



# Administrative delegation revisited: Experimental evidence on the behavioural consequences of public service motivation and risk aversion

International Review of Administrative  
Sciences

2023, Vol. 89(3) 613–631

© The Author(s) 2022



Article reuse guidelines:

[sagepub.com/journals-permissions](https://sagepub.com/journals-permissions)

DOI: 10.1177/00208523211073259

[journals.sagepub.com/home/ras](https://journals.sagepub.com/home/ras)**Markus Tepe** 

University of Oldenburg, Germany

**Susumu Shikano**

University of Konstanz, Germany

**Michael Jankowski**

University of Oldenburg, Germany

**Maximilian Lutz**

University of Oldenburg, Germany

## Abstract

Getting a grip on issues of administrative delegation is key to the performance of public organizations. The oversight game models delegation as a conflict of interest between an inspector and an inspectee to act in the interests of the former. This study tests alternative solutions to overcome ‘shirking’ in the oversight game. Specifically, we test the effect of external incentives, as implied by the game-theoretical solution, against the role of intrinsic factors, namely, public service motivation and job-related risk aversion. Evidence from a laboratory ( $N = 208$ ) and survey experiment ( $N = 794$ ) show that both the game-theoretical approach, which inspired new public management, and public

## Corresponding author:

Markus Tepe, Institute for Social Sciences, University of Oldenburg, Ammerländer Heerstraße 114–118, 26129 Oldenburg, Germany.

Email: [markus.tepe@uol.de](mailto:markus.tepe@uol.de)

service motivation, as its antithesis, fail to explain subjects' behaviour. Instead, job-related risk aversion makes oversight more and 'shirking' less likely. This finding hints towards a more differentiated view of public employees' risk attitudes to improve administrative delegation.

### **Points for practitioners**

The promise of new public management that oversight issues in administrative delegation disappear with setting appropriate extrinsic incentives is too simplistic. Public service motivation, on the other hand, which started as an antithesis of the self-interested bureaucrat, also fails to solve the issue of 'shirking' in administrative delegation. Instead, job-related risk aversion appears to improve administrative delegation, which presents a remarkable counterpoint to the popular opinion in public management that risk aversion is problematic for public organizations' performance. Rather than avoid selecting risk-averse public employees generally, more attention might be paid to the domains of administrative decision-making in which such traits can be beneficial.

### **Keywords**

laboratory experiment, oversight, public service motivation, risk attitudes, shirking, survey experiment

## **Motivation**

Administrative delegation and bureaucratic oversight are classical fields for principal-agent (PA) models (Gailmard and Patty, 2012). Since Weber (1958), asymmetries between politicians and bureaucrats, as well as senior and subordinate bureaucrats within the public administration, have characterized rational-legal authority. While authority rests with the politician or a senior bureaucrat, individual policy fields and executive instruments' expertise rests with subordinate bureaucrats (Miller, 2005: 203). PA theory provides a generic toolbox for analysing the consequences of asymmetries in such delegation relationships. In particular, PA theory is interested in finding a way to induce agents to avoid 'shirking' and take actions consistent with the principal's goals to implement public policies (Gailmard, 2010: 40).

The oversight game (OG), also referred to as the 'inspection game' (Rauhut, 2015; Tsebelis, 1990), offers a decision-theoretical model to study administrative delegation. In this game, an agent (inspectee) decides to fulfil a task at high or low quality, while higher quality requires more effort. Simultaneously, the principal (inspector) decides to control or not control the agents' behaviour, while control is costly to the inspector. The inspector's payoff is the highest if the inspectee complies without being controlled, and the inspectee's payoff is higher if they provide low quality without being controlled. The game-theoretical solution to the OG, which rests on the assumption of rational and selfish players, is to identify the optimal amount of punishment for 'shirking' at minimal

control costs for the inspector. The solution concept expects that shirking disappears if both parameters – punishment fee and control cost – are correctly set. This way of thinking about administrative delegation has inspired new public management (NPM) and provided the academic foundation for a wide range of human management reforms in public organizations.

PA theory has fundamentally challenged scholars' understanding of oversight in administrative delegation (Miller, 2005; Miller and Whitford, 2007), even though the empirical validity of the game-theoretical solution is mixed at best (Rauhut, 2015; Rauhut and Winter, 2010; Shikano et al., 2017). More importantly, however, little attention has been given to alternative explanations of inspectees' and inspectors' behaviour. This study theorizes and tests two intrinsic motives that are assumed to help overcome issues of administrative delegation, namely, public service motivation (PSM) and job-related risk aversion. First, extending on research discussing linkage between PSM, job performance (Alonso and Lewis, 2001) and the behavioural consequences of PSM (Esteve et al., 2016; Tepe, 2016), PSM is expected to affect administrative delegation. Specifically, we test Gailmard's (2010: 39) argument that PSM could help overcome 'shirking' problems. Second, risk aversion has become a significant theme in understanding public organizations' performance. Compared to PSM, it is often considered harmful, as it is associated with a lack of innovation and increasing red tape (Brown and Osborne, 2013; Chen and Bozeman, 2012). In the context of administrative delegation, however, job-related risk aversion could promote positive outcomes if risk-averse inspectees 'shirk' less and risk-averse inspectors choose to control more.

In theoretical terms, this study brings together three different strands of research dealing with administrative delegation, drawing on very different conceptualizations of the micro-motives of bureaucratic agents; rationality and self-interest; PSM; and job-related risk aversion. The goal is to bring these arguments to a rigorous experimental test. In methodological terms, this study responds to repeated calls for experimental research in public administration (James et al., 2017). Specifically, we test the robustness of experimental designs by conducting two studies with an identical parametrization of the OG: a laboratory experiment conducted on a student sample with monetary rewards; and a large-scale online survey experiment conducted on public and private sector employees with constant monetary rewards. In both experiments, we manipulated the control cost applied to the inspector and the punishment fee applied to the inspectee. Subjects' self-reported PSM (Kim et al., 2013) and job-related risk preferences (Dohmen et al., 2011) are collected in a post-treatment standardized survey.

## **Theoretical framework**

### *Oversight in PA theory*

Applying PA theory to study oversight dates back to the path-breaking contributions of Weingast and Moran (1983) and McCubbins and Schwartz (1984). The latter distinguishes between 'police-patrol' and 'fire-alarm' oversight (McCubbins and Schwartz, 1984: 161). Under 'police-patrol' oversight, politicians actively examine bureaucrats to

detect and remedy any deviating behaviour. ‘Fire-alarm’ oversight, in contrast, depends on an active public that monitors bureaucratic performance. This study focuses exclusively on ‘police-control’ oversight between a single inspector or executive in its role as principal and a single inspectee as an agent (Gailmard, 2009). The interaction structure can be formally expressed as follows (McCarty and Meiorowitz, 2007: 140): two players, the inspector and the inspectee, are involved in a sequential, zero-sum discoordination game. The inspectee can deliver a service at high or low quality,  $B_S \in \{H, L\}$ . If the inspectee delivers high quality, the effort costs  $e > 0$ , while low quality is costless. The inspectee’s expected payoff for high effort is always  $-e$ , independent of the inspector’s decision. Therefore, the inspectee prefers low-quality service with the payoff 0 for the lack of personal cost but also faces the possibility of a punishment fee  $f$  with a payoff  $-f$  for low quality if the inspector chooses to conduct oversight.

Since the inspectee and the inspector decide simultaneously, the inspector does not know about the quality of the inspectee’s service when deciding whether to control the inspectee or not:  $A_S \in \{O, N\}$ . The inspector prefers high-quality  $H$  to low-quality  $L$  with a bonus  $b$  for high-quality service and a malus  $s$  for low-quality service if no control is taken ( $N$ ). However, observing the inspectee’s effort is costly ( $c > 0$ ) and therefore preferred only in the case of low quality with payoffs of  $-c + s$  under the assumption that  $c < 2s$ . The bonus  $s$  for successfully detecting low-quality work by the inspectee represents proof of leadership abilities. Table 1 represents the strategies and payoffs in normal form.

In a strictly decision-theoretical framework, the inspectee would decide based on the expected utility of choosing high or low effort under a given punishment fee and detection probability level. This view, however, ignores the ability to anticipate the other’s action rationally. Game theory predicts that inspectors only care about the inspectees’ punishment fee, while the inspectees’ behaviour will solely depend on the inspectors’ control costs. The unique Nash equilibrium (NE) of this game is in mixed strategies (Avenhaus, 2004), where the bureaucrat (inspectee) delivers high effort with probability  $P(\text{high}) = (2s - c)/2s$ . The best response for the inspector is to make the inspectee indifferent towards choosing high or low quality. Thus, the inspector mixes pure strategies with probability  $P(\text{control}) = e/f$ . Under the assumption that  $f > c$  (effective punishment), there are no NEs in pure strategies.

The game-theoretical solution offers a straightforward prediction of how changing the control costs applying to the inspector and changing the punishment fee applying to the inspectee will affect the outcome of the OG. The punishment fee and control

**Table 1.** Strategies and payoffs.

		Inspector (principal)	
		Control (O)	No control (N)
Inspectee (agent)	High quality (H)	$-e, b - c$	$-e, b$
	Low quality (L)	$-f, -c + s$	$0, -s$

costs are the two parameters of main theoretical interest for public management research since they can be changed through legislation. Therefore, the first pair of hypotheses states the following:

H1a: Higher punishment fees decrease inspectors' propensity to choose control.

H1b: Higher control costs decrease inspectees' propensity to deliver high effort.

The game-theoretical solution rests on the assumption of rational and self-interested individuals, and it inspired multiple NPM reforms aiming to improve public organizations' performance. However, despite its practical impact, there is little evidence that this solution can adequately predict human behaviour in the OG. Growing evidence from behavioural economics and behavioural public administration indicates that human behaviour follows a multitude of motives. In the following, we will focus on two main alternative explanations in public administration research – PSM and risk aversion – and theorize why these motives could help overcome the delegation issue.

## PSM

PSM highlights individuals' intrinsic motivation. Perry and Wise (1990: 368) define PSM as 'an individual's predisposition to respond to motives grounded primarily or uniquely in public institutions' and thereby position it against the idea of selfish bureaucrats omnipresent in the NPM framework. Although PSM does not determine public sector employment, it is presumed that individuals with high levels of PSM have 'a desire to serve the public interest, loyalty to duty and to the government as a whole, and social equity' (Perry and Wise, 1990: 369). A few laboratory and quasi-experimental studies have been able to identify the behavioural consequences of self-reported PSM. These studies suggest, for example, a positive relationship between PSM and cooperation (Esteve et al., 2016), trust behaviour (Tepe, 2016), altruism (Tepe and Vanhuyse, 2017) and voluntary self-sacrifice (Heine et al., 2021). The behavioural relevance of PSM for oversight in administrative delegation has not yet been addressed.

Gailmard (2010: 40) points out that PSM offers some unique benefits to public organizations, as it allows them to staff themselves with dedicated individuals at lower cost and fosters the self-selection of individuals that comply with public service goals. Francois (2000: 296) developed a formal model showing that PSM improves the delegation of specific responsibilities to public administration. Expanding on these arguments, PSM can be helpful in public management to the extent that it mitigates delegation problems – shirking and oversight. More specifically, drawing on Le Grand (2003) and Gailmard (2010: 40), we argue that PSM alleviates the shirking problem in administrative delegation by bringing in individuals who want to work on the mission of a public organization and are self-sacrificing and altruistic. Choosing the appropriate punishment fee and control is one device to solve the OG; PSM advocates another, substitute approach.

A second perspective on why PSM could play a role in the OG is its relationship with rule enforcement. PSM entails two interrelated motives: a public interest orientation and a tendency to comply with social rules (Perry and Vandenabeele, 2015: 696). This latter aspect of PSM embraces the willingness to sanction the disloyal and those that do not

adhere to societal rules. For example, Prokop and Tepe (2020) test the effect of PSM on subjects' behaviour in a repeated public goods game with a costly option to punish free riders. They find that PSM has a twofold effect, as compassion is associated with higher contributions and attraction to policymaking is associated with punishing free riders. According to this logic, higher levels of PSM could lead to higher levels of compliance and control.

H2a: Higher levels of PSM increase inspectors' propensity to choose control.

H2b: Higher levels of PSM increase inspectees' propensity to deliver high quality.

### *Risk preferences*

It is a widespread belief that public bureaucrats are more risk averse than their private sector counterparts. Several studies provide robust evidence of public–private sector differences in self-reported risk attitudes. These studies indicate that risk-averse individuals are more likely to join the public sector (Buurman et al., 2012) and more likely to assign more importance to job security (Houston, 2011). Nevertheless, behavioural evidence on the role of risk preferences among public servants is mixed. Tepe and Prokop (2018) show that public administration students report higher levels of subjective risk aversion but do not behave in a more risk-averse way in a compound lottery choice experiment. Klijn et al. (2020: 5) show that street-level bureaucrats' enforcement style is affected by the risk of being blamed for poor performance. They found that when inspectors perceive more blame risk, they act less formally and less coercively.

Risk aversion has been a long-standing source of concern in PA theory (Miller and Whitford, 2007) and public management research (Chen and Bozeman, 2012). PA theory promises to substitute costly monitoring with an effective incentive structure that reduces shirking. However, a key hurdle in obtaining such an incentive system is risk aversion among agents. Hölmstrom (1979) is very pessimistic on this matter, as he suggests that agents' risk aversion, combined with information asymmetry, makes it impossible to replace monitoring with an equally efficient system of incentives (Livnat-Lerer et al., 2018: 1). In public management research, risk aversion is seen chiefly as an obstacle to organizational reform and performance enhancement. Advocates of NPM thus consider enhancing risk taking a cornerstone of public management reform efforts (Chen and Bozeman, 2012: 378).

This study, on the contrary, argues that risk aversion might be beneficial to overcoming the administrative delegation issue as depicted in the OG. The behavioural consequences of self-reported risk attitudes for the OG are straightforward. Those who report higher levels of risk aversion should be more likely to behave in a way that minimizes the maximum loss. This is the case when the agent chooses high effort to avoid punishment for shirking and when the principal chooses to control to avoid the malus for being a poor inspector. Thus, the effect of risk aversion on subjects' behaviour in the OG is summarized as follows:

H3a: Job-related risk aversion increases inspectors' propensity to choose control.

H3b: Job-related risk aversion increases inspectees' propensity to deliver high quality.

## Research design

The set of hypotheses derived earlier are tested by conducting a laboratory and survey experiment. While laboratory experiments with real-time interaction and monetary rewards are presumed to create stronger treatment effects, conducting such experiments on representative or theoretically relevant subpopulations is challenging. Online survey experiments offer ways to address this difficulty. Both experiments are based on the same OG. However, the experiment had to be adjusted by how preferences can be elicited in the laboratory and survey environment.

The parametrization of the OG is summarized in Table 2. We only manipulated the parameters control costs (denoted by  $c$ ; ‘high’ versus ‘low’) and punishment fee (denoted by  $f$ ; ‘high’ versus ‘low’), while keeping the other three parameters constant, namely: the inspectees’ effort  $e$ ; the inspectees’ fixed-wage or the inspectors’ benefit from high quality  $b$ , respectively; and the popularity bonus or malus  $s$  for inspectors who managed to, respectively, detect or could not detect a shirking inspectee. Using a factorial design ( $2 \times 2$ ), we have a total of four different scenarios. Table 2 also reports the NE in mixed strategies for the inspector to choose control and for the inspectee to choose high quality in each scenario (e.g. in Scenario 1, the inspector is predicted to choose control with a probability of 62.5%, while the inspectee is predicted to choose high quality with a probability of 80%).

### Study 1: laboratory experiment

The laboratory experiment was programmed in Z-Tree (Fischbacher, 2007). Each subject was confronted with four scenarios (see Table 2).<sup>1</sup> For each scenario, subjects were asked to type in their randomization strategy. Specifically, we applied a strategy method design to the OG (Rauhut and Winter, 2010). For the strategy elicitation, subjects were asked how often they want to choose ‘high’ versus ‘low’ quality, or ‘control’ versus ‘no control’, in ten interactions with the same opponent. Once subjects confirmed their decisions, the computer randomly matched each of the paired players’ ten decisions and determined their payoffs.<sup>2</sup> Before the experiment started, a test round of the OG was played to allow participants to familiarize themselves with the decision mechanism. After each

**Table 2.** Parametrization.

Treated parameter		Fixed parameter			NE	
Control costs $c$	Punishment fee $f$	Malus/ bonus $s$	Effort costs $e$	Benefit $b$	Inspector (control) $Pr(C_A)$	Inspectee (high quality) $Pr(H_B)$
2	8	5	5	10	62.5	80
8	8	5	5	10	62.5	20
2	15	5	5	10	33.3	80
8	15	5	5	10	33.3	20

period, the participants were informed about their payoff and the payoff of their partner. After each scenario, subjects were rematched randomly. Each participant had to make 40 decisions in total (four scenarios  $\times$  10 decisions per scenario). The sequence of scenarios was the same in all sessions (for further information on the experimental proceedings, see Figure 2 in the Online Appendix).

We conducted 12 experimental sessions at two German universities with 208 subjects (104 inspectors; 104 inspectees). All subjects were students recruited via the Hamburg Registration and Organization Online Tool (Hroot) online recruitment system (Bock et al., 2014). After the last scenario, subjects completed a post-experimental questionnaire, which included the PSM item battery (Kim et al., 2013) and a measure of job-related risk aversion (Dohmen et al., 2011). All subjects were paid anonymously before leaving the laboratory. The average payoff was €12.68, including the €5 show-up fee. An average experimental session took about 45 minutes.

### *Study 2: survey experiment*

The social interaction and complexity of the laboratory experiment are not directly transferable to an online survey experiment. Most importantly, online surveys may only take a short time. The survey design reduced the experiment's complexity while retaining the basic structure of the OG. The online survey was conducted among 794 respondents. The online sample was provided by respondi. Respondents received a monetary incentive for participating in the survey, but contrary to the laboratory experiment, respondents' payoff was constant and not dependent on their behaviour in the experiment. We aimed to sample approximately 500 respondents currently employed in the public sector and 250 respondents employed in the private sector. This enables us to test the effect of sector affiliation on subjects' behaviour in the OG.

The experiment was conducted by presenting tables describing the decision task (see Figure 4 in the Online Appendix). These eight situations were identical to the situations presented in the laboratory experiment described earlier. In contrast to the laboratory experiment, however, we simply asked survey respondents whether they would control or not control when playing as the inspector, or deliver high quality or low quality when playing as the inspectee. Survey participants played four rounds as the inspector and four rounds as the inspectee. The order of the roles being played was randomized to avoid systematic bias due to order effects. The situations within the two roles were also randomized. After the experiment, the survey contained the same post-experimental items on PSM and job-related risk aversion as those used in the laboratory experiment (for descriptive statistics, see Tables 1 and 2 in the Online Appendix).

## **Empirical results**

### *Study 1: laboratory experiment*

In the laboratory experiment, the dependent variable measured the share of times a participant indicated that they would control or deliver high quality in a total of ten rounds.



Thus, the dependent variable is a proportion and bounded by 0 and 1. To account for the boundedness of the dependent variable, we use a ‘fractional logit’ regression model (Papke and Wooldridge, 1996) to estimate the effect of the model parameters, PSM and subjects’ risk preference on choosing control as the inspector or high quality as the inspectee. When standard ordinary least squares (OLS) regression analysis is applied to proportions, it might predict values outside of the range from 0 to 1 (for robustness, see Table 3 in the Online Appendix). It might also not capture the non-linearities

**Table 3.** Fractional logit regression: laboratory experiment.

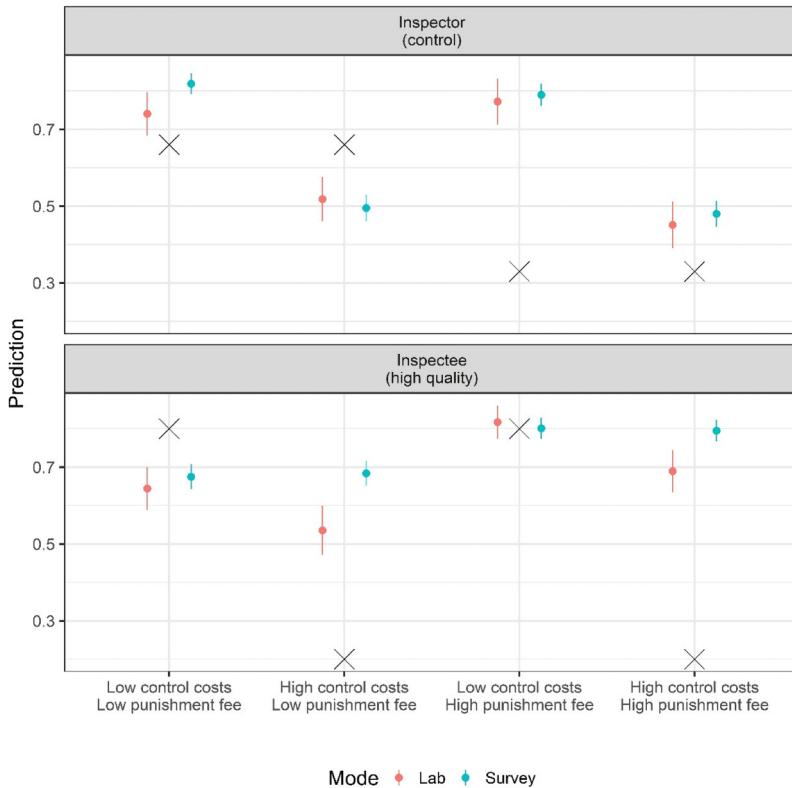
	Role of respondent					
	Inspector (principal) control			Inspectee (agent) high quality		
	(1)	(2)	(3)	(4)	(5)	(6)
<i>c</i> (high control costs)	−0.98*** [0.17]	−0.987*** [0.17]	−0.988*** [0.17]	−0.461*** [0.16]	−0.463*** [0.16]	−0.465*** [0.16]
<i>f</i> (high punishment fee)	0.173 [0.16]	0.174 [0.17]	0.174 [0.17]	0.920*** [0.17]	0.922*** [0.17]	0.927*** [0.17]
<i>c</i> * <i>f</i>	−0.45** [0.20]	−0.448** [0.20]	−0.448** [0.20]	−0.251 [0.24]	−0.251 [0.24]	−0.252 [0.24]
PSM	0.256 [0.42]		0.397 [0.45]	0.679 [0.43]		0.828** [0.40]
Risk aversion		0.724* [0.41]	0.787* [0.43]		0.824* [0.44]	0.942** [0.43]
No neutral frame	−0.240 [0.20]	−0.236 [0.20]	−0.225 [0.20]	−0.0554 [0.17]	−0.0999 [0.17]	−0.101 [0.16]
Correct test question	0.0658 [0.23]	0.0297 [0.23]	0.0188 [0.24]	0.138 [0.16]	0.0983 [0.17]	0.117 [0.16]
Age	0.0155 [0.03]	0.0171 [0.03]	0.0153 [0.03]	−0.00480 [0.02]	0.0136 [0.02]	0.00933 [0.02]
Female	0.281 [0.19]	0.243 [0.19]	0.249 [0.19]	−0.567*** [0.18]	−0.575*** [0.18]	−0.595*** [0.18]
Constant	0.442 [0.73]	0.243 [0.72]	0.0124 [0.83]	0.571 [0.52]	0.202 [0.56]	−0.260 [0.58]
Observations	416	416	416	416	416	416
AIC	533.0	531.1	532.6	516.7	515.8	515.6
BIC	569.3	567.4	572.9	553.0	552.1	555.9
LL	−257.5	−256.6	−256.3	−249.4	−248.9	−247.8

Notes: Fractional logit regression. Dependent variable is the proportion of times the inspector chooses control or the inspectee chooses high quality. AIC = Akaike-Information-Criterion. BIC = Bayesian-Information-Criterion. Standard errors in square brackets. \*  $p < 0.10$ ; \*\*  $p < 0.05$ ; \*\*\*  $p < 0.01$ .

that arise from the dependent variables' boundedness. Therefore, the fractional logit model is advantageous over OLS and is recommended for analysing proportions. Moreover, using this model increases the comparability to our second study in which logistic regression analysis is used, as logistic regression and fractional logit are very similar models.

The independent variables of interest are the dummy variables for high control costs and punishment fees, denoted by  $c$  and  $f$ ; PSM<sup>3</sup>; and job-related risk aversion (for descriptive statistics, see Figure 1 in the Online Appendix). In addition, the regression models include a dummy variable controlling for the framing of the experiment and another dummy variable that controls whether respondents answered a series of test questions correctly. Additional control variables are gender and age.

The results of the laboratory experiment are reported in Table 3 and Figure 1. High control costs  $c$  is a relevant predictor in both cases. High control costs lead to a substantial decrease in the probability of choosing to control when playing the OG as an inspector.

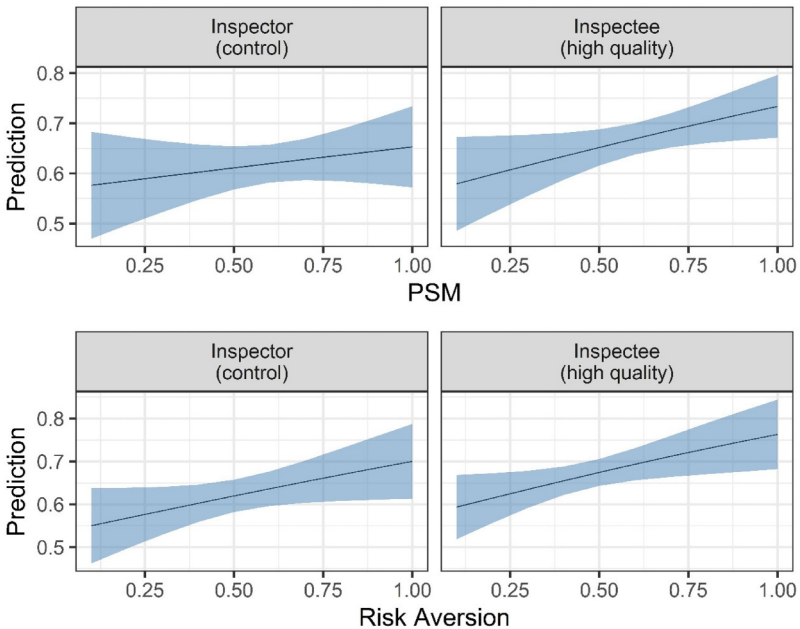


**Figure 1.** Behaviour in the laboratory and survey experiment compared to NE. Notes: x displays NE. Points denote predicted behaviour in the laboratory and survey experiment based on Models 3 and 6 from Tables 3 and 4. Vertical lines denote 95% confidence intervals.

High control costs substantially decrease the probability of providing high quality when playing the OG as an inspectee. The results suggest that participants are more likely to deliver low quality when they know that the costs of the inspector are high. A high punishment fee leads to a higher probability of delivering high quality, but only if the control costs are low.

Figure 1 summarizes subjects' behaviour in the laboratory experiment in the four treatment conditions and compares the empirically observed behaviour to the NE in mixed strategies. Contrary to H1a, the punishment fee does not affect inspectors' behaviour, while in line with H1b, the higher control cost decreases inspectees' propensity to choose high quality. Contrary to the game-theoretical prediction, higher punishment fees are associated with a higher propensity to provide high quality. Thus, inspectees do not seem to play a randomization strategy consistent with the predicted NE in mixed strategies. Both inconsistencies have been observed in previous laboratory experiments on the OG (Nosenzo et al., 2014; Rauhut, 2015) and foster doubt about the empirical relevance of the NE in mixed strategies. Furthermore, there is a clear tendency to control more often than in the rational prediction.<sup>4</sup>

When turning to the effect of PSM, Models 1, 3 and 4 indicate that the level of PSM has a positive but statistically insignificant effect on subjects' behaviour in the OG (for



**Figure 2.** Predicted behaviour in the laboratory experiment based on PSM and risk aversion. Notes: Predictions are based on Models 3 and 6 from Table 3. Lines denote point estimates, and shaded areas denote 90% confidence intervals. PSM and risk aversion are standardized to a 0–1 scale.

predicted probabilities, see Figure 2). Only in Model 6 are higher levels of PSM associated with higher proportions of providing high quality. Aside from this finding, there is no evidence that PSM impacts subjects' behaviour in the OG (H2a and H2b). Job-related risk preferences, on the contrary, have a robust effect on subjects' behaviour. In line with H3a, higher levels of job-related risk aversion lead to an increase in the probability of choosing to control. In line with H3b, higher levels of job-related risk aversion lead to an increase in the probability of providing high quality.

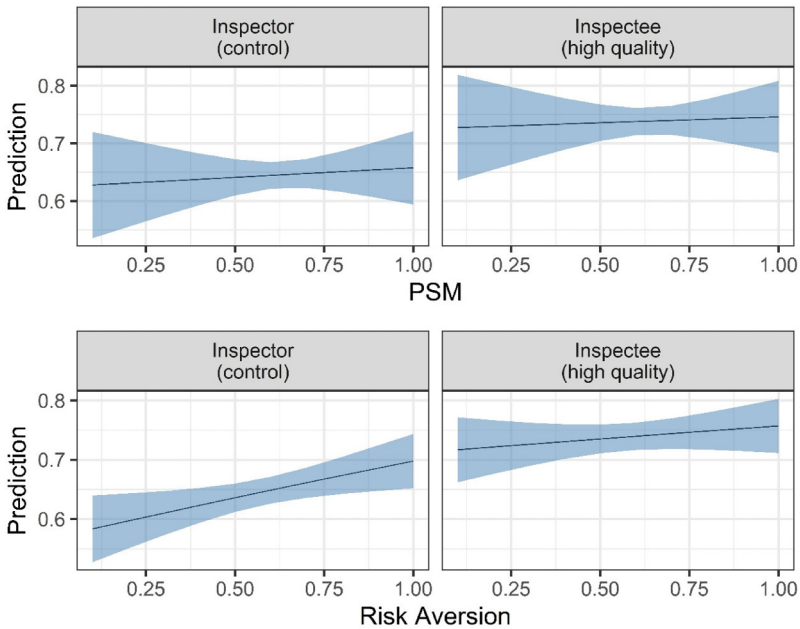
In addition, we test whether providing a more mundane reality to the otherwise neutrally framed decision situation in the laboratory environment activates or alters subjects' behaviour in the OG. In the neutral frame, subjects were assigned the roles 'A' and 'B' and chose between the two options 'C' and 'D'. In the stylized framing, subjects were assigned to the roles 'politician' and 'bureaucrat' and chose between 'control/no control' and 'high

**Table 4.** Logistic regression: survey experiment.

	Role of respondent					
	Inspector (principal) control			Inspectee (agent) high quality		
	(1)	(2)	(3)	(4)	(5)	(6)
<i>c</i> (high control costs)	-1.54*** [0.10]	-1.544*** [0.10]	-1.544*** [0.10]	0.0405 [0.07]	0.0405 [0.07]	0.0405 [0.07]
<i>f</i> (high punishment fee)	-0.185** [0.09]	-0.186** [0.09]	-0.186** [0.09]	0.662*** [0.10]	0.662*** [0.10]	0.662*** [0.10]
<i>c * f</i>	0.124 [0.11]	0.124 [0.11]	0.124 [0.11]	-0.0796 [0.11]	-0.0796 [0.11]	-0.0796 [0.11]
PSM	-0.0344 [0.41]		0.165 [0.42]	0.0340 [0.43]		0.109 [0.44]
Risk aversion		0.610** [0.25]	0.628** [0.26]		0.223 [0.26]	0.235 [0.27]
Public sector	-0.130 [0.13]	-0.176 [0.13]	-0.184 [0.13]	0.0465 [0.13]	0.0314 [0.13]	0.0266 [0.14]
Correct test question	-0.0725 [0.13]	-0.0988 [0.13]	-0.102 [0.13]	-0.0114 [0.14]	-0.0195 [0.14]	-0.0217 [0.14]
Age	0.0120** [0.01]	0.0110** [0.01]	0.0108** [0.01]	0.00154 [0.01]	0.00119 [0.01]	0.00107 [0.01]
Female	-0.142 [0.12]	-0.177 [0.12]	-0.176 [0.12]	-0.117 [0.13]	-0.131 [0.13]	-0.130 [0.13]
Constant	1.225*** [0.36]	0.973*** [0.30]	0.872** [0.38]	0.685* [0.40]	0.617* [0.32]	0.551 [0.43]
Observations	3176	3176	3176	3176	3176	3176
AIC	3766.4	3753.3	3755.0	3605.8	3604.1	3606.0
BIC	3820.9	3807.9	3815.7	3660.3	3658.7	3666.6
LL	-1874.2	-1867.7	-1867.5	-1793.9	-1793.1	-1793.0

Notes: Logistic regression. Dependent variable is: 1 if inspector chooses control, otherwise 0; or 1 if the inspectee chooses high quality, otherwise 0. AIC = Akaike-Information-Criterion. BIC = Bayesian-Information-Criterion.

Standard errors in square brackets. \*  $p < 0.10$ ; \*\*  $p < 0.05$ ; \*\*\*  $p < 0.01$ .



**Figure 3.** Predicted behaviour in the survey experiment for PSM and risk aversion.

Notes: Predictions are based on Models 3 and 6 from Table 4. Lines denote point estimates, and shaded areas denote 90% confidence intervals. PSM and risk aversion are standardized to a 0–1 scale.

quality/low quality’. This treatment can be considered as a variant of issue framing, which involves emphasizing a subset of potentially relevant considerations when constructing opinions that lead to behaviour (Druckman, 2004: 672). The issue frame used in this experiment emphasizes the political and administrative character of the decision situation. However, in none of the models summarized in Table 3 did the ‘no neutral frame’ dummy exert a substantive or statistically significant effect on subjects’ behaviour.

### Study 2: survey experiment

In the survey experiment, the respondents had to indicate whether they would control or deliver high quality, not how often they chose a specific option. As the dependent variable in the survey experiment is binary, we estimate binary logit regression models. The survey experiment did not contain a framing treatment, as was the case in the laboratory. Instead, we add a variable that controls for public sector affiliation. The survey sample consists of respondents who are currently employed in either the private or public sector. The results of the survey experiment are reported in Table 4 and Figure 1.

The results of the survey experiment are very much comparable to the laboratory experiment, with only a few notable differences. High control costs are only relevant

when playing the OG as the inspector. When playing the OG as the inspectee, the control costs of the inspector are irrelevant. This result was different in the laboratory experiment, where the control costs mattered in both cases. Thus, the effect of control costs on inspectees' behaviour is inconsistent with H1a. In line with H1b, the punishment fee has a negative and significant effect on the inspector's behaviour. This effect is, however, relatively small, as can also be seen from Figure 1. This weaker effect of the manipulation on inspectees' behaviour than on inspectors' is consistent with previous studies conducted by Rauhut (2015). Unfortunately, Rauhut manipulated the incentives of inspectors and inspectees separately in each study and did not comparatively discuss the effect size. Contrary to H1b, a high punishment fee leads to a higher probability of providing high quality. Figure 1 demonstrates that the predictions from the experiment and the theoretical predictions from the NE follow similar patterns for the inspector's role. For the role of the inspectee, however, the patterns are again very different from the theoretical predictions. As can be seen, respondents only paid attention to the level of the punishment fee, without considering the control costs of the inspector.

Concerning the effect of PSM and job-related risk aversion on survey participants' behaviour in the OG, the findings are broadly consistent with previous results from the laboratory experiment (for predicted probabilities, see Figure 3). Contrary to H2a and H2b, PSM does not affect subjects' behaviour in the OG, while higher levels of job-related risk aversion increase the probability of controlling as a politician and providing high quality as a bureaucrat (H3a and H3b). Concerning the effect of sector affiliation and behaviour in the OG, Table 4 shows that the 'public employment' dummy does not exert a substantive or statistically significant effect in any model.

## **Discussion and conclusions**

Many arrangements in public administration are arrangements of delegation, whereby a principal (politician or senior bureaucrat) delegates a task to an agent (subordinate bureaucrat) and the welfare of the former is affected by the choices made by the latter. If agents act in a rational and self-interested way, principals are confronted with 'shirking'. The OG is a game-theoretical model specifically designed for capturing the essence of oversight in administrative delegation. Getting a grip on 'shirking' agents is key to public organizations' performance. This study conducts a laboratory and survey experiment with identical parametrizations to test three alternative solution approaches to administrative delegation. First, the game-theoretical solution assumes that 'shirking' disappears once extrinsic incentives are set appropriately, specifically, the punishment fee that applies to the agent and the control costs that apply to the principal. Second is the role of PSM in overcoming shirking through intrinsic motives (Gailmard, 2010: 40), and third is job-related risk aversion, which is generally perceived as a negative trait (Chen and Bozeman, 2012) but might bring about positive results to the delegation issue. Experimental results can be summarized in three points.

First, contrary to H1a and H1b, the game-theoretical solution – an NE of mixed strategies – is a poor predictor of human behaviour in the OG. In this respect, our experimental results

corroborate prior empirical research which indicates that the NE has only a limited capacity to predict human behaviour (McCabe et al., 2000; Rauhut, 2015). Although inconsistent with the game-theoretical solution, inspectors' and inspectees' behaviour is responsive to changes in the control costs and punishment fee. However, rather than anticipating strategic interdependence, inspectors bluntly respond to higher control costs with less control, and inspectees provide higher effort with a higher punishment fee. Other experiments on the OG also find that inspectors care about the inspection costs and inspectees about the punishment fee (Nosenzo et al., 2014; Shikano et al., 2017).

Second, contrary to H2a and H2b, self-reported PSM does help to overcome the delegation issue. This is an important finding, as it contradicts the argument derived from Gailmard (2010) and may also point towards the conceptual limits of PSM (Bozeman and Su 2014; Perry and Vandenabeele, 2015). PSM's behavioural consequences appear to be restricted to the domain of cooperative games and might be less relevant in the zero-sum discoordination game that has been used to model the administrative delegation here.

Third, mostly in line with H3a and H3b, higher levels of job-related risk aversion increase inspectors' propensity to choose control in both the laboratory and survey experiments. Job-related risk aversion increases the inspectees' propensity to deliver high quality in the laboratory experiment but not in the survey experiment. These findings are affected neither by subjects' sector affiliation nor by the framing of the decision context (neutral frame versus politician/bureaucrat frame). Risk-averse attitudes and behaviour of bureaucratic agents have primarily been associated with public sector inefficiencies. Contrary to the negative image of risk aversion in public management, the experimental results suggest that job-related risk aversion can be beneficial to maintaining oversight in administrative delegation.

In methodological terms, we are confident that the findings presented in this study are not biased by experimental design choices, as both the laboratory and survey experiments offer largely consistent results. This may present good news for laboratory experiments in public administration research, which are still scant even though they offer the highest possible degree of experimental control and internal validity. Although this study is the first to test the role of PSM (Gailmard, 2010) and job-related risk aversion (Chen and Bozeman, 2012) for oversight in administrative delegation, it also comes with limitations. This study utilized a two-person OG. Future research might utilize a version incorporating a third passive player representing the public, and doing so could provide more comprehension tests of PSM in administrative delegation.

The broader implications from this study for public management are threefold. First, the promise of NPM – that delegation issues disappear with appropriate extrinsic incentives – appears to be a false promise. This study fosters the conclusion that subjects behave less strategically – or less rationally – than game theory presumes. Concerning administrative delegation as modelled in the OG, game theory does not provide a valid basis for administrative reforms in the real world. Second, PSM, which started as an antithesis of the self-interested bureaucrat, also fails to solve the issue of 'shirking'. This study cannot confirm the expectation of Gailmard (2010: 43) and others (Le Grand, 2003) that PSM can be considered a tool to solve the principal-agent problems that the public faces concerning bureaucrats. Neither a one-sided focus on extrinsic incentives nor the recruitment by and cultivation of PSM alone will be sufficient to overcome

‘shirking’ and improve public organizations’ performance. Third, most surprisingly, job-related risk aversion helps to improve administrative delegation, which presents a remarkable counterpoint to the popular opinion in public management that risk aversion is problematic. These findings should certainly not lead us to regard risk aversion as a public sector virtue. We fully agree with Chen and Bozeman (2012: 378) that public sector managers should be willing to take well-considered risks. However, the results suggest that a healthy dose of risk aversion can produce better results in stylized delegation situations like the one depicted in the OG. This suggests that rather than avoid selecting risk-averse public employees generally, more attention might be paid to the domains of administrative decision-making in which such traits can be beneficial.


### **Declaration of conflicting interests**

The authors declared no potential conflicts of interest with respect to the research, authorship and/or publication of this article.

### **Funding**

This work was supported by the Deutsche Forschungsgemeinschaft (grant number FP517/13).

### **ORCID iD**

Markus Tepe  <https://orcid.org/0000-0002-6679-0346>

### **Supplemental material**

Supplemental material for this article is available online.

### **Notes**

1. The laboratory experiment includes nine scenarios (four OGs with high payoff, four OGs with low payoff and one prisoner’s dilemma). For the purpose of this study, we only rely on the four high-payoff scenarios, which have the same parametrization of the OG as the four scenarios played in the survey experiment. In the laboratory experiment, the high- versus low-payoff treatment had no substantive effect on subjects’ behaviour in the OG.
2. The theoretical rationale and the operationalization of the strategy method is explained in further detail in Online Appendix A.
3. We use a simple additive index to aggregate the 33 PSM items into a single variable. As part of the robustness analysis, we also used factor scores resulting from a series of subdimension-specific principal component factor analyses and an additive index for the Attraction to Public Policy-making (APP) dimension of PSM. Using PSM factor scores yields similar results (see Tables 4 and 5 in the Online Appendix). Higher APP scores positively affect control and high effort in the laboratory but not in the survey experiment (see Tables 6 and 7 in the Online Appendix).
4. In addition, we explored a less sophisticated strategy: the minimax decision rule. Applied to the OG, it predicts that inspectors always choose control and inspectees always choose high quality. Both predictions are inconsistent with subjects’ behaviour in the two experiments.



## References

- Alonso P and Lewis G (2001) Public service motivation and job performance. *American Review of Public Administration* 31(4): 363–380.
- Avenhaus R (2004) Applications of inspection games. *Mathematical Modelling and Analysis* 9(3): 179–192.
- Bock O, Baetge I and Nicklisch A (2014) Hroot: Hamburg registration and organization online tool. *European Economic Review* 71(October): 117–120.
- Bozeman B and Su X (2014) Public service motivation concepts and theory: A critique. *Public Administration Review* 75(5): 700–710.
- Brown L. and Osborne SP (2013) Risk and innovation. *Public Management Review* 15(2): 186–208.
- Buurman M, Dur J and Van den Bossche S (2012) Public sector employees: Risk averse and altruistic? *Journal of Economic Behavior & Organization* 83(3): 279–291.
- Chen CA and Bozeman B (2012) Organizational risk aversion: Comparing the public and non-profit sectors. *Public Management Review* 14(3): 377–402.
- Dohmen T, Falk A, Huffman D et al. (2011) Individual risk attitudes: Measurement, determinants, and behavioral consequences. *Journal of the European Economic Association* 9(3): 522–550.
- Druckman J (2004) Political preference formation: Competition, deliberation, and the (ir)relevance of framing effects. *American Political Science Review* 98(4): 671–686.
- Esteve M, Urbig D, van Witteloostuijn A et al. (2016) Prosocial behavior and public service motivation. *Public Administration Review* 76(1): 177–187.
- Francois P (2010) Public service motivation as an argument for government provision. *Journal of Public Economics* 78(3): 275–299.
- Fischbacher U (2007) Z-Tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics* 10(2): 171–178.
- Gailmard S (2009) Multiple principals and oversight of bureaucratic policy-making. *Journal of Theoretical Politics* 21(2): 161–186.
- Gailmard S (2010) Politics, principal–agent problems, and public service motivation. *International Public Management Journal* 13(1): 35–45.
- Gailmard S and Patty J (2012) Formal models of bureaucracy. *Annual Review of Political Science* 15(1): 353–377.
- Heine F, van Witteloostuijn A and Wang TM (2021) Self-sacrifice for the common good under risk and competition: An experimental examination of the impact of public service motivation in a volunteer’s dilemma game. *Journal of Public Administration Research and Theory* 32(1): 217–232.
- Hölmstrom B (1979) Moral hazard and observability. *The Bell Journal of Economics* 10(1): 74–91.
- Houston D (2011) Implications of occupational locus and focus for public service motivation: Attitudes toward work motives across nations. *Public Administration Review* 71(5): 761–771.
- James O, Jilke S and Van Ryzin G (2017) *Experiments in Public Management Research: Challenges and Contributions*. Cambridge: Cambridge University Press.
- Kim S, Vandenabeele V, Wright BE et al. (2013) Investigating the structure and meaning of public service motivation across populations. *Journal of Public Administration Research and Theory* 23(1): 79–102.
- Klijn EH, Eshuis J, Opperhuizen A et al. (2020) Blaming the bureaucrat: Does perceived blame risk influence inspectors’ enforcement style? *International Review of Administrative Sciences*, epub ahead of print, 24 February: 1–19. Available at: <https://doi.org/10.1177/0020852319899433>

- Le Grand J (2010) Knights and knaves return: Public service motivation and the delivery of public services. *International Public Management Journal* 13(1): 56–71.
- Livnat-Lerer M, Sulitzeau-Kenan R and Kogut T (2018) Foresighted outcome effect: A micro-foundation of agents' risk aversion in principal–agent relations. *Journal of Behavioral Public Administration* 1(1): 1–10.
- 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. Analytical Methods for Social Research*. Cambridge: Cambridge University Press.
- McCubbins M and Schwartz T (1984) Congressional oversight overlooked: Police patrols versus fire alarms. *American Journal of Political Science* 28(1): 165–179.
- Miller G (2005) The political evolution of principal–agent models. *Annual Review of Political Science* 8(1): 203–225.
- Miller G and Whitford A (2007) The principal's moral hazard: Constraints on the use of incentives in hierarchy. *Journal of Public Administration Research and Theory* 17(2): 213–233.
- Nosenzo D, Offerman T, Sefton M et al. (2014) Encouraging compliance: Bonuses versus fines in inspection games. *Journal of Law, Economics, and Organization* 30(3): 623–648.
- Papke L and Wooldridge J (1996) Econometric methods for fractional response variables with an application to 401(k) plan participation rates. *Journal of Applied Econometrics* 11(6): 619–632.
- Perry J and Vandenberg W (2015) Public service motivation research: Achievements, challenges, and future directions. *Public Administration Review* 75(5): 692–699.
- Perry J and Wise L (1990) The motivational bases of public service. *Public Administration Review* 50(3): 367.
- Prokop C and Tepe M (2020) Do future bureaucrats punish more? The effect of PSM and studying public administration on contributions and punishment in a public goods game. *International Public Management Journal* 23(1): 84–112.
- Rauhut H (2015) Stronger inspection incentives, less crime? Further experimental evidence on inspection games. *Rationality and Society* 27(4): 414–454.
- Rauhut H and Winter F (2010) A sociological perspective on measuring social norms by means of strategy method experiment. *Social Science Research* 39(6): 1181–1194.
- Shikano S, Stoffel M and Tepe M (2017) Information accuracy in legislative oversight: Theoretical implications and experimental evidence. *Rationality and Society* 29(2): 226–254.
- Tepe M (2016) In public servants we trust? A behavioral experiment on public service motivation and trust among students of public administration, business sciences and law. *Public Management Review* 18(4): 508–538.
- Tepe M and Prokop C (2018) Are future bureaucrats more risk averse? The effect of studying public administration and PSM on risk preferences. *Journal of Public Administration Research and Theory* 28(2): 182–196.
- Tepe M and Vanhuysse P (2017) Are future bureaucrats more prosocial? *Public Administration* 94(4): 957–975.
- Tsebelis G (1990) Penalty has no impact on crime: A game theoretic analysis. *Rationality and Society* 2(3): 255–286.
- Weber M (1958) *Max Weber: Essays in Sociology*. New York, NY: Oxford University Press.
- Weingast B and Moran M (1983) The congressional-bureaucratic system: A principal–agent perspective. *Journal of Political Economy* 91(5): 765–800.

**Markus Tepe** is Professor of the Political System of Germany at the University of Oldenburg, Germany. His research centres on comparative public policies, behavioural public administration and political sociology.

**Susumu Shikano** is Professor of Political Methodology at the University of Konstanz, Germany. He is doing research on diverse topics in political behaviour.

**Michael Jankowski** is Interim Professor for Comparative Political Economy at the University of Bremen and a post-doctoral researcher at the University of Oldenburg, Germany. His research focuses on electoral systems, party politics and social science research methods.

**Maximilian Lutz** holds a doctorate in political economics from the University of Oldenburg and works as a data specialist in the private sector. His research focuses on decision theory, political justice and data science.