

# Information accuracy in legislative oversight: Theoretical implications and experimental evidence

Rationality and Society  
2017, Vol. 29(2) 226–254  
© The Author(s) 2017  
Reprints and permissions:  
[sagepub.co.uk/journalsPermissions.nav](http://sagepub.co.uk/journalsPermissions.nav)  
DOI: [10.1177/1043463117700610](https://doi.org/10.1177/1043463117700610)  
[journals.sagepub.com/home/rss](http://journals.sagepub.com/home/rss)



**Susumu Shikano**

University of Konstanz, Germany

**Michael F Stoffel**

University of Konstanz, Germany

**Markus Tepe**

University of Oldenburg, Germany

## Abstract

The relationship between legislatures and bureaucracies is typically modeled as a principal–agent game. Legislators can acquire information about the (non-) compliance of bureaucrats at some specific cost. Previous studies consider the information from oversight to be perfect, which contradicts most real-world applications. We therefore provide a model that includes random noise as part of the information. The quality of provided goods usually increases with information accuracy while simultaneously requiring less oversight. However, bureaucrats never provide high quality if information accuracy is below a specific threshold. We assess the empirical validity of our predictions in a lab experiment. Our data show that information accuracy is indeed an important determinant of both legislator and bureaucrat decision-making.

---

## Corresponding author:

Susumu Shikano, Department of Politics and Public Administration, University of Konstanz, Postfach 92, Universitaetsstrasse 10, 78457 Konstanz, Germany.  
Email: [susumu.shikano@uni-konstanz.de](mailto:susumu.shikano@uni-konstanz.de)

## Keywords

Inspection game, laboratory experiment, legislative oversight, noisy information, principal agent

## Introduction

The bureaucracy plays a decisive role in every political system. Bureaucrats prepare legislation and implement the outcomes of the legislative process. Yet, since the seminal work of Weber (1958), the relationship between politicians and bureaucrats has been characterized as asymmetric. Bureaucrats have expertise in specific policy fields, know the exact proceedings of individual cases, and have executive instruments at their command (cf. Miller, 2005: 203). Hence, bureaucrats can exploit their informational advantage to their own benefit. To cope with this situation, legislators and the executive's political leaders in their role as principals use a variety of tools to come to grips with their agents.

Positive political theory has provided various insights into the workings of public administration and how politicians deal with it (see Gailmard and Patty, 2012 for an extensive overview). Legislators not only use oversight to control bureaucratic behavior but also employ ex post sanctions to create ex ante incentives for bureaucrats to comply with their principal (Miller, 2005: 209). Thus, if legislators do not spend much resource for monitoring bureaucrats, this could mean that there is an effective incentive system in place that induces bureaucrats to take actions conforming to the legislator's goals (Weingast and Moran, 1983: 768).

Yet, even though monitoring in all of its forms has been subject to investigation, we still know little about the outcome from the monitoring process: the information and its consequences. More specifically, does the quality or accuracy of information affect the way bureaucrats act and, if so, how? This study tackles this question, focusing on a specific type of legislative oversight, the so-called "police-patrol" (McCubbins and Schwartz, 1984: 166), that is, direct monitoring of the bureaucracy through the legislators. We base our results on a dual way of investigation. We first determine the consequences of different levels of information accuracy on bureaucratic slack in a mathematical model. Then, we assess whether our predictions are in accord with human behavior in a laboratory experiment.

We model the strategic interaction of bureaucrats and legislators by extending an inspection game (Dresher, 1962; Rauhut, 2009; Rauhut and Winter, 2012; Tsebelis, 1990).<sup>1</sup> In our game, a legislator can choose to invest into oversight and, in return, gets informed about the behavior of the inspected bureaucrat. As no observation technique is free from human or technical failure and bureaucratic implementation can have far-reaching impact on society with

uncertainty, we deviate from previous deterministic models and include random noise. For example, the accuracy of information from oversight can be affected by technical issues in evaluating the quality of a certain good or service produced or delivered by the bureaucrat (e.g. legal expertise) as well as the rigor of an observation (e.g. investigating only a sample of deliveries).

We find that the willingness of bureaucrats to provide high-quality service weakly increases with information quality. Moreover, also the size of a possible punishment improves the quality of bureaucratic service, whereas the latter decreases with the costs of observation for the legislator. We test these propositions in a series of laboratory experiments. The results from these experiments by-and-large support the findings from the theoretical model on the role of information accuracy, both at the aggregate and at the micro-level. The remainder of this article proceeds as follows. The next section provides a short review of the literature on legislative oversight and argues why it is important to include noisy information into the inspection game. The section “Model” introduces and solves a full and a noisy information model of legislative oversight. Subsequently, we present our experimental setting and provide the results. In the final section, we summarize our findings and discuss their theoretical and practical implications.

## **Literature overview**

### *Legislative oversight*

As in every principal–agent relationship, bureaucrats and legislators can differ in terms of their preferences, which beg the problem for the principal that the agent pursues goals that differ from her own. The agent, moreover, has an informational advantage that she can use to her own benefit. Being aware of this situation, legislators can employ different means to keep the bureaucracy under their control. At the most general level, one can distinguish between two strategies: statutory control and legislative oversight (Bawn, 1997: 102). Statutory control refers to the fact that legislators deliberately design bureaucratic structures and processes *ex ante* in a way that favors some goals or policies over others. Legislative oversight, in contrast, is an *ex post* method to control the task fulfillment of bureaucrats. In this case, politicians directly monitor bureaucratic output or processes to gain the information necessary for deciding whether to punish and/or correct undesirable behavior. Depending on the circumstances, a vote-seeking politician will choose one strategy or the other.

If a legislator decides to rely on oversight in controlling the bureaucracy, she may use “police-patrol” or “fire alarm” forms of oversight (McCubbins and Schwartz, 1984: 168). Under police-patrol oversight,

politicians themselves selectively examine bureaucrats to detect and remedy any deviating behavior (McCubbins and Schwartz, 1984: 161). In contrast, fire alarm oversight is ad hoc and depends crucially on third party actors monitoring bureaucratic performance. This study focuses exclusively on police-patrol oversight that takes place between a single politician and a single bureaucrat.

Police-patrol oversight in the legislative context has been modeled in the *oversight game* (Gailmard, 2009; Maor, 1999; Shapiro, 1994; Whitford, 2008), which is also referred to as the *law enforcement game* (Scholz, 1991, 1984, 1986). In its basic form, it describes the following situation. A bureaucrat-type agent has to choose a regulatory enforcement level or level of quality at which she provides public goods and services: high-quality provision is costly and low-quality provision is free. The legislative principal has to decide simultaneously whether she wants to control the bureaucrat's behavior or not. As both actors do not know the choice of their opponent when making their own decision, the characteristic feature of this game is imperfect information (McCarty and Meirowitz, 2007: 148).

In the other fields of social sciences, there exists an analogous decision-making structure, the so-called *inspection game* (Dresher, 1962).<sup>2</sup> In its general form, this game is a mathematical model of a situation where an inspection authority, called inspector, verifies that another party, called inspectee, adheres to certain legal rules (Avenhaus, 2004: 179). For example, social researchers used this model to explore whether more severe penalties are helpful in reducing the crime rate in a society (e.g. Rauhut, 2009; Rauhut and Junker, 2009; Tsebelis, 1990). In this context, the game consists of police officers (inspectors) and criminals (inspectees). Their goals and choice sets are similar to those of politicians and bureaucrats in the oversight situation. The conceptual proximity between the inspection game and police-patrol oversight becomes evident in the original definition of police-patrol oversight by McCubbins and Schwartz (1984: 166), which states that "analogous to the use of real police patrols, police-patrol oversight is comparatively centralized, active, and direct." Thus, our theoretical model in the next section relies on this decision-making structure.

### *Legislative oversight with noisy information*

Even though the acquisition of information about agents is at the heart of both oversight and inspection games, the information itself and, more important, its quality have not been subject to scholarly interest.<sup>3</sup> Instead, most studies assume information to be either perfect (after oversight) or completely hidden (no oversight). However, since no observation technique is free from human or technical failure, it is more realistic to assume that the

information an inspector obtains from controlling an inspectee's behavior involves random noise. In case of the original inspection game situation, it may be hard to prove that inspected citizens have committed a crime (Rizzolli and Stanca, 2012) or to detect doping substances in inspected athletes (Kirstein, 2014). In the case of legislative oversight, the outcome of implementation by bureaucrats can have more far-reaching impact in society, which leads to further uncertainty. That is, even after monitoring, one can never be perfectly informed about the consequences of the implementation. This kind of uncertainty for political actors is analogous to that of the existing formal models of the relationship between political actors and bureaucrats (e.g. Epstein and O'Halloran, 1994). Thus, we extend the "police-patrol" oversight by introducing noisy information, which legislators obtain from observing bureaucratic behavior.

In the context of legislative oversight, noisy information needs to be distinguished from intentionally manipulated information. The manipulation of information—in particular by bureaucrat-type actors—has been at the center of expertise and agenda control experiments conducted by Eavey and Miller (1984) and Altfeld and Miller (1984). Another type of information manipulation in an oversight setting has been explored by Lupia and McCubbins (1998). They consider situations where principals cannot directly oversee the behavior of agents but have to rely on third-party assessments. The third parties, in their case interest groups, pursue their own goals when sending signals about bureaucratic behavior. To punish a shirking agent, the principal thus needs to assess the quality of the signal given by the sender about the quality of the agents' behavior.

In this study, we consider noise as lack in the accuracy of information that the legislator obtains from inspecting the bureaucrat's behavior. This notion of noisy information refers to the limited capacities and imperfect techniques that legislators have at their disposal to observe the behavior of bureaucrats. In particular, in cases where policy measures require time to take effect, information about bureaucratic behavior comes with a considerable amount of uncertainty. The degree of uncertainty is given by a parameter.

Taking the perspective of a policymaker who is confronted with scarce fiscal resources, the question is which policy measure is most effective in making bureaucrats' actions consistent with her goal. There are three policy options at hand: (1) reducing the control costs for the legislator, (2) increasing the punishment fee for the bureaucrat, (3) investing into inspection techniques to increase the accuracy of information obtained from observing the bureaucrats' behavior. Which of these options is most likely to make the bureaucrats' actions consistent with the legislator's goal? The next section provides a game-theoretical answer to this question.

## Model

### *Oversight game with full information*

There are two players, a bureaucrat  $B$  and a legislator  $L$ , who interact in the sequential game  $\Gamma$ . The bureaucrat can decide to deliver service at high or low quality,  $A_B \in \{H, -H\}$ . Without knowing the action taken by the bureaucrat, the legislator then chooses whether or not to observe the bureaucrat's behavior  $A_O \in \{O, -O\}$ . If she chooses to observe, she receives a signal  $i \in \{i_H, i_{-H}\}$  that indicates whether or not the bureaucrat has performed high. The signal fully reveals the strategy taken by the bureaucrat, and the legislator can now punish the bureaucrat if she desires,  $A_P \in \{P, -P\}$ . The set of strategies of the legislator is given by  $S_L = A_O \times A_P$ . For convenience, we summarize all strategy combinations in which the legislator chooses not to observe the action of the bureaucrat under the label  $-O$  since these are payoff equivalent. Moreover, as the remaining strategies of the legislator require her to observe the behavior of the bureaucrat, we drop the  $O$  when referring to these strategies and focus on the punish/not punish decision. Hence  $S_L = \{-O; P|i_H, P|i_{-H}; P|i_H, -P|i_{-H}; -P|i_H, P|i_{-H}; -P|i_H, -P|i_{-H}\}$

The payoff structure of the players is as follows. The bureaucrat has an initial benefit  $b$ . If she moves to deliver a high-quality good, she has to invest effort  $e$ . In case that the legislator decides to punish her, she has to pay the corresponding punishment fee  $f$ , regardless of whether she has provided a high- or low-quality good. The legislator on her side wants the bureaucrat to deliver a high-quality good, the case in which she earns a policy rent of size  $b$ , for example, in terms of votes in the election. If she chooses to observe the bureaucrat's behavior, the legislator has to bear the costs  $c$  of monitoring. Furthermore, the legislator may gain or lose in terms of popularity when dealing with the bureaucracy. She appears as being incompetent in the eyes of the electorate and loses popularity  $s$  if she falsely punishes a bureaucrat who delivers a high-quality good or does not punish a bureaucrat who delivers a low-quality good. In contrast, she gives proof of her leadership abilities and gains popularity  $s$  if she punishes a slacking bureaucrat. The entire game  $\Gamma$  is visualized in Figure 1.

We make two assumptions. First, we require that the punishment fee for the bureaucrat be greater than the effort necessary for producing a high-quality good, that is,  $f > e$  (A1, effective punishment). Second, we require that voters be sufficiently responsive to the way in which the legislator deals with the bureaucracy. Stated differently, the gains for the legislator from good oversight are sufficiently high to outweigh her costs of observation, that is,  $2s > c$  (A2, responsive electorate).

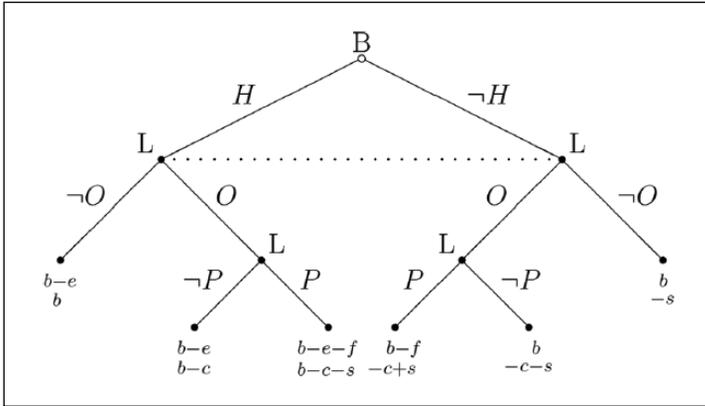


Figure 1. Full information game  $\Gamma$ .

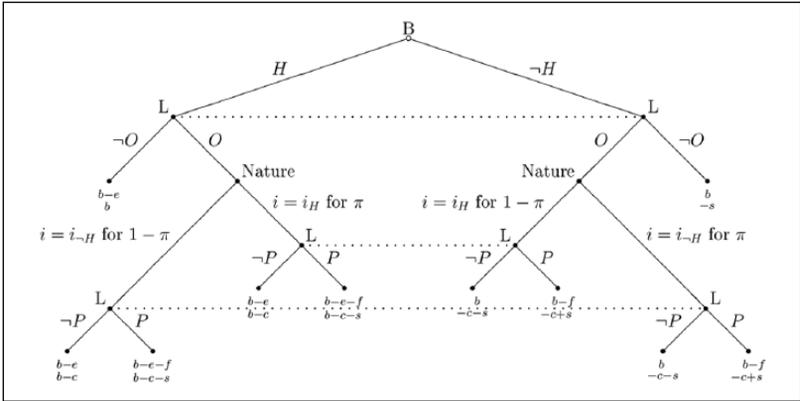
The unique Nash equilibrium of this game is in mixed strategies, where the bureaucrat delivers high effort with probability  $(2s - c) / 2s$ . To hold the bureaucrat indifferent, the legislator mixes over her pure strategies  $\neg O$  with probability  $(f - e) / f$  and  $\neg P|_{i_H}, P|_{i_{\neg H}}$  with probability  $e / f$

$$Pr(H) = 1 - \frac{c}{2s}; \quad Pr(O) = \frac{e}{f} \tag{1}$$

With regard to comparative statics, this implies that higher the control costs for the legislator are compared to the popularity gains from good oversight, the lower is her incentive to provide oversight. In turn, the lower is the incentive of the bureaucrat to deliver high-quality service. The legislator's level of observation increases in the ratio of the bureaucrat's effort for high-quality service over the punishment fee. If this ratio is low, punishment is not deterrent for the bureaucrat and more oversight is necessary.

### Oversight game with noisy information

With the baseline estimate at hand, we assess whether and how an imprecise signal affects the equilibrium outcome. The setup of the noisy information game  $\Gamma_n$  is similar to the full information game  $\Gamma$  (see Figure 2). Random noise is introduced into the game tree through a move of nature after the legislator's decision to observe the bureaucrat's behavior. The accuracy of signal  $I$  is common knowledge and denoted by  $\pi$ . For example, if the level of accuracy is  $\pi = 0.8$  and the signal is  $i = i_H$ , then the bureaucrat has



**Figure 2.** Noisy information game  $\Gamma_n$ .

delivered high-quality service in 8 out of 10 cases and low-quality service in 2 out of 10 cases. As in the full information game, the legislator can decide whether to punish the bureaucrat or not after receiving the signal.

Regarding the information accuracy, we add a third assumption to those already given in  $\Gamma$ . We require that the accuracy of the signal received from observing the bureaucrat’s behavior  $\pi \in (1/2, 1]$ . That is, the signal is strictly more informative than a random toss (A3, informative signal).<sup>4</sup> Please note that the noisy information game  $\Gamma_n$  reduces to the full information game  $\Gamma$  for  $\pi = 1$ .

The number and types of equilibria resulting from the noisy information game are dependent upon the accuracy of the signal. Thus, we distinguish between three scenarios in the following: a signal with high accuracy ( $\pi > 2/3$ ), a signal with low accuracy ( $\pi < 2/3$ ), and the borderline between the two ( $\pi = 2/3$ ). The derivation of the results appears in “Decision to punish and information accuracy” of Appendix 1.

If information accuracy is high ( $\pi > 2/3$ ), the equilibrium is similar to the full information case. The unique Nash equilibrium of the game is in mixed strategies. To hold the bureaucrat indifferent, the legislator mixes over her strategies  $\neg P|i_H, P|i_{\neg H}$ . Similarly, the bureaucrat mixes between delivering high- and a low-quality service, where

$$\Pr(H) = \frac{2\pi s - c}{(1 + \pi)s}; \quad \Pr(O) = \frac{e}{(2\pi - 1)f} \tag{2}$$

All else being equal, the level of oversight decreases with information accuracy, whereas the level of high-quality service increases with

information accuracy in expectation. Moreover, the level of oversight increases in the ratio  $e/f$ , whereas the level of high-quality service decreases in the ratio  $c/s$ . The first ratio indicates that the bureaucrat is more likely to shirk if the punishment fee is not high enough when compared to the effort of producing high-quality service. The second ratio shows that the legislator is not motivated to provide extensive oversight if the electorate is not responsive enough.

If information accuracy is low ( $\pi < 2/3$ ), the equilibrium is entirely different. The signal is too imprecise for the legislator to make her decision based on it. Thus, she would rather like not to observe the bureaucrat's behavior. However, this induces the bureaucrat to never provide high-quality service. In turn, the legislator always observes and strictly punishes the bureaucrat, no matter what the signal  $i$  is ( $P|i_H, P|i_{-H}$ ). Still, the bureaucrat has no incentive to provide high-quality service, and both are stuck in a race to the bottom.

$$Pr(H) = 0; \quad Pr(O) = 1 \quad (3)$$

Further changes of information accuracy or the other variables do not affect the equilibrium level of high-quality service and oversight. Finally, if information accuracy is on the borderline ( $\pi = 2/3$ ), both equilibria exist. The comparative statics depend on the particular equilibrium.

## Experimental setup

We put the predictions of our oversight game with noisy information to a test in a controlled laboratory experiment. Despite the impressive amount of formal research on the relationship between political and bureaucratic actors, there are only a handful studies which make use of experimental designs for testing the empirical validity of those models (Margetts, 2011: 190). This is all the more surprising given the impact that laboratory experiments have had on political research in recent times (Druckman et al., 2006; Webster and Sell, 2007).

There are good reasons for being skeptical about the predictions of game-theoretic models for legislative-bureaucracy interactions and putting them to a test. Even though the concept of mixed strategies is at the heart of formal modeling, its empirical relevance has repeatedly been met with skepticism (Chiappori et al., 2002: 1138). It is not clear whether and, if so, how individuals actually come to generate a lottery over their pure strategies to make their opponent indifferent. Evidence from laboratory experiments on this issue is rather mixed (e.g. Brown and Rosenthal, 1990; McCabe et al., 2000; Ochs, 1995; O'Neill, 1987). Moreover, the study of

**Table 1.** Parameterization and mixed strategy solutions.

	Low cost	Low cost	High cost	High cost	Noisy	Accurate
	Low fee	High fee	Low fee	High fee		
$b$	10	10	10	10	10	10
$e$	5	5	5	5	5	5
$s$	5	5	5	5	5	5
$c$	3	3	7	7	7	7
$f$	10	20	10	20	20	20
$\pi$	0.8	0.8	0.8	0.8	0.6	0.9
$Pr(H)$	56%	56%	11%	11%	0%	21%
$Pr(O)$	83%	42%	83%	42%	100%	31%
$n$	40	36	40	40	40	40

Bloomfield (1994) indicates that subjects might use strategies other than the Nash solution to determine their optimal choice.

### Parameterization

In our experimental application, we compare a variety of parameter combinations for the costs of observation  $c$ , the punishment fee  $f$ , and the accuracy of information  $\pi$  in six different treatments. In the first four treatments, we combine low and high observation costs ( $c = 3$  or  $c = 7$ ) with a low and high punishment fee ( $f = 10$  or  $f = 20$ ). In all of these treatments, we set  $\pi = 0.8$ . In the fifth and sixth treatment, we fix observation costs and punishment fee at their higher level but alter the accuracy of information  $\pi$ . That is,  $\pi$  is set to 0.6 in the fifth and to 0.9 in the sixth treatment. Table 1 presents an overview over the different treatments. It also includes the Nash predictions for the probability that the bureaucrat delivers high-quality service,  $Pr(H)$ , and the probability that the legislator investigates the bureaucrat's behavior,  $Pr(O)$ .

We chose to vary the parameters above because it is rather easy to change them in real-world applications. Instead, the remaining parameters are tied to specific institutional designs or inherent to the actors, which is why these are held constant across treatments. More precisely, both bureaucrat and legislator benefits  $b$  are set to 10 across the treatments. The effort  $e$  that the bureaucrat has to invest to provide a high-quality good is set to 5, as is the popularity penalty or bonus  $s$  applying to the legislator for (not) detecting a bureaucrat that provides low quality.

### Procedures

Subjects were given instructions about the rules of the game before the experiment. These were provided as a handout and read to the experimentees

by the instructor (see section “Information provided to participants” in ). All subjects were informed about their possible earnings, which comprised a €3 show-up fee and the sum of all tokens gained from playing 30 rounds of the oversight game. The exchange rate to convert tokens into Euros was 0.01. The average earning plus show up-fee was €9.82. An average session took about 45 minutes including instructions and disbursement. We conducted a total of 12 sessions, two for each treatment and each time inviting 20 subjects. Since four subjects did not attend a session for the low cost and high fee treatment, we have 236 subjects in total, half of which were bureaucrats and legislators, respectively. Most subjects were students from the University of Oldenburg, who had been invited via the Online Recruitment System for Economic Experiments (ORSEE) by Greiner (2004). The experiment was conducted at the lab of the University of Oldenburg using the computer program z-Tree (Fischbacher, 2007).

At the beginning of the game, subjects were randomly assigned to play either the role of a legislator or the role of a bureaucrat and then kept their role throughout the entire experiment. Subjects were randomly paired in each round, not receiving any identifying information about their opponent. Moreover, we used a neutral framing, referring to player *A* and *B* rather than “legislator” and “bureaucrat” to preclude any confounding connotation. To give subjects an opportunity to learn the structure of the game and the consequences of their decisions for their payoff, experimentees played three training rounds with no effect on their eventual payoff. The play of the game proceeded according to the game-theoretical model in the previous section. After each round, both players were informed about their own actions, their partner’s actions, and the corresponding payoffs.

## Results

In this section, we present our experimental results about bureaucrat decision-making. We begin with comparing the aggregate-level results with the predictions of our theoretical model. We then investigate bureaucrat and legislator decision-making at the individual-level.

### *Aggregate-level results*

Table 2 presents the aggregate-level results for how bureaucrats and legislators react to the different cost-fee treatments, the parameters of the conventional oversight game. We begin with inspecting bureaucrat decision-making. According to the theoretical model, bureaucrats should provide high-quality service more often under the low cost than under the high cost treatments, while changes in the punishment fee should have no effect. The empirical

**Table 2.** Aggregate-level results of different cost-fee treatments.

Bureaucrats' probability to provide high-quality service			
	Control cost		
	High (11%)	Low (56%)	
<i>Punishment fee</i>			
High	42%	<	51%
	∨		∥
Low	29%	<	48%
Legislators' probability to observe the bureaucrat's behavior			
	Control cost		
	High	Low	
<i>Punishment fee</i>			
High (42%)	43%	<	67%
	∧		∥
Low (83%)	55%	<	70%

≈ means no significant difference at  $\alpha = 5\%$ .

Predicted value from the theoretical model is given in parentheses.

results are in line with the prediction concerning the cost-treatment effect. In terms of the punishment fee, however, the treatment has a significant effect if the legislator's control cost is high. In particular, the probability for high-quality performance in the high punishment treatment is much higher (42%) than predicted (11%). One possible explanation is that experimentees needed more periods to adjust their behavior. This high cost and high fee treatment is the only treatment in which bureaucrat behavior shows a certain time trend: The subjects decided to provide high-quality performance much more frequently in the first 10 periods (46%) than in the last 10 periods (36%).<sup>5</sup> While the punishment fee still has a significant effect in the analysis including only the last 10 periods, the corresponding t-value is much smaller than in the analysis including all periods. We may speculate that the effect can be insignificant after repeating more periods.

In contrast to bureaucrats, legislators apparently deviated from the predictions. According to the model, legislators should control bureaucrats more frequently in the low than in the high punishment fee treatments, while the control costs should have no effect. However, legislators overall seem to have reacted to their costs of control. Furthermore, among the low cost treatments, the punishment fee did not lead to significant differences in legislators' decisions. This unexpected pattern also remains even if one considers only the last 10 periods.

**Table 3.** Aggregate-level results of different information accuracy treatments.

		$\pi = 60\%$		$\pi = 80\%$		$\pi = 90\%$	
B's $Pr(H)$	Forecast	0%		11%		21%	
	Results	21%	<	42%	<	52%	
L's $Pr(O)$	Forecast	100%		42%		31%	
	Results	62%	>	43%	$\approx$	38%	

$\approx$  means no significant difference at  $\alpha = 5\%$ .

We will now turn to the effect of information accuracy (Table 3). Our theoretical model predicts that bureaucrats provide high-quality service more often if the information from inspection is more accurate. Our results are in line with this prediction, with the caveat that bureaucrats provided high-quality service much more frequently than predicted in general. Concerning the legislators, their tendency to observe the bureaucrat's behavior should decrease as the information becomes more accurate. Even though there is no significant difference between the  $\pi = 80\%$  and  $\pi = 90\%$  treatment, the relative relationship supports the theoretical model.<sup>6</sup>

In the analysis of the legislator's decision thus far, we have only been interested in whether or not to observe the bureaucrat. Now, we assess whether there are differences across treatments in the way legislators react to the information they obtain from oversight. From a theoretical point of view, the effect of information accuracy on the decision of the legislator whether or not to punish the bureaucrat stems from her own expected benefit. In particular, under the noisy information treatment ( $\pi = 60\%$ ), she is predicted to punish the bureaucrat irrespective of the information obtained from control. Under the medium ( $\pi = 80\%$ ) and high ( $\pi = 90\%$ ) information treatment, she should punish according to the signal. To test this, we compare the probability of punishment dependent on the information content (Table 4). In line with the prediction, legislators punish bureaucrats much more frequently under the noisy information treatment, even if the signal indicates high-quality service from the bureaucrat (58%). In contrast, there is no significant difference among the other treatments as expected (with the exception of the pair of the low cost and low fee and high cost and low fee treatments).

To summarize, the aggregate-level results of bureaucrat decision-making mostly confirm the directional predictions from our game-theoretic model concerning the role of information accuracy. That is, higher information accuracy of oversight makes bureaucrats produce high-quality products and services. Legislator decision-making is also in line with the theoretical prediction in terms of information accuracy. Particularly, legislators strongly tend to punish at low-levels of information accuracy, irrespective of whether or not

**Table 4.** Information content and punishment decision.

	Low cost	Low cost	High cost	High cost	Noisy	Accurate
	Low fee	High fee	Low fee	High fee		
Punishment in case of information for high quality						
Forecast	0%	0%	0%	0%	100%	0%
Result	10%	17%	26%	16%	58%	24%
Punishment in case of information for low quality						
Forecast	100%	100%	100%	100%	100%	100%
Result	97%	96%	98%	98%	95%	90%

the signal from oversight indicates good quality service by the bureaucrat. In terms of the conventional oversight model, the results are more mixed, since not only lower control cost but also a higher punishment fee make bureaucrats choose high effort. With regard to legislators, a higher punishment fee makes them choose less oversight but also higher control costs.

### *Individual-level results*

To refine and further substantiate our results, we investigate bureaucrat and legislator behavior at the individual decision level. Since persons can have different and inherent levels of cooperation, we estimate multi-level logistic regression models with random intercepts at the subject level (Shikano et al., 2012). Note that standard errors and t-statistics for the treatment effects are therefore not based on the total number of observed decisions, but on the number of individuals in the experiment (118 bureaucrats/legislators) because treatments do not vary within subjects (cf. Snijders, 2005). Since only 20 bureaucrats/legislators were assigned to each treatment, we regard the significance level of 10% as solid evidence for an effect.

Beyond the treatments which exogenously set up based on our theoretical model, it is reasonable to assume that each individual decision has affected further factors which are endogenous in the experimental process. As such variables, we added the own lagged decision and the partner's lagged decision. For the bureaucrat's decisions, we also considered the interaction effect whether the high-quality performance was punished in the last period. It is reasonable to expect its strong disappointing effect on the bureaucrat. For the legislator decision on punishment, we added two variables concerning the information which the legislator obtains in case of observation. One is whether the information signals a high-quality performance. The other is the interaction of the last variable with the noisy information treatment. They are based on the theoretical predictions that

**Table 5.** Multi-level logit models of bureaucrat and legislator decision-making.

	Bureaucrat	Legislator	
	High performance	Observe	Punish
	(1)	(2)	(3)
<i>Treatment-level effects</i>			
Low fee	-0.72 (0.50)	0.68* (0.39)	0.72 (0.52)
Low cost	0.26 (0.48)	1.41*** (0.38)	0.13 (0.50)
Low cost × low fee	0.70 (0.69)	-0.59 (0.54)	-0.59 (0.70)
Low information accuracy ( $\pi=60\%$ )	-1.16** (0.49)	0.85** (0.37)	0.09 (0.60)
High information accuracy ( $\pi=90\%$ )	0.49 (0.48)	-0.15 (0.38)	0.44 (0.53)
<i>Individual situation-level effects</i>			
Signal $i=i_H$			-5.62*** (0.27)
Signal $i=i_H \times$ low information accuracy			2.46*** (0.48)
$t - I$ observe	0.22* (0.13)	0.41*** (0.09)	
$t - I$ high performance	1.51*** (0.12)	-0.46*** (0.08)	
$t - I$ punish	0.87*** (0.16)		
$t - I$ high performance × punish	-1.94*** (0.20)		
Intercept	-1.39*** (0.35)	-0.41 (0.27)	3.46*** (0.41)
<i>Random effects</i>			
Individuals (standard deviation)	1.42	1.10	1.16
Decisions	3422	3422	1919
Individuals	118	118	116

Standard errors in parentheses.

\* $p < 0.10$ ; \*\*  $p < 0.05$ ; and \*\*\*  $p < 0.01$ .

legislator responds to the information content only in the case that the information level is high enough (Table 5).

We begin with the decision of the bureaucrat whether to provide high- or low-quality performance (1 = high quality, 0 = low quality; Model 1). In line with the prediction, the low fee treatment does not influence bureaucrats' tension to provide high quality service. Yet, neither do the legislator's costs

of observation, which is against the prediction. Finally, the coefficients for the information treatments are in the expected direction where the low information accuracy treatment is significant with a large effect size. In terms of the high information accuracy, our analysis may be underpowered for the effect of 10%-point difference ( $\pi = 80\%$  and  $90\%$ ).

The insignificant effect of the low cost treatment may seem to be puzzling if one considers the aggregate-level results. It is because the effect in Model 1 is absorbed by the endogenous factors. In particular, the presence of legislative oversight does influence bureaucrats' behavior. If a bureaucrat has been observed the last time through a legislator, she is more likely to perform high this time. This effect exists even though subjects are randomly matched each round and the probability of a bureaucrat for being matched to exactly the same legislator as before is only 10%. The legislator's decision on observation, in turn, is influenced by its control cost as in Model 2. Therefore, bureaucrats seem to learn to respond to the cost-treatments through the legislative oversight in an adaptive way. This is a different mechanism from the prospective adjustment of strategies assumed in the Nash solution; however, it has similar consequences.

Besides the presence of legislative oversight, bureaucrats also react to a punishment for slack in the previous round in the same way. However, the most drastic result can be found in the interaction effect at the bottom of the table. If a bureaucrat has been punished in the last round even though she was cooperative, her willingness to cooperate in the present round drops massively. This has clear implications for real-life delegation relationships: When supervising the bureaucracy, the political leadership should really be sure that misbehavior has occurred because being falsely punished absolutely demotivates the agent. As we can see below in Model 3, such situation is more likely if the information level is quite low where the legislator tends to punish the bureaucrat even if the information signals a high-quality performance.

We now turn to the decision-making of the legislators and begin with their choice whether or not to observe the bureaucrat (Model 2). We see that legislators adjust their behavior to the size of the punishment for the bureaucrat. When the fee is low, they observe the bureaucrat more often as they know that the latter will exploit the situation otherwise. Against the expectation, the own costs of observation do play a role: higher costs deter legislators from providing oversight. The results with regards to the information accuracy are similar to those in Model 1. In line with the theoretical model, legislators control bureaucrats more often if accuracy is low. However, there is no difference for high-levels of information accuracy, even though the coefficient is in the expected direction.

The two added endogenous lags are also significant and allow us an intuitive interpretation. First, legislators have a different tendency to control the bureaucrat, which is why legislators who provided oversight in the previous

round are more likely to provide oversight in the present round. Second, legislators observe the bureaucrat less often if the bureaucrat in the previous period provided performance at a high-level.

Finally, we provide results for the legislators' decision to punish the bureaucrat (Model 3). The findings are in line with the predictions of our theoretical model. That is, the strongest determinant of a legislator's decision to punish is whether or not the signal indicates high performance. However, the effect of the signal is less pronounced at low-levels of information accuracy. That is, legislators sometimes punish bureaucrats despite a signal indicating high performance as the interaction effect shows.

To conclude, the individual-level results above give additional support for the role of information accuracy in our theoretical model. In particular, the low-level of information accuracy can have wide reaching effects on both bureaucrat and legislator's behavior. In such situations, the legislator tends to observe and punish the bureaucrat. Knowing about the legislator's likely behavior, the bureaucrat provides only low-level service. Even well-intended bureaucrats may be strongly disappointed after getting punished for their high-level performance.

## **Discussion**

This article investigates an important aspect of principal-agent relationships between legislators and the bureaucracy that has not received scholarly attention before: the accuracy of information gained from oversight. Our game-theoretical model shows that information accuracy not only affects the location but also the nature of the equilibrium in oversight games. Bureaucrats never provide high-quality service if information accuracy is below a certain threshold. Legislators, on their part, are aware of this fact and strictly control and punish bureaucrats. If information accuracy exceeds this threshold, bureaucrats take a mixed strategy in which the quality of their service increases with the accuracy of information and decreases with the control costs legislators have to bear from oversight. However, bureaucrats are not responsive to the extent of punishment for delivering low-quality services. Legislators exhibit less control if the punishment for bureaucratic slack is high and the information from oversight is very accurate, but do not consider their own costs of control when making their decision.

The theoretical predictions are quite robust to some reasonable changes in the model assumptions. For example, we can extend our model by including a positive reward to the legislator for non-punishment of a non-shirking bureaucrat. This reward does not exceed the reward for punishment of a shirking bureaucrat. A model with this modification generates different threshold values but identical results concerning the other parameters, particularly the

effect of information accuracy on bureaucrat/legislator decision-making. In this modified model, the threshold is lower ( $\pi = 2/3$ ) but still remains above 50%. The threshold reaches 50% and the low information equilibrium vanishes only if the reward for non-punishment of non-shirking is identical with the reward for punishment of shirking. This, however, appears to be an unrealistic assumption.

In a series of laboratory experiments, we assessed whether the behavior of individuals actually is in accordance with these predictions. Overall, we find strong support for the theoretical model, particular on the important role of information accuracy. The key findings can be summarized as follows. First, on the aggregate-level, experimental evidence shows that the behavior of bureaucrats and legislators follow the expected directions. Most important, the data show that increasing information accuracy makes it more likely that bureaucrats provide high-quality goods and legislators waive the control. Second, a micro-level level analysis confirms our findings.

Our results carry important implications for policy-makers in practice. To increase both effectiveness and efficiency of legislative oversight, they can pursue two strategies. First, they can reduce the costs of oversight, for example, through developing routines. Moreover, legislators can install experts in the corresponding policy field or can appoint former bureaucrats to controlling positions. Second, regardless of other aspects of an oversight situation, legislators can always raise the accuracy of information, thus, enabling them to engage themselves in oversight less frequently while having a higher level of compliance at the same time. In this way, legislators not only increase their own payoff but also the payoff of bureaucrats, since more accurate information makes the decisions of legislators more predictable and reliable for bureaucrats. In fact, policy-makers in Organization for Economic Co-operation and Development (OECD) countries have already invested time and money to establish a broad range of performance monitoring systems at various levels of bureaucratic activity (Organization for Economic Co-operation and Development (OECD), 2011). Such monitoring systems might be steps in the right direction if they are capable to improve the level of information accuracy in legislative oversight.

## Acknowledgements

Earlier versions of this paper were presented at the annual meetings of the Midwest and European Political Science Association in 2012. We thank Bernhard Kittel, Heiko Rauhut, and John Patty for helpful comments on an earlier version of this manuscript. We are indebted to Steffen Felix Bandlow, Karin Becker, Philipp Bosch, Amélie Dupuy-Seltmann, Jana Keller, and Verena Mack for their assistance in conducting experiments and preparing the manuscript. The first author also thanks the Hanse Institute for Advanced Study in Delmenhorst for supporting this project.

## Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by the German Science Foundation under grant #FP517/13.

## Notes

1. Oversight games have been applied to a variety of different contexts, such as arms control (Avenhaus, 2004), police control (Andreozzi, 2004), business administration (Fandel and Trockel, 2011), and smuggling (Hohzaki, 2011).
2. See Avenhaus et al. (2002) for an overview of applications and extensions of the inspection game.
3. An exception is Gailmard (2002) who investigates a specific form of noisy information. However, the signal in his model gives information about a random shock affecting policy outcomes and not about the behavior of bureaucrats. Moreover, there is research on the so-called “coarseness” of information in other fields of agency research, which has a slightly different focus (e.g. Dasgupta and Williams, 1995). Information is divided into a finite number of intervals, a higher number of partitions indicating “finer” information. In contrast, we are interested in how likely a signal gives the correct information.
4. The requirement that the signal be informative allows us to rule out implausible strategies of the legislator according to which she does not punish the bureaucrat after receiving the signal  $i = i_{-H}$ .
5. Aggregate-level results over periods appear in Appendix 3.
6. The difference between the medium and high information treatment becomes more pronounced if we consider only the intermediate and the ending phase of the experiment. However, the difference still fails to achieve conventional levels of statistical significance.

## References

- Altfeld MF and Miller GJ (1984) Sources of bureaucratic influence: expertise and agenda control. *Journal of Conflict Resolution* 28(4): 701–730.
- Andreozzi L (2004) Rewarding policemen increases crime. Another surprising result from the inspection game. *Public Choice* 121(1–2): 69–82.
- Avenhaus R (2004) Applications of inspection games. *Mathematical Modelling and Analysis* 9(3): 179–192.
- Avenhaus R, von Stengel B and Zamir S (2002) Inspection games. In: Aumann RJ and Hart S (eds) *Handbook of Game Theory*, vol. 3. Amsterdam: North-Holland, pp. 1947–1987.
- Bawn K (1997) Choosing strategies to control the bureaucracy: statutory constraints, oversight, and the committee system. *Journal of Law, Economics, and Organization* 31(1): 101–126.
- Bloomfield R (1994) Learning a mixed strategy equilibrium in the laboratory. *Journal of Economic Behavior and Organization* 25(3): 411–436.

- Brown JN and Rosenthal RW (1990) Testing the minimax hypothesis: a reexamination of O'Neill's game experiment. *Econometrica* 58(5): 1065–1081.
- Chiappori P-A, Levitt SD and Groseclose T (2002) Testing mixed strategy equilibria when players are heterogenous: the case of penalty kicks in soccer. *American Economic Review* 92(4): 1138–1151.
- Dasgupta S and Williams KC (1995) Search behavior of asymmetrically informed voters: an experimental study. *Economics & Politics* 7(1): 21–41.
- Dresher M (1962) *A Sampling Inspection Problem Arms Control Agreements: A Game-theoretic Analysis*. Santa Monica, CA: RAND Corporation.
- Druckman JN, Green DP, Kuklinski JH, et al. (2006) The growth and development of experimental research in political science. *American Political Science Review* 100(4): 627–636.
- Eavey CL and Miller GJ (1984) Bureaucratic agenda control: imposition or bargaining? *American Political Science Review* 78(3): 719–733.
- Epstein D and O'Halloran S (1994) Administrative procedures, information, and agency discretion. *American Journal of Political Science* 38: 697–722.
- Fandel G and Trockel J (2011) Optimal lot sizing in a non-cooperative material manager-controller game. *International Journal of Production Economics* 133(1): 256–261.
- Fischbacher U (2007) Ztree: Zurich toolbox for ready-made economic experiments. *Experimental Economics* 10(2): 334–352.
- Gailmard S (2002) Expertise, subversion, and bureaucratic discretion. *Journal of Law, Economics, and Organization* 18(2): 536–555.
- Gailmard S (2009) Multiple principals and oversight of bureaucratic policymaking. *Journal of Theoretical Politics* 21(2): 161–186.
- Gailmard S and Patty JW (2012) Formal models of bureaucracy. *Annual Review of Political Science* 15: 353–377.
- Greiner B (2004) An online recruitment system for economic experiments. In: Kremer K and Macho V (eds) *Forschung und wissenschaftliches Rechnen 2003: GWDG Bericht 63*. Goettingen: Goettingen University, pp. 79–93.
- Hohzaki R (2011) An inspection game with smuggler's decision on the amount of contraband. *Journal of the Operations Research Society of Japan* 54(1): 25–45.
- Kirstein R (2014) Doping, the inspection game, and Bayesian enforcement. *Journal of Sports Economics* 15: 385–409
- Lupia A and McCubbins MD (1998) *The Democratic Dilemma: Can Citizens Learn What They Need To Know?* New York: Cambridge University Press.
- McCabe K, Mukherji A and Runkle D (2000) An experimental study of information and mixed-strategy play in the three-person matching-pennies game. *Economic Theory* 15(2): 421–462.
- McCarty N and Meirowitz A (2007) *Political Game Theory: An Introduction*. New York: Cambridge University Press.
- McCubbins MD and Schwartz T (1984) Congressional oversight overlooked: police patrols versus fire alarms. *American Journal of Political Science* 28(1): 165–179.
- Maor M (1999) Electoral competition and the oversight game: a transaction cost approach and the Norwegian experience. *European Journal of Political Research* 35(3): 371–388.

- Margetts HZ (2011) Experiments for public management research. *Public Management Review* 13(2): 189–208.
- Miller GJ (2005) The political evolution of principal-agent models. *Annual Review of Political Science* 8: 203–225.
- Ochs J (1995) Games with unique, mixed strategy equilibria: an experimental study. *Games and Economic Behavior* 10(1): 202–217.
- O'Neill B (1987) Nonmetric test of the minimax theory of two-person zero-sum games. *Proceedings of the National Academy of Sciences of the United States of America* 84(7): 2106–2109.
- Organization for Economic Co-operation and Development (OECD) (2011) *Government at a Glance 2011*. Paris: OECD Press.
- Rauhut H (2009) Higher punishment, less control? *Rationality and Society* 21(3): 359–392.
- Rauhut H and Junker M (2009) Punishment deters crime because humans are bounded in their strategic decision-making. *Journal of Artificial Societies and Social Simulation* 12(3): 1–23.
- Rauhut H and Winter F (2012) On the validity of laboratory research in the political and social sciences: the example of crime and punishment. In: Kittel B, Luhau WJ and Morton RB (eds) *Experimental Political Science: Practice and Principles*. London: Palgrave Research Methods Series, pp. 209–232.
- Rizzolli M and Stanca L (2012) Judicial errors and crime deterrence: theory and experimental evidence. *Journal of Law and Economics* 55(2): 311–338.
- Scholz JT (1984) Voluntary compliance and regulatory enforcement. *Law & Policy* 6(4): 385–404.
- Scholz JT (1986) Regulatory enforcement in a federalist system. *American Political Science Review* 80(4): 1249–1270.
- Scholz JT (1991) Cooperative regulatory enforcement and the politics of administrative effectiveness. *American Political Science Review* 85(1): 115–136.
- Shapiro SA (1994) Political oversight and the deterioration of regulatory policy. *Administrative Law Review* 45(4): 1–40.
- Shikano S, Bräuningner T and Stoffel M (2012) Statistical analysis of experimental data. In: Kittel B, Luhau WJ and Morton RB (eds) *Experimental Political Science: Practice and Principles*. London: Palgrave Research Methods Series, pp. 163–177.
- Snijders TA (2005) Power and sample size in multilevel linear models. In: Everitt BS and Howell D (eds) *Encyclopedia of Statistics in Behavioral Science*, vol. 3. Chichester: Wiley, pp. 1570–1573.
- Tsebelis G (1990) Penalty has no impact on crime: a game theoretic analysis. *Rationality and Society* 2(3): 255–286.
- Weber M (1958) *From Max Weber: Essays in Sociology*. New York: Oxford University Press.
- Webster M and Sell J (2007) *Laboratory Experiments in the Social Sciences*. Burlington, MA: Academic Press.
- Weingast BR and Moran MJ (1983) The congressional-bureaucratic system: a principal-agent perspective. *Journal of Political Economy* 91(5): 765–800.
- Whitford AB (2008) A test of the political control of bureaucracies under asymmetric information. *Rationality and Society* 20(4): 445–470.

## Appendix I

### Derivation of equilibria

In this section, we derive the set of Nash equilibria inherent to the game. As a first step, we show that the legislator's strategy choice in terms of punishment depends on the information accuracy.

*Decision to punish and information accuracy.* The expected utility of punishment for the legislator (given that she has decided to observe the bureaucrat before) depends on the signal  $i \in \{i_H, i_{-H}\}$ .

(a)  $i = i_H$  (signal indicates high performance of bureaucrat)

The expected utility of punishing and not punishing is

$$EU(P|i_H) = \pi(b - c - s) + (1 - \pi)(-c + s) \quad (4)$$

$$EU(\neg P|i_H) = \pi(b - c) + (1 - \pi)(-c - s) \quad (5)$$

Thus

$$EU(P|i_H) < EU(\neg P|i_H) \Leftrightarrow \pi > \frac{2}{3} \quad (6)$$

$$EU(P|i_H) = EU(\neg P|i_H) \Leftrightarrow \pi = \frac{2}{3} \quad (7)$$

$$EU(P|i_H) > EU(\neg P|i_H) \Leftrightarrow \pi < \frac{2}{3} \quad (8)$$

(b)  $i = i_{-H}$  (signal indicates low performance of bureaucrat)

The expected utility of punishing and not punishing is

$$EU(P|i_{-H}) = (1 - \pi)(b - c - s) + \pi(-c + s) \quad (9)$$

$$EU(\neg P|i_{-H}) = (1 - \pi)(b - c) + \pi(-c - s) \quad (10)$$

Thus

$$EU(P|i_{-H}) > EU(\neg P|i_{-H}) \Leftrightarrow \pi > \frac{1}{3} \quad (11)$$

$$EU(P|i_{-H}) = EU(\neg P|i_{-H}) \Leftrightarrow \pi = \frac{1}{3} \quad (12)$$

$$EU(P|i_{-H}) < EU(\neg P|i_{-H}) \Leftrightarrow \pi < \frac{1}{3} \quad (13)$$

As we consider only situations in which the signal is informative, for example,  $\pi \in (1/2, 1]$ , it follows from equation (11) that the legislator will always punish given  $i = i_{-H}$ .

The results above imply different strategy choices depending on the information accuracy

$$EU(P|i_H, i_{-H}) < EU(\neg P|i_H, P|i_{-H}) \quad \text{if } \pi > \frac{2}{3} \quad (14)$$

$$EU(P|i_H, i_{-H}) = EU(\neg P|i_H, P|i_{-H}) \quad \text{if } \pi = \frac{2}{3} \quad (15)$$

$$EU(P|i_H, i_{-H}) > EU(\neg P|i_H, P|i_{-H}) \quad \text{if } \frac{1}{2} < \pi < \frac{2}{3} \quad (16)$$

Therefore, the legislator considers the signal and does not punish given  $i = i_H$  only if the information accuracy exceeds  $2/3$ . For levels of information accuracy below this threshold, she strictly prefers to punish the bureaucrat, no matter what the signal reveals. Finally, the legislator is indifferent between punishing and not punishing the bureaucrat given the signal  $i = i_H$  and information accuracy  $\pi = 2/3$ .

**Probability of punishment.** The probability of punishment depends on the accuracy of information and the decision taken by the bureaucrat.

(a)  $\pi > 2/3$  (information accuracy is high)

The probability that the legislator (mistakenly) punishes the bureaucrat given that the latter has performed high is given by

$$Pr(P|H) = Pr(P|H, i_H) + Pr(P|H, i_{-H}) \quad (17)$$

$$Pr(P | H) = Pr(P | i_H)Pr(i_H | H) + Pr(P | i_{-H})Pr(i_{-H} | H) \quad (18)$$

$$Pr(P | H) = 0Pr(i_H | H) + 1Pr(i_{-H} | H) \quad (19)$$

$$Pr(P | H) = (1 - \pi) \quad (20)$$

Analogously

$$Pr(P | \neg H) = \pi \quad (21)$$

(b)  $\pi < 2/3$  (information accuracy is low)

The probability that the legislator (mistakenly) punishes the bureaucrat given that the latter has performed high is given by

$$Pr(P | H) = Pr(P | i_H)Pr(i_H | H) + Pr(P | i_{-H})Pr(i_{-H} | H) \quad (22)$$

$$Pr(P | H) = Pr(i_H | H) + Pr(i_{-H} | H) = 1 \quad (23)$$

Analogously

$$Pr(P | \neg H) = 1 \quad (24)$$

(c)  $\pi = 2/3$  (borderline case)

In the borderline case, the legislator is indifferent between  $P | i_H, i_{-H}$  and  $\neg P | i_H, P | i_{-H}$ . With the first strategy, the probability of punishment  $Pr(P | H) = Pr(P | \neg H) = 1$ . With the latter strategy, equations (20) and (21) in combination with  $\pi = 2/3$  imply  $Pr(P | H) = 1/3$  and  $Pr(P | \neg H) = 2/3$ .

**Nash equilibria**

(a)  $\pi > 2/3$  (information accuracy is high)

The expected utilities of performing high or low for the bureaucrat are given by

$$EU(H) = Pr(\neg O)(b - e) + Pr(O)\{Pr(P | H)(b - e - f) + Pr(\neg P | H)(b - e)\} \quad (25)$$

$$EU(\neg H) = Pr(\neg O)b + Pr(O)\{Pr(\neg P | \neg H)b + Pr(P | \neg H)(b - f)\} \quad (26)$$

The bureaucrat performs high if and only if

$$EU(H) > EU(-H) \Leftrightarrow Pr(O) \{Pr(P|-H) - Pr(P|H)\} > \frac{e}{f} \quad (27)$$

By substituting equations (20) and (21), we obtain the best response function of the bureaucrat

$$Pr(H) = \begin{cases} 1 & \text{if } Pr(O) > \frac{e}{(2\pi-1)f} \\ (0,1) & \text{if } Pr(O) = \frac{e}{(2\pi-1)f} \\ 0 & \text{if } Pr(O) < \frac{e}{(2\pi-1)f} \end{cases}$$

For the legislator, the expected utility of not observing as well as of observing and punishing only if  $i = i_{-H}$  is given by

$$EU(O) = Pr(-H) \{Pr(P|-H)(-c+s) + Pr(-P|-H)(-c-s)\} + Pr(H) \{Pr(P|H)(b-c-s) + Pr(-P|H)(b-c)\} \quad (28)$$

$$EU(-O) = Pr(-H)(-s) + Pr(H)b \quad (29)$$

The legislator decides to observe the bureaucrat's behavior if and only if

$$EU(O) > EU(-O)$$

With some algebra and by substituting equations (20) and (21), we derive the legislator's best response function

$$Pr(O) = \begin{cases} 1 & \text{if } Pr(H) < \frac{2\pi s - c}{(1+\pi)s} \\ (0,1) & \text{if } Pr(H) = \frac{2\pi s - c}{(1+\pi)s} \\ 0 & \text{if } Pr(H) > \frac{2\pi s - c}{(1+\pi)s} \end{cases}$$

Comparing the best response functions, the unique Nash equilibrium is in mixed strategies with

$$Pr(O) = \frac{e}{(2\pi - 1)f} \quad (30)$$

$$Pr(H) = \frac{2\pi s - c}{(1 + \pi)s} \quad (31)$$

(b)  $\pi < 2/3$  (information accuracy is low)

We know from equation (16) that the legislator always plays her strategy  $P|i_H, P|i_{-H}$  if she has chosen  $O$  before and  $\pi < 2/3$ . Thus,  $Pr(P|H) = 1$  from equation (23) and  $Pr(P|\neg H) = 1$  from equation (24). She chooses between this strategy and  $\neg O$ .

For the bureaucrat, in turn, it always holds that

$$\begin{aligned} EU(H) &< EU(\neg H) \\ b - e - Pr(O)Pr(P|H)f &< b - Pr(O)Pr(P|\neg H)f \\ b - e - Pr(O) \cdot 1 \cdot f &< b - Pr(O) \cdot 1 \cdot f \\ -e &< 0 \end{aligned}$$

independent of  $Pr(O)$ , which is why she always plays  $\neg H$ . As a response, the legislator will play her pure strategy  $P|i_H, P|i_{-H}$ .

(c)  $\pi = 2/3$  (borderline case)

In case that information accuracy  $\pi = 2/3$ , both of the preceding equilibria exist. That is, in the pure strategy Nash equilibrium the bureaucrat plays  $\neg H$  and the legislator responds with  $P|i_H, P|i_{-H}$ . In the mixed equilibrium, the bureaucrat plays  $H$  with probability  $(4s - c)/5$ . The legislator on her side mixes over  $\neg O$  and  $\neg P|i_H, P|i_{-H}$ , the respective probabilities being  $(f - 3e)/f$  and  $3e/f$ . The mixed equilibrium follows from inserting  $\pi = 2/3$  into equations (30) and (31).

## Appendix 2

### *Information provided to participants*

*Instructions (Example with:  $\pi = 0.80$ ,  $f = 10$ ,  $c = 7$ ).* The aim of this experiment is the study of decision behavior. You and the other participants will make decisions during the experiment for which you will earn money. The amount of your payment is determined by your own decisions as well as those of the other participants in accordance with the rules on the following pages. The instructions are used to fully inform you about the structure of the experiment and the consequences of your decisions.

*Merit.* During the experiment, you will earn tokens. At the end of the experiment, your tokens will be converted into Euros according to the following exchange rate: 1 token = 0.01 Euros. At the end of the experiment, you will receive the amount of money that you have earned over the course of the experiment in cash plus €3 allowance.

*Time.* The entire experiment will take about 50 minutes. Following the experiment, a questionnaire will appear on your screen. After you have completed the questionnaire, please wait until your desk number is called. Then you will receive your payment in Euros.

*Anonymity.* No participant will know about your identity or the identity of the other participants in the course of the experiment. The other participants will not know how much you have earned.

*Communication ban.* Throughout the experiment, communication is strictly prohibited! Please also switch off your mobile phone. In addition, we point out that you are only allowed to use those functions of the PC that are required for the experiment. Violations of these rules will lead to an exclusion from the experiment.

### *Game structure*

- There are two types of players: Type A and Type B.
- Type A has to provide a good for Type B.
- Type A decides whether to provide this good in high or low quality.
- Type B decides whether to monitor the quality of the good or not.
- For A, the provision of the good in high quality causes costs.
- For B, controlling the property of the good causes costs.
- If B decides to review the quality, she receives information about the quality of the good.
- The information on the quality of the good is correct in 80% of cases.
- After having received the information, B can charge a penalty against A.
- In each round, both players A and B will receive a basic payment.

- Type A additionally obtains a certain income, from which the expenses is subtracted if she delivers high quality. Furthermore, A has to pay penalty if he was controlled by B and punished.
- Type B receives an additional income, only if A delivers high quality. From this income, the control costs are subtracted.
- If B penalizes an A player who has provided a low quality, B receives a bonus.
- If B penalizes an A player who has delivered high quality, B will be sanctioned for her mistake.
- If B does not penalize an A player who has provided low quality, B will also be sanctioned for her mistake.

**Table 6.** Token for both types of players.

Tokens for A							
Decision A	First Decision B	Second Decision B	Endowment	Income	Effort	Punishment fee	Tokens
High quality	No control		15	10	-5	0	20
High quality	Control (Info. true with 80%)	No punishment	15	10	-5	0	20
High quality	Control (Info. true with 80%)	Punishment	15	10	-5	-10	10
Low quality	No control		15	10	0	0	25
Low quality	Control (Info. true with 80%)	No punishment	15	10	0	0	25
Low quality	Control (Info. true with 80%)	Punishment	15	10	0	-10	15

Tokens for B							
First Decision B	Second Decision B	Decision A	Endowment	Income	Control cost	Bonus/ Malus	Tokens
No control		High quality	15	10	0	0	25
No control		Low quality	15	0	0	-5	10
Control (Info. true with 80%)	No punishment	High quality	15	10	-7	0	18
Control (Info. true with 80%)	No punishment	Low quality	15	0	-7	-5	3
Control (Info. true with 80%)	Punishment	High quality	15	10	-7	-5	13
Control (Info. true with 80%)	Punishment	Low quality	15	0	-7	5	13

These tables were not only printed on the instructions but also explained verbally including the visual presentation of a reading example. The verbal Protocol and Slides are available upon request.

### Sequence of decisions

- A and B decide simultaneously which quality to deliver and whether to control or not.

- If B decides to control the quality, she receives the information and decides on the basis of an 80% probability of correct information as to whether A has to pay a penalty fee.
- At the end of each round, you will be informed about the decision of your opponent and your payoff (Table 6).

**Several rounds.** The game will be repeated 33 times. The first three rounds will be used to allow you to familiarize yourself with the structure of the game. The tokens you will make in the first three rounds will not be considered. You will be randomly assigned to a Type of Player at the beginning of the game. Your type will not change during the course of the game. Each time you will play with another person. You will not know the identity of your opponent.

**Time.** You will have 30 seconds to make your decision.

**Payment.** To calculate your payment in Euro (plus allowance), all the tokens that you will earn throughout the game of 30 rounds will be used (exchange rate: 1 token=0.01 Euros).

**Questionnaire.** After the experiment you will be asked to fill out a questionnaire. Finally, you will see your earnings plus allowance in Euro.

Click the Next button and wait until your desk number is called for payment. Write your name on the receipt that you have received at the beginning of the experiment. Take your place card, the instructions, the pen, and the completed receipt for payment.

Thank you for your participation.

## Appendix 3

### Aggregate-level results over periods

**Table 7.** Aggregate-level results over periods.

Period	Low cost	Low cost	High cost	High cost	Noisy	Accurate
	Low fee	High fee	Low fee	High fee		
<i>Pr(H)</i>						
1–10	52%	51%	31%	46%	23%	55%
11–20	44%	50%	32%	44%	21%	52%
21–30	48%	52%	26%	36%	20%	50%
Prediction	56%	56%	11%	11%	0%	21%
<i>Pr(O)</i>						
1–10	63%	68%	52%	42%	55%	41%
11–20	72%	63%	56%	42%	59%	40%
21–30	76%	71%	58%	44%	72%	35%
Prediction	83%	42%	83%	42%	100%	31%
<i>n</i>	40	36	40	40	40	40