WORKING PAPER

COMPLIANCE AND TRUTHFULNESS: LEVERAGING PEER INFORMATION WITH COMPETITIVE AUDIT MECHANISMS

Timo Goeschl
Marcel Oestreich
Alice Soldà

May 2023
2023/1069

GHENT UNIVERSITY

Department of Economics

Faculty of Economics and Business Administration
Campus Tweekerken, St.-Pietersplein 5, 9000 Ghent - BELGIUM

Compliance and Truthfulness: Leveraging Peer Information with Competitive Audit Mechanisms*

Timo Goeschl — Heidelberg University
Marcel Oestreich — Brock University
Alice Soldà — Ghent University

October 20, 2022

Accepted for pub. at Journal of the Association of Environmental and Resource Economists

Abstract

How to design audit mechanisms that harness the benefits of self-reporting for achieving compliance with regulatory targets while limiting misreporting is a pressing question in many regulatory contexts, from climate policies to public health. Contrasting random audit and competitive audit mechanisms, this paper theoretically and experimentally studies their performance in regulating socially undesirable emissions when peer information about others' emissions is present or absent. Our focus is on the compliance of emission levels with regulatory targets, going beyond existing results on truthfulness of reporting. Confirming theoretical predictions, the experiment shows that in contrast to the random audit mechanism, the competitive audit mechanism can leverage peer information for compliance: emission levels are closer to the social optimum. Yet, emission levels fall somewhat short of full compliance. The results highlight the considerable potential of competitive audit mechanisms for achieving not only more truthfulness, but also more compliance.

Keywords: Regulation; compliance; tournament theory; online experiment.

JEL classification: D62; H41; H83; L51; Q58.

*We are grateful to Tim Cason, Jay Shimshack, John Stranlund, Christian Vossler, two anonymous referees, seminar audiences at the Universities of Bayreuth, Heidelberg, Mannheim and ZEW-Mannheim, and audiences from AERE, EAERE and CEA conferences for helpful comments; to Tobias Pfrommer for excellent research assistance, and to the GATE Lab for participant pool access. This research has been funded by the Federal Ministry of Education and Science (grant no. 01LA1806A). Oestreich acknowledges financial support by the Social Science and Humanities Research Council of Canada (grant no. 435-2017-463).

1 Introduction

In many policy areas, regulated entities are obligated to self-report the level of a regulated activity to a regulator, broadly defined (Innes, 2017). Environmental policy is a prominent example. The self-reporting of emissions and release of regulated pollutants has been common practice in water, air, and toxic chemicals since the 1980s (Malik, 1993; Helland, 1998). Self-reporting acknowledges that the regulated entities tend to be better informed about their own emissions levels than the regulator, who typically has to incur considerable monitoring costs in order to observe emissions (Harford, 1987). Yet, self-reporting is obviously no perfect remedy for the information asymmetry between regulator and regulated: Entities can strategically misreport their emissions unless dissuaded to do so by the threat of a truth-revealing regulatory audit coupled with fines (De Marchi and Hamilton, 2006). How to design efficient regulatory systems that harness the benefits of self-reporting while limiting misreporting is therefore a question of interest both to the theorist and to regulators charged with enforcing policies, but constrained both by tight budgets for conducting costly audits and by limited enforcement options (Harford, 1987; Friesen and Gangadharan, 2013; Duflo et al., 2018).

Progress in designing such systems has been considerable, in particular with respect to encouraging more truthful self-reporting. Environmental regulations are a case in point (Cason et al., 2021). There, both theoretical and experimental results have demonstrated the potential of harnessing self-reporting. Gilpatric et al. (2011) were the first to show that compared to a random auditing mechanism (RAM), self-reporting is more truthful when regulators employ competitive audit mechanisms (CAMs) that condition audit probabilities on self-reported emission levels.[1] Gilpatric et al. (2015) extend this finding to a dynamic setting. Cason et al. (2016) experimentally affirm the prediction of more truthful self-reporting under CAMs.

Progress with respect to compliant emission levels has been slower. This is despite the fact that in most cases the primary regulatory concern ought to be emission levels rather than

[1]CAMs are also sometimes referred to as "relative conditional audit mechanisms" (Cason et al., 2021), a terminology that captures more directly their underlying logic.

the truthfulness of self-reporting.[2] In an authoritative review of the experimental literature, Cason et al. (2021) point out that evidence regarding emission levels is particularly scarce. A rare exception is Cason et al. (2016) who find that CAMs are conducive to emissions reduction, even though the underlying theoretical model does not predict such a reduction. A theoretical model that does comes from Oestreich (2017) who finds that in contrast to RAMs, there are CAMs that can implement the socially optimal level of emissions when firms have perfect peer information about each others' emissions. This is an extension of Oestreich (2015) who shows that CAMs can lead to higher or lower emission levels than RAMs, depending on their exact specification. There is also emerging empirical evidence that CAMs can enhance regulatory performance in the field (Earnhart and Friesen, 2021). Outside the environmental regulation context, Bayer and Cowell (2009) and Vossler and Gilpatric (2018) likewise find that CAMs can enhance tax compliance by firms and individuals. Jointly, these findings raise the possibility that CAMs help budget-restricted regulators to meet not just the secondary objective of truthful self-reporting, but also realize the primary target of socially desired emission levels as long as peers are well informed about each other.

The present paper uses a combination of theory and experiments to extend the literature in two directions. One is to address the research gap identified by Cason et al. (2021). This means focusing squarely on the impact of the mechanisms on the emission levels of regulated entities, in addition to their self-reporting. The task we set ourselves is to test experimentally the provocative theoretical prediction that a properly designed CAM is capable of inducing not just more compliance, but socially optimal emission levels Oestreich (2017). The second direction of extension is to explore explicitly the role of peer information structures in determining emission and reporting outcomes under RAMs and CAMs. Peer information structures matter not only for key results in the literature, but also differ substantially in seemingly similar settings: Self-reporting polluters may be farmers crop-spraying neighboring fields or firms that have little insight into each other's operations. A meaningful comparison

[2] See Evans et al. (2009) for a case in which information disclosure, rather than emissions reduction, is the primary source of regulatory benefits because it allows affected parties to take defensive measures.

of regulatory outcomes under different information structures requires a unifying theoretical framework and an experimental design in which information structures can be exogenously manipulated in isolation.

The paper extends the literature in these two directions to make one main contribution. This is to provide experimental evidence for the theoretical claim that a competitive audit mechanism can harness peer information to better align emission levels with the social optimum – while a random audit mechanism cannot. At the same time, it experimentally challenges the claim that the associated emission levels will be socially optimal. To make this contribution, the paper proceeds in two steps. First, it develops a parsimonious analytical framework that generates results isomorph to the main theoretical findings by Gilpatric et al. (2011), Cason et al. (2016), and Oestreich (2017). This framework is capable of reproducing the core findings of the three papers in a single model suitable for experimental implementation. These findings, derived in the context of a binary-outcome audit model, predict that in the absence of peer information, emission levels under the CAM and the RAM will be identical. In the presence of perfect information about each others' emission levels, on the other hand, the CAM will induce not only lower emission levels than the RAM, but emission levels will be socially optimal. The intuition relies on the strategic interaction between a firm's self-reporting and others' emission levels, which is observable when peer information is present. In equilibrium, the direct effect of increasing the emission level is exactly counterbalanced by the strategic effect that the observable increase has on the self-reporting of others. The first step therefore leads to experimentally testable hypotheses on specific differences in emission levels, as well as self-reporting patterns, between CAM and RAM depending on whether peer information is present or absent.

In a second step, the paper demonstrates in an interactive online experiment with 131 participants that all but one of the core hypotheses survive testing in a controlled setting. There, participants play seven rounds of a game that mimics the unifying theoretical model. In each round, participants are (re-)matched in groups of three and take two individual decisions:

emission level and self-reported level. Each self-reported unit of emissions incurs a fixed fee; each non-reported unit incurs a penalty if the participant is audited. After each round, exactly one of the three participants is audited. There are two treatment dimensions, the audit mechanism (RAM or CAM) and the participant's information about other group members' emission levels (No Information or Perfect Information). In the RAM condition, the individual audit probability is fixed and uniform. In the CAM condition, it depends on the participant's self-reported level relative to the reports of the other two participants in the group.

We show that in the treatment condition with peer information (PI), the CAM induces lower emission levels than the RAM, significantly improving compliance with regulatory targets. To our understanding this is the first experimental evidence to confirm this theoretical prediction, and the experiment generates evidence to corroborate the underlying causal mechanism. Yet, we also find that emission levels under the CAM do not align perfectly with regulatory targets and explore candidate explanations for this deviation. In the treatment condition without peer information (NI), mean emission levels in the CAM and the RAM condition are statistically indistinguishable. The prediction that the CAM does not outperform the RAM in regulating emission levels is therefore borne out, but so is the prediction that reporting under the CAM is more truthful. There, players self-report significantly higher emission than under the RAM. Finally, we replicate the results from the existing experimental literature that self-reports are more truthful under the CAM. Taken together, these findings suggest that policy-makers cannot do worse by deploying audit tournaments, should prioritize regulatory contexts with peer information for their use, and consider low-cost investments that enhance peer information.

In the following section, the paper develops the analytical framework that culminates in the derivation of three main testable hypotheses. The experimental design and its parametrization are presented in section 3 and lead to the concrete experimental predictions. Section 4 reports the results of the experiment and of our statistical tests. Section 5 discusses the experimental evidence, its policy implications, and concludes.

2 Theoretical Framework

The theoretical framework serves a dual purpose. One is integrating the existing insights from the literature in a parsimonious framework that allows emission levels to be endogenized and variations in peer information to be captured. The other is developing a framework that can bridge into the experimental laboratory by allowing testable hypotheses to be generated and by readily translating into an experimental design.

To serve this dual purpose, we present an analytical framework based on an extension of Oestreich (2017) who models CAMs through contest success functions that add noise through the probabilistic selection of the audited firms. The basic strategic interactions in this framework are similar to those described by Gilpatric et al. (2011) and Gilpatric et al. (2015) in that regulated firms self-report their emission level and the audit decision by the authorities does (CAM) or does not (RAM) depend on the relative comparison of the firms' reports. In line with Cason et al. (2016) we implicitly analyze a simultaneous move game whereby players make their choices about the emission level and self-reporting at the same time (or without observing the others' actions) in one variant of our model.[3]

While our analytical framework is rooted in the literature, we also extend it in several ways. Building on Gilpatric et al. (2011) and Gilpatric et al. (2015) we allow not only for endogenous self-reporting, but also for endogenous emission choices. Furthermore, our framework captures two variations in the peer information structure: one wherein firms have perfect peer-information about each other's emission levels as in Oestreich (2017) and another wherein firms have no peer-information as in Cason et al. (2016). To accommodate both structures, the solution concept in our paper is the equilibrium of a sequential move game whereby players observe the emission levels of their competitors first before self-reporting their own emission. This coincides with the simultaneous-move equilibrium when peer information is absent.

[3] Gilpatric et al. (2011, 2015) and Cason et al. (2016) study CAMs that involve tournaments wherein the regulator has imperfect information about firms' emissions prior to the audit.

Setting. Consider a setting in which a regulator is charged with enforcing a fee-based regulation of n regulated firms. The n firms choose a privately beneficial, but externality-generating emission whose level is denoted by e_i. The risk-neutral firms accrue benefits from the emission captured by the benefit function $g(e_i)$. All firms have the same benefit function which is motivated by the common practice of enforcement agencies to group firms according to observable characteristics such as risk, size and industry before allocating audit resources as described in Telle (2015). The benefit function is assumed to be strictly concave with a maximum at e^0. Hence, in the absence of regulation, firms choose the maximum beneficial emission level, i.e. $e_i = e^0$, for all $i \in n$ where the marginal benefits are zero, that is $g'(e^0) = 0$.

Given their emission level e_i, firms self-report emission r_i and pay a linear fee t for every self-reported unit. The fee level t is exogenously set by a higher authority. A candidate for the fee level would be one that follows Pigovian principles and equates the fee level with the social marginal cost. Accordingly, we expect the socially efficient emission level to be e^t, which is implicitly defined by $g'(e^t) = t$. Emission levels may be under-reported by the firms to reduce fee payments. If so, $e_i - r_i$ is the amount of under-reported emission level by firm i.

Regulator. The regulator is charged with enforcing the fee system. It can only observe the chosen emission level by firms after conducting a costly audit. Its operating budget is fixed, including resources allotted to conducting audits. Let K be the number of firms that the regulator can afford to audit, where $K \leq n$. Let $k \equiv K/n$ define the audit rate. If the regulator decides to increase the audit probability for one firm, it has to decrease the audit probability of at least one other firm in order to balance its budget. Specifically, the assigned audit probabilities have to add up to the number of total audits at all times: $\sum_{i=1}^{n} p_i = K$.

Let $\mathbf{r} = (r_1, ..., r_n)$ denote the vector of reported emission levels for the n firms. The regulator's problem is to decide how to spread a limited auditing budget across the n firms. Accordingly, we define an audit mechanism as a strategy for the regulator that assigns an audit probability p_i to every regulated firm i. The announced mechanism is represented by the function $p_i : (r_1, ..., r_n) \to [0, 1]^n \; \forall i \in n$ that maps the vector of firms' emission-reports into firms' audit probabilities. We assume that the regulator knows the unregulated emission level e^0, possibly from the time without the regulation, and can use it as reference value when designing the audit mechanism $p_i(\mathbf{r})$.[4] After an audit, the regulator perfectly observes the actual emission level chosen by the audited firm and levies a linear penalty θ per unit of under-reported emission level, where $\theta > t$. The results are therefore derived in the context of a binary-outcome model rather than in the context of a richer continuous-outcome model in which audits are subject to error (Gilpatric et al., 2011).

Information structure. How much firms know about each others' emission levels varies considerably in the field. Here, we consider the two limit cases of perfect peer information (Assumption PI) and of no peer information (Assumption NI) and separately analyze their effects on firms' emission levels and self-reporting.

Assumption NI Firms have *no information* about each others emission levels.

Assumption PI Firms have *perfect information* about each others emission levels.

Timing. The multistage game between regulator and firms consists of the following four stages:

- *Stage 1*: The regulator announces an audit mechanism $p_i : (r_1, ..., r_n) \to [0, 1]^n$ which maps emission-reports into audit probabilities for each firm upon receiving the reports.
- *Stage 2*: Firms choose the emission level e_i. Firms are not informed (Assumption NI) or perfectly informed (Assumption PI) about the activities of the other firms.
- *Stage 3*: Firms choose emission reports r_i that are submitted to the regulator.

[4]This assumption is without loss of generality. Specifically, the audit mechanism requires the regulator to use some arbitrary high emission level as reference value. The unregulated emission level e^0 seems to be a natural candidate for this reference value, but other values would also work.

- *Stage 4*: A subset of firms is audited according to the announced audit mechanism. A linear fine θ is levied for every unit of under-reported emission levels $[e_i - r_i]$ detected.

Firm's problem. Firm i chooses emission level e_i and emission report r_i in order to maximize expected profits:[5]

$$\max_{e_i \geq 0,\ r_i \leq e_i} \mathbb{E}\Pi_i(\mathbf{e}, \mathbf{r}(\mathbf{e})) = g(e_i) - tr_i - p_i(\mathbf{r}(\mathbf{e}))\theta[e_i - r_i] \ \forall i \in n, \tag{1}$$

where $\mathbf{e}$ denotes the vector of emission levels and $\mathbf{r}$ denotes the vector of reports chosen by all firms. Emission levels provide benefits to the firm through $g(e_i)$ and their expected cost is determined endogenously by the fee t on emission report r_i, the firm's individual audit probability p_i and penalty θ for potentially under-reported emission levels $[e_i - r_i]$.

Random audit mechanism. The RAM is the common benchmark in the literature. It allocates equal audit probabilities among all firms regardless of reports such that $p_i = k$ $\forall i \in n$. It is well known that in theory, the RAM can fully enforce fees on the regulated emission as long as the expected marginal cost of under-reporting, $k\theta$ is larger than or equal to the fee rate, t. In that case, firms can do no better than reporting truthfully. Knowing it is going to pay fees on all of its emission, a firm chooses socially efficient emission levels e^t. Thus, for sufficiently large expected fines, $k\theta \geq t$, the regulator can fully enforce truthful reporting where $r_i = e_i$ and implement the socially efficient emission level $e_i = e^t$ $\forall i \in n$

Regulatory reality is typically characterized by limited auditing budgets and capped fines. Therefore, we focus on the case in which the expected fine is below the fee, $k\theta < t$. In this case, the RAM induces neither truthful reporting nor socially efficient emission levels because it is cheaper for the regulated firms to under-report emission levels (evade fees t) and instead face the expected penalty $k\theta$.

[5] In line with regulatory practice, over-reporting is not rewarded. Therefore, rational firms never over-report, that is $r_i \leq e_i$. Hence, without loss of generality, we can set $\max\{\theta(e_i - r_i), 0\} = \theta(e_i - r_i)$, and restrict the set of reported emission levels to be $r_i \leq e_i$.

Proposition 1 (Emission and reporting under RAM) *If $k\theta < t$, the random audit mechanism induces zero emission-reporting, i.e.: $r_i = 0$ $\forall i \in n$ and per-firm emission level $e_i = e^{k\theta}$, which is implicitly defined by:*

$$g'(e^{k\theta}) = k\theta \quad \mathit{for} \quad \forall i \in n. \tag{2}$$

Emission levels and self-reports are independent of the information structure.

Proof. See Appendix A.1.

The predictions of zero emission-reporting and excessive emission levels under capped fines and insufficient auditing budgets under the RAM motivate the search for more sophisticated audit mechanisms such as CAMs that can harness strategic interactions among the regulated firms to gain auditing leverage.

Competitive audit mechanism. The literature on CAMs is steadily growing focusing on four features that CAMs tend to have in common. First, they are applied in settings in which firms self-report emission levels to a regulator. Second, CAMs *decrease* the audit probability of a firm if the firm *increases* its reported emission levels: $\partial p_i(\mathbf{r})/\partial r_i < 0$ $\forall i \in n$. Third, they *increase* the audit probability of a firm if another firm *increases* its reported emission level: $\partial p_i(\mathbf{r})/\partial r_j > 0$ $\forall j \neq i \in n$. Finally, CAMs keep the regulator budget balanced: If the regulator increases the audit probability for one firm, it has to decrease the audit probability of at least one other firm: $\sum_{i=1}^{n} \partial p_i(\mathbf{r})/\partial r_j = 0$ $\forall j \neq i \in n$.

We adopt the CAM proposed by Oestreich (2017) that shares these features and is capable of inducing the socially optimal emission level in equilibrium such that $e_i = e^t$ $\forall i \in n$. According to this CAM, the regulator allocates the audit probabilities according to the following audit mechanism:[6]

$$p_i(\mathbf{r}) = k + \lambda \ln\left(\frac{(R_i)^{n-1}}{\prod_{j\neq i}^{n}(R_j)}\right), \tag{3}$$

[6] As usual, the audit probabilities are bound by zero and one. Formally, the audit probability is

where $R_i = e^0 - r_i$ and e^0 serves as a reference value for the regulator to compare reports against.[7] Parameter λ determines the *degree of competitiveness* induced by the CAM. Specifically, the audit probability of firm i changes by λ given a one percent increase in the report of firm i. If $\lambda = 0$, random auditing results where $p_i = k \; \forall i \in n$. If $\lambda > 0$, the audit mechanism is competitive in that higher reports relative to other firms result in lower assigned audit probabilities. The CAM adopted here relies on the special case where $\lambda = (t/\theta - k)/((n-1)(2-N))$, and $N = \left(n - 2 + \sqrt{n^2 + 4n - 4}\right)/(2\,(n-1))$. This specific functional form for $p_i(\mathbf{r})$ can induce the optimal emission level e^t for all firms as shown below.

One interesting feature of this CAM is the inverse relationship between the size of relative audit budget of the regulator (measured by the difference between t/θ and k) and the degree of competitiveness induced by the CAM. This implies that for audit budgets of sufficient size ($t/\theta \leq k$), the CAM coincides with a fully enforcing RAM that implements the optimal emission level and truthful reporting.

Illustrative example. To aid intuition, figure 1 illustrates the allocation of audit probabilities under the adopted CAM for a simple example. There are three firms ($n = 3$) and a regulatory budget sufficient for a single audit ($K = 1$). In that case, the interior part of the audit function for firm 1 simplifies to:

$$p_1(r_1, r_2, r_3) = \frac{1}{3} + \lambda[\ln(\frac{R_1}{R_2}) + \ln(\frac{R_1}{R_3})],$$

with $R_1 = e^0 - r_1$, $R_2 = e^0 - r_2$ and $R_3 = e^0 - r_3$. The lines in Figure 1 trace out the audit rates p_1, p_2 and p_3 as a function of reports r_1, r_2 and r_3. The reports of firm 2 and firm 3 are fixed in this illustration at the equilibrium value $r_2 = r_2^*$ and $r_3 = r_3^*$. The report of firm 1, r_1, varies along the horizontal axis. When the reports coincide ($r_1 = r_2 = r_3$), the audit

$\min[\max[p_i, 0], 1]$.

[7]This specific contest success function adds noise through the probabilistic selection of the audited firms similar than a standard Tullock contest. Please refer to Oestreich (2015) for the formal comparison in our context of the performance for enforcement of the Tullock and the more competitive contest format, namely the all-pay

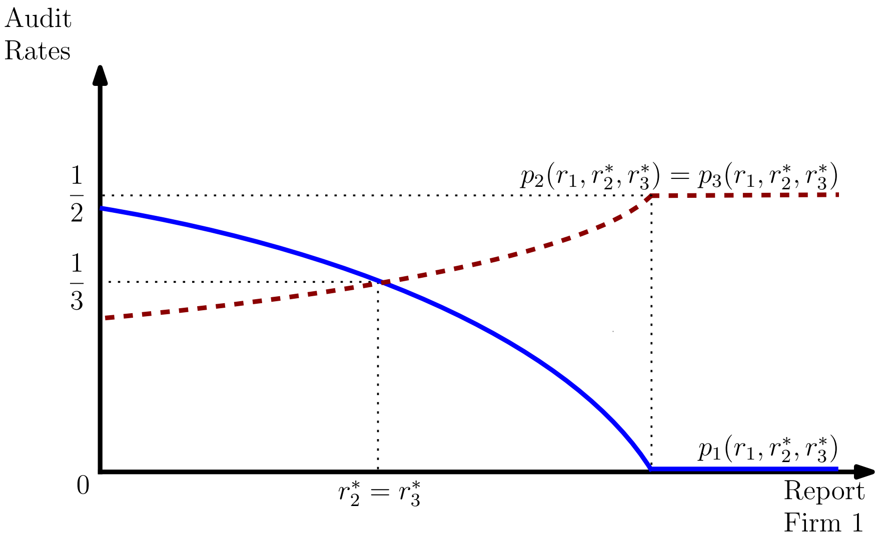

Figure 1: Optimal audit rates p_1, p_2 and p_3 for $n = 3$ and $K = 1$ as a function of report r_1 by firm 1, given equilibrium reporting $r_2 = r_2^*$ by firm 2 and $r_3 = r_3^*$ by firm 3.

probabilities also coincide ($p_1 = p_2 = p_3 = 1/3$). If r_1 is increased, p_1 decreases and at the same time p_2 and p_3 increase.

Equilibrium concept. Firms are symmetric. We therefore conjecture that there is a symmetric equilibrium in pure strategies, where $e_i = e_j$, $r_i = r_j$ and $p_i = k \; \forall j \neq i \in n$, following Oestreich (2017). Appendix B shows that the symmetric equilibrium in pure strategies exists for the case of perfect and no information between firms using the specific set of parameters of the subsequent experiment.

2.1 Stage 4: Audits and Enforcement

The implementation of the audit mechanism and the imposition of fines is automatic. There are no choices to be made by either regulator or firms.

auction.

2.2 Stage 3: Reporting Equilibrium

In this stage, firms simultaneously choose reports in order to minimize the total cost of their chosen emission level given the announced audit mechanism, their own emission level and the other firms' emission reports. Firms' reporting choices take into account both the fees on reported emission levels and the potential for audit and enforcement in stage 4.

Differentiating profit function (1) with respect to r_i yields the first-order condition (FOC) for an interior reporting solution ($\partial \mathbb{E}\Pi_i(\mathbf{e}, \mathbf{r}(\mathbf{e}))/\partial r_i = 0$) – denoted by r_1^*. This FOC can be re-written as:

$$\underbrace{p_i\theta}_{\substack{\text{direct}\\ \text{MB}}} + \underbrace{\lambda(n-1)\theta(\frac{e_i - r_i^*}{e^0 - r_i^*})}_{\text{indirect MB}} = \underbrace{t}_{\text{MC}}, \; at \; r_i = r_i^* \in [0, e_i]. \tag{4}$$

For interior solutions ($0 < r_i^* < e_i$), expression (4) has an intuitive interpretation: The marginal cost (MC) of reporting another unit of emission is the unit fee t. The marginal benefit (MB) has two components. The *direct* benefit of increasing reported emission by one unit is that it reduces under-reporting by one unit, thus decreasing the expected fine by $p_i\theta$. The *indirect* benefit is that, holding other firms' reports constant, the firm is less likely to be audited when reporting higher emission level. This lowers the expected fine for the remaining under-reported emission levels by $-(\partial p_i/\partial r_i)\theta(e_i - r_i)$, with $\partial p_i/\partial r_i = -\lambda(n-1)/(e^0 - r_i)$ under the CAM in (3). This indirect effect induces firms to report some of their emission level while they would report zero under the RAM, i.e. when $\partial p_i/\partial r_i = 0$.

Proposition 2 (Reporting under CAM) *The competitive audit mechanism (3) induces a symmetric reporting equilibrium given by:*

$$r_i^*(\mathbf{e}) = \frac{e_i^* - e^0(2-N)}{N-1} \; \forall i \in n, \tag{5}$$

where $N = \left(n - 2 + \sqrt{n^2 + 4n - 4}\right) / (2\,(n-1))$ *and* e_i^* *is the equilibrium emission level. As long as* $e_i^* - e^0(2 - N) > 0$*, the reporting equilibrium under the CAM is positive and thus larger than the reporting equilibrium under the RAM.*

Self-reports are independent of the information structure.

Proof. See Appendix A.2.

A sufficient condition for non-zero reporting in equilibrium is that the marginal benefits of emission ($g'(e)$) are high enough and/or the unit fee on emission (t) low enough such that regulation does not suppress the equilibrium emission level (e_i^*) too far below the unregulated emission level (e^0). How far regulation can suppress equilibrium emission without inducing zero reporting depends, through N, on the number of firms n: The more firms the regulator has to regulate, the higher the marginal benefit of emission (or the lower the unit fee) needs to be to ensure that firms still report any positive emission level $r_i^* > 0$. In practice, n is likely to be small, given that only firms of sufficient comparability can be subject to the CAM.

2.3 Stage 2: Emission Equilibrium

In stage 2, firms simultaneously choose emission levels and consider how their choices translate into the reporting equilibrium at stage 3 given the audit mechanism $p_i(.)$ and other firms' emission levels. It is here that assumptions about the presence or absence of information about peers' emission levels can affect firms' choices, requiring an analysis of both cases.

For ease of exposition, we start with the case of perfect peer information (Assumption PI). As in Tirole (1988), the determination of optimal emission levels relies on the total derivative of expected profits $\mathbb{E}\Pi_i(\mathbf{e}, \mathbf{r}(\mathbf{e}))$ with respect to e_i and an application of the envelope theorem. From the optimization at the reporting stage we know that $\partial \mathbb{E}\Pi_i / \partial r_i = 0$. The effect of e_i on $\mathbb{E}\Pi_i$ through the firm's own reporting choice is therefore irrelevant for the determination of optimum emission levels.

Under the conjecture of a symmetric pure-strategy equilibrium, the starting point of any deviation from equilibrium play is $e_1 = ... = e_n$. Here we consider a deviation of firm i such that $e_i \neq e_1 = .. = e_{i-1} = e_{i+1} = .. = e_n$. In this case, it must be true that the strategic effects of all other firms are identical. Thus, the total derivative is:

$$\frac{d\mathbb{E}\Pi_i}{de_i} = \underbrace{\frac{\partial E\Pi_i}{\partial e_i}}_{\substack{\text{direct} \\ \text{effect}}} + \underbrace{(n-1)(\frac{\partial E\Pi_i}{\partial r_j}\frac{\partial r_j}{\partial e_i})}_{\text{strategic effects}}, \ \forall j = k \neq i \in n. \tag{6}$$

Using (6) and the particular profit function in (1), the FOC for a profit maximum is:

$$\underbrace{g'(e_i) - p_i\theta}_{\substack{\text{direct} \\ \text{effect}}} - \underbrace{(n-1)(\frac{\partial p_i}{\partial r_j}\frac{\partial r_j}{\partial e_i}\theta(e_i - r_i^*))}_{\text{strategic effects}} = 0, \ \forall j = k \neq i \in n. \tag{7}$$

The LHS of expression (7) contains two main parts, the *direct effect* and the *strategic effect* of varying emission levels on profit. The direct effect of changing e_i consists of the marginal benefit of varying emission levels ($g'(e_i)$) net of expected fines ($p_i\theta$). The strategic effect depends on the peer information structure. Since firms observe others' emission levels (Assumption PI), a change in e_i not only changes the firm's own reporting behavior, but also the other firms' reporting behavior (via $(\partial r_j/\partial e_i)(n-1)$). The change in the other firms' reporting behavior affects the audit probability of firm i, p_i, which in turn affects firm i's expected fine of unreported emission levels (in proportion to $(\partial p_i/\partial r_j)\theta(e_i - r_i^*)$). The total effect of e_i on $E\Pi_i$ is the sum of the direct and strategic effects.

At the point of symmetry ($e_i = e_j$ and $r_i = r_j$, $p_i = k$, $\frac{\partial p_i}{\partial r_i} = \frac{\partial p_j}{\partial r_j}$ and $\frac{\partial r_j}{\partial e_i} = \frac{\partial r_k}{\partial e_i}$), we can re-write the FOC as:

$$g'(e_i) = k\theta + \frac{\partial r_j}{\partial e_i}(t - k\theta), \ \text{at} \ \forall j = k \neq i \in n. \tag{8}$$

We can learn from (8) that if $\partial r_j/\partial e_i = 1$ at $e_i = e_j$, then $g'(e_i) = t$, i.e. a necessary condition for socially efficient emission levels in equilibrium holds. In other words, if an audit mechanism induces all other firms j to increase their report by one unit when firm i increases her emission level by one unit, then this audit mechanism may implement efficient emission levels among all firms. This is precisely what the CAM in (3) achieves under Assumption PI.

Proposition 3 (Emission under CAM - perfect information) *Given that firms have perfect information about each other's emission levels (Assumption PI), the competitive audit mechanism (3) induces socially efficient emission levels for all firms, i.e.*

$$e_i = e^t \qquad \forall i \in n.$$

Proof. See Appendix A.3.

When firms have no information about each others' emission levels (Assumption NI), then the report of one firm cannot react to a change in the emission levels of other firms. Thus $\partial r_j/\partial e_i = 0$, and the strategic effect falls away. As a result, $g'(e_i) = k\theta < t$ in equilibrium, i.e. the firm equalizes marginal benefits from emission levels $g'(e_i)$ to marginal cost, $k\theta$. By concavity of the benefit function $g(e_i)$, this emission level $e^{k\theta}$ is higher than that under perfect information e^t. In fact, it coincides with the emission level under the RAM ($p_i = k$).

Proposition 4 (Emission under CAM - no information) *Given that firms have no information about each other's emission level (Assumption NI), competitive auditing leads to the same per-firm emission level as random auditing, $e^{k\theta}$, implicitly defined by:*

$$g'(e^{k\theta}) = k\theta < t.$$

emission levels are higher than the socially optimal emission level.

Propositions 2, 3, and 4 constitute the core deliverables of an analytical framework that aims to integrate previous theoretical contributions into a setting that can bridge into the laboratory. Propositions 2 and 4, in particular, are isomorph to the main theoretical findings by Gilpatric et al. (2011) and Cason et al. (2016), respectively. Proposition 3 is isomorph to the main finding by Oestreich (2017). Jointly, they provide a body of results that can be subjected to experimental testing.

Collusion. One remaining concern in the CAM centers on the possibility of collusion among regulated firms. We discuss its implication in 2.3.

3 Hypotheses, Experimental Design, and Procedures

3.1 Testable Hypotheses

A natural structure for a test of the relative performance of CAM and RAM in the presence (Assumption PI) or absence (Assumption NI) of peer information is a 2x2 design. In this design, the experimenter manipulates, along one dimension, the audit mechanism (RAM vs. CAM) and along the other, the information firms have about each others' level of emission (NI vs. PI). Reorganizing propositions 1 to 4 to fit a 2x2 design leads us to three main hypotheses derived from the theoretical framework and testable in an experimental implementation that induces a symmetric equilibrium in pure strategies.

Hypothesis 1 considers the relative performance of CAM and RAM by combining insights from propositions 1, 2, and 4 about emission and reporting in the absence of peer information.

Hypothesis 1. In a setting in which participants have *no information* about each other's emission level (NI), participants making choices under the competitive audit mechanism (CAM) will exhibit, on average, a) the same level of emission and b) a higher level of reported emission than participants making choices under the random audit mechanism (RAM).

The second hypothesis considers the relative performance of CAM and RAM by combining the insights of propositions 1, 2, and 3 derived from our theoretical framework about emission and reporting in the presence of peer information.

Hypothesis 2. In a setting in which participants have *perfect information* about each others' emission level (PI), participants making choices under the competitive audit mechanism (CAM) will exhibit, on average, a) a lower level of emission and b) a higher level of reported emission than participants under the random audit mechanism (RAM).

On emission levels, Hypothesis 2 relies on the fact that the strategic interaction among firms under the CAM creates an additional marginal cost from producing an extra unit of emission that is only present if (i) firms can observe each others' emission levels and (ii) if the audit mechanism creates a reporting competition among the firms. The experimental CAM therefore needs to be "calibrated" to induce the optimal emission level this way (see section 3.2). On reporting, both Hypothesis 1 and Hypothesis 2 rely on the direct and indirect components of reporting one more unit of emission on the expected fine. The indirect effect, in particular, is present whether or not the other firms can observe each others' emission levels.[8]

The properly calibrated CAM should, in theory, not just support better regulatory performance of the CAM compared to the RAM (Hypotheses 1 and 2). It should also implement social efficient emission levels. This leads to the final hypothesis about experimental outcomes.

Hypothesis 3. In a setting in which participants have *perfect information* about each others' emission level (PI), average emission levels under the competitive audit mechanism (CAM) do not differ from the socially efficient level of emission.

3.2 Experimental Design

In the experimental implementation, participants play seven rounds of a game that mimics the theoretical framework of section 2 under a unique audit mechanism (CAM or RAM) and

[8]Hypothesis 2 holds for functions of $g(e_i)$ that guarantee the existence of a symmetric equilibrium in pure strategies and for a binary-outcome audit model. If audits are subject to error (Gilpatric et al., 2011), interior

a unique information structure (PI or NI). In each round, participants are matched in groups of three and make two decisions individually: first, their emission level and second, their self-reported emission level.[9] Each unit of emission reported incurs a fee t. Each unit of emission not reported costs θ if an audit reveals that the participant under-reported. Audits take place at the end of each round. Per round, exactly one participant in each group is audited according to its treatment condition (RAM or CAM). To ensure independence between each round, participants are re-matched every round. Due to a perfect stranger matching procedure, a participant never encounters the same group member more than once.

The unfolding of a particular round is displayed in Figure 2: Each round is composed of 4 stages: a production stage, an information stage, a report stage and an audit stage.

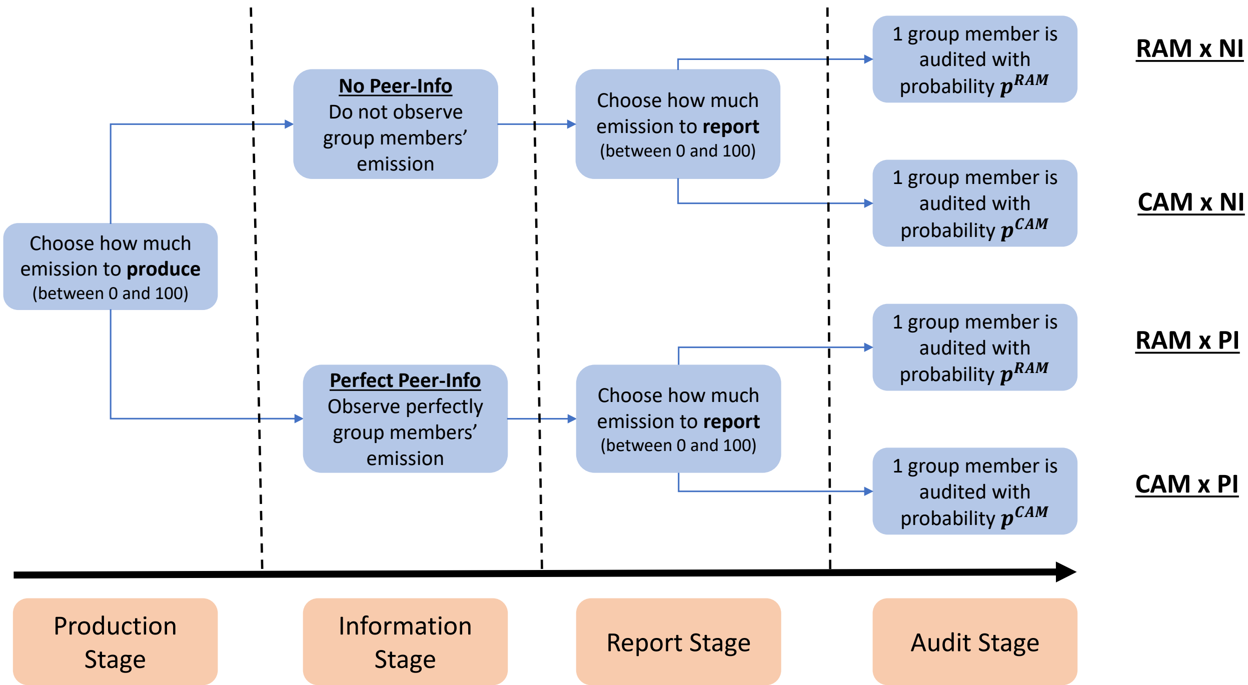

Figure 2: The four stages of an experimental round.

reporting solutions, rather than either truthful or zero reporting, may arise even in the RAM. In principle, it could then be possible, depending on $g(e_i)$, that the CAM induces such a large reduction in actual emissions that reported emissions also decrease, while reported emissions are a larger share of actual emissions than under RAM. We thank an anonymous reviewer for pointing this out.

[9] Note that in the experiment, we use the term 'output' instead of 'emission' to keep the framework neutral. Screenshots of the experimental program and an English translation of the instructions are provided in Appendix F

In the production stage, participants choose how much emission to produce on a slider. In the information stage, participants in the PI treatment are reminded of their chosen emission level and are informed about the chosen emission level of their fellow group members. In the NI treatment, participants do not receive any information about their fellow group members and are only reminded of their own emission level. In the report stage, participants choose how much emission to report on a slider. In the audit stage, at the end of each round, exactly one participant per group is selected for an audit according to the assigned audit probabilities by the audit mechanism. In the RAM treatment, each participant has a fixed probability of 1/3 of being the audited member regardless of their decisions or the decisions of their fellow group members. In the CAM treatment, the probability of being the audited member depends – for each participant – on her report relative to the reports of her fellow group members. The *higher* a participant's report relative to the reports of her fellow group members, the *lower* her probability of being audited. Conversely, the *lower* a participant's report relative to the reports of her fellow group members, the *higher* her probability of being audited. The exact audit probabilities are calculated according to the CAM algorithm presented in equation (3).[10] Participants can see their final audit probability and conditional payoffs. In addition, information about the actual and reported emission level of every group member are displayed on the screen. This information is the same in all treatments.[11] By pressing a button, participants see whether they have been audited and their earnings for this round. Before moving to the next round, participants are asked to record the information provided on the screen on their personal record sheet.

Post-experimental Questionnaire. At the end of the session, participants are asked to report their age, gender and whether French is their native language. In addition, we elicit risk attitudes following Dohmen et al. (2011) by asking participants to indicate how willing

[10] To ensure that only one participant is audited in each round in the CAM, we attribute weights to each group member on the basis of their reported emissions. In practice, this is equivalent to attributing a virtual urn with 100 balls in three colors (one per group member) to each group, with the number of balls of each color corresponding to the relative weight of each group member. In each round, one ball is drawn.

[11] This feature of the design ensures that treatment differences are not driven by differences in learning.

they are to take risks in general on a scale from 0 (not willing at all) to 10 (extremely willing).

Profit calculator. In order to help participants to make informed decisions, we provided them with a profit calculator (Healy, 2006; Requate and Waichman, 2011; Cason and Gangadharan, 2013) that allowed them to simulate how their choices would affect their expected payoffs, both in the production and the report stage. In each stage, participants were required to input at least one "guess" regarding the decisions made by other participants. In the production stage, participants were required to input their guess regarding the output produced by other participants in the *Perfect Information* (PI) treatment and the reported output by other participants in the *No Information* (NI) treatment. The profit calculator would then inform participants about their own cost-minimizing report, their probability of being audited, their probability of not being audited, their payoff in each case, as well as their expected payoffs for the current round. In the PI treatment, participants were additionally informed about the cost-minimizing reports of their fellow group members.

In the report stage, participants participants were required to input their guess regarding the output reported by other participants both in the PI treatment and the NI treatment. In the former, the slider default was set to the cost-minimizing reports of each of the two remaining group members.[12] The profit calculator would then inform participants about their probability of being audited, their probability of not being audited, their payoff in each case, as well as their expected payoffs for the current round. Participants could submit multiple scenarios, allowing them to assess how incremental changes in output produced and reported affect the aforementioned variables before logging their decision.

Parametrization. Table 1 presents the functional forms and parameters chosen for each

[12]Because the audit mechanisms differ in their complexity, we purposely provided participants in the CAM with information about the cost-minimizing reports. Since the number of interactions in the experiment is limited for procedural reasons, providing the information about the cost-minimizing reports simply serves to accelerate a learning process that would occur in any case in repeat play. The reader might be concerned that these slider presets may generate an anchor effect. We discuss this possibility in section 4.2.

variable in the experiment. With these parameter values, a symmetric equilibrium in pure strategies exists under the CAM.[13]

Table 1: Parameters of the experiment

Notation/ Functional form	**Definition**	**Parameters**
N	Number of participants per group	3
K	Number of audits per round	1
p^{RAM}	Random audit probability	0.33
e	emission level	[0, 100]
r	Reported emission level	[0, 100]
$g'(e) = 10 - 0.1e$	Marginal benefit from emission	
t	Fee on reported emission level	2.5
θ	Penalty on under-reported emission	3
e^0	Unregulated emission level	100
e^t	Optimal emission level	75

3.3 Implementation

The experimental design, hypotheses and procedure were pre-registered on the AEA RCT Registry.[14]

Participants. A total of 131 participants completed the experiment. Participants were recruited via Hroot (Bock et al., 2014) from a large pool of students, mainly from local engineering, business, and medical schools, who had previously registered to be potential participants

[13] We show in Appendix B that there is no profitable deviation for either firm from the symmetric equilibrium and that the reporting equilibrium is positive. Specifically, we show the evolution of profits if one of the players deviates from the socially optimal emission level e^t and the resulting changes to the reporting equilibrium. We show that there is a global profit maximum at $e_i = e^t$ for $i = 1, 2, 3$.

[14] RCT ID: AEARCTR-0004996

in economics experiments at GATE-lab (Ecully, France). Overall, 57% of the participants were female and the average age was 23 years (SD = 3.99).

Procedure. The experiment was programmed using oTree (Chen et al., 2016) and conducted online in a highly controlled environment that mimics the conditions of the laboratory ("interactive online experiment").[15] The experiment was carried out over a series of eight sessions varying between 15 and 21 participants during fall 2020. Digital copies of the instructions were provided to the participants, which were read aloud by the experimenter. To facilitate learning, participants were asked to answer questions about two hypothetical scenarios related to the experiment.[16] The experiment took an average of 105 minutes, including the assembly of participants in the online room and explanations on how to behave in the online environment (20 min), the reading out of the instructions (20 min), the implementation of the experiment (50 min) and the payment procedure (15 minutes).

The procedural parallels between our approach and a laboratory experiment are deliberate: The interactive online approach allowed us to emulate closely the environment and the subject pool used in earlier experiments on CAM and RAM during a period when experimental laboratories were closed due to the COVID-19 pandemic. This maximized comparability of results. One area in which the emulation could have broken down is the possibility of attrition among online participants. In anticipation of this potential threat, we pre-registered an attrition management procedure (see Appendix C) that was only triggered in four cases and whose presence demonstrably did not affect results (see Appendix D.4).

We pre-registered a sample size of 30 independent observations per treatment. With this sample size, the minimum detectable mean difference between treatments at the recommended .80 level (Cohen, 2013) is 1.47 unit of emissions, which corresponds to a tenth of the treatment difference predicted by the theory.[17]

[15] See Appendix C for more details about our online setting.

[16] The comprehension questionnaire is available in Appendix F.2.

[17] This effect size was computed based on an expected 2-unit standard deviation, which is in the range of the standard deviations reported by Cason et al. (2016) (1.7 and 3.8 units using an action space of 0 to 125

Payment. Participants were paid the sum of their earnings for four randomly selected rounds in addition to a €2 show-up fee and an additional €3 for completing the experiment.[18] The average payoff was €24.71 (SD = 7.02). Participant earnings were denominated in ECU (experimental currency), which was exchanged for euros at the end of the session.[19] At the end of the session, participants were sent a link to retrieve their payment electronically via a third-party platform.

4 Results

We first provide descriptive statistics about our participant pool and give an overview of the raw data. This is followed by the core part with tests of the main hypotheses about treatment differences in both actual and reported emission levels and a comparison between observed emission levels and the Nash predictions. We conclude with exploratory analyses that exploit some of the experimental evidence to inform future research.

4.1 Descriptive Statistics

Data Overview. Figure 3 provides an overview of our experimental data. Starting with participants' emission levels, the upper panel shows their distribution aggregated across rounds at the participants level for both RAM (in red) and CAM (in blue), with the NI treatment at the top and the PI treatment below. For each treatment, the figure provides cloud plots of the kernel density at each emissions level, box plots of the emissions levels, and scatter plots of the emission levels aggregated across all rounds for each participant. The lower panel displays, on a different scale, the same information for participants' self-reported emission levels.

compared our action space of 0 to 100).

[18]In order to reduce the variability in payoffs and avoid negative payoffs, we decided to randomly pay out 4 out of 7 rounds rather than 1 out of seven rounds. If the sum of the participant earnings from the 4 rounds turned out to be negative regardless, the participants received a minimum payoff of €5.

[19]In order to avoid large variations in payoffs between treatments, we use an exchange rate of 20 ECU equals €1 for the CAM treatments and 30 ECU equals €1 in the RAM treatments.

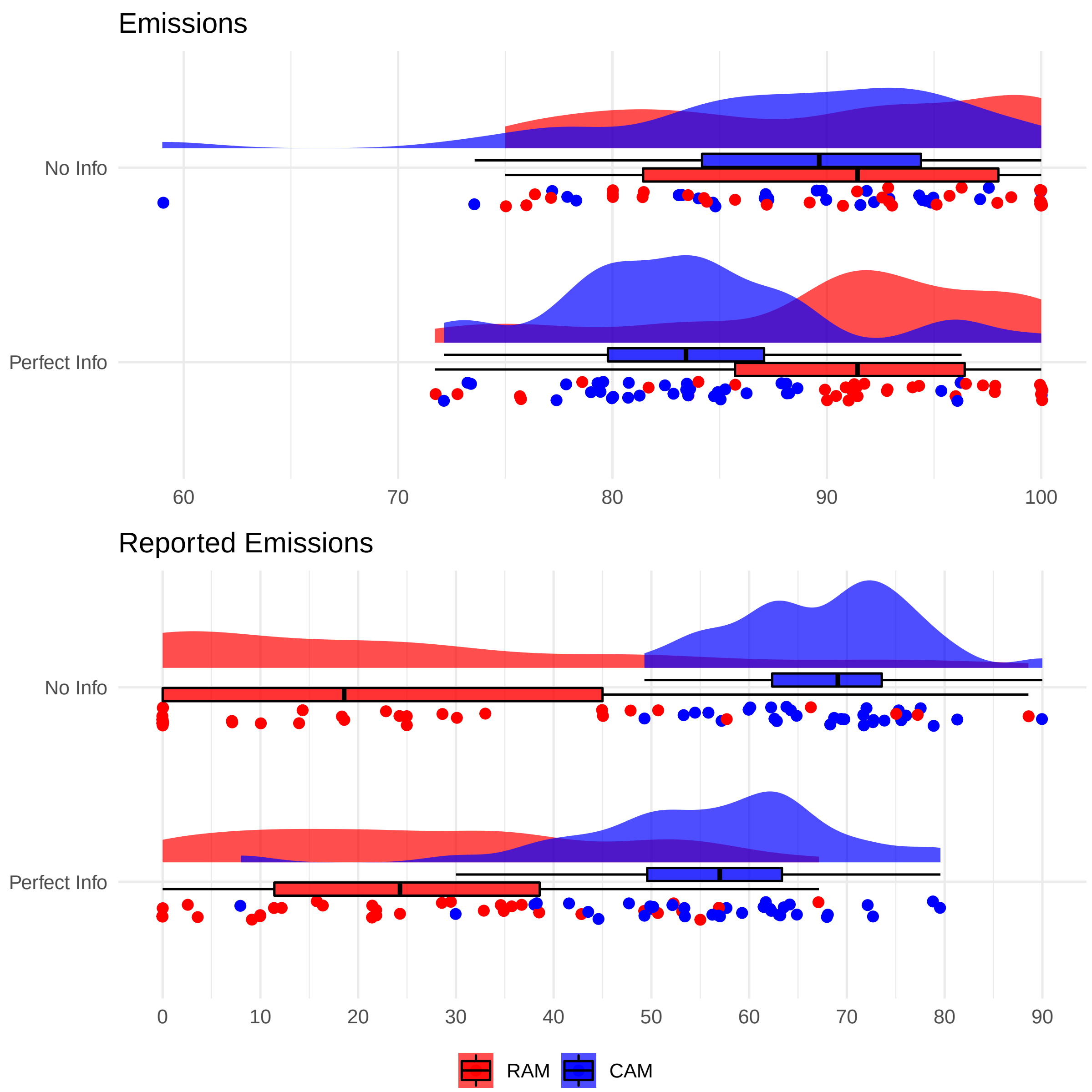

Figure 3: Distribution of emission levels (upper panel) and self-reported emission levels (lower panel) in the NI (top) and PI (bottom) condition under the RAM (red) and the CAM (blue) aggregated across rounds. Cloud plots of the kernel density on top, box plots in the middle, and scatter plots of participant-level averages across all rounds.

Inspecting the upper panel ('emission'), most observed emission levels lie in the interval of 70 to 100, with means between 83 and 91 and therefore above the social optimum at 75. In the absence of peer information (NI), the CAM (blue) and the RAM (red) produce visually similar distributions of emission levels (top). In the presence of peer information (PI), by contrast, the distributions visually diverge (bottom): emission levels under the CAM shift perceptibly to the left of those under the RAM. This is a first indication that emission levels are lower under CAM with PI. However, the mean of emission levels under CAM–PI is clearly situated to the right of the optimum at 75.

In the lower panel ('Reported emission'), reports cover a wide range between 0 and 90, but show clear patterns across treatments. Starting with the RAM treatment in both information conditions PI and NI, first recall the Nash equilibrium reporting prediction of zero (Proposition 1). While some participants indeed choose zero or near zero reporting levels, most of the observations are well away from zero and therefore more truthful than predicted. A glance at the distributions also shows that the reporting levels under the CAM treatments sit discernably to the right of those under the RAM treatments, irrespective of information structure (PI and NI). This divergence of distributions is suggestive of higher reporting levels under the CAM regardless of the information structure. Taken together, the experimental data exhibit considerable heterogeneity in participants' emission and reporting choices, but also reveal patterns that are in line with the theoretical results and invite formal testing.

4.2 Tests of Main Hypotheses

Treatment differences. The four panels of Figure 4 display the main results. They show the evolution of the average levels of actual (top panels) and reported (bottom) emission levels in each round, under both the RAM (in red) and the CAM (in blue) in the NI (left) and the PI condition (right). The theoretical predictions are visible as dotted lines in matching colors.

Consistent with Hypotheses 1 and 2, the top-right panel of Figure 4 shows that the average emission level under the CAM in each round is below the average emission level under the

RAM when peer information is present (PI). At the same time, emission levels under the CAM are above those predicted for the treatment. The top-left panel, on the other hand, shows that in the absence of peer-information, emission levels under the CAM and RAM accord with the theoretical predictions and are statistically indistinguishable.

Results on self-reporting are equally consistent with Hypothesese 1 and 2: The bottom panels of Figure 4 show that in every single round, the average reporting levels under the CAM are higher than those under the RAM. Reporting is above theoretical predictions in both information conditions.

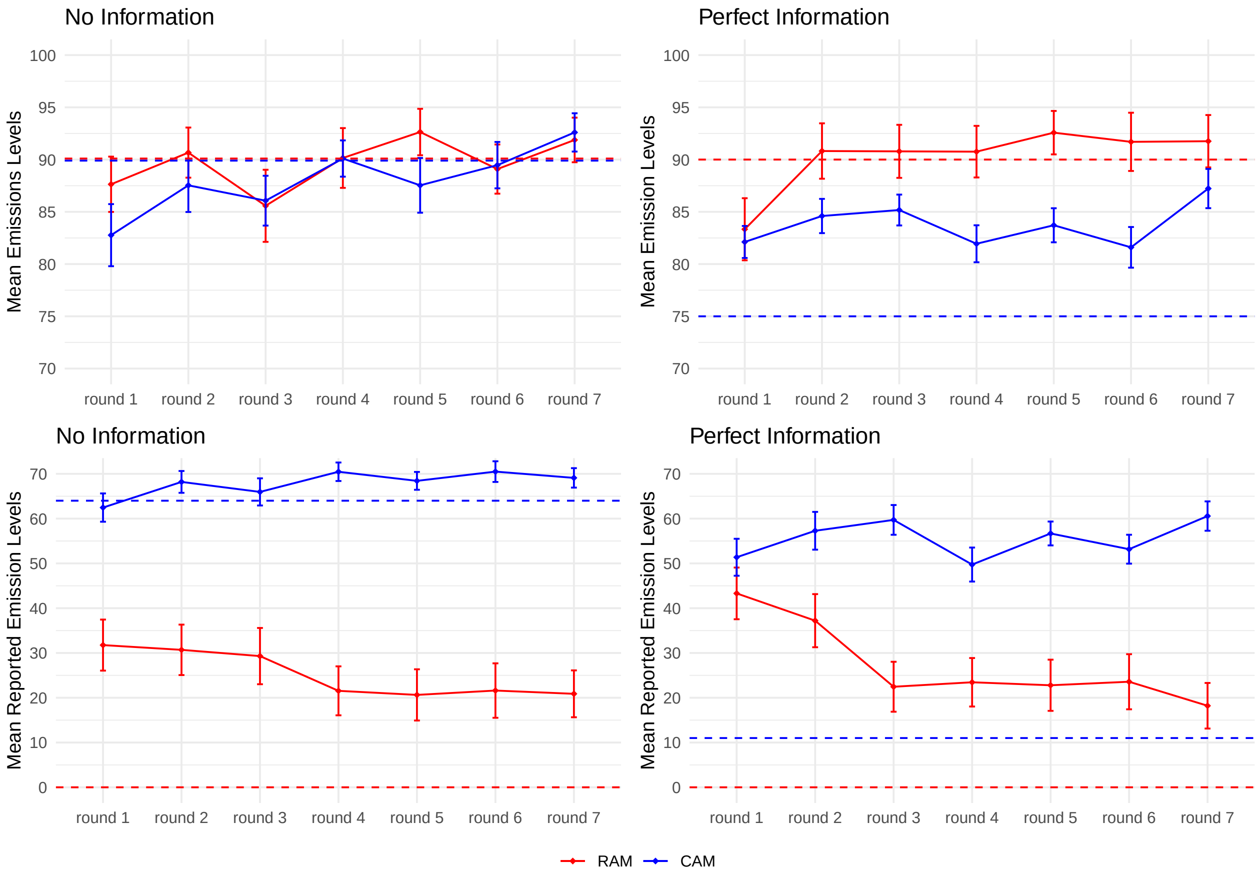

Figure 4: Evolution of the mean emission levels and mean reported emission levels under both the RAM (in red) and the CAM (in blue) across rounds, by information structures. Vertical bars indicate standard errors. Red dotted lines indicate the Nash Equilibrium for the RAM. Blue dotted lines indicate the Nash Equilibrium for the CAM.

To test our main hypotheses, we compare means across the RAM and the CAM treatments under each information condition, PI and NI. The empirical means (EM) of emission and report levels are given, by treatment, in Table 2 together with the Nash predictions derived from the model (NE). We use a participant's emission level (actual or reported) averaged across all rounds as our unit of independent observations.

In the absence of peer-information (NI), we find no significant differences in emission levels between the CAM and the RAM (two-sided Mann-Whitney tests:[20] p=0.495). However, we do find significantly higher reported emission levels under the CAM than under the RAM (MW test: p<0.001). This supports Hypothesis 1.

Table 2: Empirical means (EM) and Nash equilibrium (NE) of actual and reported levels of emission, by treatments.

Emission	RAM-NI		RAM-PI		CAM-NI		CAM-PI	
level	EM	NE	EM	NE	EM	NE	EM	NE
Actual	89.66	90	90.25	90	88.01	90	83.77***	75
	(1.481)		(1.427)		(1.594)		(1.077)	
Reported	25.20***	0	27.29***	0	67.88*	64	55.51***	11
	(4.520)		(3.291)		(1.676)		(2.376)	
Obs.	33		33		30		35	

Note: Table 2 displays the empirical means (EM) and Nash equilibrium (NE) of both actual and reported emission levels, by treatment. Standard errors in parentheses. Our unit of observation is the average across all rounds of a participant's level of emission. Stars indicates differences between EM and NE using one-sample Wilcoxon sign-ranks test. *p<0.05; **p<0.01; ***p<0.001.

Result 1 *In a setting in which participants have no information about each other's emission level (NI), the CAM leads to a) the same level of emission and b) a higher level of reported emission than the RAM (supports H1).*

In the presence of peer information (PI), we find that the CAM leads to significantly lower

[20] MW test, hereafter.

levels of emission (MW test: $p<0.001$) and higher levels of reported emission than the RAM (MW test: $p<0.001$). This supports Hypothesis 2.

Result 2 *In a setting in which participants have perfect information about each other's emission level, the CAM lead to a) a lower level of emission and b) a higher level of reported emission than the RAM (supports H2).*

To probe the above results further, we perform OLS regressions clustered at the participant level. The independent variables include a dummy variable equal to 1 if the participant was allocated to the CAM treatment and 0 otherwise plus rounds fixed effects. The OLS coefficients are displayed in columns (1) to (8) of Table 3. We use emission levels as the dependent variable in columns (1) to (4) and self-report levels in columns (5) to (8). Columns (1), (2), (5) and (6) show the results in the NI condition, columns (3), (4), (7) and (8) those for the PI condition. In columns (2), (4), (6) and (8) we additionally control for participants' demographics (gender, age, French as native language) as well as risk attitudes.[21]

Consistent with Results 1 and 2, columns (1) and (2) of Table 3 show no significant differences between the RAM and the CAM in emission levels in the NI condition ($p=0.450$ and $p=0.829$, respectively). In contrast, columns (3) and (4) show that the CAM induces lower emission levels than the RAM in the PI condition and the results are significant at the 0.1% level ($p\leq0.001$ in both specifications).[22] For self-reporting, columns (5) to (8) show that the CAM induces significantly higher levels of reported emission than the RAM, both in the NI ($p<0.001$ in both models) and the PI condition ($p<0.001$ in both specifications).[23]

[21]The missing cluster in columns (2) and (6) is due to one participant leaving the experiment without completing the post-experimental questionnaire.

[22]Note that while we find a significant treatment effect on average, the effect is not present in the final round. This could result from an attenuation of the treatment effect across rounds, erasing differences between the CAM-PI and the RAM-PI with further repetition; or it could be the result of an end-round effect. To test this, we examine the treatment differences from one round to the next in order to detect possible trending behavior. We do not find evidence for a pattern of a diminishing treatment effect across rounds. On balance, the shift observed in the final round of the CAM-PI is therefore probably attributable to an end-round effect rather than to a convergence of the emissions under the CAM-PI to the RAM-PI over time. Future research should address the question of the long-run dynamic properties of the mechanisms.

[23]Appendices D.4 and D.5 show that our main results are robust to various sample restrictions.

In summary, the experimental evidence confirms two key theoretical predictions: When participants have perfect information about each other's emission level, the CAM outperforms the RAM both with respect to the primary objective of aligning emission levels with the regulatory target and with respect to the secondary objective of encouraging more truthful self-reporting. When participants have no information about each others' emission, the CAM performs as well as the RAM on emission levels, but induces more truthful self-reporting.

Table 3: Effect of the audit mechanisms on actual and reported levels of emission, by information structure.

Dep. var:	Actual emission level				Reported emission level			
	NI		PI		NI		PI	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
CAM	-1.653	-0.442	-6.479***	-7.864***	42.67***	41.92***	28.22***	25.57***
	(2.175)	(2.037)	(1.787)	(1.676)	(4.823)	(4.932)	(4.058)	(4.212)
Round FE	X	X	X	X	X	X	X	X
Ind. controls		X		X		X		X
Const.	86.10	106.22	86.04	103.39	26.06	23.66	32.93	49.76
	(2.129)	(11.85)	(2.031)	(6.151)	(4.930)	(23.54)	(4.650)	(21.97)
Obs.	441	434	476	476	441	434	476	476
Clusters	63	62	68	68	63	62	68	68

Note: Table 3 displays the OLS coefficients clustered at the participant level. Standard errors in parentheses. Stars indicates significant differences from the RAM. $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

Point predictions. Point predictions constitute the most challenging test for the experiment given the many cognitive and affective factors that could induce a deviation in participants' behavior from the assumptions of the model. For tests of point predictions, we refer back to Table 2 to compare the empirical data against our theoretical predictions.

Comparing the predicted (NE) and observed (EM) reporting behavior in all treatments, we replicate the well-established empirical finding (Mazar et al., 2008; Goldstone and Chin, 1993) that individuals' self-reporting behavior typically deviates from pure payoff maximization (MW tests: RAM-NI: $p<0.001$; RAM-PI: $p<0.001$; CAM-NI: $p=0.035$; CAM-PI: $p<0.001$). The literature offers both moral and social preferences, such as lying-aversion (Gneezy, 2005), as candidate explanations. This result also suggest that providing information about the cost-minimizing report is unlikely to have anchored choices.[24]

Comparing predicted (NE) and observed (EM) emission levels, we find no difference for the RAM under either of the information structures (MW test: RAM-NI: $p=0.851$; RAM-PI: $p=0.376$) and for the CAM under NI (MW test: $p=0.367$). In contrast, the CAM induces higher levels of emission than predicted by theory under perfect information (MW test: $p<0.001$). As in the case of reporting behavior, this points to behavioral mechanisms that the theory does not account for. We summarize as follows.

Result 3 *In a setting in which participants have perfect information about each others' emission level (PI), average emission levels under the CAM are significantly above the socially optimal level of emission (does not support H3).*

The deviation of the experimental data from the point predictions highlights the potential value added of refining the theoretical model. The parsimonious version developed in section 2 performs more than adequately in terms of predicting the direction of the treatment effects. In order to succeed with point predictions, the theory requires a considerably richer model of firm behavior in a complex decision situation. We offer some pointers towards the underlying mechanisms that such a refinement may want to consider in section 4.3.2.

4.3 Exploratory Analyses

Results 1 through 3 technically exhaust the design of the experiment. However, the experiment has produced additional evidence that can help inform future research. We focus on three aspects: Differences in heterogeneity across mechanisms (Gilpatric et al., 2011), underlying

[24]We found that 65% of subjects in the RAM and 80% of subjects in the CAM deviate from the cost-minimizing

mechanisms, and the evidence on collusive behavior (Appendix E).

4.3.1 Heterogeneity

Compared to the RAM, the contest-like structure CAM is more complex. Higher variance in experiments involving contests and tournaments is a common finding in the literature, adding possible subtlety to comparisons based on empirical means only. One exception are Gilpatric et al. (2011) who find that the CAM leads to less variance in reporting than the RAM.

Visually, the experimental data suggests patterns in line with the findings of Gilpatric et al. (2011): In Figure 3, the box plots of emission and reporting levels exhibit less variance among observations under the CAM compared to the RAM. To explore this issue further, we first investigate treatment differences in the variance of emission choices. To do so, we follow Gilpatric et al. (2011) by estimating OLS regressions of the squared deviation from the mean emission level, that is $(e_{ij} - \bar{e_j})^2$, where $\bar{e_j}$ is the treatment-specific mean level of emission in round j, on the treatment dummies. We control for rounds fixed effects and standard errors are clustered at the individual level. The OLS coefficients are reported in column (1) of Table 4. In addition, we also investigate treatment differences in the variance of reported emission. To do so, we estimate the same OLS regression as before, using the squared deviation from the mean reported emission level as the dependent variable, that is $(r_{ij} - \bar{r_j})^2$, where $\bar{r_j}$ is the treatment-specific mean reported level of emission in round j. The OLS coefficients of this estimation are reported in column (2) of Table 4.

Column (1) in Table 4 shows no significant differences in variance between the CAM and the RAM in terms of emission levels in the absence of peer-information (p=0.216). In contrast, column (2) shows that the CAM leads to significantly less heterogeneity in emission levels under perfect peer information (p=0.017). Second, we replicate Gilpatric et al.'s (2011) finding that the CAM leads to significantly less heterogeneity in reported emission levels under both information structures (p <0.001 in both cases), as shown in columns (3) and

report in the first round.

(4). These results suggest that, in contrast to experimental findings from previous studies investigating competitive incentives, our competitive audit mechanism actually leads to not more heterogeneity in individual behavior.

Result 4 *The CAM does not induce more heterogeneity in actual or reported emission choices.*

Table 4: Effect of information structure and audit mechanism on actual and reported levels of emission.

Dep. var:	$(e_{ij} - \bar{e_j})^2$		$(r_{ij} - \bar{r_j})^2$	
	NI	PI	NI	PI
	(1)	(2)	(3)	(4)
CAM	-56.54	-115.08*	-874.12***	-600.29***
	(45.25)	(46.95)	(156.27)	(142.20)
round FE	X	X	X	X
Const.	266.15	237.36	1098.40	1124.11
	(53.85)	(51.07)	(128.68)	(117.36)
Obs.	441	476	441	476
Clusters	63	68	63	68

Note: Table 4 displays the OLS coefficients clustered at the participant level. Standard errors in parentheses. Stars indicates significant differences from the baseline (RAM-NI). *p <0.05, **p <0.01, ***p <0.001.

4.3.2 Underlying mechanisms

While the tests of Hypotheses 1 and 2 are successful, they do not interrogate the experimental data on the question whether the mechanisms underlying these results align with the theory used to derive the hypotheses. The theory makes rather specific predictions about patterns in individual-level choice data that a closer look at the evidence can uncover. This closer look comes with a disclaimer: Our experiment was not designed to test competing hypotheses about the mechanisms underlying. The exploratory analysis on the mechanism is therefore purely in the spirit of enhancing the evidence base for future experimentation in this area.

We focus on three predictions of the theory concerning reporting behavior: One is that under the RAM, own emission level as well as those of competitors do not have a significant influence on reports. The reason is that neither affect the likelihood of being audited, regardless of the information structure (see Proposition 1). The second is that under the CAM with no information, participants change their reports solely according to their own emission level and not according to their peers' emission levels. The reason is trivial: The participant cannot observe others' emission levels. The third prediction is that under the CAM with perfect information, both one's own emission level and the competitors' emission level influence self-reporting. The reason are the direct and indirect components of the marginal benefit of reporting (equation 4) that link reports in a group. If the predictions are valid, their mechanisms should give rise to recognizable patterns in the individual-level choice data.

To detect the predicted patterns, we regress participant i's reported emission level in round t on i's actual emission level in round t and the average emission level of i's fellow group members, including rounds fixed effects. The OLS coefficients are displayed in Table 5 for RAM-NI (columns (1) and (2)), RAM-PI (columns (3) and (4)), CAM-NI (columns (5) and (6) and CAM-PI (columns (7) and (8)). In models (2), (4), (6) and (8) we control for participant i's gender, age, risk attitudes and whether French is i's native language.

Table 5 shows that under the RAM, reports are independent from emission levels in the NI and PI condition ($p \geq 0.137$ in columns (1) to (4)). Under the CAM, participants condition their reports on their own emission levels in both conditions ($p \leq 0.016$ in columns (5) to (8)). In the PI condition, there is also evidence that participants condition their reports on the average emission level observed in their group ($p = 0.029$ and $p = 0.001$ in column (7) and (8), respectively). In terms of effect directions, these findings constitute patterns that would be consistent with the theoretical mechanisms that underpin Hypotheses 1 and 2.

Table 5: Effect of emission levels on reported emission levels.

Dep. var: emissions reported	RAM-NI (1)	RAM-NI (2)	RAM-PI (3)	RAM-PI (4)	CAM-NI (5)	CAM-NI (6)	CAM-PI (7)	CAM-PI (8)
Emissions	-0.294	-0.156	-0.069	-0.053	0.560***	0.529***	0.705*	0.710***
	(0.193)	(0.191)	(0.151)	(0.184)	(0.065)	(0.090)	(0.278)	(0.197)
GA emissions	-0.121	-0.048	-0.281	-0.296	0.032	0.031	0.474*	0.517*
	(0.161)	(0.138)	(0.218)	(0.207)	(0.057)	(0.069)	(0.208)	(0.207)
Round FE	X	X	X	X	X	X	X	X
Ind. controls		X		X		X		X
Const.	68.12***	22.21	72.49***	87.77	13.43	27.41	-45.48	2.830
	(18.72)	(34.52)	(23.36)	(44.21)	(6.628)	(23.78)	(30.12)	(34.71)
Obs.	231	231	231	231	209	202	245	245
Clusters	33	33	33	33	30	29	35	35

Note: Table 5 displays the OLS coefficients clustered at the participant level of i's reported emission level on i's actual emission levels and the average emission levels of i's fellow group members, including rounds fixed. We control for participant i's gender, age, risk attitudes and whether French is i's native language in columns (2), (4), (6) and (8). Standard errors in parentheses. Stars indicates significant differences from 0. *p<0.05, **p<0.01, ***p<0.001.

While consistent in terms of direction, the effect sizes are smaller than predicted. For instance, in the CAM-PI treatment, one can compare the theoretically predicted effect of a firm increasing her emission by one unit at the equilibrium on her own reporting with the empirically estimated coefficient. Using the calibration in table (1), the predicted effect size is 1.57. In contrast, the empirically estimated effect size in specifications (7) and (8) is 0.705 and 0.710, respectively (chi2 tests H0: $\beta = 1.57$: $p < 0.001$ in both models). Likewise in the CAM-PI treatment, the theoretically predicted effect of both peers increasing their emission level by one unit (*i.e.*, GA emission level increases by 1) on participant's own reporting is 2.01 units. In contrast, the empirically estimated effect in specifications (7) and (8) is only 0.474 and 0.517, respectively (chi2 tests H0: $\beta = 2.01$: $p < 0.001$ in both models). These observations are in line with Result 3: The CAM-PI treatment produces lower emission levels, but not as low as predicted by Hypothesis 3.

The failure of Hypothesis 3 points to the presence of mechanisms that the theory does not account for. Several mechanisms could be at play: For instance, if participants systematically report above the profit-maximizing level, then this can dilute the competitive effects induced by the CAM. Also, the CAM adopted in this paper is "calibrated" to implement socially optimal emission levels among risk-neutral firms. Because laboratory participants are typically risk-averse (Harrison and Rutström, 2008), deviations that reduce variation in expected profit are to be expected. Moreover, even though the experiment conceptually is a series of one-shot plays, we cannot exclude the possibility of inter-round spillovers.[25] Finally, several participants reported that their goal was to avoid being audited. Such 'audit aversion' may lead to higher emission levels in stage 1 in order to be able to report more in stage 2, which in turn leads to a lower audit probability in stage 3. One piece of evidence from the experimental data is that risk attitudes, while affecting self-reporting behavior, does not appear to be linked to the choice of emission level (see Table A2 in Appendix).

Ultimately, the analyses remain speculative since we cannot rule out alternative causal mechanisms through experimental design. Further research, however, can build on this evidence to establish the role of other drivers and to determine what features would enable the mechanism to implement full compliance.

5 Conclusion

Recent theory and experiments point to the potential of competitive audit mechanisms for harnessing the benefits of self-reporting to achieve regulatory targets while limiting misreporting. This makes such mechanisms of interest to both academics and regulators. To make progress, our paper scrutinizes this potential with respect to two dimensions. One is a comparison of

[25] Table A3 in Appendix D.2 shows that the results from Table 5 remain quantitatively the same after controlling for whether a participant has been audited in the previous round.

competitive audit mechanisms with the more conventional random audit mechanism not just in terms of truthful reporting, but also in terms of aligning emission levels with regulatory targets. The other is a better understanding of the role of peer information in the comparative performance. The present paper integrates the key results of the existing literature into a parsimonious theoretical framework and translates this framework into an experimental design in which information structures can be exogenously manipulated in isolation.

The theoretical part of our investigation generates predictions that the competitive audit mechanism will outperform the random audit mechanism in terms of truthful self-reporting irrespective of the presence or absence of peer information, but in terms of emission levels only in its presence. In the latter case, however, socially optimal emission levels are predicted to be implementable. All except one of these predictions turn out to be accurate. The online experiment that implements the theoretical framework in the form of an experimental design finds that on average, participants' emission levels do not differ between the mechanisms in the absence of peer information. When participants can observe peers' emission levels, the competitive mechanism brings emission levels closer into line with regulatory targets. The only theoretical prediction that fails its test is that the competitive mechanism implements the socially optimal emission levels: Participants still under-comply with the regulatory objectives. This could be the result of several factors, including risk attitudes, lying aversion, or a type of audit aversion. These insights are important whenever competitive auditing is considered as a viable alternative for constrained regulators charged with enforcing policies. For these regulators, it is crucial to understand the circumstances under which competitive auditing leads to advantages over random auditing with regards to their primary objective of aligning emission levels closer with the regulatory target. Finally, we replicate the previously established finding that participants self-report more truthfully under the competitive mechanism in both information conditions.

Two stylized facts that are not accounted for by the theoretical framework emerge from the experiment. First, while the equilibrium under competitive auditing may be threatened by collusion, the collusive equilibrium of no reporting is not observed in the data. In fact, we replicate the well-established finding that self-reporting is significantly higher than predicted by theory regardless of the information structure. This is consistent with the behavioral literature that people decisions can also be influenced by payoff-irrelevant motivations such as social considerations or lying-aversion. Second, the competitive audit mechanism can induce less heterogeneity in individuals' decisions, both in terms of self-reported emission levels and in terms of actual emission levels. These effects constitute unexpected co-benefits of using a competitive audit mechanism and reaffirm the benefits of test-bedding regulations.

One limitation in our current study is that we could only examine the two limiting cases of peer information. In regulatory reality, most situations will fall somewhere between the two extremes of no or perfect peer information. In addition, regulators may have access to some information about firms emissions prior to conducting an audit. It will be interesting to explore in future research just how good peer information has to be for competitive audit mechanisms to outperform a random mechanism that recommends itself to regulators through its simplicity. Despite this limitation, the theoretical considerations and experimental evidence presented in this paper strongly indicate that policy-makers should give audit tournaments more consideration. This is because a regulator looks unlikely to ever do worse by choosing an audit tournament instead of a random mechanism. This is particularly true in contexts in which reporting accuracy has particular regulatory value. There, audit tournaments do strictly better than a random mechanism. But our results extend this insight to emissions as well. Both of these points should make policy-makers interested in considering audit tournaments. When introducing audit tournaments, policy-makers should prioritize regulatory settings in which peer information is high and consider low-cost investments that enhance peer information alongside the introduction of audit tournaments.

References

Bayer, R. and Cowell, F. (2009). Tax compliance and firms' strategic interdependence. *Journal of Public Economics*, 93(11-12):1131–1143.

Bock, O., Baetge, I., and Nicklisch, A. (2014). hroot: Hamburg registration and organization online tool. *European Economic Review*, 71:117–120.

Cason, T. N., Friesen, L., and Gangadharan, L. (2016). Regulatory performance of audit tournaments and compliance observability. *European Economic Review*, 85:288–306.

Cason, T. N., Friesen, L., and Gangadharan, L. (2021). Complying with environmental regulations: experimental evidence. In *A Research Agenda for Experimental Economics*, pages 69–92. Edward Elgar Publishing.

Cason, T. N. and Gangadharan, L. (2013). Empowering neighbors versus imposing regulations: An experimental analysis of pollution reduction schemes. *Journal of Environmental Economics and Management*, 65(3):469–484.

Chen, D. L., Schonger, M., and Wickens, C. (2016). otree—an open-source platform for laboratory, online, and field experiments. *Journal of Behavioral and Experimental Finance*, 9:88–97.

Cohen, J. (2013). *Statistical power analysis for the behavioral sciences*. Academic press.

De Marchi, S. and Hamilton, J. T. (2006). Assessing the accuracy of self-reported data: an evaluation of the toxics release inventory. *Journal of Risk and Uncertainty*, 32(1):57–76.

Dohmen, T., Falk, A., Huffman, D., Sunde, U., Schupp, J., and Wagner, G. G. (2011). Individual risk attitudes: Measurement, determinants, and behavioral consequences. *Journal of the european economic association*, 9(3):522–550.

Duflo, E., Greenstone, M., Pande, R., and Ryan, N. (2018). The value of regulatory discretion: Estimates from environmental inspections in india. *Econometrica*, 86(6):2123–2160.

Earnhart, D. and Friesen, L. (2021). Use of competitive endogenous audit mechanisms by federal and state inspectors within environmental protection agencies. *Journal of Environmental Economics and Management*, in press.

Evans, M. F., Gilpatric, S. M., and Liu, L. (2009). Regulation with direct benefits of information disclosure and imperfect monitoring. *Journal of Environmental Economics and Management*, 57(3):284–292.

Friesen, L. and Gangadharan, L. (2013). Designing self-reporting regimes to encourage truth telling: An experimental study. *Journal of Economic Behavior & Organization*, 94:90–102.

Gilpatric, S. M., Vossler, C. A., and Liu, L. (2015). Using competition to stimulate regulatory compliance: A tournament-based dynamic targeting mechanism. *Journal of Economic Behavior & Organization*, 119:182–196.

Gilpatric, S. M., Vossler, C. A., and McKee, M. (2011). Regulatory enforcement with competitive endogenous audit mechanisms. *The RAND Journal of Economics*, 42(2):292–312.

Gneezy, U. (2005). Deception: The role of consequences. *American Economic Review*, 95(1):384–394.

Goldstone, R. L. and Chin, C. (1993). Dishonesty in self-report of copes made: moral relativity and the copy machine. *Basic and Applied Social Psychology*, 14(1):19–32.

Harford, J. D. (1987). Self-reporting of pollution and the firm's behavior under imperfectly enforceable regulations. *Journal of Environmental Economics and Management*, 14(3):293–303.

Harrison, G. W. and Rutström, E. E. (2008). Risk aversion in the laboratory. In *Risk aversion in experiments*. Emerald Group Publishing Limited.

Healy, P. J. (2006). Learning dynamics for mechanism design: An experimental comparison of public goods mechanisms. *Journal of Economic Theory*, 129(1):114–149.

Helland, E. (1998). The enforcement of pollution control laws: Inspections, violations, and self-reporting. *Review of Economics and Statistics*, 80(1):141–153.

Innes, R. (2017). Lie aversion and self-reporting in optimal law enforcement. *Journal of Regulatory Economics*, 52(2):107–131.

Malik, A. S. (1993). Self-reporting and the design of policies for regulating stochastic pollution. *Journal of Environmental Economics and Management*, 24(3):241–257.

Mazar, N., Amir, O., and Ariely, D. (2008). The dishonesty of honest people: A theory of self-concept maintenance. *Journal of marketing research*, 45(6):633–644.

Oestreich, A. M. (2015). Firms' emissions and self-reporting under competitive audit mechanisms. *Environmental and Resource Economics*, 62:949–978.

Oestreich, A. M. (2017). On optimal audit mechanisms for environmental taxes. *Journal of Environmental Economics and Management*, 84:62–83.

Requate, T. and Waichman, I. (2011). "a profit table or a profit calculator?" a note on the design of cournot oligopoly experiments. *Experimental Economics*, 14(1):36–46.

Telle, K. (2015). Monitoring and enforcement of environmental regulations: lessons from a natural field experiment in norway. *Journal of Public Economics*, 99:24—-34.

Tirole, J. (1988). *The theory of industrial organization.* MIT press.

Vossler, C. A. and Gilpatric, S. M. (2018). Endogenous audits, uncertainty, and taxpayer assistance services: Theory and experiments. *Journal of Public Economics*, 165:217—-229.