Indian Institute of Management Calcutta

Working paper Series

WPS No 811/August 2018

Costly State Verification and Management Control Systems: Experimental evidence

Priyodorshi Banerjee

Indian Statistical Institute Calcutta

Purusottam Sen

Indian Institute of Management Calcutta

&

P. Srikant

Indian Institute of Management Calcutta

Indian Institute of Management Calcutta
Joka, D H Road
Kolkata 700104, INDIA
Email: psen@iimcal.ac.in

URL: http://facultylive.iimcal.ac.in/workingpapers

Costly State Verification and

Management Control Systems:

Experimental evidence

Priyodorshi Banerjee

Indian Statistical Institute Calcutta

Purusottam Sen

Indian Institute of Management Calcutta

P. Srikant

Indian Institute of Management Calcutta

August, 2018

Preface

This report summarizes the progress of our experimental project, which tries to understand how managers choose specific control systems from the perspective of verification costs.

We would like to express our thanks to the Fellow Programme and Research Office, IIM Calcutta, for its support in the conduct of these experiments.

Funding for carrying out these experiments was received under the following grant:

Work Order No : 3533/RP:ESCOMCM/2012-13

"Experimental Studies on the choice and outcomes of Managerial Control Mechanisms"

We would also like to thank the the Finance and Trading Laboratory, IIM Calcutta, for providing us a convenient and well-equipped venue to carry out these experiments.

Abstract

A costly state verification or inspection situation arises when a subordinate is better informed than a manager about an outcome, and has incentives to misreport the outcome; and the manager can choose to determine, at some cost, whether the report was true or not. In this paper, we devise and conduct a series of experiments to understand how managers try to choose control systems to minimize the costs of such cheating by subordinates. For the basic inspection game, which corresponds to interactive control mechanisms in the levers of control framework, our findings are that the player behaviours are somewhat in line with the theoretical predicted mixed strategy Nash equilibrium . While theoretical predictions on the levels of misreporting and verification are not supported, most comparative static effects are in line with theoretical predictions. We frame the choice of alternative control mechanisms as follows : Managers are given a choice of making an investment in a technology that prevents misreporting, or playing the basic inspection game. We find that managers often choose to retain interactive control even when it is more efficient to invest in boundary or diagnostic control. Also, when managers choose to opt for interactive control, subordinates are likelier to misreport and managers are likelier to verify than the theoretical benchmarks.

1 Introduction

Management control systems have been defined as "the formal, information-based routines and procedures managers use to maintain or alter patterns in organizational activities" (Simons, 1994).

While management control systems play a central role in the functioning of organizations, there are still significant gaps in our understanding of what is known about what motivates the choice of a specific management control mechanism over another. The development of theories to explain such choices is complicated by several issues. Firstly, organizations operate in a variety of different environments, and have diverse firm characteristics and strategies. Secondly, multiple control mechanisms are simultaneously present in organizations and the organizational outcome depends upon the how these various control mechanisms work together . It is therefore hard to study the effect of an individual control mechanism, and to generalize the findings across a variety of firms. Thirdly, management control system are often treated as a "given", rather than as the product of conscious managerial choices.

In this article, we try to develop insights on the choice of control systems using a laboratory inspection game. To the best of our knowledge, this angle has not been explored in the prior controls literature; and we argue that this perspective can provide useful insights for the design of control systems. In a standard inspection game or costly state verification scenario, the agent stands to gain by misreporting an output that he costlessly observes, while the principal needs to expend effort to determine the output. It is not difficult to imagine several managerial contexts where such situations arise. For example, an unsupervised clerk in a store may dip into the till; or a manager may attribute the low performance of his division to weak demand. Principals may adopt several approaches to control such cheating by the agent. For instance, they may choose to engage in random checks; or they may choose to check when presented with low reports; or they may choose to invest in technologies to observe the output.

Table 1: Terminologies used in control systems

Verification approach	**Levers of Control**	**Other terms**
random	Interactive	pure / discretionary verification
always	Boundary	Preventive control
for specific reports	Diagnostic	Exception reporting; variance analysis

Several frameworks have been devised to classify control systems (Merchant and Otley, 2006) . For example, control mechanisms have been classified into action, results and personnel/cultural controls (Merchant and Van der Stede, 2007). The levers of control framework (Simons, 1994) classifies control systems into belief systems, boundary systems, interactive controls, and diagnostic controls. Here, diagnostic controls refers to management by exception, variance accounting and investigation, and corresponds to the notion of a deterministic audit regime. Interactive control is when a manager chooses to retain the discretion on when to audit, and corresponds to the notion of discretionary or probabilistic verification. And lastly, technologies to always verify prevent misreporting, and can be thought of as boundary systems.

A variety of terms are used to describe these basic approaches across different authors and disciplines. To prevent terminological confusion and clearly delineate our mapping to the levers of control framework, we present some of these alternate nomenclatures in Table 1 .

The rest of this article is structured as follows. We first present a model where we describe our basic inspection setup. We then describe an experimental implementation of a simplified version of this model. We ran two studies, which involved playing this game with different parameter combinations. We describe each of these studies along with their associated experimental results. Finally we conclude with a discussion of our findings, and its implications for the design of control systems.

2 Model

We first consider a simple two-stage pure verification game. . Inspection games were first analyzed by (Dresher, 1962). They have been applied to problems of audit starting with (Borch, 1982). See (Fudenberg and Tirole, 1991) and (Avenhaus et al., 2002) for reviews. There are two players 1 and 2 (player 1 is the manager/worker/agent, and player 2 is the firm/manager/principal). In stage 1, a divisible surplus X is realized. It is common knowledge that X can be either 0 or $x > 0$, and that the probability X takes the value 0 is $\alpha \in (0, 1)$. However, the realization of X is observed only by player 1. Player 1 then has to give a transfer$t$ to player 2. The prior understanding is that $t = \beta X$, where $\beta \in (0, 1)$. β is thus the fraction of the surplus that is claimable by player 2. There is limited liability, and the transfer cannot exceed the surplus, so if $X = 0$, player 1 has to give $t = 0$. If $X = x$, since player 1 is rational, he may have an incentive to give less thanβx, since player 2 does not know the true realization. x is the potential claimable surplus, and is a measure of the stake player 2 has in the project.

In stage 2, after receiving the transfer, player 2 has the option of accepting the transfer, in which case the payoff of player 1 is $X - t$, and that of player 2 is t. Alternatively, she can challenge/verify at a cost $c_V > 0$. c_V is the cost of verification, and can be thought of as a measure of the efficiency of the verification technology. If on verification it is found that $t \geq \beta X$, the challenge is considered a failure, and the utilities of the two parties are $X - t$ for player 1 and $t - c_V$ for player 2. However, if it is found that $t < \beta X$, the verification is considered a success and the payoffs of the two parties are $(1 - \gamma)X$ for player 1 and $\gamma X - c_V$ for player 2, where $\gamma \in (\beta, 1]$. γ is a measure of the penalty to be faced by player 1 if there is a detected misreport, or the strength of incentives to deter misreporting: if $\gamma = 1$, incentives are maximal.

We shall analyze perfect Bayesian Nash equilibrium of this game. Observe that if $\gamma x \leq c_V$, player 2 will never verify, and so player 1 will always send $t = 0$, whenever $X = x$. So we assume henceforth that $\gamma x > c_V$. Next observe that player 2 will never verify if $t \geq \beta x$, and so, since player 1's payoff is decreasing in t, he will never choose $t > \beta x$, if $X = x$.

We now examine player 2's decision in stage 2 when $t \in (0, \beta x)$. Since $t > 0$, player 2 infers that $X = x$. If she accepts the transfer, her utility is t. If she verifies, her utility is $\gamma x - c_V$. Hence a necessary condition for player 2 to verify, for any such t, is $c_V \leq \gamma x - \beta x$.

Suppose then $c_V \leq (\gamma - \beta)\, x$. Player 2 will then verify whenever $t \in (0, \beta x)$ and never verify for $t \geq \beta x$. If $t = 0$, if player 2 accepts the transfer, her utility is 0. If she verifies, given that she believes player 1 gives $t = 0$ when $X = x$ with probability σ, her payoff is $\frac{(1-\alpha)\sigma}{\alpha+(1-\alpha)\sigma}\gamma x - c_V$.

Given $X = x$, player 1 then never chooses $t \in (0, \beta x)$. If he chooses $t = \beta x$, he gets $(1 - \beta)x$. But if he chooses $t = 0$, while believing player 2 will verify with probability π, he gets $(1 - \pi)x + \pi(1 - \gamma)x$.

We cannot have an equilibrium where $\pi = 1$, i.e. player 2 always verifies a zero transfer, as then player 1 would not have an incentive to violate the agreement, leaving player 2 with no incentive to verify. We also cannot have one with $\sigma = 0$, i.e. player 1 never violates, as then player 2 would never verify, yielding an incentive to player 1 to violate. Further, if $\pi = 0$ in equilibrium, it must also be that $\sigma = 1$.

We thus find $(\pi = 0, \sigma = 1)$ is the unique equilibrium if $c_V > (1 - \alpha)\gamma x$. But if $c_V < (1 - \alpha)\gamma x$, $\sigma = 1$ can never be an equilibrium, as then player 2 would always verify a zero transfer, removing player 1's incentive to always violate.

Our interest is in situations where the prior agreement is at least sometimes upheld in equilibrium, and so we shall assume henceforth $c_V \leq min((1 - \alpha)\gamma x, (\gamma - \beta)x))$. In this case, the unique equilibrium is in purely mixed strategies

$$\pi = \frac{x - (1 - \beta)x}{x - (1 - \gamma)x} = \frac{\beta}{\gamma}, \sigma = \frac{\alpha c_V}{(1 - \alpha)(\gamma x - c_V)}$$

P_V^2, the expected net payoff of player 2 is

$$P_V^2 = \frac{\beta x\{(1-\alpha)\gamma x - c_V\}}{\gamma x - c_V}$$

P_V^1, the expected net payoff of player 1 is

$$P_V^1 = \frac{x[\{(1-\alpha)\gamma x - c_V\}(1-\beta) + \alpha c_V(1-\frac{\beta}{\gamma})]}{\gamma x - c_V}$$

To help us understand the effects of changes in c_V, γ and x, we can get the following comparative static effects in equilibrium:

$$\frac{\partial \sigma}{\partial c_V} > 0, \frac{\partial \pi}{\partial c_V} = 0, \frac{\partial \sigma}{\partial \gamma} < 0, \frac{\partial \pi}{\partial \gamma} < 0, \frac{\partial \sigma}{\partial x} < 0, \frac{\partial \pi}{\partial x} = 0$$

Thus an increase in the cost of verification c_V leads to an increase in the equilibrium probability of misreporting σ, while increases in the penalty conditional on a false report being detected γ, or the principal's stake x, lead to the misreporting probability being lowered. The verification probability π is independent of the stake size, as well as the verification cost in equilibrium, while it is reduced by an increase in the penalty.

3 Basic Experimental Design

A number of approaches have been used to address questions on the choice of control systems. For example, economic models have studied contracting relationships as a fit between the firm's task environment and firm characteristics. The recognition that firms face unique problems, and that control systems exist as packages, has inspired the use of contingency-based theorizing in studying the choice and outcomes of control systems (Chenhall, 2005). Despite various complications, a growing body of empirical work, for example (Widener, 2007), has begun to address these questions. In order to study the effect of a control mechanisms in isolation while holding other factors constant, laboratory experiments can be a useful technique(Sprinkle, 2003).

The model presented above fortunately lends itself to a straightforward experimental implementation. Our basic experimental treatment is based on a simplification of the above theoretical model, and proceeds as follows. The laboratory experimental literature surrounding inspection games remains, to our knowledge, very limited. (Glimcher et al., 2005) explore the effect of changing the cost of verification, while (Rauhut, 2009) investigates the impact of changing the penalty for misreporting. (Nosenzo et al., 2013) also use the inspection game in their study, but do not implement comparative static analyses.

The model presented above fortunately lends itself to a straightforward experimental implementation. Our basic experimental treatment is based on a simplication of the above theoretical model, and proceeds as follows. 2 players, denoted by Red and Green, have a prior agreement to share the profits of a project equally. (The Red player corresponds to Player 1 in the model described above, and the Green player corresponds to Player 2.). Both players know that the project's return is equally likely to be either 0 or x.

- In the first stage, the Red player sees the return of the project and makes a transfer
  - If the return is 0, the Red player must transfer 0.
  - If the return is x, The Red player can choose to transfer either 0 or $\frac{x}{2}$ to the Green player .
- In the second stage, the Green player receives the transfer.
  - On receiving a transfer of $\frac{x}{2}$, the Green player accepts the transfer and the game ends with each player receiving $\frac{x}{2}$.
  - On receiving a transfer of 0, the Green player can choose to either

* accept the transfer and get 0. The Red Player gets the actual return (x or 0)

* To verify the return, which involves a cost of c_v. If verification reveals that the agreement was violated, the Green Player receives $(\gamma x - c_v)$ and the Red Player receives $(1-\gamma)x$. If verification reveals that the agreement was not violated, the Green Player receives $-c_v$ and the Red Player receives 0

The key simplification made in the experimental implementation is thus that Player 1 or Red can transfer 0 or $\frac{x}{2}$ (whereas in the more general model above, he could in principle transfer anything between 0 and the actual realization). Observe, as long as the parameter restriction $c_V \leq min((1-\alpha)\gamma x, (\gamma - \beta)x))$ is satisfied, this does not affect the equilibrium. In this case, as discussed above, the strategies of the two players can be characterized by two numbers:

- σ : the probability that the Red player transfers 0 when the return is x

- π : the probability that the Green player chooses to verify when given a transfer of 0

We ran two studies which involved variations of this basic design. In study 1, subjects played the basic inspection game with different parameter combinations. Varying parameters gives rise to different treatments. The aim of study 1 was to understand how the inspection game fared in the laboratory, to see if can be a useful model to approach control problems. In study 2, subjects were given a choice between playing the basic inspection game, or choosing an alternate control mechanism. The aim of study 2 was to understand how firms choose between control mechanisms under different conditions.

Table 2: Experimental Sessions conducted

Study	***Session***	**Date**	**Time**	**Participants**	Rounds
1	1	29 Jun 2013	11 AM – 1 PM	30	25
	2	29 Jun 2013	2 PM – 4 PM	28	25
2	3	15 Sep 2013	2 PM - 3:30 PM	26	25
	4	15 Sep 2013	4 PM - 5:30 PM	36	25

4 Experimental Procedure

We conducted a total of four experimental sessions, as listed in Table 2. All sessions were conducted at the Finance and Trading Laboratory, Indian Institute of Management, Calcutta. Volunteer participants were sourced from the full-time residential MBA programs, with no subject participating in more than one session. Participants played 25 rounds of a computerized experiment, implemented using the widely-used z-Tree software (Fischbacher, 2007). Subjects were assigned a fixed role, that is, their color assignment was fixed for the duration of the experiment. In each round, they were randomly matched with a player of the opposite color, with matches dissolving at the end of each round. Anonymity was maintained, and no subject ever knew the identity of her matched partner in any round. Subjects were also not allowed to communicate during the experiment. All participants received a show-up fee of Rs. 100 and a certificate, as well as a performance-based reward of up to Rs. 500, with payments being made in private at the end of the experiment. We now describe each of these studies in detail.

5 Study 1

In our first study, we ran two sessions. In Session 1, we set $x = 100$ and vary the penalty γ and cost of verification c_v. Subjects played with fixed roles, and were randomly rematched in each round to a player of the other color to form a group. Each group in any round had one of the following parameter combinations:

Table 3: Session 1

			Payoffs: $x = 100$					
(Random)	(Red)	(Green)	1:γ=1;c_v=35		2:γ=0.8;c_v=35		3:γ=1;c_v=20	
Return	Transfer	Verification	Red	Green	Red	Green	Red	Green
0	0	Yes	0	-35	0	-35	0	-20
0	0	No	0	0	0	0	0	0
100	0	Yes	0	65	20	45	0	80
100	0	No	100	0	100	0	100	0
100	50	No	50	50	50	50	50	50
Predicted Equilibrium		$\begin{bmatrix} \sigma \\ \pi \\ P_G \\ P_R \end{bmatrix}$	$\begin{bmatrix} 0.54 \\ 0.50 \\ 11.54 \\ 25.00 \end{bmatrix}$		$\begin{bmatrix} 0.78 \\ 0.63 \\ 5.56 \\ 25.00 \end{bmatrix}$		$\begin{bmatrix} 0.25 \\ 0.50 \\ 18.75 \\ 25.00 \end{bmatrix}$	

1. $\gamma = 1;\ c_v = 35$

2. $\gamma = 0.8;\ c_v = 35$

3. $\gamma = 1;\ c_v = 20$

Subjects faced different parameter combinations over the course of the experiment. The payoff matrix and predicted equilibrium for risk-neutral agents is presented in Table 3 . P_G and P_R are the expected payoffs of the Green player and Red Player in the predicted equilibrium.

In Session 2, we set $x = 150$. Except for the change in x , sessions 1 and 2 are otherwise identical. The payoff matrix and predicted equilibrium for risk-neutral agents is presented in Table 4 .

The design therefore studies comparative static effects of changes in γ and c_V through within-subjects treatments, as the same subject faces different values of these two within a given session. Comparative static effects of a change in x by contrast are explored through between-subjects treatments, by comparing outcomes, given values of γ and c_V, across sessions 1 and 2.

Table 4: Session 2

			Payoffs: $x = 150$					
(Random)	(Red)	(Green)	1:γ=1;c_v=35		2:γ=0.8;c_v=35		3:γ=1;c_v=20	
Return	Transfer	Verification	Red	Green	Red	Green	Red	Green
0	0	Yes	0	-35	0	-35	0	-20
0	0	No	0	0	0	0	0	0
150	0	Yes	0	115	30	85	0	130
150	0	No	150	0	150	0	150	0
150	50	No	75	75	75	75	75	75
Predicted Equilibrium		$\begin{bmatrix} \sigma \\ \pi \\ P_G \\ P_R \end{bmatrix}$	$\begin{bmatrix} 0.30 \\ 0.50 \\ 26.09 \\ 37.50 \end{bmatrix}$		$\begin{bmatrix} 0.41 \\ 0.63 \\ 22.06 \\ 37.50 \end{bmatrix}$		$\begin{bmatrix} 0.15 \\ 0.50 \\ 31.73 \\ 37.50 \end{bmatrix}$	

Results from prior experimental research suggests that in strategic environments with a unique mixed strategy Nash equilibrium, actual outcomes are typically located at some distance from theoretical predictions . (see, for example, (Erev and Roth, 1998), (Selten and Chmura, 2008) etc.). The failure of subjects to conform to Nash predictions is usually attributed to the complexity of such environments. In Study 1 therefore, to give theoretical predictions some chance of being useful as a benchmark for purposes of comparison, subjects were asked additional cueing questions, meant to help them perform better calculations.The failure of subjects to conform to Nash predictions is usually attributed to the complexity of such environments. In Study 1 therefore, to give theoretical predictions some chance of being useful as a benchmark for purposes of comparison, subjects were asked additional cueing questions, meant to help them perform better calculations. The questions were

- q1: To a Red player, when the subject makes a zero transfer:

  What do you think is the likelihood (any integer between and including 0 and 100) that your matched Green player will verify ?

- q2: To a Red player, when the subject makes a transfer of $\frac{x}{2}$:

  What do you think is the likelihood (any integer between and including 0 and 100

) that your matched Green player would have verified if you had chosen t=0 ?

- q3: To a Green player, when the subject receives a zero transfer:

  What do you think is the likelihood (any integer between and including 0 and 100) that your matched Red player has violated the prior agreement ?

- q4: To a Green player, when the subject receives a transfer of $\frac{x}{2}$:

  What is the likelihood (any integer between and including 0 and 100) that you would have verified if you had received a transfer of 0 ?

Results of Study 1

We look at the values of σ and π for each parameter combination, to study how closely they match predicted equilibria, and to study comparative static effects. As shown by Table 5 , the pattern is broadly along the lines of what was predicted. The differences from predicted values are possibly due to risk aversion. Denote the likelihood that the red player transfers 0 on a high return by $\sigma_{i,j}$ for parameter combination j in treatment i . Similarly, denote the likelihood that the green player verifies when given a transfer of 0 by $\pi_{i,j}$. Our experimental data suggest the following:

- The null hypothesis that actual and predicted outcomes match is not rejected for $\sigma_{1,3}$, $\sigma_{2,3}$, $\pi_{1,1}$, $\pi_{1,3}$, $\pi_{2,1}$, $\pi_{2,2}$. The probabilities are closer to theoretical values for verification by the Green player, than for misreporting by the Red player.

- Comparative static predictions for σ within subjects match in 3 out of 4 cases, as we find $\sigma_{1,1} < \sigma_{1,2}$, $\sigma_{2,1} < \sigma_{2,2}$, and $\sigma_{2,1} > \sigma_{2,3}$. In the remaining case, we find $\sigma_{1,1} = \sigma_{1,3}$, while the predicted outcome was $\sigma_{1,1} > \sigma_{1,3}$. The statistical tests are reported in Table 6.

- Comparative static predictions for π within subjects also match in 3 out of 4 cases, as we find $\pi_{1,1} < \pi_{1,2}$, $\pi_{1,1} = \pi_{1,3}$, and $\pi_{2,1} = \pi_{2,3}$. In the remaining case, we find

$\pi_{2,1} = \pi_{2,2}$, while the predicted outcome was $\pi_{2,1} < \pi_{2,2}$. The statistical tests are reported in Table 7.

- Comparative static predictions for σ between subjects match in 2 out of 3 cases, as we find $\sigma_{1,2} > \sigma_{2,2}$, and $\sigma_{1,3} > \sigma_{2,3}$. In the remaining case, we find $\sigma_{1,1} < \sigma_{2,1}$, while the predicted outcome was $\sigma_{1,1} > \sigma_{2,1}$. The statistical tests are reported in Table 8.

- Comparative static predictions for π between subjects do not match in any of the 3 cases, as we find $\pi_{1,1} < \pi_{2,1}$, $\pi_{1,2} < \pi_{2,2}$, and $\pi_{1,3} < \pi_{2,3}$, while the respective predictions were $\pi_{1,1} = \pi_{2,1}$, $\pi_{1,2} = \pi$, and $\pi_{1,3} = \pi_{2,3}$. However, the deviations are relatively small, and in all the same direction (increased verification due to increase in x). The statistical tests are reported in Table 8.

- Except for $i = 2,\ j = 2$ and $i = 1,\ j = 3$, $\left(\sigma_{i.j}^{actual} - \sigma_{i.j}^{predicted}\right)$ and $\left(\pi_{i.j}^{actual} - \pi_{i.j}^{predicted}\right)$ have the same sign suggesting that players try to play best responses.

As mentioned, we had asked additional cueing questions in Treatments 1 and 2. We report these responses in Table 9. In addition to cueing the players to think more deeply about the game, beliefs are also the fourth lever of control in Simons' framework. The responses indicate that the Red player is fairly consistent in estimating the probability that the Green player will audit across his transfers. This conclusion follows from comparing the average responses to questions 1 and 2 across treatments and sessions: with one exception, the average response to questions 1 and 2 are the same. However, he tends to overestimates the actual likelihood of audit. We obtain this conclusion by comparing responses to questions 1 and 2 in Table *9* with actual values from Table *5*. Comparison of the two tables also show that the Green players form relatively accurate estimates of their own actions, but appear to believe that Red players will lie half the time.

Table 5: Study 1 : Comparison of actual and theoretical results

Treatment	Parameters	σ		π	
		observed	predicted	observed	predicted
1: $x = 100$	1:γ=1;c_v=35	0.31 0.00 (N=68)	0.54	0.44 0.31 (N=78)	0.50
	2:γ=0.8;c_v=35	0.65 0.02 (N=60)	0.78	0.51 0.01 (N=104)	0.63
	3:γ=1;c_v=20	0.29 0.52 (N=52)	0.25	0.48 0.75 (N=88)	0.50
2: $x = 150$	1:γ=1;c_v=35	0.41 0.07 (N=61)	0.30	0.55 0.53 (N=64)	0.50
	2:γ=0.8;c_v=35	0.53 0.05 (N=68)	0.41	0.58 0.33 (N=93)	0.63
	3:γ=1;c_v=20	0.20 0.23 (N=69)	0.15	0.59 0.19 (N=70)	0.50

σ is the probability that the Red Player reports t=0 when X > 0. π is the probability that the Green player verifies when offered a zero transfer. The null hypothesis in each row is that the observed probability matches the theoretically predicted value. The p-value is the two-sided p-value for a binomial random variable of the observed value when the null hypothesis is true.

Table 6: Study 1 : Within treatment comparative statics for σ

	$\sigma_{1,1}$	$\sigma_{1,2}$	$\sigma_{2,1}$	$\sigma_{2,2}$
$\sigma_{1,2}$	-3.845 (68, 60, 0.00)			
$\sigma_{1,3}$	0.240 (68, 52, 0.81)	3.802 (60, 52, 0.00)		
$\sigma_{2,2}$			-1.353 (61, 68, 0.18)	
$\sigma_{2,3}$			2.560 (61, 69, 0.01)	3.955 (68, 69, 0.00)

σ is the probability that the Red Player reports t=0 when X > 0. The null hypothesis in each cell the column heading is equal to the row heading. The cell entries report the Wilcoxon rank-sum statistics and number of observations in the first sample, second sample, and two-sided p-values.

Table 7: Study 1 : Within treatment Comparative statics for π

	$\pi_{1,1}$	$\pi_{1,2}$	$\pi_{2,1}$	$\pi_{2,2}$
$\pi_{1,2}$	-0.983 (78, 104, 0.33)			
$\pi_{1,3}$	-0.532 (78, 88, 0.59)	0.445 (104, 88, 0.66)		
$\pi_{2,2}$			-0.418 (64, 93, 0.68)	
$\pi_{2,3}$			-0.452 (64, 70, 0.65)	-0.065 (93, 70, 0.95)

π is the probability that the Green player verifies when offered a zero transfer. The null hypothesis in each cell the column heading is equal to the row heading. The cell entries report the Wilcoxon rank-sum statistics and number of observations in the first sample, second sample, and two-sided p-values.

Table 8: Study 1 : Between treatment comparative statics

	Parameter combinations		
	1	2	3
σ	-1.191 (68, 61, 0.23)	1.377 (60, 68, 0.17)	1.087 (52, 69, 0.28)
π	-1.312 (78, 64, 0.19)	-0.997 (104, 93, 0.32)	-1.352 (88, 70, 0.85)

σ is the probability that the Red Player reports t=0 when X $>$ 0.π is the probability that the Green player verifies when offered a zero transfer. The null hypothesis in each cell is that outcome in treatments 1 and 2 are equal. The cell entries report the Wilcoxon rank-sum statistics and number of observations in the first sample, second sample, and two-sided p-values.

Table 9: Belief elicitation questions

Treatment	Parameters	π				σ	
		q1 (Red)	q2 (Red)	q4 (Green)	observed (predicted)	q3 (Green)	observed (predicted)
1: $x = 100$	1:γ=1;c_v=35	58 (N=78, 0.00)	59 (N=47, 0.00)	49 (N=47, 0.00)	44 (50)	52 (N=78, 0.00)	31 (54)
	2:γ=0.8;c_v=35	54 (N=104, 0.00)	54 (N=21, 0.00)	44 (N=21, 0.00)	51 (63)	52 (N=104, 0.00)	65 (78)
	3:γ=1;c_v=20	66 (N=88, 0.00)	67 (N=37, 0.00)	55 (N=37, 0.00)	48 (50)	42 (N=88, 0.00)	29 (25)
2: $x = 150$	1:γ=1;c_v=35	58 (N=64, 0.00)	74 (N=36, 0.00)	53 (N=36, 0.00)	55 (50)	54 (N=64, 0.00)	41 (30)
	2:γ=0.8;c_v=35	58 (N=93, 0.00)	61 (N=32, 0.00)	47 (N=32, 0.00)	58 (63)	54 (N=93, 0.00)	53 (41)
	3:γ=1;c_v=20	74 (N=70, 0.00)	78 (N=55, 0.00)	62 (N=55, 0.00)	59 (50)	51 (N=70, 0.00)	20 (15)

The null hypothesis in each cell is that player beliefs match the predicted valuse. The p-value is the two-sided p-value of the t-statistic when the null hypothesis is true.

Table 10: Study 2: Green's investment decision

Session 3	Session 4			
$c_{boundary}$	$c_{diagnostic}$	Technology chosen	Technology not chosen	Prediction
4	8	21	18.75	Invest
8	16	17	18.75	Not invest
12	24	13	18.75	Not invest
16	32	9	18.75	Not invest

6 Study 2

In Study 2, we allow the Green players to choose whether to play the basic discretionary game with $x = 100$; $\gamma = 1$; $c_v = 20$ (Parameter combination 3 of Treatment 1); or to choose an alternate control mechanism. Hence we implement an inspection game as above, augmented with a prior mechanism choice stage. In Session 3, we allow the Green player to choose whether to invest in a technology to prevent misreporting at a cost of $c_{boundary}$. With the technology investment, Green's payoff is equally likely to be $50 - c_{boundary}$ or $-c_{boundary}$, to give an expected payoff of $25 - c_{boundary}$. The cost of the technology $c_{boundary}$ was varied across groups as shown in Table 10. Subjects faced different values of $c_{boundary}$ over the course of the experiment. The payoffs to the green player, and his likely investment decision (in the risk-neutral case), are also presented in this table.

In Session 4, we allow the Green players to choose whether to play the basic discretionary game with $x = 100$; $\gamma = 1$; $c_v = 20$ (Parameter combination 3 of Treatment 1); or to invest in a technology that commits them to a deterministic audit regime of verifying whenever they receive a transfer of 0. The cost of the technology is $c_{diagnostic}$. With the technology investment, Green's payoff is equally likely to be 50 or $-c_{diagnostic}$, to give an expected payoff of $25 - \frac{c_{diagnostic}}{2}$.

The cost of the technology $c_{diagnostic}$ was varied across groups as shown in Table 10. Subjects faced different values of $c_{diagnostic}$ over the course of the experiment. The payoffs to the green player in the risk-neutral case, and his likely investment decision, are also presented in this table. Since the model has binary outcomes, boundary and diagnostic controls systems are equivalent, and sessions 3 and 4 are essentially similar except for framing. It can be seen that the expected payoff of the Green player on choosing boundary control at a cost of c are the same as choosing diagnostic control at a cost of $2 * c$, but the spread of outcomes is wider with diagnostic control. Therefore, if agents are risk averse, we expect that boundary control will be chosen more frequently than diagnostic control with the same expected value.

Results of Study 2

Our predictions are that increasing the cost of technology will favor discretionary control being chosen over boundary / diagnostic control.

Table 11 provides a summary of the Green player's choice of control regime. The data indicate, that for a fixed cost of interactive control, green players increasingly prefer to opt for interactive control as the cost of boundary / diagnostic control increases; and that for a given expected value, boundary control is preferred over diagnostic control (except for the case where $c_{boundary} = 4$). Even for lowest technology cost, where even for risk-neutral case, prediction is people should choose technology, there is a significant proportion, about 30 - 35%, of cases where interactive control is chosen. For $c_{boundary} = \frac{c_{diagnostic}}{2} = 8$, where the payoffs are roughly equal, and so random choice should be encountered in the risk-neutral case, and preference for the deterministic outcome technology in the risk-averse case, we see strong preference for interactive control, with 60 - 65% choosing it. The preference becomes universal for higher technology costs. For each cost of technology, we are unable to reject the null hypothesis that managers choices are identical, suggesting that managers view these setups as identical.

We now focus on the instances where the Green player chooses discretionary control. The results are reported in Table 12 . The incidence of misreporting, and of verification,

Table 11: Sessions 3 and 4: Choice of regime

Session	$c_{boundary}$ or $\frac{c_{diagnostic}}{2}$			
	4	8	12	16
Risk-neutral prediction	1.000	0.000	0.000	0.000
3	0.707	0.340	0.093	0.040
4	0.660	0.400	0.064	0.020
test statistic	0.653 (0.51)	-0.923 (0.36)	0.761 (0.45)	0.784 (0.43)

Proportion of cases where technology investment is made. The null hypothesis in each cell is that the technology investments in treatments 3 and 4 are equal. The cell entries report the Wilcoxon rank-sum statistics and associated two-sided p-values.

are both much higher than the predicted values. However, the probability of verification is consistent across treatments, as shown in Table 13 . The average payoff of the Green player is more than his predicted equilibrium value. Possibly this is due to the absence of cueing questions in Treatments 3 and 4, or due to selection effects.

7 Discussion and Conclusions

Our studies demonstrate several puzzling patterns. In this section, we discuss some of the more puzzling patterns, and their possible reasons

For example, in Study 2, for the lowest technology cost, approximately one-third of managers choose to not invest in the technology. One reason could be that managers are willing to pay a premium for maintaining interactive control. Also, in a real-world situation, managers will likely have to perform several other tasks simultaneously. Forcing managers to allocate their time may likely lead to an increased preference for the technology investment as decision makers need to decide what tasks to focus their attention on (Ocasio, 1997) .

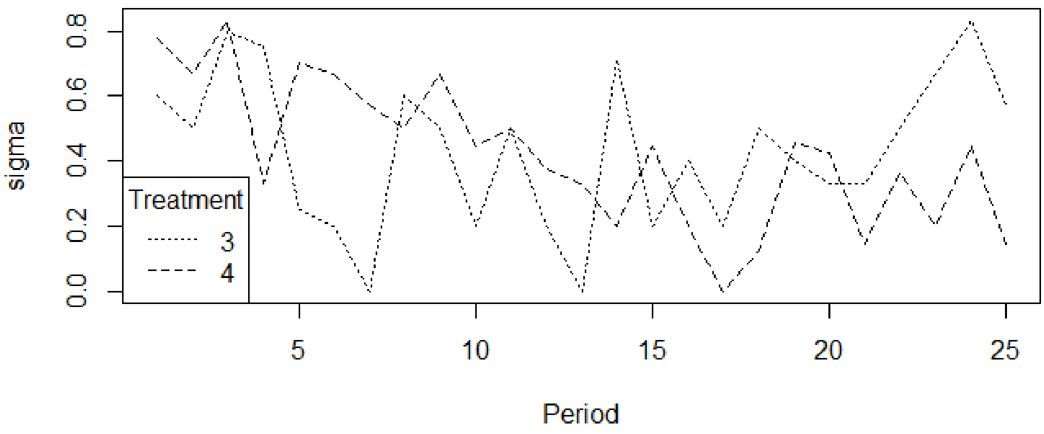
The high rate of verification in Study 2 when the technology investment is not made, relative to Study 1, suggests that the choice of control environment leads to behavioural displacement in the reporting and verifying choices of subjects. Therefore, it is important to consider such implications in the design of control mechanisms.

Table 12: Sessions 3 and 4: Discretionary regime

		$c_{boundary}$ or $\frac{c_{diagnostic}}{2}$					
Session	Variable	4	8	12	16	Overall	Predicted
3	σ	0.50 0.09 (N=12)	0.26 0.86 (N=42)	0.64 0.00 (N=28)	0.47 0.00 (N=38)	0.44 0.00 (N=120)	0.25
	π	0.938 0.00 (N=16)	0.886 0.00 (N=35)	0.862 0.00 (N=58)	0.827 0.00 (N=52)	0.86 0.00 (N=161)	0.50
	P_R	18.2	28.0	11.8	18.1	19.1	18.75
	P_G	22.7	26.2	14.7	22.8	21.4	25.00
4	σ	0.60 0.00 (N=20)	0.54 0.00 (N=41)	0.41 0.00 (N=71)	0.30 0.41 (N=50)	0.43 0.00 (N=182)	0.25
	π	0.962 0.00 (N=26)	0.804 0.00 (N=56)	0.827 0.00 (N=75)	0.905 0.00 (N=63)	0.86 0.00 (N=120)	0.50
	P_R	14.7	20.7	23.9	19.9	21.0	18.75
	P_G	29.4	22.0	26.2	19.5	23.5	25.00

σ is the probability that the Red Player reports t=0 when X $>$ 0. π is the probability that the Green player verifies when offered a zero transfer. The null hypothesis in each row is that the observed probability matches the theoretically predicted value. The p-value is the two-sided p-value of the observed value when the null hypothesis is true.

Figure 1: Reporting and verification behavior across periods



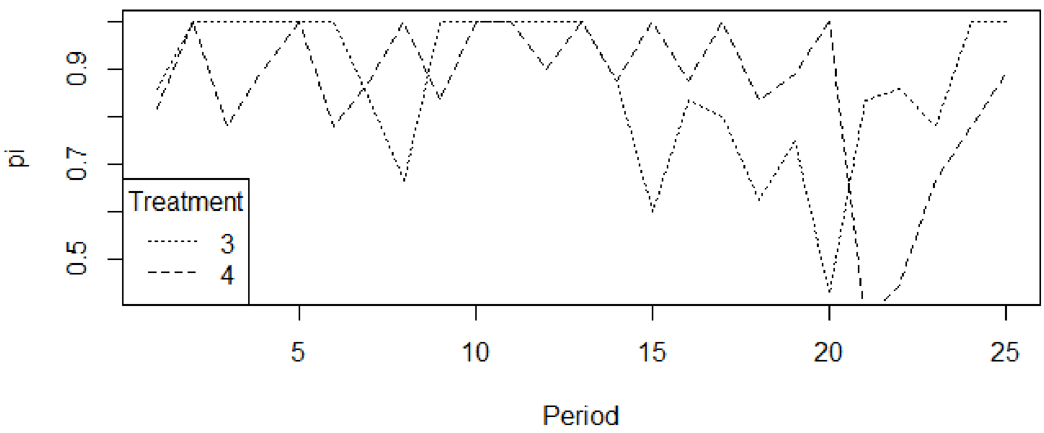


Table 13: Study 2 : Between treatment comparative statics for discretionary regime

	Parameter combinations			
	1	2	3	4
σ	-0.543 (0.59)	-2.541 (0.01)	2.093 (0.04)	1.658 (0.10)
π	-0.351 (0.73)	1.022 (0.31)	0.553 (0.58)	-1.228 (0.22)

σ is the probability that the Red Player reports t=0 when X $>$ 0.π is the probability that the Green player verifies when offered a zero transfer. The null hypothesis in each cell is that outcome in treatments 3 and 4 are equal. The cell entries report the Wilcoxon rank-sum statistics and associated p-values.

References

Rudolf Avenhaus, Bernhard Von Stengel, and Shmuel Zamir. Inspection games. *Handbook of game theory with economic applications*, 3:1947–1987, 2002.

K Borch. Insuring and auditing the auditor. In *Games, Economic Dynamics, and Time Series Analysis*, pages 117–126. Springer, 1982.

R.H. Chenhall. Integrative strategic performance measurement systems, strategic alignment of manufacturing, learning and strategic outcomes: an exploratory study. *Accounting, Organizations and Society*, 30(5):395–422, 2005.

Melvin Dresher. A sampling inspection problem in arms control agreements. 1962.

Ido Erev and Alvin E Roth. Predicting how people play games: Reinforcement learning in experimental games with unique, mixed strategy equilibria. *American economic review*, pages 848–881, 1998.

U. Fischbacher. z-tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics*, 10(2):171–178, 2007.

Drew Fudenberg and Jean Tirole. Game theory. 1991, 1991.

Paul W Glimcher, Michael C Dorris, and Hannah M Bayer. Physiological utility theory and the neuroeconomics of choice. *Games and economic behavior*, 52(2):213–256, 2005.

K.A. Merchant and D.T. Otley. A review of the literature on control and accountability. *Handbooks of Management Accounting Research*, 2:785–802, 2006.

K.A. Merchant and W.A. Van der Stede. *Management control systems: performance measurement, evaluation and incentives*. Prentice Hall, 2007.

Daniele Nosenzo, Theo Offerman, Martin Sefton, and Ailko van der Veen. Encouraging compliance: bonuses versus fines in inspection games. *Journal of Law, Economics, and Organization*, 2013.

William Ocasio. Towards an attention-based view of the firm. *Strategic management journal*, 18(S1):187–206, 1997.

Heiko Rauhut. Higher punishment, less control? experimental evidence on the inspection game. *Rationality and Society*, 21(3):359–392, 2009.

Reinhard Selten and Thorsten Chmura. Stationary concepts for experimental 2x2-games. *The American Economic Review*, 98(3):938–966, 2008.

R. Simons. *Levers of control: how managers use innovative control systems to drive strategic renewal*. Harvard Business Press, 1994.

G.B. Sprinkle. Perspectives on experimental research in managerial accounting. *Accounting, Organizations and Society*, 28(2):287–318, 2003.

S.K. Widener. An empirical analysis of the levers of control framework. *Accounting, Organizations and Society*, 32(7):757–788, 2007.