

UNIVERSITÀ COMMERCIALE “LUIGI BOCCONI”

PhD SCHOOL

PhD program in Economics and Finance

Cycle: XXXVI

Disciplinary Field: SECS-P/07

Fostering creativity, innovation, and CSR through management control systems

Advisor: Prof. Angelo DITILLO

Co-Advisor: Prof. Eddy CARDINAEELS

PhD Thesis by

Michele FUMAGALLI

ID number: 3061962

Year 2026

ABSTRACT

In the first chapter, I explore how artificial intelligence (AI) evaluators of creative outputs, instead of human evaluators, influence employees' creativity. I also investigate whether the effect depends on the type of feedback provided by the evaluator and on whether the evaluation influences employees' rewards. In an online experiment, I find that AI evaluators reduce employees' creativity compared with human evaluators when the feedback includes numeric scores and the evaluation influences employees' rewards. In contrast, the effect is not significant when the feedback is non-numeric or the evaluation does not influence employees' rewards. Employees perceive AI evaluators that generate numeric feedback as adopting a quantitative evaluation approach. This perception diminishes intrinsic motivation, thereby reducing employees' creativity. Results can be explained by a process of perspective-taking. Employees may struggle to anticipate the assessment of evaluators that they perceive as different from themselves, such as AIs adopting a quantitative evaluation approach.

In the second chapter, coauthored with Eddy Cardinaels and Naomi Soderstrom, we study how mission statements interact with compensation interdependence to influence employees' CSR engagement. When incentive plans tie employees' wealth to financial performance, employees often face conflicts between prioritizing financial performance and pursuing prosocial outcomes (e.g., external CSR). This tension amplifies under compensation interdependence, where prosocial actions that reduce financial performance may also reduce colleagues' pay. Employees may fear being viewed negatively by their peers for prosocial decisions that harm their company's financial performance and mutual bonuses when their compensation is interdependent. In an online experiment, we document that stakeholder-oriented mission statements can mitigate this tension by fostering social norms that legitimize care for external stakeholders. Under compensation interdependence, these mission statements lead to more prosocial choices benefiting external stakeholders than shareholder-oriented missions. However, when compensation interdependence is absent, the type of mission statement does not alter employees' decisions. Our research contributes to the literature on mission statements as management control systems, highlighting that their importance may increase with compensation interdependence.

In the third and final chapter, coauthored with Ariela Caglio and Angelo Ditillo, we examine the determinants and consequences of incentive plan modifications favoring CEOs (IPMs) after uncontrollable events that hinder target attainment. Using data from S&P 500 firms during 2020, the first year of COVID-19, we find that IPMs are more likely in firms that

are more severely affected by the pandemic and when CEOs have stronger departure incentives. However, such IPMs are less likely in firms with a compensation committee, suggesting that such modifications are not always in the best interest of shareholders. Examining investment decisions, we find that IPMs are associated with increases in R&D spending and environmental and social (ES) performance. This association is more pronounced for IPMs affecting long-term incentive plans rather than short-term ones. Taken together, these results suggest that, while IPMs may raise governance concerns, they can also mitigate managerial myopia by encouraging longer-term investments.

ACKNOWLEDGEMENTS

I sincerely thank my advisor, Angelo Ditillo, and my co-advisor, Eddy Cardinaels, for their constant support. I also thank Alexandra Van den Abbeele, Kristof Stouthuysen, Ariela Caglio, Clara Chen, Michael Williamson, Laura Wang, Flora Zhou, Adam Presslee, Claudia Imperatore, Roberto Vincenzi, Adriano Salerno, Claudio Ferrantino, Naomi Soderstrom, Linda Chang, and Li Yang for their insightful suggestions. Additionally, I appreciate the feedback from the participants in the 2024 Doctoral Colloquium of the European Accounting Association, the 2023 and 2025 European Accounting Association Annual Meeting, the 14th Conference on New Directions in Management Accounting, the 2025 MAS Midyear Meeting, and the brown bag seminars at Tilburg University, the University of Bern, the University of Illinois at Urbana-Champaign, the BELSS Journal Club, and Bocconi University.

TABLE OF CONTENTS

CHAPTER 1	6
1.1 INTRODUCTION.....	6
1.2 THEORETICAL BACKGROUND	11
1.2.1 Prior literature on AI and creativity.....	11
1.2.2 Hypotheses.....	12
1.3 RESEARCH DESIGN	16
1.3.1 Experiment.....	16
1.3.2 Measure of Creativity	19
1.4. RESULTS.....	19
1.4.1 Main Results	19
1.4.2 Process Evidence	22
1.4.3 Evaluation’s Influence on Rewards	25
1.5 CONCLUSIONS	26
1.6 REFERENCES.....	27
1.7 FIGURES	32
1.8 TABLES.....	36
1.9 APPENDICES.....	40
CHAPTER 2	49
2.1 INTRODUCTION.....	49
2.2. THEORETICAL FRAMEWORK AND HYPOTHESIS DEVELOPMENT.....	53
2.2.1 Mission Statements, Social Norms, and CSR.....	53
2.2.2 Setting and Hypothesis Development.....	56
2.3 EXPERIMENTAL DESIGN.....	58
2.3.1 Procedure	59
2.3.2 Manipulations	60
2.4 RESULTS.....	61
2.4.1 Main Results	61
2.4.2 Process Evidence	63
2.5 CONCLUSIONS	66
2.6 REFERENCES.....	68
2.7 FIGURES	75
2.8 TABLES.....	77

2.9 APPENDICES.....	81
CHAPTER 3	84
3.1 INTRODUCTION.....	84
3.2 PRIOR LITERATURE AND HYPOTHESES	88
3.2.1 Uncontrollable Factors and Compensation Contracts	88
3.3 RESEARCH DESIGN	92
3.3.1 Sample and Data.....	92
3.3.2 Measures	93
3.3.3 Regression Models	94
3.4 MAIN FINDINGS.....	95
3.4.1 Descriptive Statistics and Details About IPMs.....	95
3.4.2 Determinants of Favorable IPMs.....	97
3.4.3 Consequences of Favorable IPMs	98
3.5 ADDITIONAL ANALYSES	99
3.5.1 Temporal Horizon of IPMs and Investments	99
3.5.2 Favorable IPMs and Financial Performance	100
3.5.3 Consequences of Favorable IPMs for Investments Conditional on COVID-19 Impact.....	101
3.6 ROBUSTNESS CHECKS	102
3.6.1 Falsification Test	102
3.6.2 Entropy Balancing	102
3.6.3 Ruling Out Alternative Explanations	103
3.7 CONCLUSIONS.....	104
3.8 REFERENCES.....	105
3.9 TABLES.....	111
3.10 APPENDICES.....	122

CHAPTER 1

Employees' creativity with AI evaluators

1.1 INTRODUCTION

Artificial intelligence (AI) plays an increasingly important role in the evaluation and selection of creative outputs, such as new products (Eapen et al. 2023; Bell et al. 2024). For example, major movie companies invest in AI-based systems to predict the performance of movies before their release and select those to produce (Chandler 2020). These systems forecast box office revenues directly from the movie scripts, anticipate the critics' response, or reviews from the audience (Rose 2020). AI can also help evaluate advertising content (Liddicoat 2023). Given the emerging interest in using AI to evaluate creative outputs, this paper explores whether using AI evaluators rather than human ones impacts employees' creativity.¹

This study is also motivated by the debate on how to evaluate and reward creative outputs to motivate employees. Creativity is a multifaceted concept, and it is challenging to define a set of objective indicators that capture all its possible aspects (Kachelmeier, Reichert, and Williamson 2008). For this reason, prior research in management accounting acknowledges the importance of subjective evaluation by experts (Grabner 2014). However, technological advancements may challenge the role of subjective evaluation. By learning non-linear patterns from a vast amount of data and considering a multitude of factors, AI can outperform human experts in making recommendations (Castelo, Bos, and Lehmann 2019). As AI applications for evaluating and selecting creative outputs become more accurate and widely used, there is a need for research on how and why employees respond to them.

To address this gap, this study develops and experimentally tests a theory that explains how AI evaluators of creative outputs can influence employees' creativity through the process of perspective-taking (e.g., Grant and Berry 2011). As perceptions of creativity vary across individuals, employees generating creative output have incentives to take the perspective of the evaluators and try to anticipate their assessment (Ostermaier and Uhl 2020). When employees and evaluators share similar traits, they may also believe that they share similar preferences, with a positive impact on their confidence to generate a creative output that the evaluators can appreciate (Bol, LaViers, and Sandvik 2023). Therefore, employees may feel more comfortable

¹ In line with previous studies, I define creativity as: "the production of novel and useful ideas" (Amabile and Pratt 2016).

when generating creative outputs with evaluators they feel similar to, compared to a situation where they feel different from the evaluators.

The literature on human-AI interaction shows evidence of a similar behavior when people attempt to take the perspective of an AI rather than that of other humans. People tend to project their intuition and reasoning more onto other humans than AIs, because they have the impression of understanding other humans better, even when they are strangers (Bonezzi, Ostinelli, and Melzner 2022). In contrast, people view AIs as distant entities and find it challenging to anticipate their behavior (Ye et al. 2021). This phenomenon may occur because AIs and humans differ in their decision-making processes. In particular, AIs rely on data and quantitative methods, whereas humans can also make decisions based on intuition and emotions (Castelo, Bos, and Lehmann 2019; Newman, Fast, and Harmon 2020).

This evidence on perspective-taking and human-AI interaction suggests that an AI evaluator may impair employees' creativity. When employees are developing a creative output that they consider creative, they may be concerned that an AI evaluator will not concur with their view, whereas they may be more confident about the views of human evaluators. Concerns about the evaluator's assessment may cause stress (Dickerson and Kemeny 2004), especially when employees are rewarded for their creativity with tournament incentives (Cardinaels and Feichter 2021). These concerns constrain cognitive resources (Binyamin and Carmeli 2010) and reduce intrinsic motivation (Amabile and Pratt 2016), thereby negatively affecting creativity.

However, this prediction is not straightforward, as employees may appreciate an AI-based evaluation of creative outputs for different reasons. For example, an AI may alleviate concerns related to an arbitrary or distracted evaluation, with a positive impact on motivation and creativity. Moreover, AI evaluators may cause less social anxiety and evaluation apprehension than human ones, thereby making employees feel freer to express their creativity (Siemon 2023).

Given these conflicting arguments, I expect that the impact of an AI evaluator on employees' creative performance depends on contingent factors that influence the employees' perceptions of how the AI conducts the evaluation. In particular, this study focuses on the type of feedback that the evaluator provides when disclosing the results of the assessment. When the AI reports numeric scores to rate the creative outputs (numeric feedback) rather than discursive comments (non-numeric feedback), it may foster the perception that the evaluation is based on a quantitative approach. This perception may increase the concern that employees cannot anticipate the AI evaluator's perspective, potentially having a negative impact on

creative performance. In contrast, the type of feedback should not affect perceived quantification and creative performance when the evaluator is human. As humans can express their preferences by using scores (numeric feedback) without relying on quantitative methods, people may not perceive that human evaluators adopt a quantitative approach just from observing numeric rather than non-numeric feedback.

Consistent with the role of perspective-taking, I also expect that the negative effect of AI evaluators providing numeric feedback, relative to human evaluators, should emerge in situations where employees have incentives to think about how the evaluators will react to their creative output and anticipate their assessment. Specifically, when the evaluation influences employees' rewards, employees should desire a positive evaluation and, therefore, should put effort into anticipating the assessment of the evaluators when developing the creative output. In contrast, when the evaluation does not influence employees' rewards, employees should have less incentive to engage in perspective-taking and generate a creative output that the evaluators may appreciate. In this case, the type of evaluator may exert a weaker effect on creativity.

I test these predictions with an online experiment recruiting participants from Prolific. They compete in a tournament to generate the most creative idea to solve a real-world problem. The instructions state that an evaluator assesses the creativity of the ideas and determines the winner who receives a monetary bonus. Furthermore, participants know that the evaluator provides feedback about their ideas after the experiment. To create an expectation of the type of feedback, participants read two examples before starting to write their creative ideas.

I manipulate the type of evaluator and the type of feedback participants receive. In the AI-Evaluator condition, the instructions specify that the evaluator is a generative AI, whereas in the Human-Evaluator condition, it is another person recruited through Prolific. To manipulate the type of feedback, I vary the two examples that participants read before performing the creative task. In the Numeric-Feedback condition, the examples disclose an assessment of the ideas' novelty and usefulness based on numeric scores (e.g., "Novelty score: 90/100"). Conversely, in the Non-Numeric-Feedback condition, the examples contain only discursive comments (e.g., "The idea is very novel"). I also introduce two additional No-Reward-Influence conditions in the experiment, one with an AI and one with a human evaluator, that are comparable to the Numeric-Feedback conditions except that the evaluation result does not affect the participant's bonuses.

To measure the creativity of the ideas, I employ a consensus assessment technique (Amabile 1979). I recruit additional raters from Prolific, who are not involved in the creative

task or in the evaluation and are therefore blind to the experimental conditions. They independently rate the creativity of the ideas from 1 to 100. Then, I rank the ideas based on their average creativity rating.

Results show that, when participants receive numeric feedback, an AI evaluator compared to a human one decreases the likelihood of generating an idea in the top quartile of the creativity ranking (i.e., a highly creative idea). In contrast, when the feedback is non-numeric, the likelihood of generating a highly creative idea does not significantly vary depending on the type of evaluator.

I also investigate the mechanism through which an AI evaluator, combined with numeric feedback, reduces employees' creativity. I use the concept of perceived quantification from Newman et al. (2020), which is the extent to which participants believe that the evaluator focuses on quantifiable elements, thereby adopting a quantitative evaluation approach. A mediation analysis indicates that an AI evaluator, compared to a human one, increases perceived quantification, which in turn decreases intrinsic motivation, with a negative impact on creativity. Results also show that perceived quantification increases stress, and this, in turn, increases the time participants spend thinking about the criteria that the evaluator will use. Overall, this evidence is consistent with participants experiencing discomfort in taking the perspective of an AI evaluator compared to a human one when feedback is numeric, which makes the creative task less enjoyable and reduces creative performance.

As an additional test of the mechanism, I show, as expected, that the absence of rewards tied to the evaluation reduces the negative impact of an AI evaluator on employees' creativity when feedback is numeric. Results indicate that the type of evaluator does not significantly affect creativity when employees lack incentives to take the evaluator's perspective. Moreover, results show that, when an AI evaluator discloses numeric feedback, employees' creativity is higher when the evaluation does not influence rewards than when it influences rewards, suggesting that monetary incentives based on creativity can backfire with an AI evaluator perceived as using a quantitative approach.

This study contributes to the nascent literature on the implementation of AI in management control systems (Hu 2022; McDaid, Andon, and Free 2023). It answers the call of Sundström (2024) to examine how AI affects forms, practices, and infrastructures of management control systems. Algorithms can influence performance evaluation practices by monitoring, setting goals, determining compensation, and providing feedback, which may reduce employees' sense of autonomy and create an impersonal and demotivating environment (Tong et al. 2021; Parent-Rocheleau and Parker 2022). I contribute to this stream of research

by showing that AI evaluators can negatively impact creative performance and increase stress when they provide numeric feedback and influence employees' bonuses. However, non-numeric feedback and the absence of performance-based rewards mitigate this effect, such that employees' creative performance is not significantly different from a situation with human evaluators. These findings also complement Siemon (2023), who shows that submitting creative ideas to an AI rather than to humans results in lower evaluation apprehension in a setting without feedback and performance-based rewards.

This study also contributes to the literature on performance evaluation and creativity. Designing management control systems that balance the tension between the need for enabling mechanisms that stimulate employees' creativity and coercive mechanisms that constrain their behavior is a fundamental challenge in organizations (Davila 2000; Davila, Foster, and Oyon 2009; Adler and Chen 2011; Davila and Dittillo 2017b). Prior literature acknowledges that management control systems used in performance evaluation, such as budgets and incentives, play an important role in stimulating employees' creativity in line with the organization's objectives (Bisbe and Otley, 2004; Grabner 2014; Cools, Stouthuysen, and Van den Abbeele 2017). A recent stream of research shows evidence that, when creative performance is evaluated, the characteristics of the evaluator influence employees' creative outputs (Ostermaier and Uhl 2020; Bol, LaViers, and Sandvik 2023). I show that AI evaluators can negatively impact creative performance, and that the perception that AI evaluators use a quantitative approach explains, at least in part, this effect.

Finally, this study may help companies implement AI to assess and select creative outputs. The results suggest that companies should adopt AI-based systems that do not disclose numeric feedback to lessen the potential drawbacks to creative performance. Moreover, the findings indicate that linking bonuses to AI evaluation in creative tasks may backfire when contextual factors, such as the type of feedback, heighten perceived quantification in the evaluation approach. Instead, when the AI evaluator provides non-numeric feedback or does not influence employee bonuses, employees' creative performance is not significantly different from having a human evaluator. In these conditions, companies may select the type of evaluator based on pros and cons not considered in this study, such as the cost and the accuracy of the evaluation process.

1.2 THEORETICAL BACKGROUND

1.2.1 Prior literature on AI and creativity

Companies are progressively implementing AI into their creative processes as tools to enhance efficiency and, in the case of generative AI, inspire employees and provide suggestions (Pagani and Champion 2023). A growing body of research delves into the impact of AI on creativity. Jia et al. (2024) find that delegating repetitive tasks to AI allows skilled employees to focus on creative activities, ultimately enhancing performance. Samet, Williamson, and Yip (2023) show that knowledge about generative AI mitigates evaluation biases in favor of human experts. Still, to the best of my knowledge, there remains a research gap about employees' reactions to AI evaluators of creative outputs, especially when their evaluations influence employees' rewards.

Existing literature in management accounting underscores the advantages of subjectivity in evaluating and rewarding creative tasks (Grabner 2014). Formulaic incentives prove effective when managers clearly understand what the performance metrics measure and their alignment with the firm's goals (Merchant and der Stede 2007). On the contrary, the inherent nature of creativity defies simple objective definitions (Kachelmeier, Reichert, and Williamson 2008). Therefore, algorithmic performance evaluation may encourage an excessively narrow focus on maximizing specific metrics, diminishing attention on broader aspects of creative performance (Grabner 2014).

However, advances in technology may overcome this measurement problem. Unlike traditional algorithms, AI can consider a vast array of variables and autonomously learn complex relationships from data, even those that the developers do not anticipate (Alzubaidi et al. 2021). In this regard, AI's potential extends beyond merely enhancing efficiency in creative processes by delegating repetitive tasks. It can also reveal new opportunities and provide insights into the value of creative ideas (Pagani and Champion 2023).

The benefit of using AI to select from a pool of creative solutions becomes evident when objective performance measures are only available after an extended period of time (Votto et al. 2021). For example, in the context of advertising, managers can exploit AI and the richness of social media data to understand which individual creative components of past advertising campaigns drive customer attention (Liddicoat 2023). Then, they can use this information to select the features of future campaigns, reducing the need to test multiple alternatives, to rely on the opinion of experts, or to wait for data on actual consumers' responses.

Another promising application of AI is the evaluation of crowd-sourced ideas. Algorithms can evaluate large volumes of information more efficiently than humans, allowing companies to identify a subset of promising ideas from a vast pool of possible alternatives (Eapen et al. 2023). Recent papers at the intersection of marketing and computer science propose machine learning techniques to evaluate and select ideas generated by communities of consumers (Ameen et al. 2022; Just et al. 2023; Bell et al. 2024). As crowd-sourced creative contests also happen among employees (e.g., Bol, LaViers, and Sandvik 2023), these machine learning techniques have material implications for management control systems.

1.2.2 Hypotheses

Individuals have different preferences, which can determine their judgment when selecting creative ideas (Mueller et al. 2018). As it is hard to define creativity objectively, employees performing a creative task have incentives to take the evaluators' perspective and generate an output that satisfies the evaluators' preferences (Ostermaier and Uhl 2020). By doing so, employees can increase the likelihood of receiving a positive assessment. As a result of this process, the evaluators' characteristics can influence the features of a creative output. Consistent with this argument, prior research finds that performance evaluation by customers or managers encourages employees to submit creative ideas that maximize usefulness, while external experts or peers induce employees to focus on novelty (Ostermaier and Uhl 2020; Bol, LaViers, and Sandvik 2023).

The characteristics of the evaluators may also affect employees' confidence in getting a positive assessment. Individuals feel that they understand better who they resemble, feeling less uncertainty when they try to take their perspective (Epley et al. 2004). For example, Bol, LaViers, and Sandvik (2023) find that peer employee evaluators induce higher participation rates than managers in a creative contest among employees.

Prior research in human-machine interaction provides evidence consistent with this argument. People have the impression of understanding the perspective of other humans, even strangers, more readily compared to those of robots or abstract algorithms (Ye et al. 2021; Bonezzi, Ostinelli, and Melzner 2022). A possible reason for the perception of distance between humans and AI is that humans can make decisions based on intuition and gut, while AI relies on quantitative methods (Castelo, Bos, and Lehmann 2019; Newman, Fast, and Harmon 2020). Therefore, in a creative task, employees may experience greater uncertainty under an AI evaluator than a human one.

Uncertainty about the evaluation outcome can influence creative performance through increased stress (Dickerson and Kemeny 2004). Prior research establishes a connection between emotions and creativity (Amabile and Pratt 2016). Positive emotions enhance creativity by making the task more enjoyable, increasing attention, and encouraging risk-taking (Adler and Chen 2011). Conversely, negative emotions arising from evaluative stress impair creativity by inducing pressure and diminishing intrinsic motivation (Byron, Khazanchi, and Nazarian 2010). In particular, stress may impair cognitive abilities, thereby limiting resources for creative thinking (Fabio, Picciotto, and Capri 2021). Therefore, the distance that the employee perceives from the AI evaluator can lead to stress and reduce creative performance compared to a human evaluator.

Nevertheless, these predictions are not granted, as there are potential reasons why an AI evaluator may improve creative performance. First, individuals may prefer some level of objectivity when they are evaluated, even if the task involves creativity. An AI evaluator may be perceived as more objective than a human one, thereby increasing the perceived ability to evaluate creativity. Second, people might doubt that human evaluators will exert sufficient effort or attention in their assessment. Prior research finds that people are more likely to trust advice from AIs and are less confident in human judgment in situations of time pressure (Jung and Seiter 2021). Similarly, participants in a creative contest may expect that human evaluators may not diligently assess a large volume of creative ideas because of limited time and attention, whereas AI evaluators are not subject to these constraints. Then, AI evaluators may make the evaluation process more accurate and reliable, encouraging employees to exert effort that positively affects creative performance.

Given the existence of conflicting predictions, I formulate a research question for the main effect of the type of evaluator on employees' creativity:

RQ1.1: *Does the type of evaluator, AI or human, affect employees' creativity?*

If the reason for the perceived distance between humans and AI stems from the perception that AI adopts quantitative methods, the type of feedback that the evaluator generates may influence creative performance. In particular, people may perceive the AI differently depending on whether it generates numeric or non-numeric feedback (i.e., scores or discursive comments) for the evaluation.

Prior research suggests that readers distinguish between numeric and non-numeric information in managerial reports and infer different managerial skills based on the type of

information used (Stoel, Ballou, and Heitger 2017). In particular, numeric information can signal analytical competence (Kadous, Koonce, and Towry 2005). Moreover, numeric information is generally associated with mechanization, standardization, and impersonality, as well as a reduction of discretion implied in quantitative techniques (Espeland and Stevens, 2007).

While AIs are generally perceived as different from humans, these differences may be accentuated or attenuated by contextual cues. Research on anthropomorphism suggests that people attribute human-like behavior, such as intentions and emotions, to inanimate entities (Epley, Waytz, and Cacioppo, 2007). This process is more likely to occur when the object possesses features that resemble those of a human, such as a human face (DiSalvo et al., 2002). In the context of AI, the content that it generates matters. Prior research finds that when AI generates recommendations that resemble the content expected from human experts, people react to AI's recommendations more similarly to recommendations from human experts (Longoni, Bonezzi, and Morewedge 2019).

I expect that the type of feedback acts as a contextual cue that affects people's perceptions of how AI evaluates creative ideas and therefore its similarity with a human evaluator. Expecting numeric feedback from the AI evaluator (e.g., creativity scores) may reinforce the perception that it uses a quantitative approach. This may increase the distance that people perceive with an AI evaluator, as the standardization and impersonality implied in numeric information contrast with the inherently subjective evaluation of creativity that is expected by a human evaluator. Instead, the presence of non-numeric feedback may serve as a contextual cue that makes the AI perceived as more similar to a human evaluator.

On the contrary, when the evaluator is human, I do not expect the type of feedback to have different effects on creative performance because people can use heuristics to express their opinions as scores without the need for quantitative methods or formulae (e.g., Likert scales or review ratings). Subjective performance evaluation metrics can be expressed as numeric scores, with different dimensions rated separately and combined into a final score using numerical weights. However, these scores can still be perceived as depending on the discretion and personal opinions of the managers rather than on quantitative methods (e.g., Ittner et al., 2003). Therefore, observing numeric scores as feedback from a human evaluator may not foster the perception that the human evaluator adopts a quantitative approach.

Given the arguments above, I predict the following ordinal interaction effect:

H1.1: *AI (vs. human) evaluators reduce more employees' creativity under numeric than non-numeric feedback.*

Another contingent factor that can moderate the effect of the type of evaluator on employees' creativity is whether the evaluation influences employees' rewards. Following the mechanism based on perspective-taking, employees should put effort into anticipating the assessment of the evaluator when they feel that the result of the evaluation is important. This can be the case when evaluation influences employees' bonuses, as performance-based incentives make people more creative in ways that, according to their expectations, increase the likelihood of a positive performance evaluation (Speckbacher 2021). Conversely, when the evaluation does not influence rewards, employees may be less concerned about obtaining a positive evaluation and may think less about the criteria that the evaluator uses for the assessment.

In support of this argument, Cardinaels and Feichter (2021) find that tournament-based incentives, compared to more distributed reward schemes, increase stress with a negative effect on creativity. Because only one person in the group receives the reward under tournament incentives, people devote excessive attention to the evaluation process, which consumes cognitive resources that could be dedicated to generating creative output. Moreover, prior research finds that monetary incentives linked to accounting performance measures align employees' creativity with business needs, addressing the tension between artistic and commercial value of creativity (Klein and Speckbacher 2020).

In line with these arguments, I expect that, when the evaluation does not influence rewards, employees may feel less pressure to anticipate the evaluator's viewpoint, and this may alleviate the negative impact of an AI evaluator compared to a human manager when the type of feedback is numeric². Formally:

The effect of the type of evaluator (AI or human) on employees' creativity is weaker
H1.2: *when the evaluation does not influence employees' rewards than when it does influence employees' rewards.*

² Consistent with Asay et al. (2022), I study the effect of rewards as a moderating variable to test the role of perspective-taking as a process mechanism to explain why creative performance may decrease under an AI evaluator with numeric feedback compared to a human evaluator. Therefore, I do not formulate a specific hypothesis for the main effect of incentives on creativity. Similarly, I do not formulate or test a specific hypothesis for the interaction between the type of feedback (regardless of the evaluator) and incentives, or for a triple interaction effect.

1.3 RESEARCH DESIGN

1.3.1 Experiment

To test my hypotheses, I conduct a $2 \times 2 + 2$ between-subject online experiment. Prior literature acknowledges that experimental designs are useful to explore theories on how alternative designs of management control systems influence creativity (Davila and Ditillo 2017a). I manipulate the type of evaluator (AI vs. Human) and the type of feedback (Numeric vs. Non-Numeric). In addition, I add two additional conditions where feedback is numeric, but the assessment of the AI or the human evaluator does not affect rewards (No-Reward)³.

The experiment consists of two steps, each run in a separate online session. In the first one, the participants act as performers⁴, who compete in a tournament about writing a creative idea. In the second step, other Prolific users act as evaluators, ranking the ideas to determine the most creative. In both steps, I administer the experiment with Qualtrics surveys⁵.

For the first step, I recruit 138 performers from Prolific, filtering for UK residents, English speakers, and a minimum approval rate of 95% with at least 30 completed tasks⁶. Before starting the main creative task, participants read the instructions. They learn that they compete with two other Prolific participants to write a creative idea. The winner of the tournament obtains £2.5 on top of a £1.7 fixed wage for an estimated commitment of 15 minutes⁷. For the two No-Reward conditions, I recruit an additional 66 participants with the same characteristics, but the £2.5 bonus is assigned randomly to one participant in every group of 3 people, making the distribution of the payouts equivalent to the other conditions but independent from the evaluator's assessment.

The performers learn that one evaluator determines the most creative idea. The evaluator should also provide feedback about the ideas and state the motives for the evaluation outcome. To generate an expectation of the feedback format, the instructions feature two examples about an unrelated task (see Appendix 1.1).

The instructions clarify that the performers have 5 minutes to generate and write their creative idea. Sending the idea before the end of the countdown is possible, but subsequent

³ I consider these No-Reward conditions as additional because I collected the data for these conditions in September 2024, after collecting the data for the other conditions in December 2023.

⁴ In this paper, I use the term “performer” to distinguish the participants who wrote the creative ideas from the evaluators. The experimental material does not mention the term “performer”.

⁵ The study received approval from an institutional ethics committee (institutional review board).

⁶ Before collecting the data, I obtained approval from my academic institution's ethics committee (i.e., institutional review board).

⁷ The median actual completion time is 12:31 minutes.

modifications are not allowed. The instructions also include three comprehension checks, which participants must answer correctly before proceeding to the next page of the Qualtrics survey.

After answering the comprehension checks, the performers read that their task is: *“Develop and describe an idea as creative as possible to prevent alcohol abuse among teenagers, where creative means both novel and useful.”* The page displays the task, a box for writing the idea, and a countdown clock. After submitting their idea, the performers complete a post-experimental questionnaire, which also includes attention and manipulation checks.

For the evaluator type manipulation, the instructions show a different description. In the Human-Evaluator condition, the evaluator is *“another Prolific participant resident in an English-speaking country,”* whereas in the AI-Evaluator condition, the evaluator is *“a generative Artificial Intelligence (AI) based on a GPT language model.”* In both cases, the instructions specify that the evaluator can interpret English texts and write feedback. Moreover, the evaluator can search for information online to assess the novelty and usefulness of the ideas (see Appendix 1.2 for the full text). Pronoun usage also differs between conditions, with *“it”* for the AI and *“he/she”* for the human evaluator.

For the type of feedback manipulation, I vary the examples of feedback that participants read before starting to write their creative idea. The Numeric-Feedback condition displays novelty and usefulness scores, such as *“Usefulness score: 80/100”* and *“Novelty score: 40/100.”* The Non-Numeric-Feedback condition presents textual assessments, like *“This idea is very useful”* and *“This idea is not very novel.”*⁸ In both instances, the examples justify the evaluation with one or two discursive comments (see Appendix 1.1). This design choice aims to maintain constant across conditions the perceptions about the evaluator’s transparency and proficiency in English, while varying only the presence of numbers.

In both the Numeric and Non-Numeric-Feedback conditions, participants read that *“The participant submitting the most creative idea will win a bonus of £2.5 on top of the £1.7 participation reward.”* To test H1.2, I include two additional No-Reward-Influence conditions, one for the AI and one for the Human evaluator, in which the bonus is assigned randomly so that the evaluation does not influence participants’ rewards. The instructions mention: *“One participant, chosen randomly from the three who submitted the ideas, will win a bonus of £2.5 on top of the £1.7 participation reward. Therefore, the bonus is not dependent on the evaluation results.”* The decision to randomly assign the bonus aims to make the distribution of the bonus

⁸ See Appendix 1.2 for the complete text of the manipulation.

constant with the other conditions, while removing only the influence of the evaluator's assessment. In the No-Reward-Influence conditions, the type of feedback is numeric to make the context comparable with the Numeric-Feedback conditions, except for the influence of the evaluation on rewards.

The second step of the experiment consists of evaluating the creative ideas in each group of performers to determine the winner in each group. In the AI-Evaluator condition, I use Copilot (formerly branded as Bing AI) to evaluate the ideas of the participants and generate feedback. At the time the experiment was conducted, Copilot was one of the few large language models with real-time access to the Internet, increasing the ability of the AI to capture novel trends. The request starts with the following prompt: *"I will give you three creative ideas. The task was to write a creative idea about preventing alcohol abuse among teenagers. Rank these ideas based on creativity. Moreover, provide a score from 1 to 100 about their novelty and usefulness."* Then, I submit each group of three ideas separately, repeating the following prompt for every group: *"This is a group of ideas: Idea 1: ... Idea 2: ..., Idea 3: ..."*

Copilot's responses always include a numeric score and a textual assessment of creativity, novelty, and usefulness (see Appendix 1.3 for examples). Before delivering the feedback to the performers, I manually delete the unnecessary parts to conform to the examples in the experiment, which also depend on the type of feedback condition. Moreover, I check each answer from Copilot to ensure that they do not contain offensive content, but modifications are unnecessary.

For the Human-Evaluator condition, I recruit one evaluator for each group of performers on Prolific⁹. The evaluators obtain £1.6 for an expected workload of 15 minutes. First, I ask the evaluators to rank the ideas in the group with a drag-and-drop task. Then, I require them to provide more details for each idea. In the Numeric-Feedback condition, the evaluators write a number from 1 to 100 in a box to rate novelty and usefulness. In the expected Non-Numeric-Feedback condition, they answer a 7-point Likert scale question from *"Not Novel (Useful)"* to *"Extremely Novel (Useful)"*.¹⁰ Then, I combine the answers of the

⁹ In total, 47 evaluators participated, as it was necessary to replace one for submitting an incomplete answer.

¹⁰ The purpose of this decision was to standardize the evaluation across groups and make it more comparable with the examples of feedback that the performers read before the task. In any case, it was not necessary for the purpose of the experiment. The performers completed the experiment before reading the feedback of the evaluator. Therefore, differences in the actual content of the feedback across conditions could not influence the results.

evaluators using an R script to generate the text for the feedback (see Appendix 1.4 for examples), which I send as a private Prolific message to the performers¹¹.

1.3.2 Measure of Creativity

To measure the creativity of the ideas for the analysis, I employ a consensus assessment approach (Amabile 1979). I ask panels of raters to assign scores from 1 to 100 for creativity, novelty, and usefulness (e.g., Chen, Williamson, and Zhou 2012). I randomly split the ideas into eight sets to keep the task manageable in a reasonable time. For each set, I recruit eight raters on Prolific¹². These raters are blind to the experimental conditions and do not participate in the experiment as participants or evaluators. They rate around 35 ideas plus two attention checks¹³. Their compensation consists of a fixed wage of £4 for an expected workload of 30 minutes. When a rater fails an attention check, I substitute it with a new one.

For each idea, I calculate the average creativity rating of the eight assigned raters (e.g., Kachelmeier, Wang, and Williamson 2019). The Cronbach's Alpha of the raters' sets ranges from 0.728 to 0.869, indicating reasonable levels of internal consistency (Murphy and Davidshofer 1988). To avoid differences in the distribution of ratings across raters' sets, I use the rank of the ideas' average creativity rating, where 1 is the most creative idea in the set and 35 is the least creative one. Similarly to Bruggen, Feichter, and Williamson (2018), I consider an idea as highly creative if it is in the first quartile of the ranking or above. *TopCreativity* is a dummy variable equal to 1 if the idea is in the first quartile of the ranking and 0 otherwise.

1.4. RESULTS

1.4.1 Main Results

The sample consists of one observation for each participant writing a creative idea in the experiment (the performers). To ensure high quality of the data, I exclude 11 participants failing the attention checks and 1 not submitting a creative idea, resulting in 204 observations. The manipulations were successful overall, with 97% of participants correctly remembering the type of evaluator, 95% correctly remembering the type of feedback, and 92% the condition for receiving the monetary bonus. Differences in gender, age, education, and ethnicity across

¹¹ Each performer received only the portion of feedback about his or her idea. I did not disclose the ideas of the other participants in the group or their evaluation.

¹² The ratings were collected in October 2024, after obtaining the data of the additional No-Reward conditions, such that the creativity scores are comparable across all the conditions.

¹³ As one participant did not submit the creative idea, one set of raters evaluated 34 ideas instead of 35.

conditions are insignificant at conventional statistical levels ($p > 0.1$), indicating a successful randomization. Table 1.1 contains descriptive statistics of the variables, and Appendix 1.5 provides their detailed definitions.

[Table 1.1 here]

Table 1.2 and Figure 1.1 report summary statistics for the main variable of interest, *TopCreativity*, by experimental condition. The average of *TopCreativity* represents the proportion of participants submitting a top-quartile creative idea depending on the experimental conditions.

[Table 1.2 here]

[Figure 1.1 here]

To explore RQ1.1 and H1.1, I estimate a linear model considering the four conditions in which participants receive a monetary bonus contingent on their creative performance. I regress *TopCreativity* on two dummy variables related to the experimental conditions and their interaction. *AEvaluator* equals 1 for the AI and 0 for the Human-Evaluator condition. *NumericFeedback* equals 1 for participants in the Numeric-Feedback condition and 0 otherwise. I also consider three control variables. *Female* equals 1 for participants identifying as females and 0 otherwise. *Education* is the highest level of study achieved by the participant from 1 (some primary school) to 8 (graduate or professional degree). *HelpAI* is equal to 1 if the participant declares after the task that he or she used the help of a generative AI to write the creative idea.

These control variables can explain differences in creative performance not related to the experimental conditions that may introduce noise in the data¹⁴. The help of generative AI and education might improve the ideas' writing style or enable participants to write more within the limited time, potentially influencing the creative scores assigned by the raters¹⁵. *HelpAI* also mitigates the concern that results are driven by participants using AI tools. Including gender as a covariate is consistent with prior studies on creativity with tournament incentives (e.g., Chen, Williamson, and Zhou 2012; Cardinaels and Feichter 2021).

¹⁴ The control variables do not statistically differ across conditions ($p > 0.1$). Their inclusion in the models aims to improve the precision of the estimation (Piercey 2023).

¹⁵ For example, I find that participants using the help of an AI submit longer ideas (untabulated).

Table 1.3 Panel A reports the results of the estimation in column 1. The coefficient *AEvaluator* is positive but insignificant at conventional statistical levels¹⁶. *NumericFeedback* is also insignificant. Instead, the interaction term between *AEvaluator* and *NumericFeedback* is negative and significant. Column 2 reports analogous results, including the control variables in the model, while columns 3 and 4 use a logit model given the binary dependent variable. Overall, these results support H1.1, but a deeper analysis of the simple effects underlying the negative interaction is required.

[Table 1.3 here]

Table 1.3 Panel B reports the results of the linear model's marginal effects taking into account the control variables. Overall, results indicate that the effect of an AI evaluator on creative performance depends on the type of feedback. Consistent with H1.1, an AI evaluator compared to a human one significantly decreases the likelihood of generating a very creative idea by 20.5% when feedback is numeric. Instead, when feedback is non-numeric, the effect is positive but insignificant at conventional statistical levels. Therefore, the effect of an AI evaluator is more negative under numeric than non-numeric feedback. Notably, the type of feedback does not significantly affect creativity when the evaluator is a human, supporting the argument that observing numeric or non-numeric feedback from human evaluators does not significantly affect people's perceptions of the evaluation process.

Results are analogous when *TopCreativity* is defined as an idea in the top tercile of the ranking (untabulated). Moreover, a quantile regression shows that the interaction effect between the type of evaluator and feedback affects the creativity of high-creative ideas rather than low-creative ones (Appendix 1.6). For this analysis, I use a continuous creativity score standardized by the group of raters to mitigate concerns that differences in the rating distribution across raters bias the estimation. The results of the quantile regression model suggest that an AI evaluator disclosing numeric feedback reduces the creative performance of the employees who want to excel in the task and therefore have incentives to take the evaluator's perspective. Finally, to further rule out the possibility that results depend on participants relying on the help of AI to generate ideas, I remove them from the sample and obtain analogous results (untabulated)¹⁷.

¹⁶ In addition, using ANOVA, I find an insignificant main effect of the type of evaluator ($df = (1,128)$, $F = 0.367$, $p = 0.546$) and feedback ($df = (1,128)$, $F = 0.525$, $p = 0.470$).

¹⁷ I also ask participant three questions about their familiarity with AI that load on a single factor. Results are analogous when this factor is included as a control variable in the model (untabulated).

Overall, the results do not support the presence of a main effect of the type of evaluator (RQ1.1) but do support the interaction effect of H1.1. The effect of AI evaluators on creative performance is likely to depend on contingent factors. An AI evaluator disclosing numeric feedback decreases creative performance compared to a human evaluator. The concerns are mitigated when the AI evaluator discloses non-numeric feedback, such that the difference in creative performance with human evaluators is not significant at conventional statistical levels.

1.4.2 Process Evidence

The mechanism described in the theory section implies that an AI evaluator fosters the impression of adopting a quantitative approach (quantification), particularly when participants receive numeric feedback. When participants try to take the perspective of the evaluator, perceptions of quantification may increase the concern that the AI evaluator has a different evaluation approach from a human, and, therefore, it is hard to anticipate its assessment of the creative idea. This concern makes the task unpleasant, diminishing intrinsic motivation and, in turn, creativity. This effect should be more pronounced when participants expect numeric rather than non-numeric feedback from the AI evaluator, as numeric feedback may exacerbate the impression that the AI uses a quantitative evaluation approach.

To provide support for this mechanism, I estimate a process model for the impact of *Aevaluator* on *TopCreativity* with perception of quantification and intrinsic motivation as mediators¹⁸. I measure intrinsic motivation in the post-experimental question with a proxy for task attractiveness (Fessler 2003; Kachelmeier, Wang, and Williamson 2019). I ask participants to describe the task by choosing between pairs of adjectives, one representing a favorable opinion (Attractive, Exciting, Good, Interesting, Superior, Wholesome, Fun), and the other a negative one (Repulsive, Dull, Bad, Boring, Inferior, Unwholesome, Tedious) (Bailey and Fessler 2011). *Intrinsic* represents the first factor scores of these items. For perceptions of quantification in the evaluation approach, I adapt three items from Newman, Fast, and Harmon (2020)¹⁹, and compute *Quantification*, which is equal to the first factor scores of these items.

Figure 1.2 Panel A illustrates the estimated coefficients for the subsample of participants in the Numeric-Feedback condition. Results are consistent with the theorized

¹⁸ To estimate the process models in the paper, I use the R 4.3 version of the Hayes PROCESS macro with 5,000 bootstrap replications (Hayes 2017).

¹⁹ “With reference to the {Prolific participant}/{generative AI} evaluator: 1) I feel that {he/she}/{it} will just reduce my creative idea to a number. 2) I think that {he/she}/{it} will neglect some important elements of the creativity of my idea. 3) I believe that {he/she}/{it} will consider only the quantifiable elements of my creative idea.”

mechanism. Then, the indirect effect of *AIevaluator* on *TopCreativity* through *Quantification* and *Intrinsic* is negative and significant at the 90% level (bootstrapped confidence interval: [-0.611, -0.005]). I repeat the estimation for the subsample of observations in the Non-Numeric-Feedback condition. Figure 1.2 Panel B shows the results. The indirect effect of *AIevaluator* through the mediators is insignificant at the 90% level (bootstrapped confidence interval: [-0.178, 0.038]). Moreover, the process models show that AI evaluators significantly increase perception of quantification when participants expect numeric feedback, but not significantly when they expect non-numeric feedback. This result also reduces potential concerns that the effect of AI under numeric feedback is not driven by the presence of quantitative information in the feedback but by other potential confounding factors, such as the tone or the informativeness of the feedback.

[Figure 1.2 here]

To provide further support that the effect is driven by participants trying to take the perspective of the evaluator, I test whether perceptions of quantification increase stress and the time spent thinking about the evaluator's criteria. If perceptions of quantification make participants believe that an AI evaluator may have a distant mindset compared to a human one, participants should find the process of perspective-taking more stressful and cognitively demanding. Then, they should spend more time thinking about the criteria that the evaluator will use to assess the idea.

To measure stress and the time spent thinking about the evaluator's criteria, I use two 7-point Likert scale items from the post-experimental survey. Stress is the average of two questions adapted from Cardinaels and Feichter (2021) about the stress perceived for the evaluation and the stress experienced during the task. *ThinkCriteria* is the answer to an item about the time spent thinking about the evaluator's criteria.

Figure 1.3 shows the process estimated for the subsample of participants in the Numeric-Feedback conditions. As reported in Figure 1.3, an AI evaluator increases perceptions of quantification. This perception increases stress, which increases the time spent thinking about the criteria that the evaluator will use. The indirect effect of *AIevaluator* on *ThinkCriteria* through quantification and Stress is positive and significant at the 90% level [0.007, 0.227].

[Figure 1.3 here]

Finally, I test a possible alternative explanation that participants believe that the AI evaluator is not capable of providing an adequate assessment of creativity. In this case, participants may feel that the evaluation process is unfair, and spending time finding a creative idea is a waste of effort. Then, they may exert less effort in the task, which can reduce their creative performance.

I ask participants to agree whether they believe the evaluator is capable of evaluating creativity on a scale from 1 (strongly disagree) to 7 (strongly agree). *EvaluatorAbility* is the answer to this question. Then, I measure *Effort* as the time spent submitting the creative idea in seconds.

I test whether an AI evaluator affects *Effort* by estimating a process model with *Quantification* and *EvaluatorAbility* as mediators. Figure 1.4 reports the results. I find that *AIevaluator* and *Quantification* significantly reduce the perception that the evaluator is capable of assessing creativity. However, I do not find that the perceived ability to evaluate creativity significantly affects *Effort* at conventional statistical levels ($p = 0.620$), nor do I find a direct effect of *AIevaluator* ($p = 0.529$). Moreover, the indirect effect of AI through the mediators is insignificant at the 90% level $[-0.149, 0.387]^{20}$. These results do not support the alternative explanation that the decrease in creativity is due to participants being reluctant to exert effort because they believe the AI evaluator is not capable of assessing creativity.²¹

[Figure 1.4 here]

Overall, the results suggest that perspective-taking may, at least partially, explain why an AI evaluator that provides numeric feedback reduces creative performance compared to a human evaluator. The perception that the evaluator adopts a reductionistic quantitative approach may make the task more unpleasant, reducing intrinsic motivation. Results also do not support the explanation that the decline in creative performance is driven by a reduction in effort. Although participants believe that an AI evaluator is less capable of assessing creativity than a human evaluator, they do not spend less time on the task, nor write shorter ideas. Instead, consistent with the argument that taking the perspective of an AI evaluator is cognitively

²⁰ I also estimate a process model analogous to Figure 1.4, with an additional step where *TopCreativity* is the final dependent variable. I find that *Effort* is positively and significantly associated with *TopCreativity* ($p < 0.01$), but *EvaluatorAbility* is not ($p = 0.537$). The total indirect effect of the model is insignificant at the 90% level $[-1.088, 2.011]$, as well as any other indirect effect that includes *EvaluatorAbility* as mediator (untabulated). These results do not support the alternative explanation that the reduction in creative performance when feedback is numeric depends on participants' beliefs about the AI's capability to evaluate creativity compared to a human evaluator.

²¹ Results are consistent when using the number of characters written as an alternative proxy for effort (untabulated).

demanding and potentially discomfoting, participants in the AI evaluator with Numeric-Feedback condition report higher stress, which increases the time spent thinking about the criteria that the AI evaluator will use²². While this process evidence supports the role of perspective-taking and mitigates potential concerns about alternative explanations, the results should be interpreted as causal mediation, as the moderating variables were measured after participants performed the creative task (Asay et al. 2022).

1.4.3 Evaluation's Influence on Rewards

As outlined in H1.2, if results are driven by people trying to anticipate the evaluator's perspective, when the evaluation does not affect employees' rewards, they should have less incentive to take the evaluator's perspective. Therefore, it may attenuate the effect of the type of evaluator on creative performance. To test this hypothesis, I estimate a linear model, regressing *AEvaluator*, *NoRewardInfluence*, and their interaction on *TopCreativity* using the subsample of participants in the Numeric-Feedback condition and including those in the two additional No-Reward-Influence conditions. The conditions are comparable, as participants expect numeric feedback in the conditions where the evaluation does not influence rewards.

Table 1.4 Panel A reports the results. Consistent with H1.2, the interaction term is negative and significant. Results hold when including control variables. Table 1.4 Panel B shows the simple effect estimated from the model. Consistent with H1.2, the difference between AI and Human evaluators is significant when performance-based rewards are present and insignificant when they are absent.

[Table 1.4 here]

Finally, I check whether the absence of performance-based rewards reduces the time participants spend thinking about the criteria that the evaluator will adopt. Using ANOVA, I find a significant positive main effect of *NoRewardInfluence* on *ThinkCriteria* (difference in

²² Another alternative explanation is that participants in the AI evaluator condition with numeric feedback can successfully anticipate the assessment of their AI evaluator and generate ideas that are considered very creative by AIs but less creative by humans. To rule out this explanation, I ask Chat-GPT to evaluate all the ideas in the sample with a creativity score from 1 to 100. On average, the summary statistics indicate that participants in the Human Evaluator condition obtained higher creativity scores than those in the AI Evaluator condition (untabulated). Using ANOVA, the interaction term between the type of evaluator and feedback is insignificant at conventional statistical levels ($p > 0.7$), as well as the main effect of the type of evaluator ($p = 0.148$). As the main results change using Chat-GPT creativity scores, this evidence suggests that AI may evaluate creative ideas differently from human raters (the correlation coefficient between the Chat-GPT scores and *TopCreativity* is 0.518). However, these results do not support the alternative explanation that participants in the AI evaluator conditions understand the factors that drive these differences and generate ideas more likely to be evaluated favorably by an AI than by a human evaluator.

means = 0.511, $F(1,134) = 3.359$, $p = 0.069$). This test supports the expectation that the absence of performance-based rewards reduces participants' incentives to take the evaluator's perspective, thereby mitigating the negative effect of an AI evaluator on creative performance. While these results are consistent with the theoretical framework, they should be interpreted with caution, as I collect the data of the two additional conditions where the evaluation does not influence the rewards nine months after the data of the conditions where the evaluation influences rewards.

1.5 CONCLUSIONS

This paper investigates the impact of AI evaluators on employees' creative performance compared to human evaluators. In an online experiment, I find that the effect is contingent on the type of feedback generated by the evaluator and the presence of performance-based rewards. AI evaluators decrease creative performance compared to human evaluators when feedback is numeric and when the evaluation affects rewards. However, the effect becomes positive, but not significant at conventional statistical levels, when the feedback contains only discursive comments or when the evaluation does not influence rewards.

Results are consistent with the notion of perspective-taking. Anticipating the opinion of evaluators may be more challenging if individuals feel that they have different tastes and approaches to assessing creativity. AIs differ from humans in their use of quantitative methods, rather than acting based on instinct and feelings. Therefore, AI evaluators in a creative task may increase stress and make the task more unpleasant when numeric feedback reinforces the perception that the AI evaluators adopt a quantitative approach. Rewards contingent on creative performance push people to anticipate the evaluators' perspective, with a negative effect on creative performance when the evaluators are AI that generate numeric feedback.

This study has both theoretical and practical implications. It suggests that the presence of AI evaluators in creative tasks can potentially hamper creative performance, increase stress levels, and reduce intrinsic motivation. However, non-numeric feedback and the absence of performance-based rewards can mitigate these adverse effects. These insights broaden our understanding of the impact of AI on performance evaluation and provide suggestions for companies considering the implementation of AI to evaluate creative outputs. While this practice is currently limited, the technology is rapidly developing and has the potential to change how organizations select and evaluate creative ideas (Pagani and Champion 2023; Bell et al. 2024).

However, this study has several limitations. First, the experiment consists of a single round. In a multi-period setting, individuals may get familiar with their evaluator, which could reduce the effect of the evaluator's type over time. For the same reason, the effects of AI versus human evaluators and of the type of feedback are driven by participants' ex-ante perceptions rather than by real interactions with the evaluators. Second, the results may not hold in a situation where companies implement a hybrid system where a human evaluator is supported by AI tools. In the experiment, I treat AI and human evaluators as distinct extreme conditions. However, in practice, they may coexist (Yang et al. 2025). Third, sample sizes per condition are modest, diminishing the reliability of the statistical inference. Lastly, the implementation of AI for performance evaluation is a relatively recent phenomenon. As public opinion and perceptions of AI may change over time, the findings might not hold in the future. Future research can address these limitations by designing experiments that consider a multi-period setting and conditions with human evaluators assisted by AI, or by tracking how employees' perceptions of AI evaluators evolve over time.

1.6 REFERENCES

- Adler, P. S. and Chen, C. X. 2011. Combining creativity and control: Understanding individual motivation in large-scale collaborative creativity. *Accounting, Organizations and Society*, 36(2), 63–85.
- Alzubaidi, L., Zhang, J., Humaidi, A. J., Al-Dujaili, A., Duan, Y., Al-Shamma, O., Santamaría, J., Fadhel, M. A., Al-Amidie, M., and Farhan, L. 2021. Review of deep learning: Concepts, CNN architectures, challenges, applications, future directions. *Journal of Big Data*, 8, 1–74.
- Amabile, T. M. 1979. Effects of external evaluation on artistic creativity. *Journal of Personality and Social Psychology*, 37(2), 221.
- Amabile, T. M. and Pratt, M. G. 2016. The dynamic componential model of creativity and innovation in organizations: Making progress, making meaning. *Research in Organizational Behavior*, 36, 157–83.
- Ameen, N., Sharma, G. D., Tarba, S., Rao, A., and Chopra, R. 2022. Toward advancing theory on creativity in marketing and artificial intelligence. *Psychology and Marketing*, 39(9), 1802–25.
- Asay, H. S., Guggenmos, R. D., Kadous, K., Koonce, L., and Libby, R. 2022. Theory testing and process evidence in accounting experiments. *The Accounting Review*, 97(6), 23-43.

- Bailey, C. D. and Fessler, N. J. 2011. The moderating effects of task complexity and task attractiveness on the impact of monetary incentives in repeated tasks. *Journal of Management Accounting Research*, 23(1), 189–210.
- Bell, J. J., Pescher, C., Tellis, G. J., and Füller, J. 2024. Can AI help in ideation? A theory-based model for idea screening in crowdsourcing contests. *Marketing Science*, 43(1), 54–72.
- Binyamin, G. and Carmeli, A. 2010. Does structuring of human resource management processes enhance employee creativity? The mediating role of psychological availability. *Human Resource Management*, 49(6), 999–1024.
- Bisbe, J., and Otley, D. 2004. The effects of the interactive use of management control systems on product innovation. *Accounting, organizations and society*, 29(8), 709–737.
- Bol, J., LaViers, L., and Sandvik, J. 2023. Creativity contests: An experimental investigation of eliciting employee creativity. *Journal of Accounting Research*, 61(1), 47–94.
- Bonezzi, A., Ostinelli, M., and Melzner, J. 2022. The human black-box: The illusion of understanding human better than algorithmic decision-making. *Journal of Experimental Psychology: General*.
- Bruggen, A., Feichter, C., and Williamson, M. G. 2018. The effect of input and output targets for routine tasks on creative task performance. *The Accounting Review*, 93(1), 29–43.
- Byron, K., Khazanchi, S., and Nazarian, D. 2010. The relationship between stressors and creativity: A meta-analysis examining competing theoretical models. *Journal of Applied Psychology*, 95(1), 201.
- Cardinaels, E. and Feichter, C. 2021. Forced rating systems from employee and supervisor perspectives. *Journal of Accounting Research*, 59(5), 1573–607.
- Castelo, N., Bos, M. W., and Lehmann, D. R. 2019. Task-dependent algorithm aversion. *Journal of Marketing Research*, 56(5), 809–25.
- Chandler, S. 2020. Hollywood is using artificial intelligence to pick its next blockbuster. *Forbes*.
- Chen, C. X., Williamson, M. G., and Zhou, F. H. 2012. Reward system design and group creativity: An experimental investigation. *The Accounting Review*, 87(6), 1885–911.
- Cools, M., Stouthuysen, K., and Van den Abbeele, A. 2017. Management control for stimulating different types of creativity: The role of budgets. *Journal of Management Accounting Research*, 29(3), 1–21.
- Davila, A. and Dittillo, A. 2017a. Management control systems and creativity. In *The Oxford Handbook of Strategy Implementation*, 393–417. Oxford University Press.

- Davila, A. and Dittillo, A. 2017b. Management control systems for creative teams: Managing stylistic creativity in fashion companies. *Journal of Management Accounting Research*, 29(3), 27–47.
- Davila, A., Foster, G., and Oyon, D. 2009. Accounting and control, entrepreneurship and innovation: Venturing into new research opportunities. *European Accounting Review*, 18(2), 281–311.
- Davila, A. 2000. An empirical study on the drivers of management control systems' design in new product development. *Accounting, Organizations and Society*, 25(4–5), 383–409.
- Dickerson, S. S. and Kemeny, M. E. 2004. Acute stressors and cortisol responses: A theoretical integration and synthesis of laboratory research. *Psychological Bulletin*, 130(3), 355.
- DiSalvo, C. F., Gemperle, F., Forlizzi, J., and Kiesler, S. 2002. All robots are not created equal: The design and perception of humanoid robot heads. *Proceedings of the 4th Conference on Designing Interactive Systems*.
- Eapen, T. T., Finkenstadt, D. J., Folk, J., and Venkataswamy, L. 2023. How generative AI can augment human creativity. *Harvard Business Review*, 101(4), 56–64.
- Epley, N., Keysar, B., Van Boven, L., and Gilovich, T. 2004. Perspective taking as egocentric anchoring and adjustment. *Journal of Personality and Social Psychology*, 87(3), 327.
- Epley, N., Waytz, A., and Cacioppo, J. T. 2007. On seeing human: A three-factor theory of anthropomorphism. *Psychological Review*, 114(4), 864.
- Espeland, W. N. and Stevens, M. L. 2008. A sociology of quantification. *European Journal of Sociology*, 49(3), 401–36.
- Fabio, R. A., Picciotto, G., and Caprì, T. 2021. The effects of psychosocial and cognitive stress on executive functions and automatic processes in healthy subjects: A pilot study. *Current Psychology*, 1–10.
- Fessler, N. J. 2003. Experimental evidence on the links among monetary incentives, task attractiveness, and task performance. *Journal of Management Accounting Research*, 15(1), 161–76.
- Grabner, I. 2014. Incentive system design in creativity-dependent firms. *The Accounting Review*, 89(5), 1729–50.
- Hayes, A. F. 2017. *Introduction to Mediation, Moderation, and Conditional Process Analysis: A Regression-Based Approach*. Guilford.
- Hu, W. 2022. Trust versus rewards: Revisiting managerial discretion in incomplete contracts.
- Ian, L. 2023. Why AI's Creative Intelligence Will Change Advertising In 2024. *Forbes*.

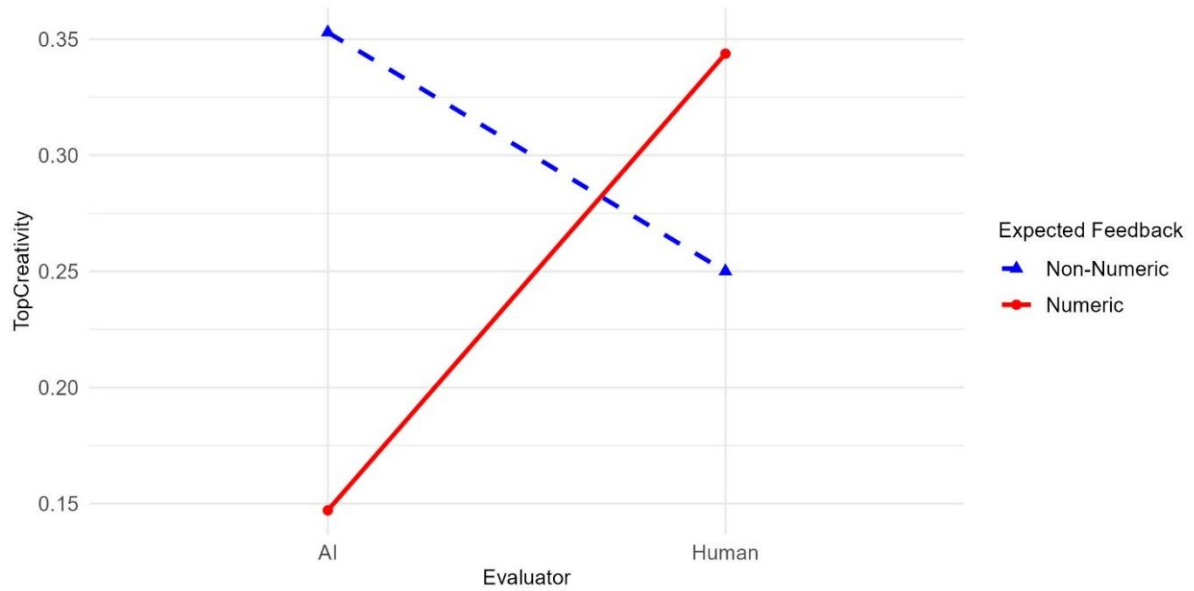
- Jia, N., Luo, X., Fang, Z., and Liao, C. 2023. When and how artificial intelligence augments employee creativity. *Academy of Management Journal* (ja).
- Jung, M. and Seiter, M. 2021. Towards a better understanding on mitigating algorithm aversion in forecasting: An experimental study. *Journal of Management Control*, 32(4), 495–516.
- Just, J., Strohle, T., Fuller, J., and Hutter, K. 2023. AI-based novelty detection in crowdsourced idea spaces. *Innovation*, 1–28.
- Kachelmeier, S. J., Reichert, B. E., and Williamson, M. G. 2008. Measuring and motivating quantity, creativity, or both. *Journal of Accounting Research*, 46(2), 341–73.
- Kachelmeier, S. J., Wang, L. W., and Williamson, M. G. 2019. Incentivizing the creative process: From initial quantity to eventual creativity. *The Accounting Review*, 94(2), 249–66.
- Kadous, K., Koonce, L., and Towry, K. L. 2005. Quantification and persuasion in managerial judgement. *Contemporary Accounting Research*, 22(3), 643–86.
- Klein, A., and Speckbacher, G. 2020. Does using accounting data in performance evaluations spoil team creativity? The role of leadership behavior. *The Accounting Review*, 95(4), 313–330.
- Longoni, C., Bonezzi, A., and Morewedge, C. K. 2019. Resistance to medical artificial intelligence. *Journal of Consumer Research*, 46(4), 629–50.
- McDaid, E., Andon, P., and Free, C. 2023. Algorithmic management and the politics of demand: Control and resistance at Uber. *Accounting, Organizations and Society*, 101465.
- Merchant, K. A. and der Stede, W. A. 2007. *Management Control Systems: Performance Measurement, Evaluation and Incentives*. Pearson.
- Mueller, J., Melwani, S., Loewenstein, J., and Deal, J. J. 2018. Reframing the decision-makers' dilemma: Towards a social context model of creative idea recognition. *Academy of Management Journal*, 61(1), 94–110.
- Murphy, K. R. and Davidshofer, C. O. 1988. *Psychological Testing: Principles and Applications*.
- Newman, D. T., Fast, N. J., and Harmon, D. J. 2020. When eliminating bias isn't fair: Algorithmic reductionism and procedural justice in human resource decisions. *Organizational Behavior and Human Decision Processes*, 160, 149–67.
- Ostermaier, A. and Uhl, M. 2020. Performance evaluation and creativity: Balancing originality and usefulness. *Journal of Behavioral and Experimental Economics*, 86, 101552.
- Pagani, M. and Champion, R. 2023. *Artificial Intelligence for Business Creativity*. Taylor and Francis.

- Parent-Rocheleau, X. and Parker, S. K. 2022. Algorithms as work designers: How algorithmic management influences the design of jobs. *Human Resource Management Review*, 32(3), 100838.
- Piercey, M. D. 2023. “Throw it in as a covariate?” Common problems using measured control variables in experimental research. *Auditing: A Journal of Practice and Theory*, 42(2), 183–205.
- Rose, S. 2020. ‘It’s a war between technology and a donkey’ – how AI is shaking up Hollywood. *The Guardian*.
- Samet, J., Williamson, M. G., and Yip, M. A. 2023. Leveling the playing field: AI-augmented design and the expertise bias in subjective evaluations of creative output. SSRN.
- Siemon, D. 2023. Let the computer evaluate your idea: Evaluation apprehension in human–computer collaboration. *Behaviour and Information Technology*, 42(5), 459–77.
- Speckbacher, G. 2021. Does performance evaluation kill creativity? A (re) interpretation of existing literature. *Pacific Accounting Review*, 33(1), 6-19.
- Stoel, M. D., Ballou, B., and Heitger, D. L. 2017. The impact of quantitative versus qualitative risk reporting on risk professionals' strategic and operational risk judgments. *Accounting Horizons*, 31(4), 53–69.
- Sundström, A. 2024. AI in management control: Emergent forms, practices, and infrastructures. *Critical Perspectives on Accounting*, 102701.
- Tong, S., Jia, N., Luo, X., and Fang, Z. 2021. The janus face of artificial intelligence feedback: Deployment versus disclosure effects on employee performance. *Strategic Management Journal*, 42(9), 1600–31.
- Votto, A. M., Valecha, R., Najafirad, P., and Rao, H. R. 2021. Artificial intelligence in tactical human resource management: A systematic literature review. *International Journal of Information Management Data Insights*, 1(2), 100047.
- Yang, C., Bauer, K., Li, X., and Hinz, O. 2025. My advisor, her AI, and me: Evidence from a field experiment on human–AI collaboration and investment decisions. *Management Science* (forthcoming).
- Ye, T., Furumi, F., da Silva, D., and Hamilton, A. 2021. Taking the perspectives of many people: Humanization matters. *Psychonomic Bulletin and Review*, 28, 888–97.

1.7 FIGURES

FIGURE 1.1

Effect of Evaluator and Feedback Type on Creative Performance

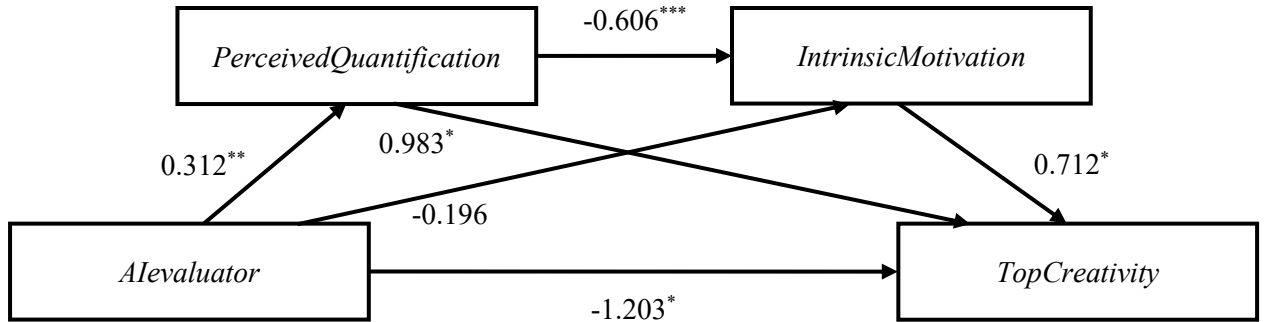


Mean of *TopCreativity* by experimental condition. *TopCreativity* is a dummy variable equal to 1 for participants submitting an idea with a top-quartile creative rating assigned from their group of raters, and 0 otherwise. The mean indicates the percentage of participants submitting a highly creative idea in the experimental condition.

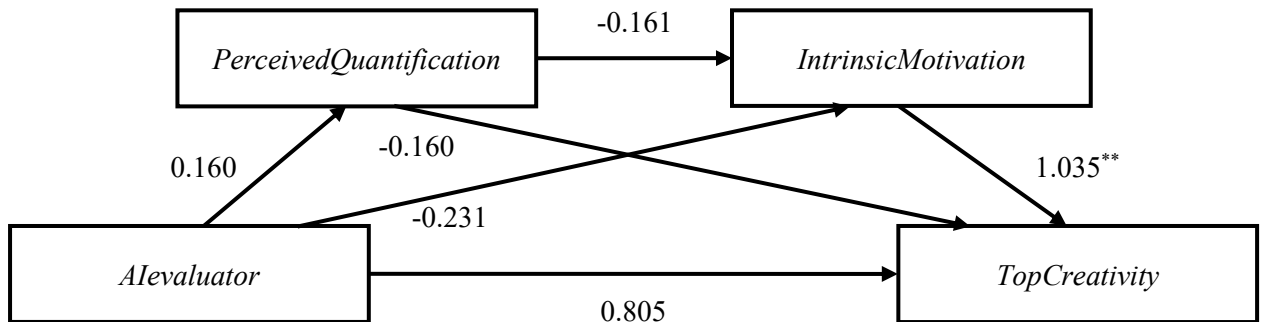
FIGURE 1.2

Perceived Quantification and Intrinsic Motivation

Panel A) Numeric Feedback Subsample



Panel B) Non-Numeric Feedback Subsample



Graphical representation of a process model estimated using Hayes (2022). Panel A and B consider, respectively, only the observations in the Numeric- and Non-numeric-Feedback conditions. The numbers on the arrows indicate the coefficients estimated from the following system of equations:

$$1) \text{ PerceivedQuantification} = \beta_1 \text{ Aevaluator} + \text{Controls} + \epsilon_1$$

$$2) \text{ IntrinsicMotivation} = \beta_2 \text{ PerceivedQuantification} + \beta_3 \text{ Aevaluator} + \text{Controls} + \epsilon_2$$

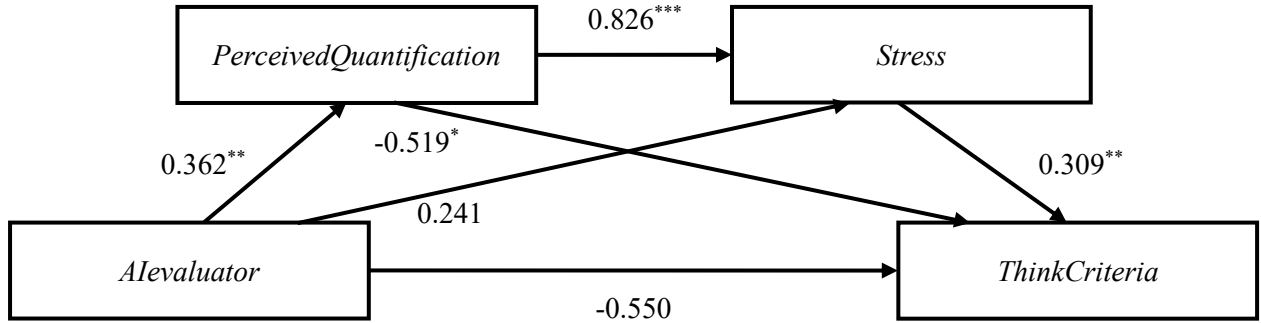
$$3) \text{ TopCreativity} = \beta_4 \text{ IntrinsicMotivation} + \beta_5 \text{ PerceivedQuantification} + \beta_6 \text{ Aevaluator} + \text{Controls} + \epsilon_3$$

Aevaluator is a dummy variable equal to 1 if the evaluator is an AI and 0 if it is a human. PerceivedQuantification is the sum of three items on the perception of quantification in the evaluator's approach. IntrinsicMotivation is a factor score related to intrinsic motivation in terms of task attractiveness. TopCreativity is a dummy variable that indicates a highly creative idea. Controls is a vector of variables including Female, Education, and HelpAI. The complete definitions of the variables are available in Appendix 1.5. The coefficients related to equation 3 are expressed in log-odds. * p<0.1; ** p<0.05; *** p<0.01 two-tailed.

FIGURE 1.3

Stress and Evaluator's Criteria

Numeric Feedback Subsample



Graphical representation of a process model estimated using Hayes (2022). This model is estimated on the subsample of observations in the Numeric-Feedback conditions. The numbers on the arrows indicate the coefficients estimated from the following system of equations:

1) $PerceivedQuantification = \beta_1 AIEvaluator + Controls + \epsilon_1$

2) $Stress = \beta_2 PerceivedQuantification + \beta_3 AIEvaluator + Controls + \epsilon_2$

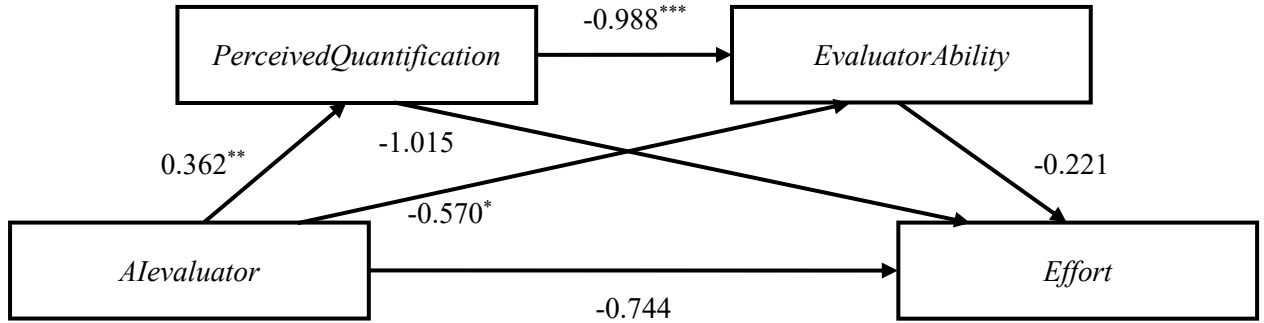
3) $ThinkCriteria = \beta_4 Stress + \beta_5 PerceivedQuantification + \beta_6 AIEvaluator + Controls + \epsilon_3$

AIEvaluator is a dummy variable equal to 1 if the evaluator is an AI and 0 if it is a human. *PerceivedQuantification* is the sum of three items about the perception of quantification in the evaluator's approach. *Stress* is a measure of self-reported stress. *ThinkCriteria* is an item on whether participants thought about the criteria that the evaluator would have used. *Controls* is a vector of variables including *Female*, *Education*, and *HelpAI*. The complete definitions of the variables are available in Appendix 1.5. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$ two-tailed.

FIGURE 1.4

Evaluator's Ability and Effort

Numeric Feedback Subsample



Graphical representation of a process model estimated using Hayes (2022). This model is estimated on the subsample of observations in the Numeric-Feedback conditions. The numbers on the arrows indicate the coefficients estimated from the following system of equations:

1) $PerceivedQuantification = \beta_1 AIEvaluator + Controls + \epsilon_1$

2) $EvaluatorAbility = \beta_2 PerceivedQuantification + \beta_3 AIEvaluator + Controls + \epsilon_2$

3) $Effort = \beta_4 EvaluatorAbility + \beta_5 PerceivedQuantification + \beta_6 AIEvaluator + Controls + \epsilon_3$

AIEvaluator is a dummy variable equal to 1 if the evaluator is an AI and 0 if it is a human. *EvaluatorAbility* is a self-reported measure of the evaluator's ability to evaluate creativity. *Effort* is the time spent writing the creative idea. *Controls* is a vector of variables including *Female*, *Education*, and *HelpAI*. The complete definitions of the variables are available in Appendix 1.5. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$ two-tailed.

1.8 TABLES

TABLE 1.1

Summary Statistics

	<i>N</i>	<i>Mean</i>	<i>sd</i>	<i>25th Pctl</i>	<i>Median</i>	<i>75th Pctl</i>
<i>TopCreativity</i>	204	0.279	0.450	0	0	1
<i>AEvaluator</i>	204	0.500	0.501	0	0.500	1
<i>NumericFeedback</i>	204	0.324	0.469	0	0	1
<i>NoRewardInfluence</i>	204	0.353	0.479	0	0	1
<i>Education</i>	204	6.324	1.401	5	7	7
<i>Female</i>	204	0.412	0.493	0	0	1
<i>HelpAI</i>	204	0.103	0.305	0	0	0
<i>PerceivedQuantification</i>	204	0.000	0.672	-0.415	0.082	0.547
<i>IntrinsicMotivation</i>	204	0.000	0.910	-0.505	0.182	0.703
<i>Stress</i>	204	3.076	1.600	2.000	3.000	4.500
<i>ThinkCriteria</i>	204	4.382	1.699	3	5	6
<i>EvaluatorAbility</i>	204	4.922	1.359	4	5	6
<i>Effort</i>	204	4.731	4.102	2.645	3.935	5.310

This table reports summary statistics for the key variables of the study. The definitions are available in Appendix 1.5. Each observation represents a participant in the experiment who generates a creative idea. Participants who fail attention checks and do not submit creative ideas are excluded from the sample.

TABLE 1.2

Creative Performance by Experimental Condition

	<u>AI Evaluator</u>			<u>Human Evaluator</u>			<u>Total</u>		
	<i>Mean</i>	<i>sd</i>	<i>N</i>	<i>Mean</i>	<i>SD</i>	<i>N</i>	<i>Mean</i>	<i>sd</i>	<i>N</i>
Non-Numeric Feedback	0.353	0.485	34	0.250	0.440	32	0.303	0.463	66
Numeric Feedback	0.147	0.359	34	0.344	0.483	32	0.242	0.432	66
No Reward Influence	0.324	0.475	34	0.263	0.446	38	0.292	0.458	72
Total	0.275	0.448	102	0.284	0.453	102	0.279	0.450	204

This table displays the mean, standard deviation, and number of observations of *TopCreativity* by experimental condition. *TopCreativity* is a dummy variable equal to 1 if the participant submits an idea with a top-quartile creative score assigned from their set of raters and 0 otherwise. The mean represents the proportion of participants submitting a highly creative idea in the experimental condition.

TABLE 1.3

Effect of Evaluator and Feedback Type on Creative Performance

Panel A: OLS and LOGIT models		<i>TopCreativity</i>			
		(1)	(2)	(3)	(4)
<i>AIevaluator</i>		0.103 (0.109)	0.093 (0.109)	0.492 (0.544)	0.469 (0.557)
<i>NumericFeedback</i>		0.094 (0.111)	0.080 (0.111)	0.452 (0.552)	0.412 (0.567)
<i>AIevaluator</i> × <i>NumericFeedback</i>		-0.300* (0.155)	-0.298* (0.154)	-1.604** (0.818)	-1.696** (0.840)
<i>Female</i>			-0.093 (0.083)		-0.531 (0.456)
<i>Education</i>			0.031 (0.028)		0.184 (0.156)
<i>HelpAI</i>			0.157 (0.121)		0.756 (0.584)
<i>Constant</i>		0.250*** (0.079)	0.079 (0.193)	-1.099*** (0.408)	-2.163** (1.095)
N		132	132	132	132
R ² or pseudo R ²		0.036	0.073	0.036	0.075
Model		OLS	OLS	LOGIT	LOGIT
Panel B: marginal (simple) effects		<i>AME</i>	<i>se</i>	<i>z</i>	<i>p</i>
Effect of AI vs. human evaluator:					
When feedback is non-numeric		0.093	0.109	0.855	0.393
When feedback is numeric		-0.205	0.109	-1.874	0.061
Effect of numeric vs. non-numeric feedback:					
When the evaluator is a human		0.080	0.111	0.724	0.469
When the evaluator is an AI		-0.218	0.107	-2.035	0.042

Panel A shows the estimation results of OLS and LOGIT models. The dependent variable *TopCreativity* is equal to 1 if the participant receives a top-quartile creativity score from their set of raters. *AIevaluator* is a dummy variable equal to 1 for the AI-Evaluator condition and 0 for the human one. *NumericFeedback* equals 1 for the Numeric-Feedback condition and 0 for the Non-Numeric-Feedback one. Standard errors in brackets. Panel B shows the marginal (simple) effects estimated from the model with control variables. AME stands for average marginal effect. The definitions of the other variables are available in Appendix 1.5. * p<0.1; ** p<0.05; *** p<0.01 two-tailed.

TABLE 1.4

Effect of Evaluator Type and Reward Influence on Creative Performance

<i>Panel A: OLS and LOGIT models</i>		<i>TopCreativity</i>			
		(1)	(2)	(3)	(4)
<i>AEvaluator</i>		-0.197* (0.109)	-0.210* (0.110)	-1.111* (0.611)	-1.223* (0.625)
<i>NoRewardInfluence</i>		-0.081 (0.106)	-0.057 (0.108)	-0.383 (0.524)	-0.258 (0.537)
<i>AEvaluator</i> × <i>NoRewardInfluence</i>		0.257* (0.151)	0.261* (0.151)	1.403* (0.802)	1.468* (0.814)
<i>Female</i>			-0.112 (0.079)		-0.613 (0.420)
<i>Education</i>			0.016 (0.028)		0.093 (0.147)
<i>HelpAI</i>			0.090 (0.127)		0.480 (0.622)
<i>Constant</i>		0.344*** (0.078)	0.278 (0.201)	-0.647* (0.372)	-1.223* (0.625)
N		138	138	138	138
R ²		0.029	0.050	0.030	0.052
Model		OLS	OLS	LOGIT	LOGIT
Panel B: marginal (simple) effects		<i>AME</i>	<i>se</i>	<i>z</i>	<i>p</i>
Effect of the type of evaluator (AI vs. Human):					
When the evaluation does not influence rewards		0.051	0.105	0.487	0.626
When the evaluation influences rewards		-0.210	0.110	-1.916	0.055
Effect of the influence on rewards (No vs. Yes):					
When the evaluator is a human		-0.057	0.108	-0.531	0.595
When the evaluator is an AI		0.204	0.109	1.879	0.060

Panel A shows the estimation results of OLS and LOGIT models. The sample consists of observations in the No Reward and in the Numeric-Feedback conditions. The dependent variable *TopCreativity* is equal to 1 if the participant receives a top-quartile creativity score from their set of raters. *AEvaluator* is a dummy variable equal to 1 for the AI-Evaluator condition and 0 for the human one. *NoRewardInfluence* equals 1 if the evaluation does not influence participants' rewards (No-Reward-Influence condition) and 0 otherwise (Numeric-Feedback condition). Standard errors in brackets. Panel B shows the marginal (simple) effects estimated from the model with control variables. AME stands for average marginal effect. The definitions of the variables are available in Appendix 1.5. * p<0.1; ** p<0.05; *** p<0.01 two-tailed.

1.9 APPENDICES

APPENDIX 1.1

Type of Feedback Manipulation

This appendix reports the instructions section related to the type of feedback manipulation. The Numeric-Feedback conditions is red, and the Non-Numeric conditions is blue. The purpose of these colors is to distinguish the text of the two conditions when visualized simultaneously. The text in the instructions that participants read on Qualtrics was black for both conditions.

Here are two examples of feedback generated for prior ideas. The topic of these ideas was to propose a creative alternative use for a business school's building. Your task will address a different problem, but the evaluator will write feedback following a similar template.

Idea ranked first: Create VR rooms in the building where students can visit companies around the world with VR technology.

- {This idea is very novel} {Novelty Score: 90/100} because there are not many VR rooms in universities or other places.
- {This idea is very useful} {Usefulness Score: 80/100} because this idea can benefit both students and companies, as they can interact and exchange ideas through VR.

Idea ranked second: Establish a museum that shows the history of the business school.

- {This idea is not very novel} {Novelty Score: 40/100} because it is a common practice to have a museum that showcases the history and achievements of an institution.
- {This idea is somewhat useful} {Usefulness Score: 60/100} because this idea can help students and visitors learn more about the university and its culture. However, this idea may not have a lot of impact or relevance for the current and future needs of the students and society.

APPENDIX 1.2

Type of Evaluator Manipulation

This is the portion of text in the instructions related to the evaluator type manipulation. The AI-Evaluator condition is in purple, the Human-Evaluator condition in green. The purpose of these colors is to distinguish the text of the two conditions when visualized simultaneously. The text in the instructions that participants read on Qualtrics was black for both conditions.

The evaluator is {a generative Artificial Intelligence (AI) based on a GPT language model}{another Prolific participant resident in an English-speaking country}. Then, {it}{he/she} is capable of interpreting English and writing feedback. Moreover, based on the content of your idea, {it}{he/she} can search related information on the internet to evaluate its novelty and usefulness.

APPENDIX 1.3

Examples of Copilot Evaluations

Here is my ranking of the ideas based on creativity (novelty and usefulness), along with a separate assessment for each criterion:

Idea 2: create a blockchain based application in the realm of web 3.0 using a token to purchase an nft to access it in a highly private and secure manner however after a kyc involving gdpr protections and adult consent and approval. following this the teenagers will be able to purchase a device to be worn based on the humane ai pin. this will leverage ai predictive models and language learning and use of smart sensors including gyro, speech analysis and parental secure recording, to detect and prevent alcohol abuse. following this its augmented reality headset component will be able to superpose notifications to the student on the microsoft hololens device to educate the teenagers and allow access for further education in the virtual reality environment of a custom platform in metaverse.

Novelty score: **100/100**. This idea is extremely novel because it uses a combination of cutting-edge technologies and concepts to address the issue of alcohol abuse among teenagers. It involves blockchain, web 3.0, nft, kyc, gdpr, humane ai, smart sensors, augmented reality, and metaverse, which are all very new and emerging fields. It creates a complex and sophisticated system that could provide a high level of privacy, security, detection, prevention, and education for the teenagers. This idea is not very common or widely used yet, and it could have some groundbreaking and revolutionary outcomes.

Usefulness score: **90/100**. This idea is very useful because it could potentially reduce the rate of alcohol abuse among teenagers by providing them with a state-of-the-art solution that could monitor their behavior, alert them of the risks, and educate them on the consequences. It could also increase their awareness and interest in the latest technologies and innovations, and help them develop digital skills and literacy. However, this idea may have some limitations, such as the cost and availability of the technologies and devices, the technical and legal challenges of the system, the compatibility and interoperability of the components, and the ethical and social implications of the intervention.

Idea 3: get real people to talk about problems using social media, young people feel this is not a problem, However alot of the stories about alcohol abuse are on traditional forms of media, older people do not post on tik tok, or snap chat - using these formats will help bring the problem to light, and highlight the problems that alcohol causes. they can also see what happened to peoples lives that drank to excess and the problems it has caused with these relationships with both partners, parents and children.

Novelty score: **75/100**. This idea is somewhat novel because it uses social media platforms to reach out to the teenagers and share real stories of alcohol abuse. It targets the platforms that are popular and relevant among the teenagers, such as tik tok and snap chat, and uses them to raise awareness and empathy. It also involves real people who have experienced or witnessed the problems of alcohol abuse, and allows them to communicate their messages and lessons.

However, this idea is not very original, as it borrows elements from existing programs and interventions that use social media or storytelling as a medium of education and prevention.

Usefulness score: **85/100**. This idea is useful because it could potentially reduce the rate of alcohol abuse among teenagers by providing them with credible and relatable information and advice from people who have faced the same issue. It could also increase their knowledge and awareness of the risks and consequences of alcohol consumption, and help them develop critical thinking and communication skills. It could also foster a sense of community and solidarity among the teenagers, and create a platform for dialogue and feedback.

Idea 1: Allow interactive zoom that meetings with former or current alcoholics so they can explain the damage alcohol can do to your health and life. Real world examples are the best to provide and since their problems may still be ongoing and could cover a variety of people from different backgrounds this would be easier for teenagers to relate to. Take for example a case where a wealthy person loses everything - their house, marriage and job because of alcohol and can explain this in person over a zoom session. This would be very compelling but also very practical to setup and arrange.

Novelty score: **65/100**. This idea is not very novel because it uses zoom meetings to connect the teenagers with former or current alcoholics who can share their experiences and insights. It allows the teenagers to hear first-hand stories of how alcohol affected their lives, and how they overcame or cope with their addiction. It also shows them that alcohol abuse can have serious and lasting consequences, regardless of their background or circumstances. However, this idea is not very innovative or original, as it is similar to other programs and interventions that use online or offline meetings or testimonies to influence behavior change.

Usefulness score: **75/100**. This idea is somewhat useful because it could potentially reduce the rate of alcohol abuse among teenagers by providing them with credible and relatable information and advice from people who have experienced the same problem. It could also increase their empathy and understanding of the challenges and struggles of alcohol addiction, and help them develop a positive attitude and a healthy lifestyle. However, this idea may have some limitations, such as the availability and suitability of the former or current alcoholics, the receptiveness and engagement of the teenagers, and the effectiveness and sustainability of the impact.

APPENDIX 1.4

Message to the Participants

AI Evaluator with Numeric Feedback

Dear participant, your idea was ranked 2 out of 3. Feedback message: "Novelty Score: 75/100. This idea uses a combination of incentives, tracking devices, and gamification to encourage teenagers to abstain from or moderate their alcohol intake. It creates a reward system that could appeal to the teenagers' interests and preferences, and also provides feedback and accountability through the monitoring device and the app. However, this idea is not very original, as it borrows elements from existing programs and interventions that use similar strategies. Usefulness Score: 85/100. This idea could potentially reduce the rate of alcohol abuse among teenagers by providing them with positive reinforcement and motivation to stay sober or drink responsibly. It could also increase their awareness of the health benefits of avoiding or limiting alcohol consumption, and help them develop self-control and discipline. However, this idea may have some limitations, such as the cost and availability of the incentives and the tracking devices, the privacy and ethical issues of monitoring the teenagers' blood alcohol levels, and the possibility of cheating or gaming the system."

AI Evaluator with Non-Numeric Feedback

Dear participant, your idea was ranked 3 out of 3. Feedback message: "This idea is not very novel. It uses an energy drink that tastes like wine but has zero degrees of alcohol. It provides a simple and convenient way for the teenagers to enjoy the taste of wine without the effects of alcohol. It also uses a medium that could be appealing and satisfying to the teenagers, and make them more willing to participate. However, this idea is not very innovative or original, as it borrows elements from existing products or beverages that use similar features or functions. This idea is somewhat useful. It could potentially reduce the rate of alcohol abuse among teenagers by providing them with a tool that could substitute or reduce their alcohol consumption. It could also increase their awareness and responsibility of the legal and social norms of alcohol consumption, and help them moderate their intake. It could also provide them with a tool that could help them enjoy their leisure time without alcohol, and access to entertainment and recreation."

Human Evaluator with Numeric Feedback

Dear participant, your idea was ranked 3 out of 3. Feedback message: "Novelty Score: 60/100. I do think it is a novel idea but i think for every 1 celebrity they see being reasonable with alcohol they will see 10 not being. Usefulness Score: 60/100. i do think it would be useful to some teenagers but certainly not all of them."

Human Evaluator with Non-Numeric Feedback

Dear participant, your idea was ranked 3 out of 3. Feedback message: "This idea is not novel. this sounds like what a policeman would do when breathalysing someone so it is not at all unique. This idea is not useful. it does not help the child as it just tells the child to use a breathalyser."

APPENDIX 1.5

Definitions

Variable	Definition
<i>TopCreativity</i>	Dummy variable equal to 1 if the participant's idea receives a creativity score equal to or above the highest quartile of the ideas having the same set of raters, 0 otherwise. I randomly assign the ideas to six set of raters. For each set, I recruited eight raters who independently evaluated the creativity of each idea in the set, assigning a score from 1 to 100. The creativity score of an idea is the average of the eight raters' scores.
<i>AEvaluator</i>	Dummy variable equal to 1 for participants assigned to the AI-Evaluator condition, 0 for the Human evaluator condition.
<i>NumericFeedback</i>	Dummy variable equal to 1 for participants in the Numeric-Feedback condition, where the evaluator provides numeric feedback, 0 otherwise.
<i>NoRewardInfluence</i>	Dummy variable equal to 1 for participants assigned the No-Reward-Influence condition, where the evaluation does not influence the participants' rewards, and 0 for those assigned to the conditions where the evaluation influences the participants' rewards.
<i>Education</i>	Answers to the question about the highest level of education obtained where 1 = "Some Primary"; 2 = "Completed Primary School"; 3 = "Some Secondary"; 4 = "Completed Secondary School"; 5 = "Vocational or Similar"; 6 = "Some University but no degree"; 7 = "University Bachelors Degree"; 8 = "Graduate or professional degree (MA, MS, MBA, PhD, JD, MD, DDS)".
<i>Female</i>	Dummy variable equal to 1 for participants who identify as females and 0 otherwise.
<i>HelpAI</i>	Dummy variable equal to 1 for participants who declare having used a generative AI to write their creative idea and 0 for the others.
<i>PerceivedQuantification</i>	First factor score of three items adapted from Newman et al. (2020) about the perception of quantification in the evaluator's approach (the wording in curly brackets depends on the evaluator condition): "With reference to the {Prolific participant}/{generative AI} evaluator: 1) I feel that {he/she}/{it} will just reduce my creative idea to a number. 2) I think that {he/she}/{it} will neglect some important elements of the

	creativity of my idea. 3) I believe that {he/she}/{it} will consider only the quantifiable elements of my creative idea". Participants answer from 1 ("Strongly Disagree") to 7 ("Strongly Agree").
<i>IntrinsicMotivation</i>	First factor score of items related to choices between a positive and a negative adjective about the attractiveness of the task, following Bailey et al. (2011). The adjective pairs are: Attractive-Repulsive, Exciting-Dull, Good-Bad, Interesting-Boring, Superior-Inferior, Wholesome-Unwholesome, Fun-Tedious. The higher the value, the more positive the adjectives selected.
<i>Stress</i>	Average score of the answers to the following items adapted from Cardinaels et al. (2021) (the wording in curly brackets depends on the evaluator condition): 1) "I felt nervous and stressed when developing the creative idea"; 2) "I feel nervous and stressed about the evaluation that the {generative AI}/{Prolific participant} will conduct." Participants answer from 1 ("Strongly Disagree") to 7 ("Strongly Agree").
<i>ThinkCriteria</i>	Answer to the question (the wording in curly brackets depends on the evaluator type condition): "While writing the creative idea, I was thinking about the criteria that the {generative AI}/{Prolific participant} evaluator will use." Participants answer from 1 ("Strongly Disagree") to 7 ("Strongly Agree").
<i>EvaluatorAbility</i>	Answer to the question (the wording in curly brackets depends on the evaluator type condition): "I think that the {generative AI}/{Prolific participant} evaluator can provide an adequate assessment of my idea in terms of creativity." Participants answer from 1 ("Strongly Disagree") to 7 ("Strongly Agree").
<i>Effort</i>	The number of characters in the creative idea divided by 100.

Appendix 6:

Quantile Regression

	<i>Creativity</i>						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
<i>AEvaluator</i>	0.035 (0.367)	0.417 (0.360)	0.473 (0.363)	0.292 (0.403)	0.224 (0.324)	0.265 (0.315)	0.419 (0.306)
<i>NumericFeedback</i>	-0.355 (0.414)	-0.141 (0.413)	0.039 (0.446)	0.141 (0.429)	0.441 (0.336)	0.374 (0.301)	0.348 (0.289)
<i>AEvaluator x NumericFeedback</i>	-0.130 (0.566)	-0.346 (0.557)	-0.498 (0.589)	-0.521 (0.531)	-0.935** (0.447)	-1.110** (0.434)	-1.117** (0.440)
<i>Female</i>	0.103 (0.300)	0.099 (0.308)	0.249 (0.302)	-0.025 (0.309)	-0.188 (0.263)	-0.261 (0.259)	-0.148 (0.265)
<i>Education</i>	0.016 (0.094)	0.031 (0.091)	0.004 (0.093)	0.01 (0.102)	0.004 (0.096)	0.028 (0.086)	0.019 (0.083)
<i>HelpAI</i>	0.211 (0.719)	0.137 (0.709)	0.800 (0.652)	0.613 (0.447)	0.836** (0.326)	0.821*** (0.288)	0.784*** (0.281)
<i>Constant</i>	-0.960* (0.562)	-0.993* (0.543)	-0.855 (0.580)	-0.152 (0.757)	0.389 (0.642)	0.493 (0.564)	0.600 (0.550)
Percentile	0.20	0.25	0.33	0.50	0.67	0.75	0.8

This table shows the results of a quantile regression estimation. The dependent variable *Creativity* is the average of the creativity scores assigned by eight raters to a creative idea from 1 to 100. The scores are standardized by group of raters to account for differences in the distribution of scores across raters. The percentiles represent the lowest quintile, the lowest quartile, the lowest tercile, the median, the highest tercile, the highest quartile, and the highest quintile. *AEvaluator* is a dummy variable equal to 1 for the AI-Evaluator condition and 0 for the human one. *NumericFeedback* equals 1 for the condition with numeric feedback and 0 for the non-numeric one. The definitions of the other variables are available in Appendix 1.5. The sample size is 132 observations. Bootstrapped standard errors with 5,000 replications in brackets. * p<0.1; ** p<0.05; *** p<0.01 two-tailed.

CHAPTER 2

Doing good for whom?

Mission statements, compensation interdependence, and employees' CSR commitment

with Eddy Cardinaels and Naomi Soderstrom

2.1 INTRODUCTION

Companies increasingly engage in prosocial actions and corporate Social Responsibility (CSR) initiatives for several reasons, including institutional pressure, regulatory requirements, and demand from various stakeholder groups (Aguilera et al. 2007; Brønn and Vidaver-Cohen 2009). CSR has become a strategic objective within companies and plays a more pivotal role in the design of Management Control Systems (MCSs), guiding employees towards more prosocial actions (Gond et al. 2012; Hansen and Schaltegger 2016; Flammer et al. 2019; George et al. 2023). This evolution is reflected in the corporate mission statements of both for-profit and non-profit organizations, which increasingly emphasize a commitment to social goals, emphasizing a broader set of stakeholders rather than solely focusing on customers or shareholders (Dhanesh 2020; Joly 2021; Grimes et al. 2020).

While mission statements are recognized for their role in external communication, their effect on employee behavior remains largely unexplored (Alegre et al. 2018; Mion et al. 2023). Some practitioners consider mission statements as useful managerial tools for motivating employees and guiding their priorities, suggesting that addressing multiple stakeholders in mission statements can affect employees' decisions (e.g., Craig 2018; Chandler 2023)²³. At the same time, prior studies show that companies may engage in activities that conflict with the social values highlighted in their mission statements (Braun et al. 2012; Grimes et al. 2019, 2020; Klein et al. 2021). This conflicting evidence raises questions regarding the effectiveness of mission statements in motivating employees to engage in CSR. We address this gap by investigating whether and under what conditions mission statements that emphasize a broad set of stakeholders (i.e., "stakeholder-oriented missions") are more effective at motivating

²³ For example, a consultant at Korn Ferry states that: "*When the entire organization is aligned to a clear set of written objectives that are visible to everyone, employees are better equipped to make the right decisions about how to allocate their time and effort. Managers are better able to set lower level objectives that support the overall goals of the company. Even the rank and file employees, when faced with a long list of to dos, will be able to choose which are most important based on the degree to which they support the mission.*" Available at: <https://www.kornferry.com/insights/featured-atopics/organizational-transformation/mission-statements-matter>.

employees to engage in prosocial actions that benefit external stakeholders (external CSR²⁴) compared to those that focus on investors (i.e., “shareholder-oriented missions”)²⁵. Specifically, we focus on the moderating role of compensation interdependence, defined as the extent to which an employee’s action influences the compensation of other employees (e.g., Lill 2020).

Besides investors, employees also have a vested interest in the company’s financial performance, as their compensation plans often tie their wealth to the company’s financial performance. Therefore, knowing how different mission statements interact with compensation interdependence is important. Interdependent compensation structures may affect employees’ perceptions about how their colleagues perceive their prosocial actions and, therefore, the decisions they tend to make in light of these perceptions. This tension between acting for employees or external stakeholders is especially relevant in today’s business environment (Douthitt et al. 2024), where organizations often face trade-offs between pursuing financial goals and more prosocial goals (Arya et al. 2019; Battilana et al. 2022).

In what follows, we theorize that the tension between pursuing external CSR at the expense of financial results intensifies when employee compensation is highly interdependent. Under such interdependence, employees who engage in external CSR may not only reduce their financial performance but also directly impact their colleagues’ bonuses, which are also tied to financial performance. Compared with low compensation interdependence settings, employees in highly interdependent systems may fear being blamed by their colleagues for prioritizing external stakeholders over profits (Raihani and Power 2021). In particular, under compensation interdependence, we predict that mission statements can guide employee behavior by reinforcing and communicating the values that top managers expect employees to follow (e.g., Christ et al. 2012; Christ 2013).

We argue that stakeholder-oriented mission statements are effective in motivating employees to pursue external CSR when compensation interdependence is high, but have less

²⁴ We distinguish between external CSR, which favors stakeholders external to the organization, and internal CSR, which favors internal stakeholders like employees. Our results may not extend to internal CSR, which may directly benefit the employees’ colleagues.

²⁵ For example, Blackstone Group emphasizes shareholder value creation on its website: “*Our mission at Blackstone is to fulfill our fiduciary duty by creating long-term value for our investors. We aim to do this by strengthening the more than 230 companies, ~12,500 real estate assets and other investments in our portfolio, equipping them to thrive amidst a changing global economy.*” On the contrary, Regions Bank mentions different groups of stakeholders: “*Regions’ mission is to achieve superior economic value for shareholders over time by making life better for customers, associates and communities and creating shared value as we help them meet their financial goals and aspirations.*”

effect when compensation interdependence is absent. Under high compensation interdependence, employees may anticipate negative judgments from their colleagues for pursuing external CSR, diminishing the positive self-image they would get from helping external stakeholders. This concern makes employees less likely to engage in external CSR compared to situations without compensation interdependence. As mission statements communicate top managers' expectations about the behavior they desire from employees (Feldman 1984), stakeholder-oriented missions signal that contributing to the well-being of external stakeholders is expected. Consequently, when compensation interdependence is present, stakeholder-oriented missions have a stronger potential to redirect employees' behavior toward external CSR compared to shareholder-oriented missions that still encourage profit maximization. When compensation interdependence is absent, mission statements have less impact. In this situation, employees can act more easily based on their personal preferences, and some may still prioritize external CSR over their own compensation to enhance their self-image (Bénabou and Tirole 2006).

To test our prediction, we conduct a 2×2 between-subjects experiment. We ask Prolific participants to imagine that they work for a fictional company where they need to choose to prioritize one of two projects. Prioritizing the financial project results in higher revenues for the company and higher pay for the participant, while prioritizing the external CSR project leads to benefits for the British Red Cross through monetary donations, but lower revenues and pay²⁶. We manipulate the type of mission statement and the presence of compensation interdependence. Participants in the Shareholder-Oriented-Mission condition read that the company's purpose is to maximize shareholders' wealth. In the Stakeholder-Oriented-Mission condition, the company's purpose also mentions supporting employees and local communities. When compensation interdependence is absent, the participants' decision to pursue the project that benefits external stakeholders only reduces their compensation. When compensation interdependence is present, the decision also affects a colleague's payout. This colleague is another randomly assigned Prolific participant facing the same decision problem, making their compensation interdependent.

In line with our theory, we find that a stakeholder-oriented mission statement, compared to a shareholder-oriented one, increases the likelihood of choosing the project that benefits

²⁶ We refer to this project as an external CSR one because it generates both revenues for the company and benefits to local communities where the firm operates through donations to the British Red Cross (the experimental participants are UK residents). Supporting local communities is often part of a company's CSR strategy. The choice between the projects reflects the trade-off employees may face when allocating effort between tasks that maximize financial performance versus tasks that also benefit external stakeholders, such as local communities.

external stakeholders when compensation interdependence is present. When compensation interdependence is absent, the effect of the mission statement's type is not significant. Additionally, consistent with the idea that employees are concerned about how their colleagues perceive them, we observe a negative effect of compensation interdependence on the likelihood of selecting the project that benefits external stakeholders in the Shareholder-Oriented-Mission condition. A planned contrast analysis supports these predictions.

We further test a process model to validate the role of the mission statement's type in altering the perceived social norms. We ask participants if they think that their colleagues would have a positive perception of them after choosing the project that benefits external stakeholders. Consistent with our arguments, we find that stakeholder-oriented missions improve the expected colleagues' perceptions when compensation interdependence is present. The expected colleagues' perceptions partially mediate the effect of stakeholder-oriented missions on the likelihood of selecting the project benefiting external stakeholders. We further document that compensation interdependence increases cognitive dissonance, thereby supporting our claim that it creates an ethical dilemma that threatens employees' self-image.

This study offers at least three important contributions. First, we contribute to the literature on the impact of mission statements on employee behavior (e.g., Alegre et al. 2018; Grimes et al. 2019, 2020; Deore et al. 2023). Even if companies can influence employees' values and beliefs through other systems, understanding the implications of mission statements' framing is critical because prior research and practice recommend formalizing a mission statement as a starting point to build corporate culture (e.g., George et al. 2023). We show that stakeholder-oriented mission statements significantly increase employees' willingness to pursue external CSR when compensation interdependence is present, but not significantly when it is absent. This contingent effect may partly explain the mixed evidence on the usefulness of mission statements in motivating employees to pursue CSR goals (Braun et al. 2012). In this context, we extend the work of Spallek et al. (2023) and Andrassy et al. (2024), who study, respectively, the effect of the substantiveness of CSR mission statements and of companies' CSR initiatives under different types of mission statements in settings without compensation interdependence. Our study provides new evidence that mission statements by themselves may not significantly affect employees' behavior, but other management control systems moderate their effectiveness. We show that mission statements can influence employees' decisions to a greater extent when features of the compensation system, such as compensation interdependence, make those decisions have direct consequences for other employees.

Second, we add to the literature on the role of MCSs in motivating employees' prosocial behavior and CSR. Although many prior studies focus on executives' decisions regarding CSR, lower-level employees can also contribute substantially to CSR performance by implementing the corporate strategy, participating in CSR initiatives, and proposing or pursuing new practices (Hejjas et al. 2019; Sendlhofer 2020; Heggen and Sridharan 2021). Given this role of employees, studying how MCSs help to align employees towards the CSR goals that companies want to pursue is crucial to avoid unintended consequences (e.g., Gond et al. 2012; Ditillo and Lisi 2016; Johnstone 2019; Beusch et al. 2022). A substantial body of research analyzes the effects of prosocial rewards on employees' motivation (e.g., Balakrishnan et al. 2011; Douthitt et al. 2022) or how MCSs affect employees' ethical decisions (e.g., Church et al. 2012; Maas and Van Rinsum 2013). However, to the best of our knowledge, little is known about how mission statements influence prosocial preferences and external CSR when there is conflict between favoring colleagues and external stakeholders. In particular, we find that mission statements, which provide guidance on what to prioritize, are more influential when compensation is interdependent.

Finally, we contribute to the literature on how MCSs interact with each other (e.g., Merchant and Otley 2006; Otley 1999). Prior literature acknowledges that MCSs act as a package, and it is important to study the overall systems rather than a single part (Grabner and Moers 2013; Bedford et al. 2016). While our experimental methodology does not allow us to manipulate the features of an entire MCS, we show evidence that cultural controls (mission statements) and features of incentive compensation systems (compensation interdependence) interact dynamically with important implications for practice. Our results suggest that firms with high compensation interdependence have more to gain by switching from a shareholder-oriented to a stakeholder-oriented mission statement, provided they consider external CSR as beneficial to their strategy. At the same time, if companies plan to adopt incentive schemes that increase compensation interdependence, they should carefully consider the implications for their external CSR initiatives, depending on their mission statement and strategy.

2.2. THEORETICAL FRAMEWORK AND HYPOTHESIS DEVELOPMENT

2.2.1 Mission Statements, Social Norms, and CSR

The management accounting literature acknowledges the role of values and beliefs as key drivers of motivation (e.g., Ouchi 1979). Within Simons' Levers of Control framework, belief systems are formal MCSs that top managers use to communicate and reinforce values that

employees are expected to align with (Simons 1995). These systems include mission statements and other tools like vision statements, credos, and statements of purpose. They are considered pivotal controls for defining company cultures, inspiring employees, and aligning them toward the company's long-term strategy (Mullane 2002; Chenhall et al. 2017).

Belief systems constitute one of the main levers that managers can use to influence employees' behavior, but they do not operate in isolation²⁷. Another important lever is diagnostic control systems. These systems provide information on the achievement of targets and critical success factors, which are also related to employees' performance evaluation and incentives (Simons 1995). The combined effect of the levers of control determines how companies manage the tension between the need to leave employees' freedom to embrace emerging opportunities and the need to avoid employees' taking excessive risks, including those related to integrating CSR into the company's strategy (Arjaliès and Mundi 2013; Kruis, Speklé, and Widener 2016). Belief systems exert a positive (enabling) force by encouraging employees to pursue emerging opportunities that align with the company's mission and values, while diagnostic control systems exert a negative (constraining) force by focusing attention on the measured performance indicators and the related targets (Simons 1995).

While diagnostic control systems influence employees by linking performance measures with evaluations and incentives, it is less clear why belief systems should influence employees' behavior. Social Norm Theory (Prentice and Miller 1993) provides a framework for explaining the motivational effect of belief systems, including mission statements. This theory posits that individuals tend to conform to the social norms in their environment and feel discomfort when they plan to deviate from them (Bernheim 1994; Robinson and O'Leary 1998; Blay et al. 2018). While various MCSs can provide cues about the presence of social norms and influence employee behavior (Tayler and Bloomfield 2011; Cardinaels and Yin 2015), belief systems play a peculiar role because they explicitly articulate the principles that firms expect their employees to follow (Feldman 1984).

Consistent with this argument, prior literature provides evidence that mission statements are instrumental in setting a tone at the top that encourages innovation (Amabile

²⁷ The Levers of Control Framework encompasses two other levels: boundary systems and interactive control systems. Boundary systems are formal rules that delimit the acceptable domain of activity (Simons 1995). An example of these systems is codes of conduct. Interactive control systems are monitoring systems characterized by frequent dialogue among top managers (Simons 1995). We do not theorize on the role of these systems, nor do we operationalize them in the experiment, because we assume that employees act within the allowed boundaries of their organization. Moreover, we assume that employees' decisions are observable by colleagues but not by managers. This assumption implies that managers cannot actively communicate with employees when they make the decisions, so that the performance measures used in compensation systems are diagnostic.

and Pratt 2016) and ethics within the company (Ostapski and Isaacs 1992; Blodgett et al. 2011). For example, statements advocating integrity norms reduce dishonest reporting in a budgetary setting (Deore et al. 2023), and mission statements emphasizing substantive CSR influence employees who already exhibit a strong preference for CSR (Spallek et al. 2023). Similarly, studies on value statements provide evidence that such statements can motivate performance, but their effect depends on the contract that firms use to motivate performance (Kachelmeier et al. 2016; Akinyele et al. 2020, 2022).

However, mission statements may not always be effective in motivating the desired behavior from employees (Mullane 2002). Companies may also implement other MCSs, such as diagnostic control systems, that can dampen the effect of mission statements. Targets and incentive compensation schemes may encourage employees to pursue activities that do not align with the company's purpose (Hyndman and Eden 2000). Moreover, mission statements may not always fit with the personal values and preferences of the employees (Chenhall 2017). Finally, when mission statements are prevalently used as an external communication device and are not endorsed by managers, employees may not feel motivated to follow them (Grimes et al. 2019, 2020).

In line with some of these criticisms, evidence about the impact of mission statements on company performance or CSR engagement is mixed (Dharmadasa et al. 2012; Alegre et al. 2018). A possible explanation is that mission statements interact with other MCSs. When the social norms that mission statements convey are implicit in other MCSs, mission statements may not produce an incremental effect on employee behavior (Deore et al. 2023). For example, firms can implement diagnostic controls, such as including CSR goals in budgets and tying monetary incentives to external CSR goals (e.g., Norris and O'Dwyer 2004; Lueg and Radlach 2016; Johnson et al. 2020). However, external CSR goals often come at the expense of other financial indicators that affect employees' pay (Church et al. 2019). In case of conflicting financial incentives or when employees are unsure about the right course of action, adopting a broader stakeholder-oriented mission may remind employees that CSR values are central to the business, motivating employees to pursue more CSR (Arjaliès and Mundi 2013).

Therefore, it is crucial to understand how such mission statements interact with other controls or incentive structures (Grabner and Moers, 2013; Bedford et al. 2016; Bedford 2020). In what follows, we study situations where individuals experience a tension between pursuing external CSR and financial performance. We propose that mission statements play a more pivotal role in guiding employees through this tension when pursuing external CSR affects the

financial wealth of other employees within the organization (i.e., when employees' compensation is interdependent).

2.2.2 Setting and Hypothesis Development

Following prior literature (e.g., Church et al. 2019; Spallek et al. 2023), we examine situations where employees must choose between an action that benefits external stakeholders (external CSR) and a profit-maximizing action that leads to higher compensation (i.e., employees' bonuses are tied to financial performance)²⁸. This decision requires employees to forgo potential monetary benefits in favor of a broader societal good. While self-interested individuals may always prioritize monetary gains, others may engage in prosocial behavior, even at personal financial cost (Bénabou and Tirole 2006). This is because prosocial preferences vary across individuals, with some employees being intrinsically motivated to pursue prosocial behavior even when it does not maximize their compensation (Shen and Benson 2016; Rupp et al. 2018; Church et al. 2019). Behavioral economics theories explain this behavior, proposing that individuals value maintaining a positive image of themselves (self-image), which also depends on how their behavior is perceived and supported by others (Bénabou and Tirole 2006).

The company's mission statement can influence how people in the organization perceive external CSR. As discussed in the previous section, mission statements signal the expected social norms and desired behaviors within an organization (e.g., Deore et al., 2023; Spallek et al., 2023). Stakeholder-oriented missions suggest that the company considers a broad range of stakeholders and the societal impact of its business decisions. These missions may support employees' decisions to prioritize external CSR, even at the expense of company profits. By contrast, shareholder-oriented missions focus more on investors and financial results and may not support external CSR at the expense of financial results.

We argue that the effectiveness of mission statements in influencing employees' behavior depends on the degree of compensation interdependence. When compensation interdependence is absent, employees can act more freely according to their own preferences for monetary gains over prosocial behavior. In this case, employees face a personal trade-off between maximizing their compensation and increasing their self-image through helping

²⁸ In some situations, it is possible that external CSR also maximizes financial performance and employees' compensation. This situation will not create any tension in the employee's decision, because the action that maximizes financial performance overlaps with the one that generates external CSR. Therefore, we assume a situation in which there is tension between maximizing financial performance and CSR.

external stakeholders. Since the colleagues who observe the decision are not affected by it, they may view external CSR positively regardless of the mission statements. Sacrificing compensation to help others might be seen as a virtuous prosocial act by the colleagues, thereby having a positive effect on self-image. Therefore, we expect mission statements to have a low effect on employees' behavior when compensation interdependence is absent.

By contrast, under high compensation interdependence, employees encounter a stronger ethical dilemma. When compensation interdependence is present, pursuing external CSR can reduce colleagues' bonuses, and colleagues may blame the employees for their choice. This concern is consistent with previous research findings that people may view prosocial actions negatively if they impose costs on others (Raihani and Power 2021), especially when individuals prioritize people outside their social group over insiders (Law et al. 2022; McManus et al. 2020). For this reason, employees may anticipate blame and a sense of guilt for reducing their colleagues' bonuses, which threatens their self-image (Battigalli and Dufwenberg 2007, 2009). Therefore, under compensation interdependence, people who value prosocial behavior more than monetary gains may be hesitant to pursue external CSR because they are uncertain about how their colleagues view their choice and may anticipate negative judgment and a potential sense of guilt.

In this situation, mission statements can clarify the (injunctive) social norm within the organization, indicating how employees are expected to behave²⁹. Under a shareholder-oriented mission, employees may be more concerned that pursuing external CSR will be viewed negatively by their colleagues because they expect that people in the organization act to maximize performance. Therefore, the benefit that CSR provides to self-image may be lower than when compensation interdependence is absent. On the other hand, a stakeholder-oriented mission signals that the firm desires actions that benefit a large array of stakeholders. In other words, the mission statement legitimizes employees' prioritizing the needs of external stakeholders over financial performance. Therefore, this mission may alleviate concerns that external CSR will be perceived negatively by the colleagues, even if it reduces their compensation, because pursuing external CSR is a behavior congruent with the norms of the organization. Then, the benefit that CSR provides to employees' self-image may be similar to

²⁹ Prior research on social norms distinguishes between injunctive and descriptive norms. Injunctive norms arise from what people approve, while descriptive norms arise from what people actually do (Smith and Louis 2008). In our theoretical framework, mission statements serve as cues of injunctive social norms as they communicate the desired behavior from the managers of the organization.

the case in which compensation interdependence is absent, where employees do not expect negative judgment from their peers for pursuing external CSR.

In sum, we predict an ordinal interaction, where we expect that the type of mission has a strong impact on motivating external CSR under compensation interdependence, but a lower impact when compensation interdependence is absent³⁰. When compensation is interdependent, employees may be less likely to pursue external CSR under a shareholder-oriented mission, whereas a stakeholder-oriented mission may mitigate the concerns of being viewed negatively by colleagues, leading them to pursue more external CSR. However, in the absence of compensation interdependence, mission statements may affect employees' decisions only to a limited extent, as employees who prefer pursuing external CSR over monetary gain may not anticipate negative judgment from their colleagues even under a shareholder-oriented mission. Formally:

When compensation interdependence is present, a stakeholder-oriented mission statement increases employees' likelihood of pursuing external CSR compared with
H2.1: *a shareholder-oriented mission. However, when compensation interdependence is absent, the type of mission statement has a more limited effect on employees' decisions.*

2.3 EXPERIMENTAL DESIGN

We conduct an online experiment to test our hypothesis H2.1. Before data collection, we obtain approval from the academic institution's ethics committee (i.e., institutional review board). We recruit 200 participants from Prolific³¹ and administer the experiment using Qualtrics. Participants receive a participation fee of 1.2£ for a 10-minute study. They can also earn a variable bonus ranging from 0.4£ to 0.8£, depending on their decisions during the experiment. Moreover, they can generate a donation of 0.4£ to the British Red Cross. The total donation for

³⁰ This prediction is not straightforward for at least two reasons. First, compensation interdependence reduces the individual cost of pursuing external CSR, as employees can shift part of the financial burden onto their colleagues. This could increase rather than decrease the willingness to pursue external CSR, potentially weakening the effect of stakeholder-oriented missions. Second, mission statements alone may not strongly influence behavior, as they do not impose explicit penalties or rewards. If employees' personal values outweigh the norms conveyed by mission statements, these statements may have a negligible impact even when compensation interdependence is present.

³¹ We filter our participants based on the following criteria: English as their primary language, residence in the UK, previous approval rate above 95 percent, and at least 30 completed tasks. One participant did not submit a response, resulting in 199 paid participants. Our sample consists of 35% males with an average age of 40 years.

the British Red Cross is 33.6£, and the total monetary bonus for the participants is 125.6£. The median completion time is 8:19 minutes.

2.3.1 Procedure

In the experiment, participants play the role of employees in a fictional advertising company. First, they individually select the name of the company they want to work for: “TEAM WHITE” or “TEAM BLACK.” This decision does not affect any subsequent features of the experiment. It aims to create a sense of identification with the fictional advertising company and with “the colleagues,” who are the other participants in the same experimental condition who choose the same company name. Following prior studies that prime artificial group identities, we use colors to name the teams and a trivial preference question to assign participants to them (e.g., Tajfel and Billig 1974; Charness et al. 2007).

After selecting the company name, participants read the instructions, which also describe the company’s mission statement. Participants then learn that they must decide which of two alternative projects to prioritize³². “Project Red Cross” consists of managing the social media presence of the British Red Cross. As this activity increases the visibility of the British Red Cross, it generates a donation of 0.4£ to the British Red Cross, which thus produces benefits for external stakeholders (external CSR). Alternatively, “Project Furniture” involves managing the marketing campaign of a furniture manufacturer. It provides more revenue for the company than “Project Red Cross” but does not imply any donation. As the fictional advertising company links employees’ compensation to the revenues their marketing campaigns generate, “Project Furniture” also maximizes the monetary bonus of the participants³³. Prioritizing “Project Furniture” generates a total bonus of 0.8£ for the employees, while “Project Red Cross” generates 0.4£ for the employees and 0.4£ for the British Red Cross³⁴.

³² We decide to ask participants to express their preference for the project rather than using a real effort task to better measure their intentions. Our theory predicts that employees will have different intentions; therefore, measuring them directly provides a more immediate test of the theory than measuring effort. We also decide to link these decisions to real monetary bonuses and donations to limit desirability bias. Since participants who prioritize Project Red Cross sacrifice real monetary bonuses to provide real donations, their decision can be considered prosocial, as it entails a personal cost and a real benefit for other people.

³³ In real situations, it is possible that external CSR also increases employees’ compensation (e.g., CSR-based incentives). We decided to adopt this stylized design to provide a clear setting for testing our theory. Our results do not extend to situations where CSR maximizes both employees’ compensation and benefits for local communities. Future research can explore this win-win scenario.

³⁴ We keep a positive variable bonus in both choices because the project that generates external CSR should generate some revenues for the company. Our experiment resembles a situation in which pursuing external CSR yields lower bonuses for the employees, but it is not a pure depletion of the firm’s resources (i.e., pursuing external CSR increases financial performance, but there exists an alternative decision that increases it more). This way,

After completing a series of attention checks, participants must choose the project to prioritize. They decide individually and anonymously, but one randomly assigned colleague will know their decision at the end of the experiment. On the Qualtrics web page, they observe a table that summarizes the payoffs generated by each project, the company's mission statement, and two buttons to click corresponding to the two projects (see Appendix 2.A). Next, they complete a series of post-experimental questions and submit the survey.

After collecting the answers, we match the participants with one colleague and compute the payoffs of both parties. Then, we use a private Prolific message to reveal the participants' decisions to their colleagues to whom they are matched. Importantly, we reveal the decision of the participants in every condition to a colleague, and the instructions explicitly mention this, to make the conditions comparable. Otherwise, participants in the conditions with compensation interdependence could infer the decision of the colleague from the payoff, while participants in the conditions without compensation interdependence could not. Participants do not take any subsequent action after receiving their compensation and learning about their colleague's decision to rule out reciprocity mechanisms, which are not the focus of the theory.

2.3.2 Manipulations

The research design is a 2×2 between-subjects experiment. Our Qualtrics survey randomly assigns participants to different conditions after they agree to the information given on the form of consent. We manipulate the type of mission statement (Shareholder-Oriented vs. Stakeholder-Oriented) and the degree of compensation interdependence (Present vs Absent).

For the type of mission statement, the participants read different information at the beginning of the instructions. The Shareholder-Oriented-Mission condition focuses only on investors and reports the mission of the company as follows: *"Our aim is to make money for our shareholders for which we do business and achieve profitable growth."* Instead, the Stakeholder-Oriented-Mission condition mentions that: *"Our aim is to make a positive difference for the world, allowing our employees, our shareholders, and all the communities in which we live to prosper."* The stakeholder-oriented mission includes, next to shareholders, the groups of employees and communities at large, which we borrowed from reading some real mission statements. We mention shareholders in both conditions to emphasize their importance

the project leading to external CSR is consistent with a stakeholder-mission because it benefits multiple stakeholders. It generates some revenues for the company, some bonuses for the employees, and some benefits for local communities by supporting the activities of the Red Cross.

in a for-profit organization and ensure that participants view profit maximization as a viable option, even with a stakeholder-oriented mission statement.

For the degree of compensation interdependence, we manipulate how the payoffs of the decision are distributed among other participants in the same condition, referred to as colleagues. Participants receive clear instructions about how pay is distributed, including the pay-off table of the condition to which they are assigned (see Appendix 2.A). In the Compensation-Interdependence-Absent condition, participants obtain the whole 0.8£ if they select Project Furniture or 0.4£ if they select Project Red Cross, and the decision does not affect any colleague's payoffs. In the Compensation-Interdependence-Present condition, the payoff generated by the decision is split with a colleague. A colleague also makes a similar decision and can influence the participant's compensation, creating interdependence³⁵ (e.g., Lill 2020). Our design choice to distribute a monetary donation to the British Red Cross and a payoff to the colleague in the Compensation-Interdependence-Present condition aims to make the ethical dilemma more salient for the participants compared to a purely based scenario experiment and mitigate the concern of social desirability bias. Importantly, the maximum and minimum possible payoffs across conditions are equivalent.

2.4 RESULTS

2.4.1 Main Results

Table 2.1 reports summary statistics for all the variables of the study. Our final sample consists of 197 participants, after excluding two observations failing one of the attention checks and one participant who did not submit a response and a completion code. We also collect additional demographic variables with our post-experimental questions: age, gender, and education (defined in Appendix 2.B). The main dependent variable of interest is *External_CSR*, which equals 1 if participants select Project Red Cross and 0 if they select Project Furniture. This variable captures the willingness to prioritize external CSR over financial results. Table 2.1 provides summary statistics of the variables of interest. Table 2.1 includes a breakdown of *External_CSR* by experimental condition.

³⁵ The experiment is single-round, removing the possibility of expecting reciprocal actions from the colleague after the choices are revealed. In addition, participants in every condition know when reading the instructions that their colleague will know their decision at the end of the experiment. This design feature aims to keep the information that the monetary bonus reveals constant across conditions. Otherwise, participants in the Compensation-Interdependence-Present condition could easily infer their colleagues' decisions based on the bonus they receive.

[Table 2.1 and Table 2.2 around here]

Given that the main dependent variable of the experiment is binary, we estimate the following logit model (Logit Model 2.1):

$$\begin{aligned}
 Prob(External_CSR_i = 1) = & \\
 & = f(\beta_1 Compensation_Interdependence_i \\
 & + \beta_2 Stakeholder_Mission_i \\
 & + \beta_3 Stakeholder_Mission_i \times Compensation_Interdependence_i \\
 & + \alpha + \epsilon_i)
 \end{aligned} \tag{2.1}$$

The independent variables are dummy indicators. *Stakeholder_Mission* equals 1 for participants assigned to the Stakeholder-Oriented-Mission condition and 0 for the Shareholder-Oriented-Mission condition. *Compensation_Interdependence* equals 1 for the Compensation-Interdependence-Present condition and 0 for the Compensation-Interdependence-Absent condition.

Table 2.3 Panel A reports the result. The coefficient *Compensation_Interdependence* is negative and significant (coefficient = -0.821, p-value < 0.10), indicating that the presence of compensation interdependence decreases the likelihood of prioritizing external CSR³⁶. The effects of the type of mission statement and the interaction term are not significant at conventional statistical levels.

[Table 2.3 around here]

To test H2.1, we use a planned contrast as it should provide greater statistical power than other methods for testing ordinal interactions in 2×2 experimental designs (Buckless and Ravenscroft 1990). Under the presence of compensation interdependence, we assign a weight of -3 to participants in the Shareholder-Oriented-Mission condition, as we expect people to prioritize the external CSR much less compared to when they would receive the stakeholder-oriented mission, for which we assign a weight of +1. The other two conditions, where compensation interdependence is absent, get weights of +1, regardless of the mission statement, as we expect the type of mission statement to have a lower impact and a considerable number of individuals to have intrinsic preferences for external CSR over profit maximization. Table 2.3 Panel B reports the results of a Chi-Square test of the contrast. This test is significant and supports H2.1 (p-value < 0.01)³⁷. Figure 2.1 represents the predicted probabilities of

³⁶ We find analogous results after including age, education, and gender as control variables (untabulated).

³⁷ Prior research acknowledges that accepting an ordinal interaction hypothesis based solely on a significant contrast test increases the risk of a Type I error, especially when the interaction term in the linear model is not

prioritizing external CSR per condition based on logit model 2.1. The pattern is visually consistent with the planned contrast and our hypothesis test.

[Figure 2.1 around here]

We now report the effect of mission statement type under the different compensation interdependence conditions (simple effects). The marginal effects reported in Panel C are consistent with H2.1. We find that, when compensation interdependence is present, stakeholder-oriented missions increase the likelihood of prioritizing the external CSR by 21.6 percent ($p\text{-value} < 0.05$) compared to shareholder-oriented missions. Conversely, the effect of the type of mission statement is not significant at conventional levels when compensation interdependence is absent ($p\text{-value} = 0.355$). Note that compensation interdependence (compared to no interdependence) also decreases the likelihood of prioritizing external CSR by 18.3 percent under the shareholder-oriented mission ($p < 0.10$). Consistent with our theory, this finding confirms that participants are more hesitant to pursue external CSR when their decisions affect their colleagues' pay.

Consistent with H2.1, the results show that stakeholder-oriented missions mitigate the adverse effect of compensation interdependence on the willingness to pursue external CSR activities. When compensation interdependence is absent, the type of mission statement has a limited effect, as it does not significantly influence participants' choices.

2.4.2 Process Evidence

This section provides additional evidence of the process described in the theoretical framework. We theorize that employees expect their colleagues to have a negative opinion of them if they prioritize external CSR over financial results when their compensation is interdependent. Expecting a negative opinion from colleagues should decrease the positive effect on self-image that people perceive from pursuing external CSR. However, when there is a stakeholder-oriented mission, the social norm implied by the mission signals that helping external

significant (Guggenmos et al., 2018). To mitigate this concern, we use the procedure proposed by Guggenmos et al. (2018). For this purpose, we estimate our linear model with ANOVA rather than LOGIT (untabulated). The p -value of the planned contrast using ANOVA is 0.005 ($F(1,193) = 8.05$). The residual between-cell variance is not significant at conventional levels ($F(2,193) = 0.440$, $p\text{-value} = 0.645$), and the proportion of between-cell variance unexplained by the contrast (q^2) is 0.108. These tests together suggest that the planned contrast explains most of the variation of *External CSR* attributable to the experimental conditions. Together with the visual assessment of Figure 2.1, they support our claim that the data fit the predicted pattern implied by the planned contrast. However, the results of these tests should be interpreted with caution, as our dependent variable is binary, while Guggenmos et al. (2018) validate their procedure with continuous dependent variables.

stakeholders is desirable for the firm. Under such a mission, employees may expect their colleagues to think more positively about their behavior when they pursue external CSR. Hence, they will be more likely to prioritize external CSR over financial results.

In the post-experimental questions, we directly measure the extent to which participants expect their colleagues to view them positively if they prioritize external CSR over financial results. We asked participants to imagine they would choose Project Red Cross (external CSR) over Project Furniture (financial results), and that their colleagues would be able to see this decision. For this scale, we adapt three Likert scale items from Hobson et al. (2021). We ask participants how much they agree (on a 1 to 7 scale) that their colleagues would view them as (1) benevolent, (2) caring, and (3) empathetic for choosing Project Red Cross (see Appendix 2.B for further details). The variable *Colleague_Perception* is the score of the first factor of these three items.

Using *Colleague_Perception* as the mediator, we estimate a mediation model using the bootstrap approach in Hayes (2022) with 5,000 replications³⁸. *External_CSR* is our key main dependent variable, and *Stakeholder_Mission* is the independent variable. We estimate this mediation model separately for the participants in the Compensation-Interdependence-Present and Compensation-Interdependence-Absent conditions. We do this because we expect that people's perceptions of their colleagues significantly differ depending on the type of mission statement they see, only when compensation is interdependent.

Figure 2.2A shows the estimated path coefficients for the Compensation-Interdependence-Present condition. As predicted, the type of mission statement has a strong effect: the effect of the stakeholder-oriented mission (relative to the shareholder-oriented mission) on *Colleague_Perception* is significantly positive (coefficient = 0.326, p-value < 0.10, untabulated). Second, *Colleague_Perception* has a significant positive effect on the likelihood of prioritizing the external CSR (coefficient = 0.532, p-value < 0.05, untabulated). For the Compensation-Interdependence-Absent condition, we do not expect mission statements to affect this perception, as people can act more freely on their preferences (caring less about what others think). Indeed, the effect of *Stakeholder_Mission* on *Colleague_Perception* is not significant when compensation interdependence is absent (Figure 2.2B: p-value = 0.290). Note that the effect of *Colleague_Perception* on the likelihood of pursuing external CSR remains positive and significant (Figure 2.2B: coefficient = 0.573, p-value < 0.05).

³⁸ We obtain analogous results using 10,000 replications (untabulated).

[Figure 2.2A and 2.2B around here]

For the compensation interdependence condition, the results in Table 2.4 show a positive indirect effect of *Stakeholder_Mission* on the likelihood of pursuing the external CSR through *Colleague_Perception* (90 percent bootstrapped confidence interval: [0.002; 0.412])³⁹. Further consistent with our theory, the indirect effect is again not significant in the absence of compensation interdependence (90 percent bootstrapped confidence interval: [-0.318; 0.120]).

[Table 2.4 around here]

The process evidence supports the idea that the fear of negative perceptions from colleagues can explain the lower likelihood of pursuing external CSR when compensation is interdependent under a shareholder-oriented mission. Expecting a more favorable opinion from the colleagues, which shareholder-oriented missions can induce, can enhance participants' willingness to pursue external CSR. In the absence of compensation interdependence, the type of mission statement does not significantly affect this perception.

Another important aspect of our theory is that compensation interdependence creates tension between acting in favor of colleagues and external stakeholders. This tension may cause discomfort in the minds of employees, who remain uncertain about what course of action is best to pursue from an ethical perspective. When compensation interdependence is present compared to when it is absent, we expect employees to experience higher cognitive dissonance (Festinger 1957), defined as the discomfort that people feel when dealing with contradictory cognitions⁴⁰. We presume that the effect is independent of the mission statement, such that compensation interdependence increases cognitive dissonance for both types of mission statements. Mission statements do not reduce this dissonance, but can offer guidance on what to pursue when cognitive dissonance is high.

To measure the participants' cognitive dissonance, we adapt seven items from Koller and Salzberger (2007). We ask participants about how confident they are and how much discomfort they face when deciding on the project they want to prioritize (see Appendix 2.B for the complete question). As expected, all the items significantly load on a single factor (eigenvalue > 1). Our variable *Cognitive_Dissonance* is equal to the first factor's score.

³⁹ Following Hayes (2022), we consider the effect significant at 90 percent level as the confidence interval does not include 0.

⁴⁰ Prior research theorizes that people experience cognitive dissonance when their positive self-image is threatened (Steele 1988; Simon et al. 1995). This situation may occur when people should make decisions that contrast with their personal beliefs or values (McGrath 2017; Hinojosa et al. 2017).

We conduct an ANOVA with *Cognitive Dissonance* as the dependent variable and the same independent variables as in logit model 2.1. Consistent with our expectations, we find a significant main effect of compensation interdependence ($F(1,193) = 5.48$; $p\text{-value} = 0.020$, untabulated), suggesting that people experience more discomfort when compensation interdependence is present than when it is absent, regardless of the mission statement. This result supports our prediction that compensation interdependence induces discomfort when pursuing external CSR, as people face a dilemma between contributing to society and caring about their colleagues' compensation.

2.5 CONCLUSIONS

We study whether stakeholder-oriented mission statements motivate employees to prioritize external CSR over financial results in a context where maximizing financial results also lead to higher employee compensation. We find that the effect of the type of mission statements depends on the degree of compensation interdependence. When compensation interdependence is absent, the type of mission statement does not affect employees' choices, as people can confidently act on their own preferences when pursuing external CSR. When compensation interdependence is present, stakeholder-oriented mission statements can motivate engagement in external CSR compared to shareholder-oriented missions. We expect that a stakeholder-oriented mission statement helps to establish social norms, indicating that the firm also values its external stakeholders. These social norms that stakeholder-oriented missions can induce diminish expectations of negative opinions from colleagues that people otherwise may have about pursuing external CSR, which decreases their monetary wealth, motivating them to pursue more external CSR.

Our findings contribute to the literature on the impact of mission statements on employees' behavior. Extending the work of Deore et al. (2023) and Spallek et al. (2023), we provide new insights regarding the mixed evidence on the effectiveness of mission statements and other belief systems in influencing people in organizations. Our results offer a rationale for these mixed findings, suggesting that the effects of mission statements depend on their interaction with other MCSs. We find that compensation interdependence is an important contextual variable for the effectiveness of mission statements when employees face a trade-off between acting for financial or prosocial goals.

While mission statements may appear to offer cheap talk, our findings suggest they can be particularly effective in organizations with high compensation interdependence. Under high compensation interdependence, a shareholder-oriented mission statement can lead employees

to neglect external stakeholders and place a strong emphasis on financial results, potentially harming companies in their pursuit of long-term-oriented CSR goals. Therefore, companies with high compensation interdependence may need to consider switching to a stakeholder-oriented mission statement, proving that external CSR supports their strategy. Moreover, our results suggest that companies that plan to implement incentive schemes that increase compensation interdependence should take into account their current mission statements and the potential consequences these mission statements have on their employees' willingness to engage in external CSR.

Our study is subject to some limitations. First, we test our theory in a clean and abstract setting, where employees read and reflect on the company's mission statement. It is worth considering that, in a real workplace, employees may adhere strongly to these mission statements but may also develop social norms that can diverge from them. Deviations from formal mission statements may arise if employees do not have a strong recall of their company's mission statement. For these reasons, prior research suggests that managers should endorse mission statements to shape an organization's culture and to communicate the values to lower-level employees (Alegre et al. 2018). In addition, we manipulate the type of mission statement such that there is a clear contrast between shareholder and stakeholder missions. However, in practice, mission statements can mix different values or be ambiguous. Future studies may test our theory in settings where there is a variation in the degree of managerial endorsement of the mission statement or in situations where mission statements are ambiguous or absent.

Second, we test our theory in a situation where CSR does not maximize financial performance. However, external CSR can be profit-maximizing for companies, especially in the longer term, creating a win-win scenario (Flammer et al. 2019). Future research can explore whether our results extend to settings where choices that maximize external CSR also maximize employees' compensation in the long term, but alternative choices lead to higher financial performance and compensation in the short term. This scenario would align the interests of employees with those of shareholders and external stakeholders in the long term. However, if employees' compensation is mainly based on short-term financial targets or if the impact of external CSR on employees' future compensation is uncertain, employees may still prefer to prioritize actions that boost short-term financial performance. Future research can address the role of the type of mission statement and compensation interdependence in this dilemma.

Finally, we operationalize external prosocial actions through charitable donations to the Red Cross. Our findings are likely to extend to other trade-offs that benefit external stakeholders but may create conflict among employees. However, it is possible that the benefits for external stakeholders and employees are not as easily quantifiable in monetary terms and comparable as in our experiment.

2.6 REFERENCES

- Andrassy, J., J. Burke, and M. T. Stern. 2024. Not Walking the Talk: The Effect of Prioritizing Profit or ESG Impact on Employee Behavior Across ESG Impact- and Shareholder Value-Focused Mission Statements. *Journal of Management Accounting Research* (forthcoming): 1-21.
- Aguilera, R. V., D. E. Rupp, C. A. Williams, and J. Ganapathi. 2007. Putting the S back in corporate social responsibility: A multilevel theory of social change in organizations. *Academy of management review* 32 (3): 836-863.
- Akinyele, K. O., V. Arnold, and S. G. Sutton. 2020. Wording and saliency matter: The impact of incentive system and organizational value statement on employees' performance. *Behavioral Research in Accounting* 32 (1): 101-118.
- Akinyele, K. O., V. Arnold, and S. G. Sutton. 2022. Motivating unrewarded task performance: The dual effects of incentives and an organisational value statement in a discretionary task setting. *Accounting & Finance* 62 (2): 2439-2466.
- Alegre, I., J. Berbegal-Mirabent, A. Guerrero, and M. Mas-Machuca. 2018. The real mission of the mission statement: A systematic review of the literature. *Journal of Management & Organization* 24 (4): 456-473.
- Amabile, T. M., and M. G. Pratt. 2016. The dynamic componential model of creativity and innovation in organizations: Making progress, making meaning. *Research in Organizational Behavior* 36: 157-183.
- Arjaliès, D. L., and J. Mundy. 2013. The use of management control systems to manage CSR strategy: A levers of control perspective. *Management Accounting Research* 24 (4): 284-300.
- Arya, A., B. Mittendorf, and R. N. Ramanan. 2019. Beyond profits: The rise of dual-purpose organizations and its consequences for disclosure. *The Accounting Review* 94 (1): 25-43.
- Balakrishnan, R., G. B. Sprinkle, and M. G. Williamson. 2011. Contracting benefits of corporate giving: An experimental investigation. *The Accounting Review* 86 (6): 1887-1907.

- Battigalli, Pierpaolo, and Martin Dufwenberg. 2007. Guilt in games. *American Economic Review* 97 (2): 170–6.
- Battigalli, Pierpaolo, and Martin Dufwenberg. 2009. Dynamic psychological games. *Journal of Economic Theory* 144 (1): 1–35.
- Battilana, J., T. Obloj, A. C. Pache, and M. Sengul. 2022. Beyond shareholder value maximization: Accounting for financial/social trade-offs in dual-purpose companies. *Academy of Management Review* 47 (2): 237-258.
- Bedford, D. S. 2020. Conceptual and empirical issues in understanding management control combinations. *Accounting, Organizations and Society* 86: 101187.
- Bedford, D. S., T. Malmi, and M. Sandelin. 2016. Management control effectiveness and strategy: An empirical analysis of packages and systems. *Accounting, Organizations and Society* 51: 12-28.
- Bem, D. J. 1972. Self-Perception Theory. *Advances in Experimental Social Psychology* 6: 1-67.
- Bénabou, R., and J. Tirole. 2006. Incentives and prosocial behavior. *American Economic Review* 96 (5): 1652-1678.
- Bernheim, B. D. 1994. A theory of conformity. *Journal of Political Economy* 102 (5): 841-877.
- Beusch, P., J. E. Frisk, M. Rosén, and W. Dilla. 2022. Management control for sustainability: Towards integrated systems. *Management Accounting Research* 54: 100777.
- Björklund, F. 2003. Differences in the justification of choices in moral dilemmas: Effects of gender, time pressure and dilemma seriousness. *Scandinavian Journal of Psychology* 44 (5): 459-466.
- Blay, A. D., E. S. Gooden, M. J. Mellon, and D. E. Stevens. 2018. The usefulness of social norm theory in empirical business ethics research: A review and suggestions for future research. *Journal of Business Ethics* 152: 191-206.
- Blodgett, M. S., C. Dumas, and A. Zanzi. 2011. Emerging trends in global ethics: A comparative study of US and international family business values. *Journal of Business Ethics* 99 (1): 29-38.
- Braun, S., J. S. Wesche, D. Frey, S. Weisweiler, and C. Peus. 2012. Effectiveness of mission statements in organizations—A review. *Journal of Management & Organization* 18 (4): 430-444.
- Brønn, P. S., and D. Vidaver-Cohen. 2009. Corporate motives for social initiative: Legitimacy, sustainability, or the bottom line? *Journal of Business Ethics* 87: 91-109.

- Buckless, F. A., and S. P. Ravenscroft. 1990. Contrast coding: A refinement of ANOVA in behavioral analysis. *The Accounting Review* 65 (4): 933-945.
- Cardinaels, E., and H. Yin. 2015. Think twice before going for incentives: Social norms and the principal's decision on compensation contracts. *Journal of Accounting Research* 53 (5): 985-1015.
- Chandler A.. 2023. Why mission and value statements matter. Available at: <https://www.forbes.com/councils/forbesbusinesscouncil/2023/08/18/why-mission-and-value-statements-matter/>
- Charness, G., L. Rigotti, and A. Rustichini. 2007. Individual behavior and group membership. *American Economic Review* 97(4): 1340-1352.
- Chenhall, R. H., M. Hall, and D. Smith. 2017. The expressive role of performance measurement systems: A field study of a mental health development project. *Accounting, Organizations and Society* 63: 60-75.
- Christ, M. H., K. L. Sedatole, and K. L. Towry. 2012. Sticks and carrots: The effect of contract frame on effort in incomplete contracts. *The Accounting Review* 87 (6): 1913-1938.
- Christ, M. H. 2013. An experimental investigation of the interactions among intentions, reciprocity, and control. *Journal of Management Accounting Research* 25 (1): 169-197.
- Church, B. K., R. L. Hannan, and X. J. Kuang. 2012. Shared interest and honesty in budget reporting. *Accounting, Organizations and Society* 37 (3): 155-167.
- Church, B. K., W. Jiang, X. Kuang, and A. Vitalis. 2019. A dollar for a tree or a tree for a dollar? The behavioral effects of measurement basis on managers' CSR investment decision. *The Accounting Review* 94 (5): 117-137.
- Craig, W. 2018. The importance of having a mission-driven company. Available at: <https://www.forbes.com/sites/williamcraig/2018/05/15/the-importance-of-having-a-mission-driven-company/>
- Dhanesh, G. S. 2020. Who cares about organizational purpose and corporate social responsibility, and how can organizations adapt? A hypermodern perspective. *Business Horizons* 63 (4): 585-594.
- Dharmadasa, P., Y. Maduraapeurma, and S. K. Herath. 2012. Mission statements and company financial performance revisited. *International Journal of Managerial and Financial Accounting* 4 (3): 314-324.
- Deore, A., S. Gallani, and R. Krishnan. 2023. The effect of systems of management controls on honesty in managerial reporting. *Accounting, Organizations and Society* 105: 101401.

- Ditillo, A., and I. E. Lisi. 2016. Exploring sustainability control systems' integration: The relevance of sustainability orientation. *Journal of Management Accounting Research* 28 (2): 125-148.
- Douthit, J. D., P. R. Martin, and M. McAllister. 2022. Charitable contribution matching and effort-elicitation. *The Accounting Review* 97 (1): 213-232.
- Douthit, J., Z. Mao, and Martin. 2024. Tend to one's own house: The effect of firm CSR on employee effort. *Contemporary Accounting Research* 41(4): 2488-2513.
- Eagly, A. H. 2009. The his and hers of prosocial behavior: An examination of the social psychology of gender. *American Psychologist* 64 (8): 644.
- Feldman, D. C. 1984. The development and enforcement of group norms. *Academy of Management Review* 9 (1): 47-53.
- Festinger, L. 1957. *A theory of cognitive dissonance*. Stanford University Press, Stanford, CA.
- Flammer, C., B. Hong, and D. Minor. 2019. Corporate governance and the rise of integrating corporate social responsibility criteria in executive compensation: Effectiveness and implications for firm outcomes. *Strategic Management Journal* 40 (7): 1097-1122.
- George, G., M. R. Haas, A. M. McGahan, S. J. Schillebeeckx, and P. Tracey. 2023. Purpose in the for-profit firm: A review and framework for management research. *Journal of Management* 49 (6): 1841-1869.
- Grabner, I., and F. Moers (2013). Management control as a system or a package? Conceptual and empirical issues. *Accounting, organizations and society*, 38(6-7), 407-419.
- Gond, J. P., S. Grubnic, C. Herzig, and J. Moon. 2012. Configuring management control systems: Theorizing the integration of strategy and sustainability. *Management Accounting Research* 23 (3): 205-223.
- Grimes, M. G., T. A. Williams, and E. Y. Zhao. 2019. Anchors aweigh: The sources, variety, and challenges of mission drift. *Academy of Management Review* 44 (4): 819-845.
- Grimes, M. G., T. A. Williams, and E. Y. Zhao. 2020. Beyond hybridity: Accounting for the values complexity of all organizations in the study of mission and mission drift. *Academy of Management Review* 45 (1): 234-238.
- Guggenmos, R. D., M. D. Piercey, and C. P. Agoglia. 2018. Custom contrast testing: Current trends and a new approach. *The Accounting Review* 93 (5): 223-244.
- Hannah, S. T., P. L. Jennings, D. Bluhm, A. C. Peng, and J. M. Schaubroeck. 2014. Duty orientation: Theoretical development and preliminary construct testing. *Organizational Behavior and Human Decision Processes* 123 (2): 220-238.

- Hansen, E. G., and S. Schaltegger. 2016. The sustainability balanced scorecard: A systematic review of architectures. *Journal of Business Ethics* 133 (2): 193-221.
- Hayes, A. F. 2022. *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach* (3rd ed.). New York: The Guilford Press.
- Heggen, C., and V. G. Sridharan. 2021. The effects of an enabling approach to eco-control on firms' environmental performance: A research note. *Management Accounting Research* 50: 100724.
- Hejjas, K., G. Miller, and C. Scarles. 2019. "It's like hating puppies!" Employee disengagement and corporate social responsibility. *Journal of Business Ethics* 157: 319-337.
- Hinojosa, A. S., W. L. Gardner, H. J. Walker, C. Coglisier, and D. Gullifor. 2017. A review of cognitive dissonance theory in management research: Opportunities for further development. *Journal of Management* 43 (1): 170-199.
- Hobson, J. L., R. D. Sommerfeldt, and L. W. Wang. 2021. Cheating for the cause: The effects of performance-based pay on socially oriented misreporting. *The Accounting Review* 96 (5): 317-336.
- Hyndman, N., and R. Eden. 2000. A study of the coordination of mission, objectives and targets in UK executive agencies. *Management Accounting Research* 11 (2): 175-191.
- Johnson, J. A., S. G. Sutton, and J. C. Theis. 2020. Prioritizing sustainability issues: Insights from corporate managers about key decision-makers, reporting models, and stakeholder communications. *Accounting and the Public Interest* 20 (1): 28-60.
- Johnstone, L. 2019. Theorising and conceptualising the sustainability control system for effective sustainability management. *Journal of Management Control* 30 (1): 25-64.
- Joly, H. 2021. Creating a meaningful corporate purpose. *Harvard Business Review*.
- Kachelmeier, S. J., T. A. Thornock, and M. G. Williamson. 2016. Communicated values as informal controls: Promoting quality while undermining productivity? *Contemporary Accounting Research* 33 (4): 1411-1434.
- Klein, S., S. Schneider, and P. Spieth. 2021. How to stay on the road? A business model perspective on mission drift in social purpose organizations. *Journal of Business Research* 125: 658-671.
- Koller, M., and T. Salzberger. 2007. Cognitive dissonance as a relevant construct throughout the decision-making and consumption process-an empirical investigation related to a package tour. *Journal of Customer Behaviour* 6 (3): 217-227.

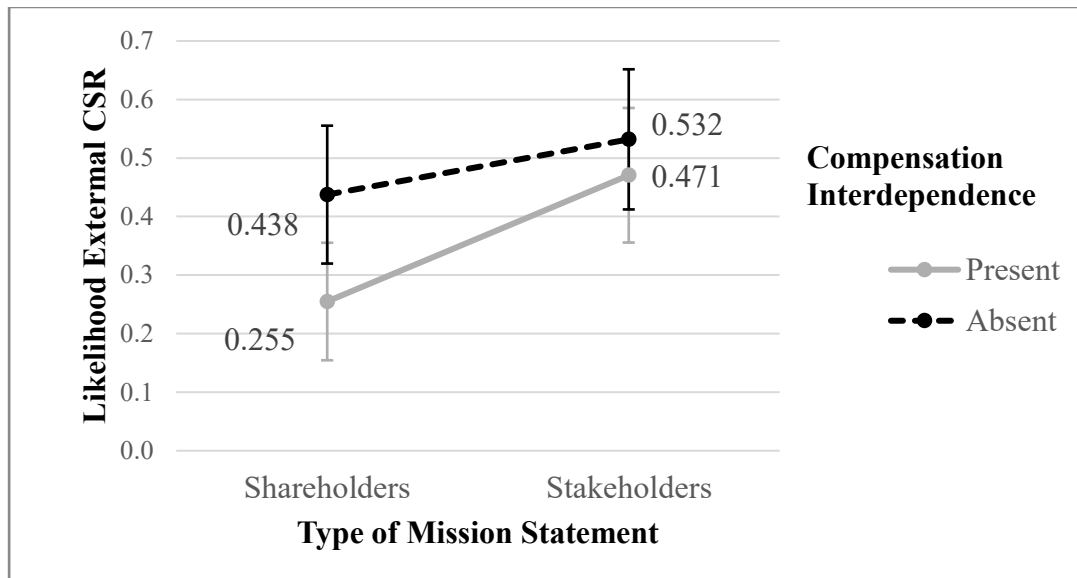
- Law, K. F., D. Campbell, and B. Gaesser. 2022. Biased benevolence: The perceived morality of effective altruism across social distance. *Personality and Social Psychology Bulletin* 48 (3): 426-444.
- Lill, J. B. 2020. When the boss is far away and there is shared pay: The effect of monitoring distance and compensation interdependence on performance misreporting. *Accounting, Organizations and Society* 86: 10114.
- Lueg, R., and R. Radlach. 2016. Managing sustainable development with management control systems: A literature review. *European Management Journal* 34 (2): 158-171.
- Maas, V. S., and M. Van Rinsum. 2013. How control system design influences performance misreporting. *Journal of Accounting Research* 51 (5): 1159-1186.
- McGrath, A. 2017. Dealing with dissonance: A review of cognitive dissonance reduction. *Social and Personality Psychology Compass* 11 (12): e12362.
- McManus, R. M., M. Kleiman-Weiner, and L. Young. 2020. What we owe to family: The impact of special obligations on moral judgment. *Psychological Science* 31 (3): 227–242.
- Mion, G., C. R. L. Adaui, A. Bonfanti, and V. De Crescenzo. 2023. Mission statements and financial and sustainability performance: An exploratory study of Benefit Corporations certified as B Corps. *Journal of Business Research* 157: 113585.
- Mullane, J. V. 2002. The mission statement is a strategic tool: when used properly. *Management Decision* 40 (5): 448-455.
- Norris, G., and B. O'Dwyer. 2004. Motivating socially responsive decision making: the operation of management controls in a socially responsive organisation. *The British Accounting Review* 36 (2): 173-196.
- Ostapski, S. A., and C. N. Isaacs. 1992. Corporate moral responsibility and the moral audit: Challenges for Refuse Relief Inc. *Journal of Business Ethics* 11 (3): 231–239.
- Ouchi, W. G. 1979. A conceptual framework for the design of organizational control mechanisms. *Management Science* 25 (9): 833-848.
- Prentice, D. A., and D. T. Miller. 1993. Pluralistic ignorance and alcohol use on campus: some consequences of misperceiving the social norm. *Journal of Personality and Social Psychology* 64 (2): 243-252.
- Raihani, N. J., and E. A. Power. 2021. No good deed goes unpunished: the social costs of prosocial behaviour. *Evolutionary Human Sciences* 3.
- Robinson, S. L., and A. M. O'Leary-Kelly. 1998. Monkey see, monkey do: The influence of work groups on the antisocial behavior of employees. *Academy of Management Journal* 41 (6): 658–672.

- Sendlhofer, T. 2020. Decoupling from moral responsibility for CSR: Employees' visionary procrastination at a SME. *Journal of Business Ethics* 167 (2): 361-378.
- Shen, J., and J. Benson. 2016. When CSR is a social norm: How socially responsible human resource management affects employee work behavior. *Journal of Management* 42 (6): 1723-1746.
- Simon, L., J. Greenberg, and J. Brehm. 1995. Trivialization: the forgotten mode of dissonance reduction. *Journal of Personality and Social Psychology* 68 (2): 247-260.
- Simons, R. 1995. *Levers of control: How managers use innovative control systems to drive strategic renewal*. Harvard Business Press.
- Smith, Joanne R., and Winnifred R. Louis. 2008. Do as we say and as we do: The interplay of descriptive and injunctive group norms in the attitude-behaviour relationship. *British Journal of Social Psychology* 47 (4): 647-66.
- Spallek, F., K. Bastini, and M. Lachmann. 2023. The more, the merrier? The behavioral effects of a firm's CSR mission and monetary CSR incentives on employee CSR engagement. *Journal of Management Accounting Research* 35 (3): 173-195.
- Steele, C. M. 1988. The psychology of self-affirmation: Sustaining the integrity of the self. In *Advances in Experimental Social Psychology* 21: 261-302.
- Tajfel, H., and M. Billic. 1974. Familiarity and categorization in intergroup behavior. *Journal of experimental social psychology* 10(2): 159-170.
- Tayler, W. B., and R. J. Bloomfield. 2011. Norms, conformity, and controls. *Journal of Accounting Research* 49 (3): 753-790.

2.7 FIGURES

FIGURE 2.1

Likelihood of prioritizing External CSR by experimental condition



This plot represents the predicted likelihood of selecting the project benefiting external stakeholders (Project Red Cross) by experimental condition. We estimate the logit model 2.1. The definitions of the variables are available in Appendix 2.B. 90 percent confidence intervals are also displayed.

FIGURE 2.2A

Effect of the type of mission statement when compensation interdependence is present

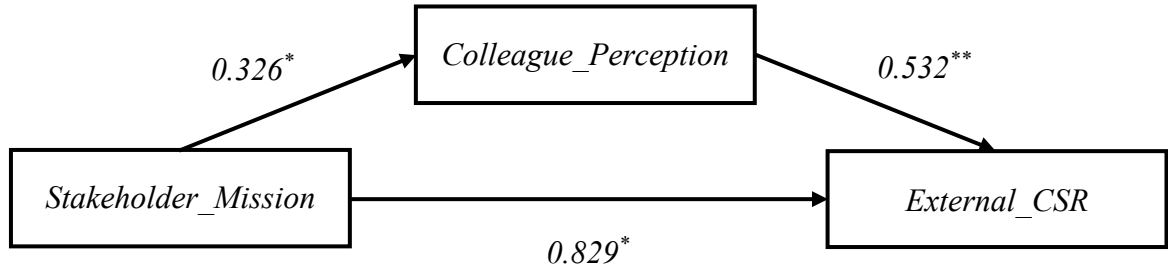
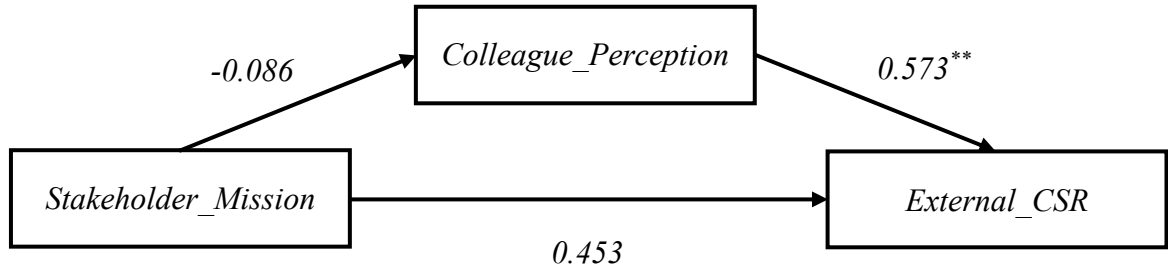


FIGURE 2.2B

Effect of the type of mission statement when compensation interdependence is absent



These figures report the results of the mediation model with *Colleague_Perception* as the mediator. The definitions of the variables are available in Appendix 2.B. Figure 2.2A considers only the observations in the Compensation-Interdependence-Present condition, while Figure 2.2B considers those in the Compensation-Interdependence-Absent condition. For the estimation, we use Hayes' (2022) PROCESS function for R, version 4.3, model 4 with 5,000 replications. The mediation model consists of two linear regressions. Regression 2.1 (OLS model): $Colleague_Perception_i = \eta_i Stakeholder_Mission + Constant_1 + \epsilon_i$. Regression 2.2 (logit model): $Prob(External_CSR_i = 1) = f(\gamma_1 Stakeholder_Mission + \gamma_2 Colleague_Perception + Constant_2 + \zeta_i)$. The numbers on the arrows represent the estimated coefficients of the model. The stars indicate significance levels * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$ (two-tailed).

2.8 TABLES

TABLE 2.1

Descriptive statistics

	n	mean	sd	p25	p50	p75
<i>External_CSR</i>	197	0.421	0.495	0.000	0.000	1.000
<i>Compensation_Interdependence</i>	197	0.518	0.501	0.000	1.000	1.000
<i>Stakeholder_Mission</i>	197	0.497	0.501	0.000	0.000	1.000
<i>Age</i>	197	40.761	13.885	30.000	40.000	50.000
<i>Education</i>	197	6.294	1.503	5.000	7.000	7.000
<i>Male</i>	197	0.350	0.478	0.000	0.000	1.000
<i>Colleague_Perception</i>	197	-0.006	0.970	-0.625	0.167	0.862
<i>Cognitive_Dissonance</i>	197	0.000	0.975	-0.786	-0.091	0.753

This table reports descriptive statistics for the variables of the study, including demographic characteristics. Each observation represents a different participant. The definitions of the variables are available in Appendix 2.B.

TABLE 2.2

Summary statistics of External_CSR by experimental condition

Compensation Interdependence	Type of Mission Statement		Overall
	Shareholder- Oriented	Stakeholder- Oriented	
Absent	0.438 (0.501) n = 48	0.532 (0.504) n = 47	0.484 (0.502) n = 95
Present	0.255 (0.440) n = 51	0.471 (0.504) n = 51	0.363 (0.483) n = 102
Overall	0.343 (0.477) n = 99	0.500 (0.503) n = 98	0.421 (0.495) n = 197

This table includes means, standard deviations (in brackets), and number of observations (n) of the main variable of interest, *External_CSR*, by experimental condition. *External_CSR* is a dummy variable equal to 1 if participants select the project that benefits external stakeholders (Project Red Cross) over the project that maximizes financial performance (Project Furniture).

TABLE 2.3

Main Results

Panel A: Logit Model		External_CSR			
Compensation_Interdependence		-0.821*			
		(0.43)			
Stakeholder_Mission		0.379			
		(0.41)			
Compensation_Interdependence × Stakeholder_Mission		0.576			
		(0.59)			
Constant		-0.251			
		(0.29)			
N		197			
Pseudo R²		0.034			
Panel B: Planned Contrast		Estimate	df	Chi2	p
Planned contrast {-3; +1; +1; +1}		2.977	1	7.52	0.006
Panel C: Marginal Effects		dy/dx	se	Z	p
Effect of a stakeholder-oriented mission:					
When compensation interdependence is present	0.216	0.093	2.325	0.020	
When compensation interdependence is absent	0.094	0.102	0.925	0.355	
Effect of compensation interdependence:					
			-		
Under a stakeholder-oriented mission	-0.061	0.101	0.608	0.543	
			-		
Under a shareholder-oriented mission	-0.183	0.094	1.941	0.052	

Panel A reports the estimation of a logit model. *External_CSR* equals 1 if participants select the project that generates benefits to external stakeholders (Project Red Cross) over the project that maximizes financial performance (Project Furniture) and 0 otherwise. *Compensation_Interdependence* and *Stakeholder_Mission* are equal to 1, respectively, for participants under the Compensation-Interdependence-Present and the Stakeholder-Oriented-Mission conditions. See Appendix 2.B for the variable definitions. Panel B includes a test for a planned contrast where the weight -3 is assigned to observations in the experimental cell that combines the Compensation-Interdependence-Present and the Shareholder-Oriented-Mission conditions. The weight +1 is assigned to the observations in the other three experimental cells. Panel C reports the marginal effects of the type of mission statement and compensation interdependence on *External_CSR*. Standard errors in panel A are in brackets. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$ indicate significance levels (two-tailed).

TABLE 2.4

Estimated Indirect Effect of Mission Statement Type on External_CSR

<u>Indirect effect of a stakeholder-oriented mission</u>	<u>Effect</u>	<u>BootSE</u>	<u>BootLLCI</u>	<u>BootULCI</u>
When compensation interdependence is present	0.173	0.130	0.002	0.412
When compensation interdependence is absent	-0.050	0.139	-0.318	0.120

This table reports the indirect effect of stakeholder missions compared to shareholder-oriented missions on the likelihood of choosing the project that benefits external stakeholders (Project Red Cross) over the project that maximizes financial results (Project Furniture) contingent on the presence of compensation interdependence. The indirect effect is estimated using the bootstrapped procedure in Hayes (2022) with his PROCESS function for R, version 4.3. The BootLLCI and BootULCI are, respectively, the lower and the upper bound of the 90 percent bootstrapped confidence intervals with 5,000 replications. The mediation model consists of two linear regressions. Regression 1 (OLS model): $Colleague_Perception_i = \eta_i Stakeholder_Mission + Constant_1 + \epsilon_i$. Regression 2 (logit model): $Prob(External_CSR_i = 1) = f(\gamma_i Stakeholder_Mission + \gamma_{2i} Colleague_Perception + Constant_2 + \zeta_i)$. The effects are expressed in log-odds ratios.

2.9 APPENDICES

APPENDIX 2.A

Task and Manipulations

Indicate which project you want to prioritize by selecting one of the options below. The table summarizes the payoffs generated by your decision.

Remember that your colleague faces the same decision problem and that the mission statement of {*name of the company depending on the participant's choice*: TEAM BLACK; TEAM WHITE} is:

{*Shareholder-Oriented-Mission condition*: “Our aim is to make money for our shareholders for which we do business and achieve profitable growth”}

{*Stakeholder-Oriented-Mission condition*: “Our aim is to make a positive difference for the world, allowing our employees, our shareholders, and all the communities in which we live to prosper.”}

{*Table for the Compensation-Interdependence-Present condition*}

Payout for:	You	Your Colleague	The Red Cross
Project Furniture	0.40£	0.40£	0.00£
Project Red Cross	0.20£	0.20£	0.40£

{*Table for the Compensation-Interdependence-Absent condition*}

Payout for:	You	Your Colleague	The Red Cross
Project Furniture	0.80£	0.00£	0.00£
Project Red Cross	0.40£	0.00£	0.40£

Please indicate your decision here:

Project Furniture

Project Red Cross

APPENDIX 2.B

Variable Definitions

<i>Variable</i>	<i>Definition</i>
<i>External_CSR</i>	Dummy variable equal to 0 if the participant selects Project Furniture (the project that maximizes financial results), which maximizes the participants' payout, equal to 1 if the participant prioritizes Project Red Cross (the project that benefits external stakeholders), which generates a lower payout for the participants but a donation for the British Red Cross.
<i>Compensation_Interdependence</i>	Dummy variable equal to 0 for the Compensation-Interdependence-Absent condition and 1 for observations in the Compensation-Interdependence-Present condition.
<i>Stakeholder_Mission</i>	Dummy variable equal to 0 for observations in the Shareholder-Oriented-Mission condition and 1 for the Stakeholder-Oriented-Mission condition.
<i>Age</i>	Answer to the question about age in ranges of 10 years. Younger than 20 is coded as 10, older than 70 is coded as 70, between 21 and 30 is coded as 20, between 31 and 40 is coded as 30, and so on.
<i>Education</i>	Answer to the question about maximum education level achieved converted in numbers such as the higher the number, the higher the education degree achieved: 1 = "Some Primary"; 2 = "Completed Primary School"; 3 = "Some Secondary"; 4 = "Completed Secondary School"; 5 = "Vocational or Similar"; 6 = "Some University but no degree"; 7 = "University Bachelors Degree"; 8 = "Graduate or professional degree (MA, MS, MBA, PhD, JD, MD, DDS)."
<i>Male</i>	Dummy variable equal to 1 for participants identifying as males and 0 for those identifying as females or as "other".
<i>Colleague_Perception</i>	Positive perceptions from the perspective of the colleague. We measure it by adapting three items from Hobson et al. (2021). We compute the main factor loading of the items. The items are about whether participants agree or disagree with the following statements on a scale from 1 to 7 (1 = Strongly disagree; 7 = Strongly agree): "Imagine

that your colleague from {TEAM WHITE / TEAM BLACK} observes you making this decision [Project Red Cross]. He/she knows the details of the situation as you do. To what extent do you agree with the following statement? 1) He/she will think that I am a benevolent person; 2) He/she will think that I am a caring person; 3) He/she will think that I am an empathetic person.”

Cognitive Dissonance

Cognitive dissonance was measured by adapting seven 7-point Likert scale items from Koller and Salzberger (2007). We take the first factor scores of the items. The questions are about whether participants agree or disagree with the following statements on a scale from 1 to 7 (1 = Strongly disagree; 7 = Strongly agree): “1) *Perhaps I should have prioritized another project.* 2) *I am not quite sure about my decision.* 3) *When thinking of the decision, I felt uncomfortable.* 4) *I don’t know whether I prioritized the right project.* 5) *When thinking of the decision, I felt uneasy.* 6) *I do not know whether this was the right choice.* 7) *I would like to undo my decision.”*

CHAPTER 3

Uncontrollable Events, Incentive Design, and Managerial Myopia:

Evidence From the COVID-19 Setting

with Ariela Caglio and Angelo Ditillo

3.1 INTRODUCTION

Compensation contracts is a field that poses significant challenges to management accounting researchers. One of these challenges is the incorporation of uncontrollable factors into contract design. On the one hand, incorporating such factors in compensation contracts may lead powerful managers to opportunistically leverage them to boost their compensation and extract rents from shareholders (Bertrand and Mullainathan 2001; Garvey and Milbourn 2006). On the other hand, ignoring uncontrollable factors may expose managers to excessive risk (Fudenberg and Tirole 1990) and lead them to perceive unfairness regarding their compensation (Giraud, Langevin and Mendoza 2008; Burkert, Fischer, and Schäffer 2011).

We contribute to this debate by examining how incentive plans are adjusted in response to uncontrollable factors that affect firm performance. Specifically, we focus on favorable modifications, i.e., adjustments that benefit the manager. These favorable adjustments may serve the interests of both principals and agents by preserving incentives while mitigating undue risk and perceived inequity (Fudenberg and Tirole 1990; Yim 2001; Şabac 2007). To address this question, we exploit the setting of the COVID-19 pandemic, an exogenous and unpredictable shock that led to wide variation in firms' responses. Specifically, we study CEOs' favorable incentive plan modifications (IPMs) during 2020, focusing on the determinants and consequences of these modifications. Among the determinants, based on prior literature, we concentrate on the financial performance effects of the COVID-19 pandemic, CEOs' incentives to depart the company, and the quality of corporate governance. Among the consequences, we specifically explore the extent to which these adjustments contribute to mitigating managerial myopia, defined as the tendency to overvalue short-term results and undervalue long-term implications (Allee, Badertscher, and Yohn 2020)⁴¹.

As for the determinants of IPMs, we have three expectations. First, firms more negatively affected financially by the uncontrollable event are more likely to introduce

⁴¹ Some papers distinguish between managerial myopia and short-termism, attributing the first to the managers' behavior and the second to the pressure from financial markets (Laverty 2004). We follow Brochet, Loumriot, and Serafeim (2015) and use myopia and short-termism interchangeably.

favorable IPMs for their CEOs, as their original performance targets may have become unachievable through no fault of management. Second, firms with higher corporate governance quality are less likely to introduce favorable IPMs as CEOs exert a more limited influence on the board (Frydman and Jenter 2010; Campbell and Thompson 2015). Third, consistently with the view that favorable pay adjustments serve a retention function (Carter and Lynch 2004; Chen 2004), we expect that firms whose CEOs face stronger incentives to leave are more likely to introduce favorable IPMs.

In terms of consequences, we expect that the introduction of favorable IPMs will mitigate managerial myopia. First, favorable IPMs may restore the motivational effect of incentives tied to long-term goals, thereby reducing the likelihood that managers cut discretionary expenditures, such as R&D and ES initiatives, in pursuit of short-term performance improvements. In addition, when compensation is reduced due to factors beyond their control, managers may perceive unfairness and dissatisfaction (Carter and Lynch 2004; Chen 2004). This unfairness may increase short-termism if managers do not feel supported and do not feel motivated to achieve the company's long-term goals. Favorable IPMs may increase perceptions of fairness and mitigate dissatisfaction, thereby reducing their myopia.

However, these predictions are not straightforward, as companies may not be willing to introduce IPMs. Favorable compensation adjustments might establish a precedent (Bol, Hecht, and Smith 2015). Conceding IPMs implies that not granting the same treatment to managers in the future could generate discontent. Moreover, if managers believe that they will be shielded from uncontrollable factors through IPMs, they may have fewer incentives to anticipate risks and mitigate their adverse effects (Merchant 1987). Similarly, managers may also be less motivated to react to uncontrollable events when they occur. Finally, IPMs may introduce excessive subjectivity in compensation contracts and set a precedent for future adjustments, increasing fairness concerns rather than reducing them (Giraud, Langevin, and Mendoza 2008).

Our starting sample consists of S&P 500 constituents in 2020. We collect and classify IPMs as described in the companies' proxy statements issued for the fiscal year 2020. Our database contains only the IPMs that are explicitly attributed to the COVID-19 pandemic in the proxy statements. We consider all the modifications mentioned in the incentive plan sections of the compensation discussion and analysis (CD&A). For the analysis of the determinants, our sample includes only one fiscal year per observation. Conversely, when assessing the consequences, we have added a pre-period of one year.

For our analysis of the determinants, we compare firms that implement favorable IPMs with those that do not implement IPMs. We focus on favorable IPMs because, as anticipated,

this aligns with our research question, namely, whether adjusting incentive compensation plans for uncontrollable events reduces managerial myopia. Moreover, in line with prior literature (Albuquerque, Bennett, Custódio, and Cvijanović 2023), we find only a few cases of firms that implement unfavorable IPMs for their CEOs⁴². Consistently with our hypotheses, we find that the greater the negative impact of COVID-19 on a firm's financial performance, the more likely the firm is to implement favorable IPMs. To capture the impact of COVID-19 on financial performance, we use the variation of sales over 2020, the buy-and-hold stock returns of March 2020, and the number of deaths attributed to COVID-19 in the state of incorporation of the company. We also find that shorter CEO tenures and age above 65 years are associated with favorable IPMs. These CEOs' characteristics, in fact, are indicative respectively of lower organizational identification and higher closeness to retirement (Abernethy, Jiang, and Kuang 2019), and represent conditions under which retention may be more difficult to promote, thus increasing the CEOs' incentives to leave the firm. Finally, considering different measures of corporate governance quality, we find that the presence of a compensation committee decreases the likelihood of introducing favorable IPMs. Moreover, the percentage of independent directors is negatively associated with introducing favorable long-term IPMs compared to short-term IPMs.

For our analysis of the consequences, we adopt a difference-in-difference approach. The treatment group consists of companies that have introduced favorable IPMs, whereas the control group is represented by companies that have not introduced them. Our consequences (the dependent variables) are proxies of managerial myopia commonly used by extant contributors and include the level of research and development expenditures scaled by lagged assets (R&D), the level of capital expenditures scaled by lagged assets (Capex), and the level of environmental and social (ES) performance (Kraft, Vashishtha, and Venkatachalam 2018; Abernethy, Jiang, and Kuang 2019). A decline in the level of these variables signals managerial myopia, as managers who are less concerned with long-term firm performance tend to cut these expenditures (Kang and Kim 2020) and exhibit weaker ES orientation. Companies that decrease these expenses and achieve a lower ES performance compared to their peers are considered more myopic, having prioritized short-term results, sacrificing long-term value. We find that favorable IPMs are positively associated with increases in R&D expenses and ES performance. However, we find that IPMs are associated with a greater reduction in Capex. Our results can be viewed in conjunction with the findings of Flammer and Ioannou (2021),

⁴² We have only 14 cases in our sample.

who show that companies increasing R&D and ES performance but decreasing Capex during the 2007-2009 crisis outperformed others during the subsequent recovery period. Therefore, our results suggest that managers receiving favorable IPMs are more oriented towards long-term results.

In the additional analyses section, we complement these findings. Firstly, we show that the positive association between favorable IPMs and R&D is more pronounced for companies implementing favorable long-term IPMs rather than short-term IPMs. Secondly, we find a negative association between long-term IPMs and an indicator of short-termism, i.e., the likelihood of beating analysts' forecasts by a small margin, but a positive association with changes in Tobin's Q as a proxy for expected long-term performance. Finally, we show that the positive association between favorable IPMs and R&D is stronger for companies whose financial performance is more negatively affected by the pandemic. Overall, our results suggest that favorably adjusting incentive plans in response to uncontrollable adverse events can reduce managerial myopia. Findings are robust to entropy balancing.

Our study contributes to the literature on compensation contract design and the controllability principle (Antle and Demski 1988; Merchant 1989; Bol and Smith 2011). Prior research documents the potential of adjusted earnings measures to reduce short-termism (Kim and Shin 2018). We add to this literature by examining a broader set of mechanisms used to react to uncontrollable factors - such as adjustment of targets, performance measures, and performance periods - introduced specifically to counterbalance the impact of such uncontrollable events. Moreover, to the best of our knowledge, our study is the first to explore ES implications of incentive plan adjustments for uncontrollable events.

We also provide insights into the literature on management control systems during economic downturns (Becker et al. 2016; Bedford, Speklé, and Widener 2022) and the design of executive incentive compensation contracts (Gilson and Vetsuypens 1993; Matolcsy 2000; Bertrand and Mullainathan 2001). This study relates to Casas-Arce, Indjejikian, and Matějka (2020), who analyze bonus plan choices during the 2007-2009 financial crisis. They find an increase in the weight of financial measures, greater perceived difficulty of targets, and reduced incentive strength⁴³. We explore the role of some determinants and consequences of favorable IPMs related to fairness concerns rather than focusing on the design of the incentive plan itself. Our study suggests that when unexpected uncontrollable events occur, companies adjust

⁴³ The percentage of compensation related to target bonuses that managers can earn if they meet all the targets (Casas-Arce, Indjejikian, and Matějka 2020).

compensation plans by taking into account potential fairness concerns of shareholders and CEOs, and that the resulting IPMs are associated with a mitigation of managerial myopia.

Additionally, we add to the recent literature about executive compensation design during the COVID-19 pandemic. We complement the work of Batish et al. (2020), who present initial descriptive statistics of compensation changes at the beginning of the pandemic. Other studies provide evidence that executives can exploit favorable IPMs to extract wealth from shareholders. Li and Yang (2023) show that CEOs who reduced their base salary during the pandemic obtained subsequent benefits by re-contracting their annual bonus plans. Similarly, Bedford, Bugeja, Ghannam, Jeganathan, and Ma (2023) find that CEOs' total compensation did not decrease during the pandemic despite voluntary salary cuts. We take a different perspective, suggesting that IPMs after unexpected uncontrollable events may provide beneficial effects for firms, such as the reduction of managerial myopia.

We also contribute to the literature about investments and reconfiguration during times of crisis (Campello, Giambona, Graham and Harvey 2011; Chakrabarti 2015; Flammer and Ioannou 2021). Our results indicate that compensation decisions may affect managers' preferences for different types of investments. Moreover, we show that changing incentive compensation plans in times of crisis implies asymmetric consequences for short-term and long-term performance.

Finally, our study can offer some insights to practitioners on how to adjust incentive compensation plans during times of uncertainty and crisis. Results suggest that the board of directors should consider adjusting incentive compensation contracts in favor of their executives if they want to reduce managerial myopia.

3.2 PRIOR LITERATURE AND HYPOTHESES

3.2.1 Uncontrollable Factors and Compensation Contracts

Firms implement different mechanisms to account for uncontrollable factors in compensation contracts before the related events occur. A possible provision is to exclude some dimensions and related costs from performance measures that are beyond the managers' control (Bloomfield, Gipper, Kepler and Tsui 2021). Alternatively, relative performance evaluation can help isolate the impact of uncontrollable events if they are common to the compensation peers (Holmstrom 1982; Yim 2001). Finally, companies can arrange insurance contracts against adverse uncontrollable events (Merchant and Van der Stede 2017), resulting in

executive contracts with lower risk premia and equity grants (Merchant and Van der Stede 2017; Armstrong, Glaeser, and Huang 2022).

However, it may be impossible or too costly to identify, price, and contractualise all the possible risks ex-ante. In this case, setting an incomplete contract and renegotiating it when uncontrollable events occur can be a preferable solution (Fudenberg and Tirole 1990; Hermalin and Katz 1991). Companies can reset performance targets to account for the mutated economic conditions (Hyun, Matějka, (Seung Hwan) Oh and Ahn 2022) or exclude extraordinary items from performance measures after the event occurs (Potepa 2020). Similarly, subjective performance evaluation and compensation adjustments can take into account uncontrollable factors and realign pay with effort (Baker, Gibbons, and Murphy 1994; Baiman and Rajan 1995; Curtis, Li, and Patrick 2021; Lillis, Malina, and Mundy 2022). Moreover, the interpretation of controllability itself can vary over time, shifting from a narrow to a broad conceptualization, allowing organizations to adapt their management control systems in response to the demands of their stakeholders (Chua, Graaf, and Kraus 2025).

Adjusting compensation contracts for uncontrollable events can have positive effects on motivation when uncontrollable factors more negatively impact performance. Employees and managers feel that not adjusting compensation is unfair when adverse uncontrollable events occur and performance decreases for reasons beyond their control (Burkert, Fischer, and Schäffer 2011; Asay, Garrett, and Tayler 2023). While traditional Agency Theory considers preferences for risk and monetary wealth as the primary determinants of efficient contracts, more recent literature in economics acknowledges that fairness is a component of the agent's utility function (e.g., Rabin 1993; Fehr and Schmidt 1999). In a contracting setting, fairness can reduce transaction costs, favoring the alignment of interests between principals and agents (Luft 1997).

Consistently with fairness concerns, prior literature provides evidence that managers are more likely to favorably adjust their subordinates' compensation in case the uncontrollable event has a more negative impact on performance ("bad luck") than adjust compensation unfavorably when the uncontrollable event has a more positive impact on performance ("good luck") (Bertrand and Mullainathan 2001; Bol and Smith 2011; Albuquerque et al. 2023). Since executives' compensation is typically tied to firm performance, we expect a similar logic to apply to incentive plan modifications (IPMs) for CEOs:

H3.1 *The higher the negative impact of the uncontrollable event on the company's performance, the higher the likelihood of introducing favorable IPMs.*

However, prior literature documents that companies do not always introduce favorable IPMs for uncontrollable factors because of potential drawbacks (Merchant 1987; Giraud, Langevin, and Mendoza 2008). First, adjusting compensation for external shocks may reduce the reactivity of managers and employees. Bearing some risk for uncontrollable events favors the alignment of interests between principals and agents, as managers have incentives to anticipate unfavorable conditions (Antle and Demski 1988; Ittner and Larcker 2002). Second, modifying compensation plans may introduce excessive subjectivity, exacerbating fairness concerns rather than reducing them (Giraud, Langevin, and Mendoza 2008). Third, the concern of setting a precedent and related fairness concerns reduces the willingness to adjust performance for uncontrollable factors (Bol, Hecht, and Smith 2015).

In addition to these issues, executive incentive compensation contracts pose a peculiar governance challenge, as CEOs can affect their own pay (Frydman and Jenter 2010). For example, prior research shows that powerful CEOs can influence the board to use highly paid executives as a benchmark for their compensation (Shin 2016). Consequently, favorable IPMs may result in excessive pay unrelated to performance, rather than restoring fair and motivating compensation. Compensation systems must be perceived as fair by all the parties affected (Libby 2001). This also includes the shareholders who act as principals.

Strong corporate governance can prevent companies from implementing favorable IPMs that can be potentially viewed as unfair by shareholders through the activity of independent directors and compensation committees (Bugeja, Matolcsy, and Spiropoulos 2016). Therefore, strong corporate governance may, overall, decrease the likelihood of implementing favorable IPMs. Formally:

H3.2 *The quality of corporate governance reduces the likelihood of introducing favorable IPMs in response to an uncontrollable event.*

On the other hand, we expect that firms whose CEOs have more outside options may benefit from favorable IPMs and are more likely to introduce them. Prior literature suggests that applying the controllability principle enhances fairness perceptions and supports the selection and retention of more motivated employees (Waller and Chow 1985; Shields and Waller 1988). If a CEO perceives that their compensation is unfair, due to performance declines caused by adverse uncontrollable events, and the company does not introduce favorable IPMs, the CEO may be more willing to move to another company in search of fairer compensation contracts. For this reason, we hypothesize:

H3.3 *CEOs incentives to leave the company increase the likelihood of introducing favorable IPMs in response to an uncontrollable event.*

Managers may have incentives to undertake actions that maximize profits and other performance measures in the short-term, but reduce the long-term value of the firm (Stein 1989). A prime example of these myopic actions is underinvesting in long-term value drivers, such as R&D, physical assets (Kraft, Vashishtha, and Venkatachalam 2018) and initiatives for improving Environmental and Social (ES) performance (Abernethy, Jiang, and Kuang 2019). This underinvestment may threaten the firm's survival in the long run (Agarwal, Vashishtha, and Venkatachalam 2018).

The structure of incentive compensation contracts is one of the critical factors that may determine managerial myopia (Edmans, Fang, and Lewellen 2017; Ladika and Sautner 2020). Prior literature shows that managers cut R&D expenditures and Capex to meet short-term performance goals (Bennett, Bettis, Gopalan and Milbourn 2017) or to boost short-term stock market prices just before equity awards vest (Edmans, Fang, and Lewellen 2017; Ladika and Sautner 2020). The incentive for these cuts is strong because missing performance targets can lead to reputational damage and career concerns for executives (Bennett et al. 2017; Bloomfield et al. 2021). For this reason, compensation plans with long-term incentives, such as equity awards with a long duration, can strengthen managers' attention towards long-term goals (Dikolli 2001; Matějka, Merchant, and Van der Stede 2009; Dikolli, Kulp, and Sedatole 2013).

In this paper, we argue that IPMs, both short-term and long-term, can mitigate managerial myopia. While favorable long-term IPMs are more directly aligned with reducing managerial myopia, favorable short-term IPMs can also play a role. When adverse uncontrollable events negatively affect performance, such that short-term performance targets are more difficult to achieve (Kim, Matějka, and Park 2023), failing to adjust short-term incentives may lead managers to focus narrowly on saving short-term results. Favorable short-term IPMs can help restore motivational balance and reduce myopic responses. Second, when adverse uncontrollable events occur, managers may feel dissatisfied if their compensation is not adjusted (Burkert, Fischer, and Schäffer 2011). Short-term and long-term incentive plans do not operate in isolation: if short-term incentives are perceived as unfair and fail to motivate managers, long-term ones may also be perceived as less motivating, especially in a crisis situation. Managers who feel supported in the short term are more likely to stay engaged with long-term goals. This is consistent with prior research theorizing and finding that fairness perceptions reduce opportunistic and unethical behavior (Cohen et al. 2007; Langevin and

Mendoza 2013; Su, Baird, and Nuhu 2022). For these reasons, we predict a negative association between IPMs and managerial myopia:

H3.4 *Favorable IPMs in response to uncontrollable events are associated with a reduction of managerial myopia.*

3.3 RESEARCH DESIGN

3.3.1 Sample and Data

The initial sample consists of S&P 500 companies at the beginning of February 2020. We then include the companies that entered the index by the end of December 2020, resulting in 510 companies⁴⁴. We conduct our analysis on large US-listed firms as the SEC requires detailed mandatory disclosure of executive compensation plans. The sources of IPMs are the companies' definitive proxy statements (Form DEF 14A)⁴⁵. We collect modifications of short- and long-term incentive compensation plans from the *compensation discussion and analysis* sections. To account for IPMs implemented during the pandemic and discretionary bonuses granted at the end of the performance period, we use the definitive proxy statements available for the fiscal year 2020⁴⁶. The authors recorded and classified the IPMs.

Table 3.1 summarizes the sample selection process. From the initial 510 companies, we remove 37 firms missing at least one variable of interest in 2019 or 2020. Financial statement variables were sourced from Compustat, while CEO compensation data and characteristics were obtained from Execucomp. ESG scores and governance variables were downloaded from Refinitiv via DataStream. Analyst forecasts and mean earnings surprises were sourced from I/B/E/S. Finally, we collected the casualties attributed to COVID-19 from the Johns Hopkins Coronavirus Resource Center and population data from the US Census Bureau or the World Bank. Appendix 3.1 indicates the source of each variable and its description.

[Table 3.1 here]

⁴⁴ We retrieve the S&P 500 constituents from CRSP.

⁴⁵ For some observations, the modifications of the incentive plan disclosed in the proxy statements were not exhaustive. Therefore, we collect supplemental information from 8K or 10K forms.

⁴⁶ More precisely, for companies with fiscal year end between April 2020 and December 2020, we use the first definitive proxy statement filled after the fiscal year ended in 2020. For companies with fiscal year end between January 2020 and March 2020, we use the first definitive proxy statement filled after the fiscal year ended in 2021.

3.3.2 Measures

The primary variable of interest, *Favorable*, is a dummy variable equal to 1 if the company modifies favorably the short-term incentive plan or the long-term incentive plan in 2020 and 0 otherwise. We consider only the modifications that the compensation discussion analysis attributes to the COVID-19 pandemic. In line with previous contributions, we focus on CEOs as they are the executives who can directly influence the firm's strategy and investments.

To establish whether IPMs are in favor of the CEOs, we use four alternative criteria. First, we check whether the IPMs increase the CEO's compensation or reduce the risk of not reaching payout targets or thresholds (69.68% of the companies with favorable IPMs). However, proxy statements do not always specify what the counterfactual compensation would be without IPMs. In those cases, we consider IPMs favorable if proxy statements indicate that the aim is retention and motivation or that the previous incentive plan was no longer effective for this purpose (9.66% of the companies with favorable IPMs). Alternatively, we classify IPMs as favorable if they are justified by the need to counterbalance the negative impact of the risks and the uncertainty of the Pandemic (6.82% of the companies with favorable IPMs). Finally, we classify as favorable modifications those intended to reflect new strategic priorities or a performance evaluation based on the response to the Pandemic (13.64% of the companies with favorable IPMs), as it would be unfair to evaluate the CEO on goals that no longer reflect the priorities of the company. Appendix 3.2 contains examples for each case.

We further classify the type of IPMs according to their Temporal Horizon. *LT_Change* is a dummy variable equal to 1 if the company modifies the long-term incentive plan and 0 otherwise. *ST_Change* is a dummy variable equal to 1 if the company modifies only the short-term bonus plan and 0 otherwise⁴⁷. When the proxy statement is not explicit about the temporal horizon of rewards, we classify cash grants as short-term modifications and equity or stock option grants as long-term.

For our determinants' analysis, we use three variables that capture the impact that the COVID-19 pandemic had on the company's financial performance. The determinants are: *Sales_%var* is the yearly percentage change in revenues. This measure accounts for the variation of cash inflows due to the crisis. *ReturnsCovid* is the buy-and-hold return for the month of March 2020. We select this month as the S&P 500 index reached its minimum after

⁴⁷ Since only 17 observations modify the long-term incentive plan and not the short-term one (after removing observations with missing values in the control variables), we do not separately analyze companies that modify only long-term plans.

the worldwide spread of COVID-19 (Ramelli and Wagner 2020). *CovidDeaths* is the number of casualties attributed to COVID-19 over 1,000 inhabitants at the end of the year in the state where the company is incorporated. The headquarters of some firms in the sample are located outside the US. For these observations, we use the number of casualties in the country of incorporation. This variable is related to shutdowns and other health and safety restrictions that may have impacted the operations of the company as well as the decision to modify incentive plans.

Then, we consider three indicators of corporate governance quality of the company: the presence of a compensation committee (*CompComm*), the proportion of independent directors (*BdIndep*), and whether the CEO is also the Chairman of the board (*Duality*). Finally, as personal characteristics, *Age* and *Tenure* are respectively the age and the number of years the CEOs have held their position in the firm.

For the analysis of IPMs' consequences, we use three proxies for managerial myopia related to investment decisions: R&D expenditures (Abernethy, Jiang, and Kuang 2019), purchases of fixed tangible assets (Capex) (Kraft, Vashishtha, & Venkatachalam, 2018), and the effort in corporate social responsibility as captured by the ES performance of the company (Abernethy, Jiang, and Kuang 2019). *RnD* is R&D expenditures scaled by lagged total assets. *Capex* is capital expenditures scaled by lagged total assets. Finally, *EsScore* is the simple average of Refinitiv's Environmental Pillar Score and Social Pillar Score.

3.3.3 Regression Models

To test our hypotheses regarding the determinants, we estimate a cross-sectional logit model with *Favorable* as the dependent variable and the aforementioned determinants as the independent variables. We include only the fiscal year 2020. So, as a baseline, we use companies that do not introduce IPMs. We also estimate the same model using *LT_Change* as the dependent variable. In this case, the baseline of the model consists of companies modifying only the short-term incentive plans.

As controls, we add variables related to the firm's financial situation before the Pandemic: firm's size as the logarithm of total assets (*LogAssets*), leverage ratio (*Leverage*), book-to-market ratio (*BTM*), dividends on assets (*DividendsTotal*), and return on assets (*ROA*) as EBIT on total assets. Moreover, we consider some characteristics of the compensation structure of the CEO. *TotCompRank* is the rank of the total compensation (as reported to the SEC) for the previous fiscal year among the CEOs in our sample. *CashComp* is the proportion

of total compensation paid in cash in the previous year. *Salary_%var* is the percentage yearly change in base salary.

For our consequences analysis, we estimate the following difference-in-difference model with OLS:

$$Y_{it} = \text{Favorable}_i \times \text{After}_t + \beta_2 \text{Favorable}_i + \beta_3 \text{After}_t + B \cdot \text{Determinants}_{it} + B \cdot \text{Controls}_{it} + \text{IndustryFE} + \alpha + \epsilon_{it} \quad (1)$$

The treatment group is the companies that modify the 2020 incentive plan in favor of the CEO, represented by the dummy variable *Favorable*. *After* is a dummy variable equal to 1 after the end of March 2020⁴⁸. For each company, we consider two years: 2019 (pre-period) and 2020 (post-period).

The coefficient of interest is β_1 , which compares the average change in the dependent variable from the pre- to the post-treatment period between the treatment and the control groups. The dependent variable Y_{it} is one indicator of managerial myopia. In addition to industry fixed effects based on SIC 2 digits, we include the determinants and the controls of the previous analysis to account for differences between the treatment and the control group.

3.4 MAIN FINDINGS

3.4.1 Descriptive Statistics and Details About IPMs

We find 194 companies on 473 (41.01%) disclosing IPMs attributed to the COVID-19 pandemic. This proportion is higher than the findings of Batish et al. (2020), who indicate a percentage of 17%. We reconcile this difference as their work considers early changes in compensation plans following the stock-market turbulence in March 2020. Instead, our analysis includes the whole 2020 and modifications that companies disclose after the end of 2020 in their proxy statements⁴⁹.

Table 3.2 Panel A reports the frequency of the adjustments classified according to favorableness for the CEO. 161 companies modify the incentive plan in favor of the CEO (*Favorable*). Only 14 firms introduce changes that penalize the CEOs (*Unfavorable*), while in 19 cases the CEOs voluntarily forgo their incentive compensation (*Forfeit*). This finding is

⁴⁸ For companies with fiscal year end in January, February, and March the post-treatment period is in fiscal year 2021 and the pre-treatment period in year 2020.

⁴⁹ Another potential difference is that Batish et al. (2020) use a sample of Russell 3000 companies, while we focus on companies in the S&P 500 index.

consistent with prior evidence that unfavorable adjustments are less common (Bol and Smith 2011). We further decompose the favorable modifications according to their time horizon. Most firms in our sample modify only the short-term incentive plans (100). 17 firms modify only the long-term plan, while 44 both.

[Table 3.2 here]

Table 3.2 Panel B presents a more detailed description of the favorable IPMs. Consistently with prior literature, we find a substantial heterogeneity in how companies adjust their incentive compensation plans (Merchant 1989). The most common modification is an ex-post discretionary adjustment of the short-term incentive payout (34.62% of companies implementing favorable IPMs). Special awards, instead, are the most common long-term modification (9.89%), while it was less frequent for short-term plans (4.4%).

Other companies prefer to estimate the impact of COVID-19 on their performance measure and adjust them by adding back some costs or projecting revenues (13.74% for short-term plans and 5.49% for long-term plans). Others modify the performance period, excluding some months from the calculation of short-term performance measures or the whole 2020 fiscal year for long-term incentives. Alternatively, they divide the performance period into two or more parts (e.g., first and second semesters of 2020 or separate years for long-term incentives) to establish different performance measures or targets for each part (13.19% for short-term and 8.24% for long-term incentive plans).

Among the adjustments before the end of the performance period, introducing new objective measures or changing their weights (24.18% for short-term and 4.4% for long-term plans) is relatively more frequent than introducing new subjective measures or changing their weight (19.23% for short-term and 1.1% for long-term plans). Other IPMs involve adjusting the targets (14.29% for short-term and 4.95% for long-term plans) or the minimum payout thresholds (4.95% for short-term plans), thereby reducing the risk of not reaching targets or obtaining a zero payout. Only a few companies introduce adjustments by increasing the weight of relative performance evaluation (1.65% for short-term and 3.3% for long-term plans) despite previous studies highlighting the usefulness of this practice in case of uncontrollable events (Holmstrom 1982; Albuquerque 2014). 4 companies adjust by increasing the weights of group or corporate performance measures (2.2% for short-term plans), while 4 adopt an opposite perspective and increase the weights of individual measures (2.2% for short-term plans). This decision is related to the subjective evaluation of leadership or response to the Pandemic. Other

companies recognize the difficulty in defining or changing incentive plans in an uncertain environment and postpone their definition to better understand the changed environmental conditions (8.24% for short-term and 12.3% for long-term plans).

Finally, a few favorable IPMs are specific to the long-term incentive plans. 8 companies adjust by increasing the weight of restricted stock units (RSU) over performance-based incentives. They motivate the decision with the difficulty of setting targets under uncertain periods and the need to favor retention. Only one company increases the weight of stock options, disclosing a similar reason. 2 firms change the quantity or the value of the equity awards to account for the volatility of the stock market during the pandemic, and one company increases the timing of the outstanding stock options.

Table 3.3 presents the summary statistics of the variables used in the analyses. Appendix 3.3 presents additional details about the industry composition of the sample and IPMs.

[Table 3.3 here]

3.4.2 Determinants of Favorable IPMs

We report the results of the IPMs' determinants analysis in Table 3.4. Consistently with H3.1, our three measures of COVID-19 impact are significantly related to the probability of introducing favorable IPMs. Accordingly, to the sign of the measure, *Sales_%var* and *ReturnsCovid* exhibit a negative and significant correlation (Column 1, $p < 0.1$, $p < 0.01$), while *CovidDeaths* is positive and significant (Column 1, $p < 0.05$). Then, results suggest that companies negatively affected by the pandemic are more likely to introduce favorable IPMs.

Regarding the quality of the governance, we find that only *CompComm* is negative and significant. This result is consistent with H3.2, so that companies with better governance are less willing to introduce favorable IPMs. Nevertheless, we do not find results for CEO duality. For this reason, it is possible that our findings capture an effect specific to the compensation committees rather than the overall quality of governance. For example, compensation committees may increase the procedural complexity of modifying incentive plans, thereby reducing the flexibility of renegotiation in case of uncontrollable events. Finally, we do not find that board independence increases the likelihood of introducing favorable IPMs, but we find that it is negatively and significantly correlated with *LT_Change* (Column 2: $p < 0.05$). Among the companies that introduce favorable IPMs, companies with a higher percentage of independent directors are less likely to modify the long-term incentive plans and prefer to

change only the short-term ones. Since long-term incentives often involve equity-based compensation, independent directors may be particularly attentive to protecting shareholders' wealth.

We also find that the coefficient of *Tenure* is negative and significant, while the one of *Age65* is positive and significant (Column 1: $p < 0.1$, $p < 0.05$). This result supports H3.3. Companies are more likely to introduce favorable IPMs when CEOs have higher incentives to leave, either because they have relatively short tenures (and thus weaker organizational attachment) or because they are approaching retirement age, making retention more uncertain.

[Table 3.4 here]

For robustness, we use the number of IPMs rather than a dummy variable to proxy for the intensity of the changes. Appendix 3.4 shows the results estimated with OLS, which are qualitatively equivalent.

3.4.3 Consequences of Favorable IPMs

Table 3.5 presents the estimation of Equation 1 using the three types of strategic investments as dependent variables. The interaction term *After x Favorable* is positive and significant for R&D (Column 1: $p < 0.01$) and the ES score (Column 3: $p < 0.05$). On the contrary, the interaction term for Capex is negative and significant (Column 2: $p < 0.05$).

On average, companies implementing favorable IPMs are associated with a change in scaled R&D 0.982 percentage points higher than companies not adjusting incentive plans. This association can be economically significant since the mean of scaled R&D in the sample is 3.977 (Table 3.3). Moreover, companies that modify incentive plans are associated with an average change in their ES score of 1.907 points higher than the control group (Column 3).

These findings suggest that companies introducing favorable IPMs for the COVID-19 pandemic sustain investments in R&D and ES but limit Capex compared to companies not modifying incentive plans⁵⁰. Consistently with the investment strategy described by Flammer and Ioannou (2021), we interpret these results as suggesting lower managerial myopia during a crisis. Therefore, they support H3.4.

⁵⁰ Appendix 3.5, we estimate the model including companies that implement unfavorable IPMs or whose CEO forfeit compensation. We also add the dummy variables *Unfavorable* and *Forfeit* as well as their interaction term with *After*. We do not find significant interactions with these variables, while *Favorable x After* remains consistent with the findings in Table 5 for all the dependent variables. This suggests that the association we document is specific to favorable IPMs.

[Table 3.5 here]

3.5 ADDITIONAL ANALYSES

3.5.1 Temporal Horizon of IPMs and Investments

With our additional analysis, we disentangle the effects of the IPMs according to their temporal horizon. For this purpose, we modify equation 1 and substitute *Favorable* with the two dummy variables *LT_Change* and *ST_Change*. Companies that do not introduce IPMs remain the control group:

$$\begin{aligned} Y_{it} = & \beta_1 LT_Change_i \times After_t + \beta_2 ST_Change_i \times After_t + \\ & \beta_3 LT_Change_i + \beta_4 ST_Change_i + \beta_5 After_t + \\ & B \cdot Determinants_{it} + B \cdot Controls_{it} + IndustryFE_i + \alpha + \epsilon_{it} \end{aligned} \quad (2)$$

Table 3.6 reports the estimation of Equation 2. We find that the interaction term *After x LT_Change* is positive and significant for R&D (Column 1: 1.897, $p < 0.01$) and ES (Column 3: 2.803, $p < 0.1$), while it is negative and significant for *Capex* (Column 2: -0.798, $p < 0.05$). Conversely, the interaction terms *After x ST_Change* are insignificant for every dependent variable.

We also test the difference between the two interaction terms using Wald tests. We find that *After x LT_Change* is significantly different from *After x ST_Change* for R&D ($F(1, 249) = 5.61$, $p = 0.019$). This finding suggests that preserving the motivational effect of long-term incentive plans during uncontrollable events can lead managers to invest relatively more in R&D. However, we do not find statistically significant differences between the interactions for *Capex* ($F(1, 439) = 2.04$, $p = 0.154$) and ES ($F(1, 434) = 0.79$, $p = 0.374$).

[Table 3.6 here]

A potential caveat of the approach we use to classify long- and short-term IPMs is that some equity or option grants of the long-term incentive plan may work as short-term incentives when they are about to vest (Edmans, Fang, and Lewellen 2017; Edmans, Fang, and Huang 2022)⁵¹. For this reason, we adopt an alternative definition of long-term IPM based on the

⁵¹ For example, in the middle of 2020 the board may change the targets of a 2018-2020 performance stock grant that vest at the beginning of the next year. This modification affects the compensation of the CEO in 2021 but not afterward. Some proxy statements do not clearly indicate the vesting period of the equity grants modified. If not explicitly mentioned, we assume that performance stocks having 2020 as the last year of the performance period vest before the end of 2021. On the contrary, we assume that retention stocks granted in 2020 vest after 2021.

payout time for cash incentives or the vesting dates of equity awards and stock options, and find analogous results (untabulated).

3.5.2 Favorable IPMs and Financial Performance

To complement our analysis of managerial myopia, we test whether modifying incentive plans is associated with financial performance. If IPMs reduce managerial myopia, treated firms should present a negative association with short-term measures of financial performance and a positive association with long-term ones. Moreover, consistently with our theoretical framework, we expect that long-term favorable IPMs are more effective than favorable short-term ones.

We consider an indicator of managerial myopia related to short-term financial results. *BeatAnalyst* is a dummy variable equal to 1 if the company meets or surpasses the annual EPS analyst consensus forecast by a small margin. The intuition behind this measure is that managers reporting results slightly better than the analysts' forecasts are more likely to receive a positive short-term market reaction (He and Tian 2013; Kraft, Vashishtha, and Venkatachalam 2018). We use a range between 0.00 and +0.05 dollars from the consensus forecast⁵². Moreover, we use Tobin's Q as an indicator of expected long-term performance (Ferrell, Liang, and Renneboog 2016).

Table 3.7 reports the estimated coefficients of equation 2 using the three measures of financial performance as dependent variables⁵³. The interaction term *After x LT_Change* is negative and significant for our measure of short-termism (Column 1: $p < 0.1$). Companies modifying long-term incentives, on average, change the probability to closely beat analyst consensus forecasts 0.152 points lower compared to the control group. On the contrary, the interaction term is positive and significant for Tobin's Q (Column 3: 29.86, $p < 0.1$).

The coefficient *After x ST_Change* is not statistically significant at conventional levels for all the dependent variables. We perform a Wald test for the difference of the two interaction terms, and we obtain significant results for *TobinQ* ($F(1, 439) = 4.46$, $p = 0.035$), but not for *BeatAnalyst* ($F(1, 429) = 1.58$, $p = 0.209$).

⁵² This range is larger than the difference of 1 dollar cent adopted in some previous studies (e.g., Kraft, Vashishtha, and Venkatachalam 2018). We motivate this decision with the small size of our sample: only 21 companies fall in the range between 0 and +0.01 dollars during fiscal year 2020. Similarly, we do not consider companies reporting small positive EPS. Only one company in our sample reports EPS slightly higher than the zero earnings threshold in 2020.

⁵³ The dependent variables are multiplied by 100 to increase readability.

Overall, our findings suggest that companies modifying long-term incentives during the pandemic are less oriented towards short-term results than companies modifying only short-term incentives and the control group. This finding indicates a lower degree of managerial myopia after the modification of the long-term incentive plan and provides further support to H3.4.

[Table 3.7 here]

3.5.3 Consequences of Favorable IPMs for Investments Conditional on COVID-19

Impact

In this section, we provide additional evidence that companies negatively affected by the uncontrollable event drive the positive association between favorable IPMs and changes in R&D. These results further support H3.4 and our theoretical framework that assumes a negative impact of the uncontrollable event.

To test this cross-sectional association, we modify equation 1 adding interaction terms with the impact of COVID-19, where *Impact* is either *Sales_%var* or *ReturnsCovid*:

$$RnD_{it} = \beta_1 Favorable_i \times After_t \times Impact_i + \beta_2 Favorable_i \times After_t + \beta_3 After_t \times Impact_i + \beta_4 Favorable_i + \beta_5 Impact_i + \beta_6 After_t + B \cdot Determinants_{it} + B \cdot Controls_{it} + IndustryFE_i + \alpha + \epsilon_{it} \quad (3)$$

The coefficient of interest is β_1 . A negative coefficient suggests that, the more negative the impact of the pandemic on financial performance, the more positive is the association between favorable IPMs and changes in R&D.

Table 3.8 reports the results. We find that *Favorable x After x Sales_%var* (Column 1: -8.724, $p < 0.05$) and *Favorable x After x Sales_%var* (Column 2: -6.040, $p < 0.1$) are negative and significant for R&D expenditures. This result suggests that favorable IPMs reduce managerial myopia in companies negatively impacted by uncontrollable events. However, we do not find that the triple interaction term is significant for Capex or ES score (untabulated), suggesting that R&D expenditures are more sensitive to favorable IPMs than other forms of investments.

[Table 3.8 here]

3.6 ROBUSTNESS CHECKS

3.6.1 Falsification Test

A potential concern of our analyses of the consequences of favorable IPMs on investment decisions is that the associations between favorable IPMs and the investment variables capture a trend that exists before the pandemic, which is also correlated to the implementation of favorable IPMs. We address this potential issue by performing a falsification test. Table 3.9 presents the results of estimating equation 1 in the period 2018-2019. We set *After2019* equal to 1 in 2019 and *After2019* equal to 0 in 2018. We do not find any significant interaction term (*After2019* $\times$ *Favorable*), suggesting that our main analysis captures correlations between events that happened during the COVID-19 period rather than pre-existing differences in trends between the treatment and the control group.

[Table 3.9 here]

3.6.2 Entropy Balancing

In our analysis of the consequences, we compared investments of companies that introduce favorable IPMs in response to COVID-19 with companies that do not alter their compensation plans for COVID-19. However, the decision to change compensation contracts is not exogenous and may be correlated to differences between the treatment and the control group. Although we control for several company characteristics, our results may suffer from an omitted variable bias (Shipman, Swanquist, and Whited 2017).

To reduce this concern, we provide further analyses using entropy balancing. This technique aims to increase the comparability of the treatment and the control group (Hainmueller 2012). It assigns weights to the observations such that the treatment and the control group present similar post-weighting distribution moments (e.g., mean, variance, and skewness) of the covariates. This approach reduces the concern that latent omitted variables that cause differences in the covariates bias the estimation of the association (Çolak and Öztekin 2021). For this reason, recent studies in accounting adopt this technique (McMullin and Schonberger 2020; Ege, Seidel, Sterin and Wood 2022; Godsell 2022).

We balance the sample according to three moments: mean, variance, and skewness. Appendix 3.6 includes descriptive statistics of the full sample before (panel A) and after the procedure (panel B). We balance the sample, including only companies that report R&D, to avoid missing values generating differences between the treatment and the control groups.

Table 3.10 contains the estimation of Equation 1 using the entropy-balanced samples. In line with the findings of Table 3.5, the coefficient *After x Favorable* is positive and significant for R&D (Column 1: 1.454, $p < 0.01$) and the ES score (Column 3: 4.886, $p < 0.05$). On the contrary, the interaction term for Capex is negative and significant (Column 2: -0.686, $p < 0.1$).

[Table 3.10 here]

3.6.3 Ruling Out Alternative Explanations

A potential concern of our analysis on the association between favorable IPMs and investment decisions is that companies with high levels of R&D, Capex, or ES performance before the pandemic could be more likely to introduce favorable IPMs. Therefore, the positive association we document in our consequences' analysis may be explained by the companies that invested more before the pandemic. To rule out this possible explanation, we estimate the logit model of Table 3.4 Column 1 by adding the lagged investment variables as additional determinants. Table 3.11 panel A reports the results. None of the coefficients of the investment variable is statistically different from zero, reducing this concern⁵⁴.

Another related concern is that the association we document is due to companies that frequently change their incentive plan. In other words, the association between favorable IPMs and investment variables does not occur because companies change their incentive plan in response to COVID-19, but because they tend to adopt a more flexible compensation approach in general.

To rule out this alternative explanation, we collect data about performance measures used for CEOs' compensation targets from 2016 to 2019 from ISS Incentive Lab. We check whether companies modified their set of performance measures from year to year and create two dummy variables. *GoalChanger* is equal to 1 if the company changed performance measures in the CEO bonus plans (e.g., revenues, EPS...) at least twice in the previous 3 years and 0 otherwise.

Table 3.11 Panel B reproduces our determinants analysis, adding *GoalChanger* as a predictor. From Column 1 to Column 3, we use *Favorable*, *LT_Change*, and *ST_Change* as dependent variables, keeping companies that do not introduce IPMs as the baseline. The

⁵⁴ We also use the industry-based classification by Kile and Phillips (2009) to identify high-technology firms and include it as an additional binary determinant in the model of Table 3.4 Column 1. The estimated coefficient is positive but insignificant at conventional statistical levels (untabulated, $p = 0.355$).

coefficients of *GoalChanger* are not statistically significant for every dependent variable. Similarly, we do not find statistically significant coefficients if we consider frequent goal changers, the companies that changed performance measures in the CEO bonus plans three or more times over the previous four years (untabulated). These results reduce the concern that favorable IPMs for COVID-19 reflect a more general tendency of companies to change frequently incentive plans from year to year.

[Table 3.11 here]

3.7 CONCLUSIONS

This study examines CEO incentive plan modifications (IPMs) following the COVID-19 pandemic. We identify several determinants of favorable IPMs and find that the likelihood of such modifications increases with the severity of COVID-19's negative impact on firm performance. Favorable IPMs are also more likely when CEOs have shorter tenures or are over 65, suggesting that potential early departure due to weaker organizational attachment or retirement plays a role in deciding about such modifications. Moreover, we find that the presence of a compensation committee and independent directors reduces the likelihood of implementing favorable and long-term IPMs, indicating that weaker governance structures may allow for modifications less aligned with shareholder interests. We further examine whether CEO IPMs adopted in response to uncontrollable events affect managerial myopia. Overall, our evidence suggests that IPMs help mitigate managerial myopia. We find a positive association between IPMs and changes in R&D and ES performance but a negative association with Capex. These effects are stronger for long-term IPMs than for short-term ones. Finally, consistently with a reduction of managerial myopia, we observe that firms introducing long-term IPMs are less likely to narrowly beat analysts' earnings forecasts and exhibit higher Tobin's Q.

This study presents some limitations. First, the introduction of favorable IPMs is not exogenous. It reflects firm-specific responses to the impact of COVID-19 and other characteristics. Despite the control variables and the robustness check with entropy balancing, our analysis does not establish causality. Findings should be interpreted as indicative associations. Second, our sample is limited to US companies with large market capitalization and we focus on CEO incentive plans. Findings may not be generalizable to smaller companies, non-US contexts, or non-executive compensation contracts, as the power dynamics and external opportunities of managers may differ. Moreover, the COVID-19 pandemic is a

peculiar context as it is an unprecedented unexpected shock that introduces a significant level of uncertainty. Predictable adverse uncontrollable events may produce different results. Finally, some elements of the incentive plan classification are subject to measurement error. We cannot observe counterfactual compensation, making it difficult to identify favorable IPMs without assumptions. Our classification is largely based on the reasons reported in the Compensation Discussion and Analysis, which may reflect strategically manipulated narratives. Future research could explore these dynamics further.

3.8 REFERENCES

- Abernethy, M. A., L. Jiang, and Y. F. Kuang. 2019. Can organizational identification mitigate the CEO horizon problem? *Accounting, Organizations & Society* 78: 101056.
- Agarwal, V., R. Vashishtha, and M. Venkatachalam. 2018. Mutual fund transparency and corporate myopia. *Review of Financial Studies* 31 (5): 1966–2003.
- Albuquerque, A. M. 2014. Do growth-option firms use less relative performance evaluation? *Accounting Review* 89 (1): 27–60.
- Albuquerque, A., B. Bennett, C. Custódio, and D. Cvijanović. 2023. CEO compensation and real estate prices: pay for luck or pay for action? *Review of Accounting Studies* 28 (4): 2401–2447.
- Allee, K. D., B. A. Badertscher, and T. L. Yohn. 2020. Private versus public corporate ownership: implications for future changes in profitability. *Journal of Management Accounting Research* 32 (2): 27–55.
- Antle, R., and J. S. Demski. 1988. The controllability principle in responsibility accounting. *Accounting Review* 63 (4): 700–718.
- Armstrong, C. S., S. A. Glaeser, and S. Huang. 2022. Contracting with controllable risk. *Accounting Review* 97 (4): 27–50.
- Asay, H. S., J. B. Garrett, and W. B. Tayler. 2023. Fairness and luck: contract away or come what may? *Journal of Management Accounting Research* 35 (1): 1–21.
- Baiman, S., and M. V. Rajan. 1995. The informational advantages of discretionary bonus schemes. *Accounting Review* 70 (4): 557–579.
- Baker, G., R. Gibbons, and K. J. Murphy. 1994. Subjective performance measures in optimal incentive contracts. *Quarterly Journal of Economics* 109 (4): 1125–1156.
- Batish, A., A. Gordon, D. F. Larcker, B. Tayan, E. M. Watts, and Y. Courtney. 2020. *Sharing the pain: How did boards adjust CEO pay in response to covid-19?* Rock Center for

- Corporate Governance at Stanford University Closer Look Series: Topics, Issues and Controversies in Corporate Governance no. CGRP-86.
- Becker, S. D., M. D. Mahlendorf, U. Schäffer, and M. Thaten. 2016. Budgeting in times of economic crisis. *Contemporary Accounting Research* 33 (4): 1489–1517.
- Bedford, A., M. Bugeja, S. Ghannam, D. Jeganathan, and N. Ma. 2023. Were CEO pay cuts during the COVID-19 pandemic merely symbolic? Shareholders' reaction and outrage. *Pacific-Basin Finance Journal* 79: 101993.
- Bedford, D. S., R. F. Speklé, and S. K. Widener. 2022. Budgeting and employee stress in times of crisis: evidence from the Covid-19 pandemic. *Accounting, Organizations & Society* 101: 101346.
- Bennett, B., J. C. Bettis, R. Gopalan, and T. Milbourn. 2017. Compensation goals and firm performance. *Journal of Financial Economics* 124 (2): 307–330.
- Bertrand, M., and S. Mullainathan. 2001. Are CEOs rewarded for luck? the ones without principals are. *Quarterly Journal of Economics* 116 (3): 901–932.
- Bloomfield, M., B. Gipper, J. D. Kepler, and D. Tsui. 2021. Cost shielding in executive bonus plans. *Journal of Accounting & Economics* 72 (2-3): 101428.
- Bol, J. C., and S. D. Smith. 2011. Spillover effects in subjective performance evaluation: bias and the asymmetric influence of controllability. *Accounting Review* 86 (4): 1213–1230.
- Bol, J. C., G. Hecht, and S. D. Smith. 2015. Managers' discretionary adjustments: the influence of uncontrollable events and compensation interdependence. *Contemporary Accounting Research* 32 (1): 139–159.
- Brochet, F., M. Loumiot, and G. Serafeim. 2015. Speaking of the short-term: disclosure horizon and managerial myopia. *Review of Accounting Studies* 20 (3): 1122–1163.
- Bugeja, M., Z. Matolcsy, and H. Spiropoulos. 2016. The association between gender-diverse compensation committees and CEO compensation. *Journal of Business Ethics* 139 (2): 375–390.
- Burkert, M., F. M. Fischer, and U. Schäffer. 2011. Application of the controllability principle and managerial performance: the role of role perceptions. *Management Accounting Research* 22 (3): 143–159.
- Campbell, T. C., and M. E. Thompson. 2015. Why are CEOs paid for good luck? an empirical comparison of explanations for pay-for-luck asymmetry. *Journal of Corporate Finance* 35: 247–264.

- Campello, M., E. Giambona, J. R. Graham, and C. R. Harvey. 2011. Liquidity management and corporate investment during a financial crisis. *Review of Financial Studies* 24 (6): 1944–1979.
- Carter, M. E., and L. J. Lynch. 2004. The effect of stock option repricing on employee turnover. *Journal of Accounting & Economics* 37 (1): 91–112.
- Casas-Arce, P., R. J. Indjejikian, and M. Matějka. 2020. Bonus plan choices during an economic downturn. *Journal of Management Accounting Research* 32 (2): 85–105.
- Chakrabarti, A. 2015. Organizational adaptation in an economic shock: the role of growth reconfiguration. *Strategic Management Journal* 36 (11): 1717–1738.
- Chen, M. A. 2004. Executive option repricing, incentives, and retention. *Journal of Finance* 59 (3): 1167–1199.
- Cohen, J. R., L. Holder-Webb, D. J. Sharp, and L. W. Pant. 2007. The effects of perceived fairness on opportunistic behavior. *Contemporary Accounting Research* 24 (4): 1119–1138.
- Çolak, G., and Ö. Öztekin. 2021. The impact of COVID-19 pandemic on bank lending around the world. *Journal of Banking & Finance* 133: 106207.
- Chua, W. F., J. Graaf, and K. Kraus. 2025. “What is this thing called controllability?” A field study of the integration of the controllability principle in the redesign of a performance measurement system. *Accounting Review*: 1–28.
- Curtis, A., V. Li, and P. H. Patrick. 2021. The use of adjusted earnings in performance evaluation. *Review of Accounting Studies* 26 (4): 1290–1322.
- Dikolli, S. S. 2001. Agent employment horizons and contracting demand for forward-looking performance measures. *Journal of Accounting Research* 39 (3): 481–494.
- Dikolli, S. S., S. L. Kulp, and K. L. Sedatole. 2013. The use of contract adjustments to lengthen the CEO horizon in the presence of internal and external monitoring. *Journal of Management Accounting Research* 25 (1): 199–229.
- Edmans, A., V. W. Fang, and A. H. Huang. 2022. The Long-Term consequences of Short-Term incentives. *Journal of Accounting Research* 60 (3): 1007–1046.
- Edmans, A., V. W. Fang, and K. A. Lewellen. 2017. Equity vesting and investment. *Review of Financial Studies* 30 (7): 2229–2271.
- Ege, M., T. A. Seidel, M. Sterin, and D. A. Wood. 2022. The influence of management’s internal audit experience on earnings management. *Contemporary Accounting Research* 39 (3): 1834–1870.

- Fehr, E., and K. Schmidt. 1999. A theory of fairness, competition, and cooperation. *Quarterly Journal of Economics* 114 (3): 817–868.
- Ferrell, A., H. Liang, and L. Renneboog. 2016. Socially responsible firms. *Journal of Financial Economics* 122 (3): 585–606.
- Flammer, C., and I. Ioannou. 2021. Strategic management during the financial crisis: how firms adjust their strategic investments in response to credit market disruptions. *Strategic Management Journal* 42 (7): 1275–1298.
- Frydman, C., and D. Jenter. 2010. CEO compensation. *Annual Review of Financial Economics* 2 (1): 75–102.
- Fudenberg, D., and J. Tirole. 1990. Moral hazard and renegotiation in agency contracts. *Econometrica* 58 (6): 1279–1319.
- Garvey, G. T., and T. T. Milbourn. 2006. Asymmetric benchmarking in compensation: executives are rewarded for good luck but not penalized for bad. *Journal of Financial Economics* 82 (1): 197–225.
- Gilson, S. C., and M. R. Vetsuypens. 1993. CEO compensation in financially distressed firms: an empirical analysis. *Journal of Finance* 48 (2): 425–458.
- Giraud, F., P. Langevin, and C. Mendoza. 2008. Justice as a rationale for the controllability principle: A study of managers' opinions. *Management Accounting Research* 19 (1): 32–44.
- Godsell, D. 2022. Financial reporting consequences of sovereign wealth fund investment. *Contemporary Accounting Research* 39 (3): 2090–2129.
- Hainmueller, J. 2012. Entropy balancing for causal effects: A multivariate reweighting method to produce balanced samples in observational studies. *Political Analysis* 20 (1): 25–46.
- He, J. (J.), and X. Tian. 2013. The dark side of analyst coverage: the case of innovation. *Journal of Financial Economics* 109 (3): 856–878.
- Hermalin, B. E., and M. L. Katz. 1991. Moral hazard and verifiability: the effects of renegotiation in agency. *Econometrica* 59 (6): 1735–1753.
- Holmstrom, B. 1982. Moral hazard in teams. *Bell Journal of Economics* 13 (2): 324–340.
- Hyun, J. H., M. Matějka, P. (Seung Hwan) Oh, and T. S. Ahn. 2022. Performance targets and ex post incentive plan adjustments. *Contemporary Accounting Research* 39 (2): 863–892.
- Ittner, C. D., and D. F. Larcker. 2002. Determinants of performance measure choices in worker incentive plans. *Journal of Labor Economics* 20 (S2): S58–90.
- Kang, J. K., and J. Kim. 2020. Do family firms invest more than nonfamily firms in employee-friendly policies? *Management Science* 66 (3): 1300–1324.

- Kile, C. O., and M. E. Phillips. 2009. Using industry classification codes to sample high-technology firms: Analysis and recommendations. *Journal of Accounting, Auditing & Finance*, 24(1), 35-58.
- Kim, S., and J. Y. Shin 2018. Subjective Adjustments to Objective Performance Measures in Executive Annual Bonus Contracts. *AAA 2018 Management Accounting Section (MAS) Meeting*, Available at SSRN: <https://ssrn.com/abstract=3018199> or <http://dx.doi.org/10.2139/ssrn.3018199>
- Kim, S., M. Matějka, and J. Park. 2023. Economic determinants and consequences of performance target difficulty. *Accounting Review* 98 (2): 361–387.
- Kraft, A. G., R. Vashishtha, and M. Venkatachalam. 2018. Frequent financial reporting and managerial myopia. *Accounting Review* 93 (2): 249–275.
- Ladika, T., and Z. Sautner. 2020. Managerial short-termism and investment: evidence from accelerated option vesting. *Review of Finance* 24 (2): 305–344.
- Langevin, P., and C. Mendoza. 2013. How can management control system fairness reduce managers' unethical behaviours? *European Management Journal* 31 (3): 209–222.
- Laverty, K. J. 2004. Managerial myopia or systemic short-termism? the importance of managerial systems in valuing the long term. *Management Decision* 42 (8): 949–962.
- Li, T., and E. B. Yang. 2023. *CEO compensation, stakeholders, and fairness: Evidence from the COVID-19 pandemic*. Available at SSRN 4097179. <https://ssrn.com/abstract=4097179>
- Libby, T. 2001. Referent cognitions and budgetary fairness: a research note. *Journal of Management Accounting Research* 13 (1): 91–106.
- Lillis, A. M., M. A. Malina, and J. Mundy. 2022. The role of subjectivity in mitigating incentive contracting risks. *Accounting Review* 97 (1): 365–388.
- Luft, J. L. 1997. Fairness, ethics and the effect of management accounting on transaction costs. *Journal of Management Accounting Research* 9: 1–20.
- Matějka, M., K. A. Merchant, and W. A. Van Der Stede. 2009. Employment horizon and the choice of performance measures: empirical evidence from annual bonus plans of loss-making entities. *Management Science* 55 (6): 890–905.
- Matolcsy, Z. P. 2000. Executive cash compensation and corporate performance during different economic cycles. *Contemporary Accounting Research* 17 (4): 671–692.
- McMullin, J. L., and B. Schonberger. 2020. Entropy-balanced accruals. *Review of Accounting Studies* 25 (1): 84–119.
- Merchant, K. A. 1987. How and why firms disregard the controllability principle. *Accounting & Management: Field Study Perspectives*: 316–338.

- Merchant, K. A. 1989. *Rewarding Results : Motivating Profit Center Managers* / Kenneth A. Merchant, Boston: Harvard Business School Press.
- Merchant, K. A., and W. A. Van der Stede 2017. *Management Control Systems: Performance Measurement, Evaluation, and Incentives* / Kenneth A. Merchant (University of Southern California), wim A. van der Stede (London School of Economics), Harlow, England: Pearson.
- Potepa, J. 2020. The treatment of special items in determining CEO cash compensation. *Review of Accounting Studies* 25 (2): 558–596.
- Rabin, M. 1993. Incorporating fairness into game theory and economics. *American Economic Review* 83 (5): 1281–1302.
- Ramelli, S., and A. F. Wagner. 2020. Feverish stock price reactions to COVID-19. *Review of Corporate Finance Studies* 9 (3): cfaa012.
- Şabac, F. 2007. Dynamic agency with renegotiation and managerial tenure. *Management Science* 53 (5): 849–864.
- Shields, M. D., and W. S. Waller. 1988. A behavioral study of accounting variables in performance—incentive contracts. *Accounting, Organizations & Society* 13 (6): 581–594.
- Shin, T. 2016. Fair pay or power play? pay equity, managerial power, and compensation adjustments for CEOs. *Journal of Management* 42 (2): 419–448.
- Shipman, J. E., Q. T. Swanquist, and R. L. Whited. 2017. Propensity score matching in accounting research. *Accounting Review* 92 (1): 213–244.
- Stein, J. C. 1989. Efficient capital markets, inefficient firms: A model of myopic corporate behavior. *Quarterly Journal of Economics* 104 (4): 655–669.
- Su, S. X., K. Baird, and N. Nuhu. 2022. Controllability of performance measures and managerial performance: the mediating role of fairness. *Meditari Accountancy Research* 30 (2): 313–341.
- Waller, W. S., and C. W. Chow. 1985. The self-selection and effort effects of standard-based employment contracts: A framework and some empirical evidence. *Accounting Review* 60 (3): 458–476.
- Yim, A. T. 2001. Renegotiation and relative performance evaluation: why an informative signal may be useless. *Review of Accounting Studies* 6 (1): 77–108.

3.9 TABLES

Table 3.1

Sample Selection

S&P 500 constituents in 2020 with proxy statements	510
- Companies missing one or more variables of interest	(37)
- Companies that implemented unfavorable IPMs or whose CEO forfeited compensation	(33)
Unique companies	440
Total firm-year observations	880

Table 3.2***Classification of IPMs******Panel A: Classification of companies according to IPMs***

No IPMs	Favorable IPMs			Forfeit	Unfavorable IPMs	Total
	Long- and Short- Term	Only Long- Term	Only Short- Term			
279	44	17	100	19	14	473

Panel B: Absolute and Relative Frequencies of Favorable IPMs by Type

Type of Modification	Short-Term Plan		Long-Term Plan	
	N	% of Favorable IPMs	N	% of Favorable IPMs
Discretionary Payout Modification	63	34.62%	10	5.49%
Special Awards	8	4.40%	18	9.89%
Adjust for the COVID Impact	25	13.74%	10	5.49%
Modified Performance Period	24	13.19%	15	8.24%
Modified Objective Measures	44	24.18%	8	4.40%
Increased Discretion	35	19.23%	2	1.10%
Modified Targets	26	14.29%	9	4.95%
Modified Minimum Payout Threshold	9	4.95%	0	0.00%
Increased Relative Performance Evaluation	3	1.65%	6	3.30%
Increased Group/Company Measures	4	2.20%	0	0.00%
Increased Individual Measures	4	2.20%	0	0.00%
Postponed Incentive Plan Definition	15	8.24%	7	3.85%
Increased RSU	0	0.00%	9	4.95%
Modified Timing of The Award	0	0.00%	1	0.55%
Modified Stock Options Target Price	0	0.00%	2	1.10%
Increased Stock Options	0	0.00%	1	0.55%
Increased RSU	0	0.00%	9	4.95%

This table presents frequencies of IPMs for the fiscal year 2020 in response to the COVID-19 pandemic. In Panel A we classify the modification in terms of favorableness. No IPMs indicates companies that did not modify incentive plans. Favorable IPMs is the frequency of modifications in favor of the CEOs. Forfeit the companies whose modification stems from the CEO voluntarily forgoing their compensation. Unfavorable IPMs indicates the companies that adopt modifications against the interests of the CEOs. We further classify favorable IPMs according to their temporal horizon. LT & ST is the frequency of companies that modify both the long-term and the short-term incentive plan. Only LT and Only ST are the frequency of companies that modify only the long-term and only the short-term plan respectively. The second dimension is Favorableness. In panel B we present frequencies about specific types of IPMs by companies that we classified as favorable. Some companies implemented more than one modification simultaneously.

Table 3.3
Summary Statistics

	N	Mean	Std. Dev.	min	p25	p75	max
RnD	502	4.009	5.657	0	.082	5.581	30.791
Capex	880	3.426	3.409	0	.977	4.996	16.82
ESscore	875	62.711	19.082	14.92	49.635	77.815	92.91
BeatAnalyst	866	.263	.441	0	0	1	1
TobinQ	880	270.34	213.992	91.2	131.025	321.593	1174.732
Favorable	880	.366	.482	0	0	1	1
LT_Change	880	.139	.346	0	0	0	1
ST_Change	880	.227	.419	0	0	0	1
After	880	.5	.5	0	0	1	1
GoalsChanger	753	.635	.482	0	0	1	1
Sales_%var	880	.016	.165	-.636	-.047	.083	.626
ReturnsCovid	880	-.176	.15	-.639	-.266	-.065	.107
CovidDeaths	880	.588	.687	0	0	1.011	2.144
LogAssets	880	10.003	1.335	7.302	9.069	10.823	13.745
Leverage	880	.282	.184	0	.149	.388	.967
BTM	880	.377	.325	-.103	.144	.541	1.586
ROA	880	.113	.082	.004	.049	.149	.388
CompComm	880	.959	.198	0	1	1	1
BdIndep	880	84.646	8.283	55.56	81.25	90.91	93.75
Tenure	880	7.682	6.542	1	3	10	32
Duality	880	.678	.467	0	0	1	1
TotCompRank	880	.762	.205	.012	.679	.914	.997
CashComp	880	.121	.098	0	.076	.128	.676
Salary_%var	880	.003	.185	-.731	0	.045	.982
Age65	880	.127	.333	0	0	0	1

The sample consists of S&P 500 firms during 2020. We exclude firms with unfavorable IPMs, whose CEO voluntarily forfeits incentive compensation, and with missing values in some variables. The sample period includes fiscal years 2019 and 2020. Continuous variables are winsorized at 1% level.

Table 3.4
Determinants of Favorable IPMs

	(1) Favorable	(2) LT_Change
Sales_%var	-1.399* (0.738)	-1.478 (1.208)
ReturnsCovid	-3.770*** (1.004)	0.608 (1.782)
CovidDeaths	0.505** (0.216)	-0.303 (0.408)
LogAssets	-0.206* (0.120)	0.329 (0.253)
Leverage	0.214 (0.606)	0.149 (1.275)
BTM	-0.830 (0.528)	-0.382 (1.093)
ROA	-2.954* (1.791)	2.452 (3.237)
CompComm	-1.030* (0.594)	1.143 (0.795)
BdIndep	-0.0142 (0.0150)	-0.0477** (0.0231)
Tenure	-0.0340* (0.0195)	-0.0156 (0.0332)
Duality	0.245 (0.237)	-0.402 (0.414)
TotCompRank	0.568 (0.723)	-1.565 (1.737)
CashComp	-0.379 (1.456)	-1.726 (3.218)
Salary_%var	-0.772 (0.684)	-2.352** (1.024)
Age65	0.742** (0.373)	0.0155 (0.537)
Observations	440	161
Pseudo R-squared	0.123	0.122

The sample consists of 440 US listed firms with large market capitalization in 2020. We exclude firms with unfavorable IPMs or whose CEO forfeits. The sample period includes only fiscal year 2020. Companies with missing values in the control variables were discarded. In Column 1, the dependent variable Favorable is a dummy variable equal to 1 if a favorable IPM occurs and zero otherwise. The control group consists of the companies not making any modifications. In Column 2, the dependent variable LT_Change is a dummy variable equal to 1 if a favorable IPM occurs in the long-term incentive plan and zero otherwise. The control group consists of the companies making only favorable short-term IPMs. All continuous variables are winsorized at 1% level. Robust standard errors in brackets. The definitions of the other variables are in Appendix 3.1. * p<0.1, **p<0.05, ***p<0.01.

Table 3.5
Consequences of favorable IPMs

	(1) RnD	(2) Capex	(3) ESscore
After x Favorable	0.982*** (0.360)	-0.426** (0.203)	1.907** (0.899)
Favorable	-0.781 (0.589)	0.685** (0.283)	0.628 (1.802)
After	0.943 (0.599)	-0.303 (0.286)	-1.702 (1.520)
Sales_%var	7.809*** (1.913)	1.869*** (0.680)	-1.986 (4.838)
ReturnsCovid	6.550** (2.989)	1.390 (1.639)	9.529 (7.857)
CovidDeaths	-0.860* (0.518)	-0.111 (0.225)	2.753** (1.283)
LogAssets	-0.528 (0.460)	0.159 (0.194)	6.559*** (0.948)
Leverage	-5.617** (2.190)	-1.148 (0.881)	5.316 (5.495)
BTM	-2.714** (1.357)	-0.547 (0.705)	-0.714 (3.713)
ROA	0.120 (5.582)	7.006** (2.974)	19.70 (13.10)
CompComm	2.963*** (0.597)	-0.301 (0.606)	-1.431 (3.757)
BdIndep	-0.0440 (0.0400)	-0.0148 (0.0176)	0.233** (0.0941)
Tenure	0.0174 (0.0579)	-0.00382 (0.0193)	-0.0591 (0.128)
Duality	-0.904 (0.696)	-0.162 (0.235)	1.791 (1.615)
TotCompRank	2.574 (2.603)	-1.472 (1.225)	15.50** (6.246)
CashComp	-0.00924 (3.809)	-1.628 (2.057)	12.00 (9.826)
Salary_%var	1.024 (0.746)	1.517*** (0.552)	0.0365 (3.203)
Age65	0.355 (1.083)	0.00103 (0.387)	-5.803** (2.384)
Observations	500	880	870
R-squared	0.492	0.521	0.446
Industry FE	YES	YES	YES

The sample consists of 440 firms in 2020. We exclude firms with unfavorable IPMs or whose CEO forfeits incentive compensation. The sample period includes fiscal years 2019 and 2020. Companies with missing values for at least 1 year were discarded. The sample of column 1 includes only companies reporting R&D. Favorable is a dummy variable equal to 1 if the company implements an IPM in favor of the CEO. All continuous variables are winsorized at 1% level. Standard errors clustered by firm in brackets. Industry fixed effects are computed according to SIC2 codes. The definitions of the other variables are in Appendix 3.1. * p<0.1, **p<0.05, ***p<0.01.

Table 3.6***Additional analyses: Temporal Horizon of IPMs and Investments***

	(1) RnD	(2) Capex	(3) ESscore
After x LT_Change	1.897*** (0.600)	-0.798** (0.367)	2.803* (1.511)
After x ST_Change	0.555 (0.351)	-0.216 (0.220)	1.409 (0.931)
LT_Change	-0.141 (0.864)	1.109** (0.469)	-3.084 (2.845)
ST_Change	-0.985 (0.659)	0.460 (0.322)	2.464 (2.012)
After	0.944 (0.595)	-0.294 (0.286)	-1.684 (1.524)
Observations	500	880	870
R-squared	0.497	0.522	0.450
Industry FE	YES	YES	YES
Controls	YES	YES	YES

The sample consists of S&P 500 firms in 2020. We exclude firms with unfavorable IPMs or whose CEO forfeits. The sample period includes fiscal years 2019 and 2020. Companies with missing values for at least 1 year were discarded. The sample of column 1, 4, and 6 include only companies reporting R&D. LT_Change (ST_Change) is a dummy variable equal to 1 if the company implements a favorable IPM for the CEO's long-term (only for the short-term) incentive compensation plan. All continuous variables are winsorized at 1% level. Standard errors clustered by firm in brackets. Controls are omitted for clarity. Industry fixed effects are computed according to SIC2 codes. The definitions of the variables are in Appendix 3.1. * p<0.1, **p<0.05, ***p<0.01.

Table 3.7***Additional analyses. Favorable IPMs and Financial Performance***

	(1) BeatAnalyst	(2) TobinQ
After x LT_Change	-0.152* (0.0869)	29.86* (17.13)
After x ST_Change	-0.0323 (0.0660)	-4.928 (10.72)
LT_Change	0.192** (0.0787)	18.34 (23.30)
ST_Change	0.115* (0.0607)	20.65 (16.58)
After	-0.188*** (0.0590)	58.13*** (17.93)
Observations	860	880
R-squared	0.155	0.654
Industry FE	YES	YES
Controls	YES	YES

The sample consists of S&P 500 firms in 2020. We exclude firms with unfavorable IPMs or whose CEO forfeits. The sample period includes fiscal years 2019 and 2020. Companies with missing values for at least 1 year were discarded. BeatAnalyst is a dummy variable equal to one if the company's actual EPS falls in a range between 0 and +0.05 from the analysts' annual consensus forecast. TobinQ is total is the Tobin's Q using the total market capitalization at the end of the year. All continuous variables are winsorized at 1% level Standard errors clustered by firm in brackets. Controls are omitted for clarity. Industry fixed effects are computed according to SIC2 codes. The definitions of the other variables are in Appendix 3.1. * p<0.1, **p<0.05, ***p<0.01.

Table 3.8

Additional analyses. Consequences of Favorable IPMs for Investments Conditional on COVID-19 Impact

	(1) RnD	(2) RnD
After x Favorable x ReturnsCovid	-8.724** (3.992)	
After x Favorable x Sales_%var		-6.040* (3.090)
After x Favorable	-0.791 (0.728)	0.702** (0.335)
After x ReturnsCovid	-0.981 (2.295)	
After x Sales_%var		-1.001 (3.215)
Favorable	-0.614 (0.592)	-0.737 (0.600)
After	0.728 (0.648)	1.031* (0.604)
Observations	500	500
R-squared	0.499	0.498
Industry FE	YES	YES
Controls	YES	YES

The sample consists of S&P 500 firms in 2020 reporting R&D expenditures. We exclude firms with unfavorable IPMs or whose CEO forfeits. The sample period includes fiscal years 2019 and 2020. Companies with missing values for at least 1 year were discarded. All continuous variables are winsorized at 1% level. Standard errors clustered by firm in brackets. Controls are omitted for clarity. Industry fixed effects are computed according to SIC2 codes. The definitions of the variables are in Appendix 3.1. * p<0.1, **p<0.05, ***p<0.01.

Table 3.9***IPMs and Investment Decisions Falsification Test***

	(1) RnD	(2) Capex	(3) ESscore
After2019 x Favorable	-0.307 (0.393)	0.0547 (0.158)	0.287 (0.820)
Favorable	-0.525 (0.663)	0.287 (0.294)	0.858 (1.895)
After2019	0.713*** (0.244)	0.186* (0.110)	2.783*** (0.620)
Observations	528	926	924
Model	OLS	OLS	OLS
R-squared	0.435	0.597	0.466
Industry FE	YES	YES	YES
Controls	YES	YES	YES

The sample consists of S&P 500 firms in 2020. We exclude firms with unfavorable IPMs or whose CEO forfeits. The sample period includes fiscal years 2018 and 2019. In this table, After2019 equals 1 in fiscal year 2019 and 0 in 2018. All the controls are included, except for CovidDeaths that is always equal to 0 in this sample. Companies with missing values for at least 1 year were discarded. Column 1 considers only companies reporting R&D. All continuous variables are winsorized at 1% level. Standard errors clustered by firm in brackets. Industry fixed effects are computed according to SIC2 codes. The definitions of the variables are in Appendix 3.1. * p<0.1, **p<0.05, ***p<0.01.

Table 3.10***IPMs and Investment Decisions. Entropy Balancing***

	(1) RnD	(2) Capex	(3) ESscore
After x Favorable	1.454*** (0.477)	-0.686* (0.349)	4.886** (2.456)
Favorable	-0.493 (0.584)	1.070*** (0.373)	0.826 (2.830)
After	-0.351 (0.692)	0.355 (0.626)	-5.146 (3.288)
Observations	500	500	492
R-squared	0.534	0.515	0.549
Industry FE	YES	YES	YES
Controls	YES	YES	YES

The sample consists of S&P 500 firms in 2020 reporting R&D expenditures. The sample period includes fiscal years 2019 and 2020. Companies with missing values for at least 1 year were discarded. Columns 1, 2 and 3 include only companies reporting R&D. All continuous variables are winsorized at 1% level. The sample is entropy balanced on all the covariates after winsorization on mean, variance, and skewness. Standard errors clustered by firm in brackets. Industry fixed effects are computed according to SIC2 codes. The definitions of the variables are in Appendix 3.1. * p<0.1, **p<0.05, ***p<0.01.

Table 3.11***Alternative Explanations******Panel A: Investment Decisions as Determinants of Favorable IPMs***

	(1) Favorable	(2) LT Change	(3) ST Change
RnD_SC	-2.111 (3.456)		
Capex_SC		-0.995 (3.146)	
ESscore			0.005 (0.007)
Observations	251	440	439
Model	LOGIT	LOGIT	LOGIT
Pseudo R-squared	0.135	0.123	0.130
Controls	YES	YES	YES

Panel B: Frequent Goal Changers as Determinants of Favorable IPMs

	(1) Favorable	(2) LT Change	(3) ST Change
GoalChanger	-0.102 (0.246)	0.050 (0.370)	-0.219 (0.283)
Observations	381	290	327
Model	LOGIT	LOGIT	LOGIT
Pseudo R-squared	0.123	0.179	0.117
Controls	YES	YES	YES

In Panel A, The sample consists of S&P 500 firms in 2020. We exclude firms with unfavorable IPMs or whose CEO forfeits. Companies with missing values for at least 1 year were discarded. RnD_SC and Capex_SC are respectively R&D and capital expenditure scaled by lagged assets one fiscal year before 2020. ESScore is the average of the environmental and the social Refinitiv score one year before 2020. In Panel B, GoalChanger is a dummy variable equal to one if the goals in the CEO compensation plan changed twice in the previous three years. Standard errors in brackets. All continuous variables are winsorized at 1% level. The definitions of the other variables are in Appendix 3.1. * p<0.1, **p<0.05, ***p<0.01.

3.10 APPENDICES

APPENDIX 3.1

Definitions of the Variables

Variable	Description	Source
<u>Independent Variables:</u>		
<i>Favorable</i>	1 if the modification of the incentive plan overall is in favor of the CEO.	Collected and classified from the companies' proxy statements
<i>ST_Change</i>	1 if only short-term incentives were modified, 0 otherwise.	Collected and classified from the companies' proxy statements
<i>LT_Change</i>	1 if long-term incentives were modified, 0 otherwise.	Collected and classified from the companies' proxy statements
<i>Unfavorable</i>	1 if the modification of the incentive plan aims to reduce the compensation of the CEO and the CEO does not voluntarily forfeit his or her compensation. 0 otherwise.	Collected and classified from the companies' proxy statements
<i>Forfeit</i>	1 if the proxy statement indicates that the CEO voluntarily forfeits his or her incentive compensation, 0 otherwise. If the CEO forfeits part of his or her incentive compensation but receives simultaneously a favorable modification, the modification is classified as Favorable.	Collected and classified from the companies' proxy statements
<i>After</i>	Indicator equal to 1 after year 2020 for companies with fiscal year end after March. It is equal to 1 after 2021 for companies with fiscal year end before March 2020, 0 otherwise.	Compustat (for fiscal year ends)
<u>Dependent Variables:</u>		
<i>Capex</i>	Capital Expenditures scaled by previous year total assets multiplied by 100.	Compustat
<i>RnD</i>	R&D expenses scaled by previous year total assets multiplied by 100.	Compustat
<i>EsScore</i>	Mean of SocialScore and EnvScore	Refinitiv
<i>EsgScore</i>	Refinitiv's overall ESG score.	Refinitiv
<i>EnvScore</i>	Refinitiv's environmental pillar score.	Refinitiv
<i>SocialScore</i>	Refinitiv's social pillar score.	Refinitiv
<i>GovScore</i>	Refinitiv's governance pillar score.	Refinitiv
<i>ROA</i>	Earnings before interest and taxes on previous year assets multiplied by 100.	Compustat
<i>BeatAnalyst</i>	1 if the company's actual EPS falls in a range between 0 and +0.05 from the analysts' annual	I/B/E/S

	consensus forecasts, 0 otherwise. Multiplied by 100 for rescaling.	
<i>TobinQ</i>	Tobin's Q calculated as total assets minus stockholder's equity plus market capitalization at the end of the year all divided by total assets. The variable is multiplied by 100.	Compustat
<u>Determinants and controls:</u>		
<i>Sales_var%</i>	Percentage yearly variation of revenues.	Compustat
<i>Returns Covid</i>	Buy-and-hold returns in March 2020 for the primary stock of the company.	CRSP
<i>CovidDeaths</i>	Number of casualties attributed to COVID-19 over 1000 habitants in the state where the company is incorporated at the end of 2020. For companies incorporated abroad, we use the country of incorporation. Missing are substituted with zeros.	COVID-19 casualties: John Hopkins University. State Population: United States Census Bureau. Country Population: World Bank
<i>Salary_%var</i>	Percentage yearly salary variation.	Execucomp
<i>LogAssets</i>	Lagged logarithm of total assets.	Compustat
<i>Leverage</i>	Lagged leverage ratio, computed as long-term debt on assets.	Compustat
<i>BTM</i>	Lagged book-value of assets ("Stockholders Equity – Total") on market value ("Price Close - Annual – Calendar" times "Common Shares Outstanding").	Compustat
<i>ROA</i>	Lagged earnings before interest on previous year assets.	Compustat
<i>CompComm</i>	1 If the company has a compensation committee in the previous year, 0 otherwise.	Refinitiv
<i>BdIndep</i>	Percentage of independent directors in the previous year.	Refinitiv
<i>TotComprank</i>	Percentile of the total compensation received by the CEO compared to the compensation of the other CEOs in the sample in that fiscal year (as reported to the SEC). CEOs are identified according to the CEO flag for the fiscal year. This variable is a lagged value.	Execucomp
<i>CashComp</i>	Percentage of cash compensation on total compensation (as reported to the SEC). CEOs are identified according to the CEO flag for the fiscal year. This variable is a lagged value.	Execucomp
<i>Duality</i>	1 if the CEO is the Chairman of the board 0 otherwise.	Refinitiv
<i>Tenure</i>	Number of years the executive served as CEO in the company.	Execucomp
<i>Age65</i>	Dummy variable equal to 1 if the CEO is 65 years old or above.	Execucomp

<i>ReturnsCovid</i>	Buy-and-hold returns from the 1st to the 31st of March 2020.	CRSP
<i>GoalChanger</i>	Dummy variable equal to 1 if the companies changed performance measures (e.g., revenues, EPS...) in the CEO compensation package at least twice in the previous three years.	IncentiveLab
<i>IndustryFE</i>	Dummy indicators for industries based on SIC2-digit industry codes.	Compustat

APPENDIX 3.2

Examples of Favorable IPMs

IPMs that increase compensation or decrease the risk of not reaching payout thresholds:

LAMB WESTON HOLDINGS INC:

This resulted in a weighted average payout at 0% of target award opportunity driven primarily by the decline in global restaurant traffic and foodservice demand in our fiscal fourth quarter following government-imposed stay-at-home orders and other restrictions that were implemented in response to the COVID-19 pandemic. However, as discussed above under “—Overview—Fiscal 2020 Compensation Highlights,” the Compensation Committee elected to recognize the performance achieved during the period before our business was negatively impacted by the effects of the COVID-19 pandemic, or our first three fiscal quarters, and then prorated the target award opportunity by three-fourths (i.e., 75% of target) to reflect the condensed period for which fiscal 2020 performance was evaluated.

IPMs aimed at retaining and motivating:

DENTSPLY SIRONA Inc.:

The HumanResources Committee adopted the 2020 Annual Incentive Plan in February 2020 (the “Original Bonus Plan”); however, due to the disruption the global dental services markets from the COVID-19 pandemic and the resulting severe impact on the demand for the Company’s products, the human Resources Committee determined that the performance targets in the Original Bonus Plan were no longer appropriate to properly motivate our key executives to achieve the Company’s goals and objectives.

IPMs addressing the risks and the uncertainty of the Pandemic:

SKYWORKS SOLUTIONS INC:

The performance goals established for the second half of fiscal year 2020 were based on the Company’s outlook in May 2020 for the remainder of the fiscal year and reflected the significant economic uncertainty associated with the COVID-19 outbreak, including an expectation of revenue for the second half of fiscal year 2020 lower than the Company’s original operating plan for the fiscal year.

IPMs that reflect new strategic priorities and evaluate the response to the Pandemic:

HESS CORP:

In addition, the Committee approved adjustments to weightings and targets for certain annual enterprise-level metrics to reflect the Company’s shift in priorities given the economic crisis. Revised Plan targets are intended to preserve the rigor of initial 2020 incentive targets while responding decisively to the current disruptions in the global oil markets. [...] The committee’s change in weightings and targets from early March 2020 to June 2020 were to reflect increased focus on capital discipline and controlling cash costs in response to the sharp decline in oil prices due to COVID-19, as provided in the chart below.

APPENDIX 3.3

IPMs by Industry

	IPM		
	NO	YES	Total
Amusement & Recreation Services	1	3	4
Apparel & Accessory Stores	1	3	4
Apparel & Other Textile Products	1	5	6
Automotive Dealers & Service Stations	2	1	3
Building Materials & Gardening Supplies	2	1	3
Business Services	31	15	46
Chemical & Allied Products	23	10	33
Communications	6	4	10
Depository Institutions	13	9	22
Eating & Drinking Places	3	2	5
Electric, Gas, & Sanitary Services	27	6	33
Electronic & Other Electric Equipment	14	4	18
Engineering & Management Services	1	4	5
Fabricated Metal Products	4	1	5
Food & Kindred Products	15	4	19
Food Stores	0	1	1
Furniture & Fixtures	1	0	1
Furniture & Homefurnishings Stores	0	1	1
General Building Contractors	3	1	4
General Merchandise Stores	4	3	7
Health Services	1	4	5
Holding & Other Investment Offices	15	12	27
Hotels & Other Lodging Places	0	2	2
Industrial Machinery & Equipment	16	9	25
Instruments & Related Products	17	19	36
Insurance Agents, Brokers, & Service	3	2	5
Insurance Carriers	14	8	22
Leather & Leather Products	0	1	1
Lumber & Wood Products	1	0	1
Metal, Mining	1	1	2
Miscellaneous Manufacturing Industries	1	1	2
Miscellaneous Retail	1	3	4
Motion Pictures	1	0	1
Non-Classifiable Establishments	1	2	3
Nondepository Institutions	3	2	5
Nonmetallic Minerals, Except Fuels	2	0	2
Oil & Gas Extraction	6	9	15
Paper & Allied Products	4	3	7
Personal Services	1	1	2
Petroleum & Coal Products	4	1	5
Primary Metal Industries	1	0	1
Printing & Publishing	0	1	1
Railroad Transportation	1	3	4
Real Estate	1	0	1
Rubber & Miscellaneous Plastics Products	2	1	3
Security & Commodity Brokers	13	3	16
Special Trade Contractors	1	0	1
Textile Mill Products	0	1	1
Tobacco Products	0	2	2
Transportation Equipment	6	8	14
Transportation Services	1	1	2
Transportation by Air	1	5	6
Trucking & Warehousing	3	1	4
Water Transportation	0	3	3
Wholesale Trade of Durable Goods	3	5	8
Wholesale Trade of Nondurable Goods	2	2	4
Total	279	194	473

The sample consists of US listed firms with large market capitalization in 2020. Companies with missing values for at least 1 year were discarded. In this table, we consider all types of IPMs.

APPENDIX 3.4

Determinants Per Number of Modifications

	(1) NumModaLL	(2) NumModLT	(3) NumModST
Sales_%var	-1.262*** (0.420)	-0.589*** (0.185)	-0.745** (0.301)
ReturnsCovid	-2.527*** (0.639)	-0.568** (0.284)	-1.884*** (0.459)
CovidDeaths	0.314** (0.140)	0.039 (0.064)	0.247** (0.0985)
LogAssets	-0.140** (0.0563)	0.009 (0.022)	-0.150*** (0.044)
Leverage	-0.119 (0.312)	-0.0493 (0.131)	-0.0597 (0.250)
BTM	-0.271 (0.352)	-0.223 (0.156)	-0.066 (0.267)
ROA	-0.098 (1.164)	0.010 (0.485)	-0.140 (0.810)
CompComm	-0.434 (0.301)	0.0815 (0.080)	-0.525* (0.280)
BdIndep	-0.010 (0.010)	-0.010** (0.004)	-0.001 (0.007)
Tenure	-0.007 (0.013)	-0.006 (0.006)	-0.001 (0.009)
Duality	-0.084 (0.138)	-0.034 (0.058)	-0.0263 (0.105)
TotCompRank	0.498 (0.360)	0.037 (0.146)	0.458* (0.271)
CashComp	0.116 (1.206)	-0.064 (0.515)	0.328 (0.894)
Salary_%var	-0.529 (0.407)	-0.414** (0.186)	-0.111 (0.278)
Age65	0.531** (0.269)	0.109 (0.126)	0.406** (0.193)
Constant	2.441** (1.109)	0.772 (0.475)	1.690** (0.819)
Observations	440	440	440
R-squared	0.206	0.146	0.184
Model	OLS	OLS	OLS
IPM	Favorable	Favorable	Favorable

The sample consists of US listed firms with large market capitalization in 2020. We exclude firms with unfavorable IPMs or whose CEO forfeits. The sample period includes only fiscal year 2020. Companies with missing values in the control variables were discarded. In Column 1, the dependent variable NumModaLL is the number of IPMs related to COVID-19 for the company. In Column, 2 NumModLT is the number of IPMs for the long-term incentive plan, while NumModST in Column 3 is the number of IPMs for the short-term plan. All continuous variables are winsorized at 1% level. Robust standard errors in brackets. The definitions of the other variables are in Appendix 3.1. * p<0.1, **p<0.05, ***p<0.01.

APPENDIX 3.5

Favorableness of IPMs and Investment Decisions

	(1) RnD	(2) Capex	(3) ESscore
After x Favorable	0.972*** (0.360)	-0.411** (0.201)	1.941** (0.886)
After x Unfavorable	0.0673 (0.617)	-0.385 (0.430)	1.985 (1.223)
After x Forfeit	1.261 (0.848)	0.239 (0.373)	-1.943 (2.254)
Favorable	-0.807 (0.586)	0.650** (0.283)	0.607 (1.778)
Unfavorable	1.986 (2.091)	-0.338 (0.760)	-2.348 (4.486)
Forfeit	0.0845 (1.635)	-0.0998 (0.496)	2.387 (3.801)
After	0.933 (0.586)	-0.335 (0.272)	-1.521 (1.470)
Observations	532	946	934
R-squared	0.496	0.530	0.445
Industry FE	YES	YES	YES
Controls	YES	YES	YES

The sample consists of US listed firms with large market capitalization in 2020. The sample period includes fiscal years 2019 and 2020. Companies with missing values for at least 1 year were discarded. Column 1 includes only companies reporting R&D. All continuous variables are winsorized at 1% level. Standard errors in brackets. Standard errors are clustered by firm. Industry fixed effects are computed according to SIC2 codes. The definitions of the variables are in Appendix 3.1. * p<0.1, **p<0.05, ***p<0.01.

APPENDIX 3.6

Data Before Entropy Balancing, Full Sample

Panel A: Before entropy balancing

Variables	Treated Group				Control Group			
	Obs	Mean	Variance	Skewness	Obs	Mean	Variance	Skewness
Sales_%var	184	-0.016	0.034	-0.654	318	0.049	0.021	0.294
ReturnsCovid	184	-0.206	0.022	-0.519	318	-0.119	0.017	-0.854
CovidDeaths	184	0.591	0.469	0.739	318	0.564	0.457	0.897
LogAssets	184	9.550	1.122	0.413	318	9.702	1.496	0.372
Leverage	184	0.318	0.036	0.866	318	0.289	0.030	0.557
BTM	184	0.300	0.062	0.885	318	0.261	0.059	2.301
ROA	184	0.118	0.006	0.910	318	0.135	0.006	1.011
CompComm	184	0.967	0.032	-5.263	318	0.969	0.031	-5.370
BdIndep	184	82.861	65.557	-0.689	318	85.199	60.759	-1.797
Tenure	184	7.092	49.352	1.939	318	7.110	31.631	1.837
Duality	184	0.750	0.189	-1.155	318	0.664	0.224	-0.692
TotCompRank	184	0.751	0.042	-1.379	318	0.749	0.056	-1.444
CashComp	184	0.121	0.010	4.288	318	0.118	0.011	3.369
Salary_%var	184	-0.012	0.051	1.050	318	0.025	0.027	2.226
Age65	184	0.120	0.106	2.345	318	0.075	0.070	3.214

Panel B: After entropy balancing

Variables	Treated Group				Control Group			
	Obs	Mean	Variance	Skewness	Obs	Mean	Variance	Skewness
Sales_%var	184	-0.016	0.034	-0.654	316	-0.016	0.034	-0.654
ReturnsCovid	184	-0.206	0.022	-0.519	316	-0.206	0.022	-0.520
CovidDeaths	184	0.591	0.469	0.739	316	0.591	0.469	0.740
LogAssets	184	9.550	1.122	0.413	316	9.548	1.121	0.420
Leverage	184	0.318	0.036	0.866	316	0.318	0.036	0.868
BTM	184	0.300	0.062	0.885	316	0.300	0.062	0.886
ROA	184	0.118	0.006	0.910	316	0.118	0.006	0.911
CompComm	184	0.967	0.032	-5.263	316	0.967	0.032	-5.256
BdIndep	184	82.861	65.557	-0.689	316	82.839	65.540	-0.681
Tenure	184	7.092	49.352	1.939	316	7.091	49.339	1.940
Duality	184	0.750	0.189	-1.155	316	0.749	0.188	-1.151
TotCompRank	184	0.751	0.042	-1.379	316	0.751	0.042	-1.376
CashComp	184	0.121	0.010	4.288	316	0.121	0.010	4.290
Salary_%var	184	-0.012	0.051	1.050	316	-0.012	0.051	1.050
Age65	184	-0.016	0.034	-0.654	316	0.120	0.106	2.340

Summary statistics using entropy balancing on three moments. The definitions of the variables are in Appendix 3.1.