 = -3.624$) and under discretionary use ($\beta_1 + \beta_3 = -4.384$, both $p < 0.020$, one-tailed). Overall, our results provide evidence that, irrespective of use, employees' allocation proposals are less distorted for the self-excluding design than for the self-including design.

4.2.3. Final Bonus Allocations (Stage 3). Next, we examine the deviation of final bonuses from fair bonuses due to the two potential sources of deviation in stage 3, namely, efficiency loss and managerial bias. First, we focus on the efficiency loss for the self-excluding design, which arises from taking the simple average of all three proposed bonuses for a given employee—including the self-proposed bonus of zero—to derive final bonuses. While

---

[37] Of all 600 self-including proposals, 352 proposals (59%) exhibit opportunistic allocation behavior in that self-proposed bonuses are larger than justified by effort contributions (including 74 proposals with self-proposed bonuses comprising the entire bonus pool), 52 proposals (9%) exhibit altruistic allocation behavior (i.e., self-proposed bonuses are smaller than justified), while the remaining 196 proposals (33%) exhibit non-distorted, self-proposed bonuses (24 of which are, by default, non-distorted, because the bonus pool is zero).

[38] For a given employee i, we define *Distortion* (at the decision level) as the absolute deviation of the *adjusted* average proposed bonus for i from i's fair bonus. For the self-including design, the adjusted average proposed bonus is equal to the average proposed bonus, while, for the self-excluding design, the adjusted average proposed bonus is derived by (1) scaling each proposed bonus by the percentage of the pool created by i and the other non-proposing employee and (2) taking the average of the *two* scaled proposed bonuses, ensuring comparability with i's fair bonus: $Distortion_{i,t,r} = \begin{cases} \left|\sum_k (p_{k,i,t,r})/3 - 1.5\ e_{i,t,r}\right| & \text{for the self} - \text{including design} \\ \left|\sum_{k \neq i} (p_{k,i,t,r} \cdot \frac{B_{t,r} - 1.5\ e_{k,t,r}}{B_{t,r}})/2 - 1.5\ e_{i,t,r}\right| & \text{for the self} - \text{excluding design} \end{cases}$, where, in each team t and round r, (1) $p_{k,i,t,r}$ denotes the proposed bonus by employee k for employee i ($i,k \in \{A, B, C\}$), (2) $e_{i,t,r}$ denotes i's *Effort*, (3) $1.5\ e_{i,t,r}$ describes i's *Fair Bonus*, and (4) $B_{t,r} = \sum_k (1.5\ e_{k,t,r})$ describes the bonus pool. For the self-excluding design, scaling ensures that, absent distortion in the proposed bonus, we obtain the fair bonus.

the efficiency loss dampens the benefits of the self-excluding design relative to the self-including design, we do not expect them to be eliminated, given its properties (i.e., transitivity). To test this expectation, we create a measure, *Distortion & Efficiency Loss*, that captures the absolute deviation of the average proposed bonus—without any adjustment—from the fair bonus.[39] Table 5, panel C shows the descriptive results. Note that, under formulaic use, this measure is identical to the deviation of final bonuses from fair bonuses in table 4, panel B, as distortion and efficiency loss are the only two sources of deviation in conditions [1] and [2].

Table 6, column 3 shows the results of regressing *Distortion & Efficiency Loss* on indicator variables capturing our manipulations. Consistent with our expectation, the self-excluding design leads to less *Distortion & Efficiency Loss* than the self-including design both under formulaic use ($\beta_1 = -2.840$) and under discretionary use ($\beta_1 + \beta_3 = -3.135$, both $p < 0.050$, one-tailed). Thus, we provide evidence that the deviation of final bonuses from fair bonuses due to *both* distortion in proposals (stage 2) *and* efficiency loss (stage 3) is smaller for the self-excluding design than for the self-including design. In other words, the efficiency loss does not eliminate the benefits of the self-excluding design. Table 7, panel A illustrates *Efficiency Loss* isolated from *Distortion.*[40]

Next, we focus on managerial bias under discretionary use as a potential source of deviation of final bonuses from fair bonuses. Recall that we predict managerial bias to be more pronounced for the self-excluding design than for the self-including design. To test this prediction, we derive *Managerial Bias* as the difference between (1) the absolute deviation of the final bonus from the fair bonus (see table 4, panel B) and (2) *Distortion & Efficiency Loss* (see table 5, panel C). That is, *Managerial Bias* captures any deviation of the final bonus from the fair bonus beyond the deviation that arises when final bonuses are determined based on a formulaic use.[41] Table 7, panel B shows the descriptive results. Under formulaic use, *Managerial Bias* is, by default, zero. Under discretionary use, *Managerial Bias* is positive, showing that, relative to average proposed bonuses, managers' bonus allocations do not reduce the deviation from

[39] For a given employee i, we derive *Distortion & Efficiency Loss* (at the decision level) as $|\sum_k (p_{k,i,t,r})/3 - 1.5\ e_{i,t,r}|$ where, in each team t and round r, (1) $p_{k,i,t,r}$ denotes the proposed bonus by employee k for employee i ($i,k \in \{A, B, C\}$), (2) $\sum_k (p_{k,i,t,r})/3$ is the *Average Proposed Bonus* for employee i, (3) $e_{i,t,r}$ denotes i's *Effort*, and (4) 1.5 $e_{i,t,r}$ denotes i's *Fair Bonus.* We capture efficiency loss also under discretionary use, as if managers derived bonuses formulaically.

[40] Note that for the self-including design, *Efficiency Loss* is, by default, zero. Focusing on the self-excluding design. Table 8, column 1 shows *Efficiency Loss* significantly differs from zero under both formulaic use ($\beta_1 = 0.784$) and discretionary use ($\beta_1 + \beta_3 = 1.250$, both $p < 0.010$, one-tailed), without significant difference between the two uses.

[41] Put differently, *Managerial Bias* in condition [4] does not include the efficiency loss, although managers could avoid the efficiency loss. Thus, we provide a conservative test for the predicted mechanism, which we confirm with the measure *Efficiency Loss & Managerial Bias* (see table 7, panel C and table 8, column 3).

TABLE 7

Efficiency Loss and Managerial Bias: Descriptives Results

	Self-Including Design			Self-Excluding Design		
Panel A: Efficiency loss: Descriptive results						
Formulaic use (allocation by the team)		0.000			0.784	
	(0.000)	(0.000)	(0.000)	(2.276)	(0.958)	(0.798)
	[$n = 288$]	[$n = 36$]	[$n = 12$]	[$n = 288$]	[$n = 36$]	[$n = 12$]
Discretionary use (allocation by the manager)		0.000			1.250	
	(0.000)	(0.000)	(0.000)	(3.325)	(1.770)	(1.545)
	[n = 312]	[$n = 39$]	[n = 13]	[$n = 312$]	[$n = 39$]	[$n = 13$]
Panel B: *Managerial Bias*: Descriptive results						
Formulaic use (allocation by the team)		0.000			0.000	
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
	[$n = 288$]	[$n = 36$]	[$n = 12$]	[$n = 288$]	[$n = 36$]	[$n = 12$]
Discretionary use (allocation by the manager)		1.526			5.263	
	(11.743)	(5.616)	(4.370)	(13.140)	(7.190)	(6.705)
	[$n = 312$]	[$n = 39$]	[$n = 13$]	[$n = 312$]	[$n = 39$]	[$n = 13$]
Panel C: *Efficiency Loss & Managerial Bias*: Descriptive results						
Formulaic use (allocation by the team)		0.000			0.784	
	(0.000)	(0.000)	(0.000)	(2.276)	(0.958)	(0.798)
	[$n = 288$]	[$n = 36$]	[$n = 12$]	[$n = 288$]	[$n = 36$]	[$n = 12$]
Discretionary use (allocation by the manager)		1.526			6.513	
	(11.743)	(5.616)	(4.370)	(13.634)	(7.298)	(6.819)
	[$n = 312$]	[$n = 39$]	[$n = 13$]	[$n = 312$]	[$n = 39$]	[$n = 13$]

This table shows average *Efficiency Loss* (in panel A), average *Managerial Bias* (in panel B), and average *Efficiency Loss & Managerial Bias* (in panel C)—per employee and round—under formulaic use (i.e., allocation by the team) and discretionary use (i.e., allocation by the manager) for both the self-including design and the self-excluding design. See the appendix for variable definitions. For each condition, standard deviations are depicted in parentheses at three levels with the respective numbers of observations in brackets (left: decision level with one observation per round and employee; middle: employee level with one observation per employee averaged across all eight rounds; right: team level with one observation per team averaged across all three employees and all eight rounds).

fair bonuses.[42] Focusing on discretionary use only, table 8, column 2 depicts the results of regressing *Managerial Bias* (at the employee level)

[42] Of all 624 bonuses under discretionary use, 154 bonuses (25%) are identical to the average proposed bonus. Importantly, when proposals are self-including, managers' bonus allocations are similarly often harmful (44%: *Managerial Bias* is positive) as beneficial (41%: *Managerial Bias* is negative); whereas, when proposals are self-excluding, managers' bonus allocations are more frequently harmful (48%) than beneficial (18%).

TABLE 8
Efficiency Loss and Managerial Bias: Regression Results

	(1) *Efficiency Loss*	(2) *Managerial Bias*	(3) *Efficiency Loss & Managerial Bias*
Self-excluding (β_1)	0.784***		0.784***
	(3.46)		(3.47)
Discretionary (β_2)		1.526	1.526
		(1.28)	(1.28)
Self-excluding × *Discretionary* (β_3)	0.466	3.737* †	4.203** †
	(0.97)	(1.71)	(1.89)
# Observations	75	78	114
# Teams	25	26	38
F_{Model}	10.33***	4.93**	8.66***
R^2	0.3546	0.2706	0.3519
Simple Effects of …			
Self-excluding when *Discretionary* = 1 ($\beta_1 + \beta_3$)	1.250***		4.987** †
	(2.95)		(2.26)
Discretionary when *Self-Excluding* = 1 ($\beta_2 + \beta_3$)		5.263***	5.729*** †
		(2.87)	(3.06)

This table shows the results of regressing *Efficiency Loss* (column 1), *Managerial Bias* (column 2), and *Efficiency Loss & Managerial Bias* (column 3) – at the employee level (i.e., one observation per employee) – on indicator variables capturing our experimental manipulations (i.e., *Self-excluding* and *Discretionary*) and their interaction term. See the appendix for variable definitions. Column 1 focuses on the self-excluding design only, column 2 focuses on discretionary use only, and column 3 focuses on the self-excluding design and/or discretionary use. The table also shows the simple effects of the indicator variables *Self-excluding* and *Discretionary* when the other variable equals 1.

*, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively, referring to two-tailed tests, unless marked with a †, in which case they refer to one-tailed tests based on a directional prediction. All *t*-statistics (reported in parentheses) are obtained using standard errors clustered within teams (see # Teams).

on indicator variables capturing our manipulations, without including a constant term.[43] Consistent with our prediction, *Managerial Bias* is larger for the self-excluding design than for the self-including design (β_3 = 3.737, p = 0.050, one-tailed). Additional analyses show that managers' bonus allocations exacerbate cross-subsidization only for the self-excluding design but not for the self-including design. Specifically, managers allocate less (more) to high (low) contributors than justified by average proposed bonuses, when the design is self-excluding (both $p < 0.050$, one-tailed) but not when it is self-including (both $p > 0.500$, untabulated).

In sum, the supplemental analyses support the mechanism underlying our hypothesis. That is, (1) the self-excluding design leads, irrespective of use, to less distortion in employees' allocation proposals than the self-

[43] We do not include a constant term to facilitate comparisons across columns in table 8. Absent the constant term, $\beta_2 = 1.526$ displays average *Managerial Bias* in condition [3], discretionary use and self-including design.

including design (withstanding efficiency loss), and (2) managers' bonus allocations counteract this effect under discretionary use, introducing more bias when proposals are self-excluding than self-including. Section B of the online supplement provides additional process evidence by comparing the relations between (1) effort and (adjusted) average proposed bonus and (2) effort and final bonus within and across conditions.

5. *Conclusions*

This study examines how the design and use of peer evaluations for incentive compensation influences employees' motivation to contribute effort to group outcomes. Focusing on a mutual monitoring setting, we study peer evaluations in the form of bonus allocation proposals and compare formulaic use against discretionary use and self-including design against self-excluding design. These different uses and designs have been examined in the bonus pool literature in accounting (e.g., Arnold, Hannan, and Tafkov [2018, 2020]) and the common pool literature in economics (e.g., Dong, Falvey, and Luckraz [2019], Abbink, Dong, and Huang [2022]). Yet, we are not aware of research comparing their effectiveness in motivating employee effort. Our study fills this gap, thereby speaking to both literatures.

Consistent with our hypothesis, we find that the self-excluding design leads to more effort than the self-including design under formulaic use (i.e., allocation by the team) but not under discretionary use (i.e., allocation by the manager). We also provide evidence for the mechanism underlying this finding. First, the self-excluding design results in less distortion in employees' proposals than the self-including design, irrespective of use, suggesting that employees neglect complex possibilities of how managers could use discretion to derive fair bonuses. Second, this benefit of the self-excluding design is counteracted under discretionary use, as managerial biases enter the transformation of proposals into final bonuses more when proposals are self-excluding than when they are self-including. Overall, final bonuses best reflect relative contributions to the pool under formulaic use and self-excluding design. Thus, the combination of formulaic use and self-excluding design provides employees with the strongest incentives to contribute effort.

Our findings show that the effective use and design of subjective peer evaluations can help inform bonus allocations and, therefore, enable organizations to benefit from mutual monitoring. The results imply that organizations may want to (1) limit self-interested behaviors by separating employees' self-evaluations from assessments of their peers and (2) also limit managerial biases by reducing managers' discretion in the allocation of rewards. As such, our study contributes to research on calibration committees (Demeré, Sedatole, and Woods [2019], Grabner, Künneke, and Moers [2020]). Our study also adds to prior work on bonus pools (Baiman and Rajan [1995], Rajan and Reichelstein [2006, 2009]), in that formulaic use of self-excluding proposals extends the solution of the moral hazard problem between managers and employees to the newly arising moral hazard problem among the employees. Importantly, our study provides a theoretical and empirical underpinning for the increasingly popular practice of peer-to-peer bonus schemes. These schemes not only promote employee involvement within organizations but, when used and designed effectively, can also play a critical role in shaping organizational culture.

In a broader context, our study addresses the challenge of designing institutions that evoke high contributions (i.e., effort) to shared outcomes. In such settings (e.g., the PGG), free riding is a notorious problem, as it typically leads to a collapse of contributions. Prior research has studied institutions that can alleviate the free-rider problem by providing individuals with opportunities to reward and/or penalize each other. Our study focuses on institutions that inform the allocation of shared outcomes among individuals (here, bonus pools) via peer evaluations. We demonstrate that the design and use of peer feedback are critical in motivating effort contributions, as they affect the distortion in peer evaluations and fairness of bonus allocations. Thus, our research question can be interpreted more broadly in that we study which institutional design in the form of the design and use of peer evaluations (1) best motivates effort, (2) reduces distortion in peer evaluations, and (3) improves the fairness of bonus allocations. Overall, we create a better understanding of how the design and use of peer evaluations for compensation purposes can enable cooperation within organizations.

APPENDIX

Variable Definitions

Variable	Definition
Main variables of interest	
Effort	Effort contribution (per employee and round), ranging from 0 to 40.
	At the decision level, *Effort* captures the effort contribution of a given employee i in team t and round r:
	$Effort_{i,t,r} = e_{i,t,r}$, where $e_{i,t,r} \in \{0,\ 2,\ 4, \ldots, 40\}$.
	At the employee level, *Effort* is the average effort contribution of a given employee i in team t (across all eight rounds):
	$Effort_{i,t} = \sum_r (Effort_{i,t,r})\ /\ 8$.
	At the team level, *Effort* is the average effort contribution of a given team t (across all three employees and all eight rounds):
	$Effort_t = \sum_i (Effort_{i,t})\ /\ 3$.

Variable	Definition
Distributive Fairness	Extent to which a given employee i perceives bonuses to be fair based on responses to the following four PEQs, indicating their agreement from 1 ("*to a small extent*") to 7 ("*to a large extent*"): (1) do your individual bonuses reflect your contributions, (2) are your individual bonuses appropriate for the outputs you have generated, (3) do your individual bonuses reflect your contributions to team output, and (4) are your individual bonuses justified, given your performance? The factor variable is calculated (at the employee level) based on the loadings of the PEQs derived from a factor analysis.
Experimental manipulation variables	
Self-excluding	Indicator variable that equals 0 (1) when the design of employees' allocation proposals is self-including (self-excluding).
Discretionary	Indicator variable that equals 0 (1) when the use of employees' allocation proposals is formulaic (discretionary).
Round	Ordinal variable that equals 0 (1, 2, …, 7) for round 1 (2, 3, …, 8).
Bonus variables	
Final Bonus	Final bonus allocated to a given employee i (per round). At the decision level (for employee i in team t and round r): $\textit{Final Bonus}_{i,t,r} = b_{i,t,r}$. At the employee level (for employee i in team t across all rounds): $\textit{Final Bonus}_{i,t} = \sum_r (\textit{Final Bonus}_{i,t,r})\ /\ 8$.
Fair Bonus	Amount of bonus pool generated by a given employee i (per round). At the decision level (for employee i in team t and round r): $\textit{Fair Bonus}_{i,t,r} = 1.5\ \textit{Effort}_{i,t,r}$. At the employee level (for employee i in team t across all rounds): $\textit{Fair Bonus}_{i,t} = \sum_r (\textit{Fair Bonus}_{i,t,r})\ /\ 8$.
Average Proposed Bonus	Average of proposed bonuses for a given employee i (per round). At the decision level (for employee i in team t and round r): $\textit{Average Proposed Bonus}_{i,t,r} = \sum_k (p_{k,i,t,r})\ /\ 3$, where $p_{k,i,t,r}$ denotes the bonus proposed by employee k for employee i. At the employee level (for employee i in team t across all rounds): $\textit{Average Proposed Bonus}_{i,t} = \sum_r (\textit{Average Proposed Bonus}_{i,t,r})\ /\ 8$.
Adjusted Average Proposed Bonus	Adjusted average of proposed bonuses for a given employee i (per round). For the self-including design, this measure is identical to the *Average Proposed Bonus*. For the self-excluding design, this measure is derived (at the decision level) by scaling each proposed bonus for i—by the percentage of the bonus pool created by i and the other non-proposing employee—and taking the average of the two scaled proposed bonuses. At the decision level (for employee i in team t and round r): $\textit{Adjusted Average Proposed Bonus}_{i,t,r} = \begin{cases} \sum_k (p_{k,i,t,r})/3 & \text{for self}-\text{including design} \\ \sum_{k\neq i} \left(p_{k,i,t,r} \cdot \frac{B_{t,r}-1.5\ e_{k,t,r}}{B_{t,r}}\right)/2 & \text{for self}-\text{excluding design} \end{cases}$, where $p_{k,i,t,r}$ denotes the proposed bonus by employee k for employee i, 1.5 $e_{k,t,r}$ describes k's *Fair Bonus*, and $B_{t,r}$ is the bonus pool. At the employee level (for employee i in team t across all rounds): $\textit{Adjusted Average Proposed Bonus}_{i,t} = \sum_r (\textit{Adjusted Average Proposed Bonus}_{i,t,r})\ /\ 8$.

Variable	Definition
Supplemental process variables	
Absolute Deviation of Final Bonus from Fair Bonus	Absolute deviation of final bonus from the corresponding fair bonus for a given employee i (per round). At the decision level (for employee i in team t and round r): $\lvert \mathit{Final\ Bonus}_{i,t,r} - \mathit{Fair\ Bonus}_{i,t,r} \rvert$. At the employee level (for employee i in team t across all rounds): $\sum_r \lvert \mathit{Final\ Bonus}_{i,t,r} - \mathit{Fair\ Bonus}_{i,t,r} \rvert \ /\ 8$.
Proposal Distortion	Average distortion per proposed bonus in a given employee i's proposal (per round). For the self-including design, distortion in each proposed bonus is the absolute deviation from the fair bonus. For the self-excluding design, each proposed bonus is scaled by the percentage of the bonus pool created by the other employees and only the two proposed bonuses for others are considered. At the decision level (for employee i in team t and round r): $\mathit{Proposal\ Distortion}_{i,t,r} = \begin{cases} \sum_j \lvert p_{i,j,t,r} - 1.5\ e_{j,t,r} \rvert / 3 & \text{for self} - \text{including design} \\ \sum_{j \neq i} \lvert p_{i,j,t,r} \frac{B_{t,r} - 1.5\ e_{i,t,r}}{B_{t,r}} - 1.5\ e_{j,t,r} \rvert / 2 & \text{for self} - \text{excluding design'} \end{cases}$ where $p_{i,j,t,r}$ is the proposed bonus by employee i for employee j, $1.5\ e_{j,t,r}$ describes j's *Fair Bonus*, and $B_{t,r}$ *is* the bonus pool. At the employee level (for employee i in team t across all rounds): $\mathit{Proposal\ Distortion}_{i,t} = \sum_r (\mathit{Proposal\ Distortion}_{i,t,r})\ /\ 8$.
Distortion	Absolute deviation of the *Adjusted Average Proposed Bonus* from the *Fair Bonus* for a given employee i (per round). At the decision level (for employee i in team t and round r): $\mathit{Distortion}_{i,t,r} = \lvert \mathit{Adjusted\ Average\ Proposed\ Bonus}_{i,t,r} - \mathit{Fair\ Bonus}_{i,t,r} \rvert$. At the employee level (for employee i in team t across all rounds): $\mathit{Distortion}_{i,t} = \sum_r (\mathit{Distortion}_{i,t,r})\ /\ 8$.
Distortion & Efficiency Loss	Absolute deviation of the *Average Proposed Bonus* from the *Fair Bonus* for a given employee i (per round). At the decision level (for employee i in team t and round r): $\mathit{Distortion\ \&\ Efficiency\ Loss}_{i,t,r} = \lvert \mathit{Average\ Proposed\ Bonus}_{i,t,r} - \mathit{Fair\ Bonus}_{i,t,r} \rvert$. At the employee level (for employee i in team t across all rounds): $\mathit{Distortion\ \&\ Efficiency\ Loss}_{i,t} = \sum_r (\mathit{Distortion\ \&\ Efficiency\ Loss}_{i,t,r})\ /\ 8$.
Efficiency Loss	Difference between *Distortion & Efficiency Loss* and *Distortion* for a given employee i (per round). At the decision level (for employee i in team t and round r): $\mathit{Efficiency\ Loss}_{i,t,r} = \lvert \mathit{Distortion\ \&\ Efficiency\ Loss}_{i,t,r} - \mathit{Efficiency\ Loss}_{i,t,r} \rvert$. At the employee level (for employee i in team t across all rounds): $\mathit{Efficiency\ Loss}_{i,t} = \sum_r (\mathit{Efficiency\ Loss}_{i,t,r})\ /\ 8$.
Managerial Bias	Difference between (1) *Absolute Deviation of Final Bonus from Fair Bonus* and (2) *Distortion & Efficiency Loss* for a given employee i (per round). At the decision level (for employee i in team t and round r): $\mathit{Managerial\ Bias}_{i,t,r} = \lvert \mathit{Final\ Bonus}_{i,t,r} - \mathit{Fair\ Bonus}_{i,t,r} \rvert - \mathit{Distortion\ \&\ Efficiency\ Loss}_{i,t,r}$. At the employee level (for employee i in team t across all rounds): $\mathit{Managerial\ Bias}_{i,t} = \sum_r (\mathit{Managerial\ Bias}_{i,t,r})\ /\ 8$.

Variable	Definition
Efficiency Loss & Managerial Bias	Difference between (1) *Absolute Deviation of Final Bonus from Fair Bonus* and (2) *Distortion* for a given employee i (per round). At the decision level (for employee i in team t and round r): $\textit{Efficiency Loss \& Managerial Bias}_{i,t,r}$ $= \lvert \textit{Final Bonus}_{i,t,r} - \textit{Fair Bonus}_{i,t,r} \rvert - \textit{Distortion}_{i,t,r}$. At the employee level (for employee i in team t across all rounds): $\textit{Efficiency Loss \& Managerial Bias}_{i,t}$ $= \sum_r (\textit{Efficiency Loss \& Managerial Bias}_{i,t,r}) \;/\; 8$.

REFERENCES

Abbink, K.; L. Dong; and L. Huang. "Talking Behind Your Back: Communication and Team Cooperation." *Management Science* 68 (2022): 5187–200.

Abeler, J.; D. Nosenzo; and C. Raymond. "Preferences for Truth-telling." *Econometrica* 87 (2019): 1115–53.

Adams, J. S. "Towards an Understanding of Inequity." *Journal of Abnormal and Social Psychology* 67 (1963): 422–36.

Adams, J. S. "Inequity in Social Exchange," in *Advances in Experimental Social Psychology*, edited by L. Berkowitz. New York: Academic Press, 1965: 267–99.

Alexander, R. "Inside Goldman's Notoriously Intensive Performance Review Process: Employees Detail the Stress, the Politics, and the Tactics They've Used to Secure the Best Rating." *Business Insider*, 2022. Available at: https://www.businessinsider.com/how-goldman-sachs-annual-performance-reviews-work-wall-street-layoffs-2022-7.

Alonso, R., and N. Matouschek. "Optimal Delegation." *Review of Economic Studies* 75 (2008): 259–93.

Ambrus, A., and S. Takahashi. "Multi-sender Cheap Talk with Restricted State Spaces." *Theoretical Economics* 3 (2008): 1–27.

Andreoni, J. "Why Free Ride?" *Journal of Public Economics* 37 (1988): 291–304.

Appelo, J. "The Peer-to-Peer Bonus System." *Forbes*, July 8, 2015. Available at: https://www.forbes.com/sites/jurgenappelo/2015/07/08/the-peer-to-peer-bonus-system/?sh=54571e844329.

Arnold, M. C., and I. D. Tafkov. "Managerial Discretion and Task Interdependence in Teams." *Contemporary Accounting Research* 36 (2019): 2467–93.

Arnold, M. C.; R. L. Hannan; and I. D. Tafkov. "Team Member Subjective Communication in Homogeneous and Heterogeneous Teams." *The Accounting Review* 93 (2018): 1–22.

Arnold, M. C.; R. L. Hannan; and I. D. Tafkov. "Mutual Monitoring and Team Member Communication in Teams." *The Accounting Review* 95 (2020): 1–21.

Bailey, W. J.; G. Hecht; and K. L. Towry. "Dividing the Pie: The Influence of Managerial Discretion Extent on Bonus Pool Allocation." *Contemporary Accounting Research* 28 (2011): 1562–84.

Baiman, S., and M. V. Rajan. "The Informational Advantages of Discretionary Bonus Schemes." *The Accounting Review* 70 (1995): 557–79.

Baranski, A. "Voluntary Contributions and Collective Redistribution." *American Economic Journal: Microeconomics* 8 (2016): 149–73.

Battaglini, M. "Multiple Referrals and Multidimensional Cheap Talk." *Econometrica* 70 (2002): 1379–401.

Bol, J. C. "The Determinants and Performance Effects of Managers' Performance Evaluation Biases." *The Accounting Review* 86 (2011): 1549–75.

Bolton, G. E., and A. Ockenfels. "ERC: A Theory of Equity, Reciprocity, and Competition." *American Economic Review* 91 (2000): 166–93.

Bolton, G. E., and A. Ockenfels. "Inequality Aversion, Efficiency, and Maximin Preferences in Simple Distribution Experiments: Comment." *American Economic Review* 96 (2006): 1906–11.

Cardinaels, E., and H. Yin. "Think Twice Before Going for Incentives: Social Norms and the Principal's Decision on Compensation Contracts." *Journal of Accounting Research* 53 (2015): 985–1015.

Charness, G., and M. Rabin. "Understanding Social Preferences with Simple Tests." *Quarterly Journal of Economics* 117 (2002): 817–69.

Chaudhuri, A. "Sustaining Cooperation in Laboratory Public Goods Experiments: A Selective Survey of the Literature." *Experimental Economics* 14 (2011): 47–88.

Che, Y.-K., and S.-W. Yoo. "Optimal Incentives for Teams." *American Economic Review* 93 (2001): 525–41.

Coletti, A. L.; K. L. Sedatole; and K. L. Towry. "The Effect of Control Systems on Trust and Cooperation in Collaborative Environments." *The Accounting Review* 80 (2005): 477–500.

Deb, J.; J. Li; and A. Mukherjee. "Relational Contracts With Subjective Peer Evaluations." *The RAND Journal of Economics* 47 (2016): 3–28.

Deloitte. "The Deloitte Global 2021 Millennial and Gen Z Survey." 2021. Available at: https://www2.deloitte.com/content/dam/Deloitte/global/Documents/2021-deloitte-global-millennial-survey-report.pdf.

Demeré, B. W.; K. L. Sedatole; and A. Woods. "The Role of Calibration Committees in Subjective Performance Evaluation Systems." *Management Science* 65 (2019): 1562–85.

Dong, L.; R. Falvey; and S. Luckraz. "Fair Share and Social Efficiency: A Mechanism in Which Peers Decide on the Payoff Division." *Games and Economic Behavior* 115 (2019): 209–24.

Drouvelis, M.; D. Nosenzo; and M. Sefton. "Team Incentives and Leadership." *Journal of Economic Psychology* 62 (2017): 173–85.

Ederhof, M.; M. V. Rajan; and S. Reichelstein. "Discretion in Managerial Bonus Pools." *Foundations and Trends in Accounting* 5 (2011): 243–316.

Edwards, M., and A. Ewen. "How to Manage Performance and Pay with 360-Degree Feedback." *Compensation Benefits Review* 28 (1996): 41–46.

Eichenseer, M. "Leading-by-Example in Public Goods Experiments: What Do We Know?" *The Leadership Quarterly* (2023): 101695.

Emmonds, J. "What Are Peer to Peer Bonuses, and Do They Work?" *Vestd*, October 16, 2019. Available at: https://www.vestd.com/blog/what-are-peer-to-peer-bonuses-and-do-they-work.

Engelmann, D., and M. Strobel. "Inequality Aversion, Efficiency, and Maximin Preferences in Simple Distribution Experiments." *American Economic Review* 94 (2004): 857–69.

Fehrler, S., and M. Janas. "Delegation to a Group." *Management Science* 67 (2021): 3714–43.

Fehr, E., and S. Gächter. "Fairness and Retaliation: The Economics of Reciprocity." *Journal of Economic Perspectives* 14 (2000a): 159–81.

Fehr, E., and S. Gächter. "Cooperation and Punishment in Public Goods Experiments." *American Economic Review* 90 (2000b): 980–99.

Fehr, E., and K. M. Schmidt. "A Theory of Fairness, Competition, and Cooperation." *Quarterly Journal of Economics* 114 (1999): 817–68.

Fehr, E.; M. Naef; and K. M. Schmidt. "Inequality Aversion, Efficiency, and Maximin Preferences in Simple Distribution Experiments: Comment." *American Economic Review* 96 (2006): 1912–17.

Fiala, L., and S. Suetens. "Transparency and Cooperation in Repeated Dilemma Games: A Meta Study." *Experimental Economics* 20 (2017): 755–71.

Fischbacher, U.; S. Gächter; and E. Fehr. "Are People Conditionally Cooperative? Evidence from a Public Goods Experiment." *Economics Letters* 71 (2001): 397–404.

Fisher, J. G.; L. A. Maines; S. A. Peffer; and G. B. Sprinkle. "An Experimental Investigation of Employer Discretion in Employee Performance Evaluation and Compensation." *The Accounting Review* 80 (2005): 563–83.

Gächter, S.; E. Renner; and M. Sefton. "The Long-run Benefits of Punishment." *Science* 322 (2008): 1510.

Galbraith, J. K. *The Great Crash, 1929*. Boston: Houghton Mifflin Company, 1997.

Giamattei, M.; K. S. Yahosseini; S. Gächter; and L. Molleman. "LIONESS Lab: 'A Free Web-Based Platform for Conducting Interactive Experiments Online'." *Journal of the Economic Science Association* 6 (2020): 95–111.

Grabner, I.; J. Künneke; and F. Moers. "How Calibration Committees Can Mitigate Performance Evaluation Bias: An Analysis of Implicit Incentives." *The Accounting Review* 95 (2020): 213–33.

Graham, J. R.; C. R. Harvey; and M. Puri. "Capital Allocation and Delegation of Decision-Making Authority Within Firms." *Journal of Financial Economics* 115 (2015): 449–70.

Hannah, F. "Peer-to-Peer Rewards: Why I Tip My Colleagues at Work." *BBC News*, May 3, 2019. Available at: https://www.bbc.com/news/business-48139518.

Isaac, R. M., and J. M. Walker. "Group Size Effects in Public Goods Provision: The Voluntary Contributions Mechanism." *The Quarterly Journal of Economics* 51 (1988): 179–200.

Karakostas, A.; M. G. Kocher; D. Matzat; H. A. Rau; and G. Riewe. "The Team Allocator Game: Allocation Power in Public Goods Games." *Games and Economic Behavior* 140 (2023): 73–87.

Kim, O., and J. M. Walker. "The Free Rider Problem: Experimental Evidence." *Public Choice* 43 (1984): 113–49.

Kvaløy, O., and T. E. Olsen. "Team Incentives in Relational Employment Contracts." *Journal of Labor Economics* 24 (2006): 139–69.

Lawler, E. E. "Equity Theory as a Predictor of Productivity and Work Quality." *Psychological Bulletin* 70 (1968): 596.

Maas, V., and M. Van Rinsum. "How Control System Design Influences Performance Misreporting." *Journal of Accounting Research* 51 (2013): 1159–86.

Maas, V. S.; M. Van Rinsum; and K. L. Towry. "In Search of Informed Discretion: An Experimental Investigation of Fairness and Trust Reciprocity." *The Accounting Review* 87 (2012): 617–44.

Majerczyk, M., and T. F. Thomas. "Superiors' Discretionary Allocations When Agents Face Disparate Performance Risk." *Journal of Management Accounting Research* 33 (2021): 151–66.

Moers, F. "Performance Measure Properties and Delegation." *The Accounting Review* 81 (2006): 897–924.

Moore, J., and R. Repullo. "Subgame Perfect Implementation." *Econometrica* 56 (1988): 1191–220.

Murphy, K. J., and P. Oyer. "Discretion in Executive Incentive Contracts." Unpublished paper, University of Southern California and Stanford University, 2003. Available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=294829.

Nagar, V. "Delegation and Incentive Compensation." *The Accounting Review* 77 (2002): 379–95.

Ostrom, E.; J. Walker; and R. Gardner. "Covenants with and Without a Sword: Self-governance Is Possible." *American Political Science Review* 86 (1992): 404–17.

Rabin, M. "Incorporating Fairness into Game Theory and Economics." *American Economic Review* 83 (1993): 1281–302.

Rajan, M. V., and S. Reichelstein. "Subjective Performance Indicators and Discretionary Bonus Pools." *Journal of Accounting Research* 44 (2006): 585–618.

Rajan, M. V., and S. Reichelstein. "Objective Versus Subjective Indicators of Managerial Performance." *The Accounting Review* 84 (2009): 209–37.

Rand, D. G.; A. Dreber; T. Ellingsen; D. Fudenberg; and M. A. Nowak. "Positive Interactions Promote Public Cooperation." *Science* 325 (2009): 1272–75.

Rubinstein, A. *Modeling Bounded Rationality*. Cambridge, MA: MIT Press, 1998.

Schuhmacher, K., K. L. Towry, and J. Zureich. "Leading by Example in Socially Driven Organizations: The Effect of Transparent Leader Compensation Contracts on Following." *The Accounting Review* 97 (2022): 373–93.

Sefton, M.; R. Shupp; and J. M. Walker. "The Effect of Rewards and Sanctions in Provision of Public Goods." *Economic Inquiry* 45 (2007): 671–90.

Silverman, R., and L. Kwoh. "Peer Performance Reviews Take Off." *Wall Street Journal*, July 31, 2012.

Stoddard, B.; C. A. Cox; and J. M. Walker. "Incentivizing Provision of Collective Goods: Allocation Rules." *Southern Economic Journal* 87 (2021): 1345–65.

Stoddard, B.; J. M. Walker; and A. Williams. "Allocating a Voluntarily Provided Common-property Resource: An Experimental Examination." *Journal of Economic Behavior & Organization* 101 (2014): 141–55.

Tayler, W. B., and R. J. Bloomfield. "Norms, Conformity, and Controls." *Journal of Accounting Research* 49 (2011): 753–90.

Thöni, C., and S. Volk. "Conditional Cooperation: Review and Refinement." *Economic Letters* 171 (2018): 37–40.

Van Der Heijden, E.; J. Potters; and M. Sefton. "Hierarchy and Opportunism in Teams." *Journal of Economic Behavior & Organization* 69 (2009): 39–50.

Worldatwork and Compensation Advisory Partners. "Incentive Pay Practices Survey: Publicly Traded Companies." 2021a. Available at: https://www.worldatwork.org/resources/research/surveys/incentive-pay-practices.

Worldatwork and Compensation Advisory Partners. "Incentive Pay Practices Survey: Nonprofit/Government Organizations." 2021b. Available at: https://www.worldatwork.org/resources/research/surveys/incentive-pay-practices.

Worldatwork and Compensation Advisory Partners. "Incentive Pay Practices Survey: Private Companies." 2021c. Available at: https://www.worldatwork.org/resources/research/surveys/incentive-pay-practices.

Yang, C.-L.; Z. Boyu; G. Charness; C. Li; and J. W. Lien. "Endogenous Rewards Promote Cooperation." *Proceedings of the National Academy of Sciences* 115 (2018): 9968–73.