How does organizational task matter for the reputation of public agencies?

How does Organizational Task matter for the Reputation of Public Agencies?

Koen Verhoest, Department of Political Science, University of Antwerp

Jan Boon, Department of Political Science, University of Antwerp

Stefan Boye, Department of Political Science, Aarhus University

Heidi Salomonsen, Department of Management, Aarhus University

***This article is published in Regulation & Governance. Please cite accordingly: Verhoest, K., Boon, J., Boye, S. and Salomonsen, H.H. (2021), How does organizational task matter for the reputation of public agencies?. Regulation & Governance. https://doi.org/10.1111/rego.12441 ***

Abstract: The study of organizational task for understanding how organizations behave and evolve has been one of the classic topics in organization theory and public administration. Reputation scholarship has appeared as a promising perspective to understand internal and external organizational dynamics. Reputation scholars, too, emphasize the critical importance of task. Despite this recognition, the literature is characterized by a lack of theorization, and large-scale comparative analyses on how task characteristics are related to reputational dynamics. This study aims to address these concerns, relying on an extensive longitudinal dataset on the media reputation of 40 agencies in two countries to explain organizations' likelihood of attracting reputational threats (both in general, and targeting specific reputational dimensions) through different task characteristics. Our main finding is that as agencies perform tasks of a more coercive and authoritative nature (regulatory tasks and, to a lesser extent, redistributive tasks), they are more likely to attract reputational threats (both in general, and to all dimensions).

Key words: bureaucratic reputation; organizational task; public agencies

Introduction

A central tenet in organizational theory, early public administration literature and current public management literature is that the task public and private organizations perform matters significantly for how these organizations behave and evolve, in terms of both internal control and external relations (Wilson, 1989).

Originally formulated by Daniel Carpenter (2001; 2010), bureaucratic reputation theory (BRT) has evolved into a powerful and increasingly popular theory to explain the behavior of public organizations in general and regulatory agencies in particular (Carpenter, 2001, 2010; Waeraas & Maor, 2015). BRT sees agency task as a crucial element in the quest for a unique identity and reputation of these organizations. The reputational dimensions that public organizations will stress towards their audiences are assumed to relate to their task. Carpenter and Krause (2012, p. 26) state that "audience members' behaviors toward government agencies are a function of their beliefs regarding what tasks government agencies can and cannot perform effectively".

However, in current BRT studies the knowledge of how organizational task matters for the development of reputations is less extensively developed both in theoretical and empirical terms. Borrowing insights from the literature of public administration on task (Pollitt et al.,2004; Van Thiel & Yesilkagit, 2014; Verhoest et al., 2010), this study seeks to contribute to gaps in scholarly understanding of how different aspects related to organizational tasks (service delivery, regulatory, or redistributive nature of activities; task multiplicity and policy issue multiplicity as task-related control variables) affect agency reputations. A particular interest of this study is to see how agencies with a regulatory task, which have formed the

empirical background for the bulk of reputation-based studies (Boon et al., 2021; Waeraas & Maor, 2015), compare to agencies with other tasks. In doing so, we contribute to our understanding of how the activities that organizations perform feed into the reputational judgment they receive. These insights inform the growing literature on reputation as a key motivator in the behavior and control of agencies.

Our first main contribution to the theorization of the effects of agency task – being the activities performed by organizations –for bureaucratic reputation. This study theorizes and tests task to signal certain common norms that are applied in reputational judgment across otherwise different organizational contexts. For instance, regardless of policy domain, the coercive nature of regulatory activity leads these agencies to held to strict moral and procedural standards (Rimkutė, 2018). The wider organizational field of organizations performing similar functions also matters, as agencies may engage in reputation-based emulation with such other organizations (Carpenter & Krause, 2012). To date, however, it remains rather implicit in the BRT literature what aspects of organizational tasks matter for reputational judgment.

Second, existing BRT research is predominantly based on single case studies of regulatory agencies (Boon et al., 2021; Waeraas & Maor, 2015), which puts limits on the possibility to generalize findings to wider populations of government agencies that vary in terms of their task portfolio. This study relies on an extensive dataset that includes longitudinal data (10 years) on the media reputation of 40 agencies with different tasks in two European countries (Denmark and Flanders [Belgium]). We examine how task-related characteristics influence the likelihood of agencies receiving reputational threats (dependent variable 1), and the likelihood of distinct reputational dimensions (performative, technical, moral, procedural) being targeted (dependent variable 2). The argument for limiting our focus on reputational threats is two-fold. To start, reputation-based studies have so far typically focused on negative judgments that are voiced towards agencies (i.e. reputational threats), since these have been shown to be strong

motivators of organizational behavior (Carpenter, 2004; Maor & Sulitzeanu-Kenan, 2016). While previous studies have demonstrated how reputation-conscious agencies strategically respond to reputational threats, far less is known about what explains agencies' likelihood to receive reputational threats in the first place. In addition, the theoretical mechanisms explaining the likelihood of attracting reputational threats are likely to be different than those of reputational praises. The present study, therefore, will focus its attention on carefully developing a theoretical argument for the former.

Third, the notion that reputations are multidimensional and dynamic concepts is well recognized in political science and public administration (Carpenter, 2010; Carpenter & Krause, 2012; Maor, 2015). Empirical comparative contributions, however, have been scarce and mostly focused on agencies' own external communication (e.g. Christensen & Gornitzka, 2019; Christensen & Lodge, 2018; Rimkutė, 2020).

Fourth, public organizations often perform multiple tasks (Rolland & Roness, 2010; Van Thiel & Yesilkagit, 2014) and are dealing with multiple policy issues, which adds complexity, both related to how an agency's manifold audiences form their reputational judgment, but also to how agencies define the reputational image they want to project towards their audiences (Carpenter & Krause, 2012; Waeraas & Byrkjeflot, 2012). This study accounts for the role of task multiplicity and policy issue multiplicity as control variables.

Fifth, examining how task characteristics are related to reputational judgment may *indirectly* shed light on one of the more pressing concerns in contemporary BRT literature; that is: how do audiences arrive at reputational judgment (Maor, 2016). This study discusses two routes towards (negative) reputational judgment – one involving a cognitive assessment of agency results, and one involving an affective assessment that involves emotions – and theorizes how task characteristics may reflect these routes.

In terms of practical implications, this study fits in a contemporary line of research that addresses the legitimacy crisis and declining levels of trust in which public organizations now operate (Waeraas & Byrkjeflot, 2012). Negative reputations are important antecedents of distrust in organizations (Verhoest et al., 2014), particularly so when criticisms are voiced publicly in the media. Studying how their task environment predisposes certain organization towards reputational threats therefore offers important insights for practitioners to take into account in their reputation management strategies.

The article is structured as follows: first, some conceptual clarification on task-related concepts is provided; second, theory and hypotheses are developed in which task characteristics are related to the likelihood of receiving reputational threats (both in general and to specific reputational dimensions); third, the data and methods are introduced after which the results are presented and discussed.

Task as a concept

This article builds on previous studies that have applied a task-related perspective to the study of public organizations (see e.g. Laegreid et al., 2011; Pollitt et al., 2004; Verhoest et al., 2012). These studies indicate that there are significant variations in organizations' behavior depending on the characteristics of their tasks and their corresponding technical environment (Christensen et al., 2020). The tasks of an agency is to be understood as the activities and instruments an agency employs in order to achieve its goals, i.e. *what an organization does*. In this paper we use a rather simple classification of agency tasks, based on taxonomies in the literature on public organizations (Rolland & Roness, 2010), policy instruments (NATO classification by Hood, 1983), and policy types (Lowi, 1972); that is: (a) direct provision of general services (service delivery tasks), (b) regulatory tasks, and (c) redistributive tasks,

which in essence refer to the use of money transfer from and to government as policy instrument and hence encompasses granting subsidies and personal benefits, as well as levying taxes and fines. In the remainder of the article we will use the generic term 'task characteristics' to denote relevant aspects of the task-related technical environment of an organization.

This conceptualization distinguishes agency tasks clearly from other elements which also relates to the mission of an agency (Rolland & Roness, 2010; Van Thiel & Yesilkagit, 2014; Wilson, 1989), being both (a) the purposes or goals an agency seeks to achieve ("*why*" an agency exists; e.g. alleviate poverty, maintain the stability of the financial system); and (b) the values agencies seek to uphold while performing these activities ("*how*" agencies perform activities; e.g. emphasizing effectiveness, robustness, compassion, etc.). Although the BRT literature recognizes the critical importance of mission, goals, and tasks for understanding reputational dynamics, reputation scholars have dealt with the conceptualization and operationalization of these features rather ambiguously (see Boon et al., 2021 in which this argument is elaborated).

Before theorizing how task-related characteristics affect reputational judgment, we need to discuss an underlying broader, yet unresolved question: how do audiences actually form their opinion about the reputation of an agency?

Theory

Cognitive vs. affective reputational judgment

How do audiences make reputational judgments? Maor (2016, p. 82) highlights the complexity of this question when elaborating upon how audiences form their opinion about agencies. "When they do so, they bring into play a variety of factors including prior knowledge, goals, mental frames, heuristics,

distraction, motivation, emotion, and others. Ultimately, issues arise concerning what the audience will remember about the agency, what information regarding the agency will be suppressed (…), and how information that audiences rely upon in shaping their attitudes contributes to their judgment regarding the agency".

So which factors do audiences implicitly or explicitly take into account? On the one hand, literature on reputation in the private sector highlights the relevance of reputational judgment as a primarily cognitive exercise (Lange et al., 2011). An important dimension of reputation relates to the extent to which stakeholders perceive an organization as being able to produce quality outputs. Reputational judgments by stakeholders are tightly coupled to consequences and tangible organizational outputs (Lange et al., 2011). Likewise, most early reputation studies in public administration perceived the formation of an aggregate reputational judgment as a mainly cognitive exercise (cf. Maor, 2016 for an overview). On the other hand, reputation scholars are increasingly recognizing reputation formation as an affective endeavor. Maor (2016, p. 82) emphasizes the role of emotions as according to him "audience perceptions are not about physically objective reality. What they see in agencies is largely inferred, assumed and/or felt". Capelos et al. (2016) include confidence and trust as distinct affective components of reputation that require systematic evaluation alongside more cognitively based evaluations.

This study assumes that reputational judgment by audiences results from the interplay between cognitive and affective mental processes. The analytical distinction between cognitive and affective elements therefore does not reflect the real-live complexity of reputational judgment, but allows to theorize how agencies sharing certain task features will be more likely to receive reputational threats based on the presence (and interplay) of cognitive and affective processes. In the next section, we first discuss the mechanisms through which cognitive and affective processing occurs in an organizational task setting

(focusing on observability/attributability and authoritativeness/ coerciveness). Second, we relate these concepts to our concrete operationalization on the nature of task, i.e. why and how agencies performing service delivery, regulatory and redistributive tasks may be exposed to more reputational threats (both in general and particularly oriented towards the different reputational dimensions – see Figure 1).

[Please add Figure 1 here]

Cognition refers to the mental activity of processing information and using that information in judgment (Sternberg, 2011). For cognitive processes to occur, the availability of information is therefore central. In organization sciences and public administration, the notion that tasks differ in the extent to which they hold the potential for cognitive-based evaluation is linked to their observability and attributability. First, tasks differ in the extent to which both behavior and results are easily *observable* with their quantity and quality being easy assessable for audiences (Van Dooren et al., 2015; Wilson, 1989). In case of tasks with high observability, audiences can more easily gather information related to the quantity and quality of services, or they can use information provided by third parties (like audit office reports) or by the agencies (e.g. annual reports). Second, the ease to which certain results can be *attributed* to specific agencies also matters, as audiences can make stronger negative reputation claims about an agency when they are more sure that a certain result or outcome is (predominantly) due to the behavior of a specific agency (Wilson 1989; Van Dooren et al. 2015). This refers to the two features that according to Bertelli (2016) determine the extent to which bureaucratic actors can be hold democratically accountable: the identifiability of responsible actors and the evaluability of their actions and results.

Affective reputational judgment, in turn, is based on emotions. From this perspective, agencies are primarily expected to face reputational threats when they evoke negative emotions. The range of potential negative emotions is widespread, as public agencies have been described as wasteful, slow, too big, rigid,

not sufficiently transparent, and inefficient (Waeraas & Byrkjeflot, 2012). Yet the *authoritativeness* of public organizations, linked to their ability to *coerce* societal actors – effectively limiting these actors' control over their lives, is expected to be a particularly likely trigger for reputational threats. Perceived loss of control is known to be an antecedent of a range of negative emotions (Lazarus & Folkman, 1984). Agencies performing authoritative tasks often need to take unpopular decisions, which will affect certain individuals or groups negatively and will constrain them from doing things they wanted to do or force them to behave in a certain way. Agencies performing such tasks will hence elicit relatively more negative emotions, which may lead to more reputational threats (Waeraas and Byrkjeflot 2012).

How the nature of task affects reputational threats

In a last step, we use the observability and attributability (cognitive judgment) and authoritativeness and coerciveness (affective judgment) to theorize how task characteristics increase the likelihood for agencies to receive reputational threats. We use the following categorization of organizational task (Rolland & Roness, 2010; Verhoest et al., 2010):

- service delivery tasks, which refers to the direct provision or delivery of services to citizens or enterprises. Usually these agencies mainly use their own employed manpower to carry out most policy implementation, and therefore directly interact with the individual, end users when providing the service (e.g., bus driving, operation of sport infrastructure, processing a passport);
- regulation tasks, which are primarily concerned with limiting and/or controlling the behavior of individuals, enterprises, and other public agencies through the monitoring of the compliance of actors to rules, law, contract, or agreements (by supervising, scrutinizing, controlling, auditing and inspecting);

- redistributive tasks, i.e. the distribution or transfer of funds like subsidies to individuals and organizations (e.g., subsidies for green energy) and monetary benefits to individuals (e.g., unemployment benefits).

In the theorization below, we contrast (a) agencies with a service delivery task as most likely cases to observe reputational threats that originate from a strong observability and attributability of tasks and low authoritativeness and coerciveness with (b) agencies with a regulatory task as most likely cases to observe reputational threats that originate from a strong authoritativeness and coerciveness and low observability and attributability of tasks. Alternative hypotheses will be developed for the expectation that agencies with a service delivery task (H1a) and agencies with a regulatory task (H1b) will receive more threats. Agencies with a redistributive task will be conceptualized in between the "service delivery – regulatory task continuum", both in terms of cognition-based elements (observability/attributability) and affect-based elements (authoritativeness/coerciveness). Therefore, no explicit hypotheses will be formulated for agencies with a redistributive task.

Agencies with a service delivery task – Taking a cognitive perspective first, agencies with a service delivery task have been associated with behavior and results which are rather easily observable and attributable, as many would fall in Wilson's (1989) category of production organizations (see e.g. Van Dooren et al. 2015; Verhoest et al. 2010). As users receive their services directly from these agencies with a service delivery task (like garbage collection or water distribution agencies), they can assess quality and quantity more easily, and attribute negative results more straightforwardly to the agencies' behavior. Furthermore, the literature on media logic refers to rather similar criteria which increase the likelihood of organizations to be covered in the news (Strömbäck & Esser, 2014). Service-delivery agencies are often in direct contact with users when delivering their services, which makes their behaviors

more visible and recognizable to the general public (Boon et al., 2019). Because agencies with a service delivery task are more likely to be covered in the media (Schillemans, 2012), and because the media tend to prioritize negative stories over positive stories (Strömbäck & Esser, 2014), we these agencies to be particularly vulnerable for reputational threats.

Furthermore, service delivery tasks have the potential to face affect-driven assessments. While we generally expect authoritative/coercive tasks to evoke more negative emotions linked to perceived loss of agency on behalf of regulatees and citizens, the *absence* of an authoritative relation may well create its own reputational challenges. Compared to the vertical relationship rooted in authority that regulators have with their stakeholders, the position of agencies with a service delivery task vis-à-vis their stakeholders is more horizontal (Van Dooren et al. 2015; Verhoest et al., 2010). New Public Management sought to make agencies more customer-oriented. Being in a provider-customer relationship, stakeholders may *feel* as clients whose individual wishes deserve to be respected and met. Agencies with a service delivery task are more likely to have direct contact with citizens and users (Van Thiel & Yesilkagit, 2014), contributing to expectations of a more personalized and tailor-made approach. Because agencies with a service delivery task tend to be better known and be associated with expectations of a more personalized, customer-oriented and tailor-made approach, these agencies adhere to criteria of newsworthiness related to the familiarity, personalization and conflict (Schillemans, 2012); all factors which make agencies with a service delivery task more likely to appear in media articles by frames that present these agencies in a particular, incomplete and often negative light (cf. negativity bias in media reporting). Thorbjornsrud (2015), for instance, reports how agencies with a service delivery task are vulnerable to stories with frames that portray users as victims of the public sector. Furthermore, we mentioned before how the distinction between cognitive and affective judgment is mainly an analytical

one, and that "real-life" reputational judgment will likely involve an interplay between cognitive and affective processes. To give but one illustration, observable outputs can also be 'political'; that is: the outputs that an agency achieves may not be liked by everyone, even if it is what the agency is supposed to do. Based on these arguments, we formulate the following hypothesis:

> H1a: Agencies that perform service delivery tasks will be more likely to face reputational threats in the media compared to agencies which do not perform service delivery tasks.

We should at this point note that the hypotheses are not formulated in absolute but in relative terms; that is: when saying that we expect agencies with a service delivery task to be *more* likely to have observable and attributable results (as a basis for receiving reputational threats), we are not implying that agencies performing other tasks will never be held to account on the basis of performance information (indeed, failure in the control and inspection activities of regulators may well lead to substantial criticism). However, we do expect these instances to be relatively less likely to occur compared to agencies with a service delivery task (in which case information-based grounds for arriving at a reputational judgment are more readily available).

Agencies with a regulatory task – taking a cognitive perspective first, regulatory tasks are difficult to observe and attribute. First, negative regulatory outcomes (e.g. an unstable financial market or food crisis) are often not straightforwardly attributable to the actions of the agency, as these outcomes are dependent on regulatees' behavior (e.g. compliance by the regulatees) (Coglianese, 2012). Contributing to this attribution problem, the behavior of regulatees (and hence the outcomes of regulatory activities) is further influenced by numerous environmental factors beyond the regulators' control (Maggetti, 2012), including the decisions and behavior made by other regulators at different government levels or in different sectors. Agencies with a regulatory task do not operate in a void, as they are themselves

dependent upon actions of other actors in the regulatory chain or regulatory arrangement (Jordana & Sancho, 2004; Mathieu et al., 2016). Second, the outcomes of regulatory actions often fully manifest themselves only after a considerable time lag, resulting in a disconnection between immediate outputs of regulatory behavior, intermediate outcomes (e.g. compliance), and long-term outcomes (e.g. improved market functioning).

However, the absence from observable and attributable tasks creates ambiguity. Rather than protecting these agencies from information-based assessments, it could be that that it is *because* of this ambiguity that regulatory and redistributive activities will be prone to more reputational threats. In the absence of more objective grounds for evaluation, agencies with a regulatory task face the risk of being blamed for exogenous problems even when they are performing well. After all, reputations exist in the minds of the beholder, and the "information" that audiences rely upon may well be subjective – particularly as agencies' results are more ambiguous.

The main reason for expecting more reputational threats voiced towards agencies with a regulatory task follows from the affective perspective on reputational judgment. Regulatory tasks are characterized by their authoritativeness and coerciveness. After all, regulatory tasks imply the responsibility for agencies to decide which private actors or citizens are conducting 'right' or 'wrong' behavior, and which actors should be sanctioned or not (Baldwin et al., 2013). Lowi (1972) also discusses how regulatory policies focus on *immediate* coercion of individual conduct with costs concentrated with *specific* groups (regulatees), whereas expected benefits are more diffusely spread across society. The interplay of these factors is expected to spark reputational threats based on affective grounds from regulatees who directly see and feel the costs of regulation without being the main beneficiaries of the benefits.

As Waeraas and Byrkjeflot (2012) argue, this creates a charisma problem for agencies with a regulatory task, more so than for user-oriented activities, which makes them more susceptible for reputational threats. Likewise, Luoma-Aho (2008) finds that agencies with such authoritative and coercive tasks are considered to be more 'bureaucratic' than agencies with a service delivery task, evoking negative perceptions and attitudes towards them. She further argues that "where there is no responsibility, there is also less criticism" (p. 457), referring to service delivery organizations which do not need to take such sensitive decisions and which are more focused on serving users and be responsive to their individual demands.

Based on these arguments, we formulate the following alternative hypothesis:

> H1b: Agencies that perform regulatory activities will be more likely to face reputational threats in the media compared to agencies which do not perform regulatory tasks.

No explicit hypotheses are formulated for agencies with a redistributive task, who are situated somewhere between agencies with a service delivery task and agencies with a regulatory task. In terms of observability and attributability, their behaviors and results are usually more easily observable compared to those of agencies with a regulatory task (Verhoest et al. 2010). In terms of attributability, however, redistributive activities are again more akin to regulatory tasks. As agencies with a redistributive task aim to change private actors' behavior often by levying taxes or granting subsidies or personal benefits, these agencies are dependent upon the consecutive actions of these private actors to yield a positive result. In terms of authoritativeness and coerciveness, agencies with redistributive tasks need to decide which actors are entitled for certain grants or benefits or need to pay a certain amount of taxes and fines. Such agencies need to emphasize more the application of rules and decision-making standards, in order to justify their decisions. Indeed, the use of monetary instruments (levying taxes or granting financial

benefits and subsidies) aims to steer behavior in terms of desired and undesired behavior but does so in a less intrusive way, compared to the use of regulations, by leaving some decision making discretion with the involved private actor whether or not to pose that behavior. Monetary instruments mainly change the cost-benefit considerations of certain behavioral options, whereas regulation in terms of rules forcing/prohibiting certain behavior steer behavior more directly and hence evoke relatively more negative emotions (Hood, 1983).

How the nature of task affects reputational threats towards specific dimensions

In this section, we theorize how the nature of task activities makes agencies more prone to receive reputational threats towards specific reputational dimensions. While the first round of hypotheses was interested in how task affects organizations' likelihood to receive reputational threats, the focus on dimensions goes one step further as it allows to explore on which grounds these agencies are criticized. Daniel Carpenter's (2010) four-dimensional framework is widely accepted, and consists of the *performative dimension* (judgments on the quality, efficiency and effectiveness of agency outputs and outcomes); the *moral dimension* (the ability to meet the normative expectations posed to public organizations, such as protecting citizens and ensuring transparency); the *technical(-professional) dimension* (referring to the expertise and professional qualifications of the organization); and the *procedural dimension* (whether the organization conforms with set procedures and legislation).

Agencies with a service delivery task – The performative dimension centers on an evaluation of the quality of outputs and outcomes of agency behavior. Taking a cognitive perspective, we previously argued that the results of service delivery activities are more visible, observable and attributable. From this reasoning follows that we expect these agencies to be more likely to be evaluated on the performative dimension of their reputation. From an affective perspective, the more horizontal relationship between

agencies with a service delivery task (as providers) and service users (as customers) was referred to, which is expected to direct the focus of agency behavior on the quality of outputs. We therefore expect that:

> H2: Agencies that perform service delivery tasks will be more likely to face reputational threats in the media towards their performative reputation compared to agencies which do not perform service delivery tasks.

Agencies with a regulatory task – the more authoritative and coercive in nature their task is, the more agencies will be scrutinized on the extent to which they have strictly followed the rules and decision-making procedures that need to justify their behavior (i.e. procedural dimension) (Baldwin et al., 2013). The quality of decisions and processes of agencies with a regulatory task are harder to evaluate in terms of substantive accuracy, and regulatory actions are often appealed before court by operators on formal, procedural grounds (Coglianese 2012). Therefore, these agencies have been said to be mainly accountable to the extent that they have adhered to the decision making steps (like consultation of specific actors) as prescribed in regulations and in norms of sound decision-making (Baldwin et al., 2013). In addition, as agencies with a regulatory task take far-reaching decisions on individuals and groups, their audiences will expect the agency's behavior to align with high moral and ethical standards (i.e. moral dimension). For example, agencies that regulate liberalized markets will be scrutinized by both market operators and user organizations, to check whether they do offer no preferential treatment to one of the market providers. Lastly, and related to the technical-professional dimension, regulators have been shown to be judged on their technical-instrumental use of scientific knowledge and technical data, and to develop evidence-based regulation rooted in scientific expertise (Rimkuté, 2020).

We formulate no expectation concerning the relation between (a) agencies performing service delivery tasks and dimensions other than the performative dimensions, and (b) agencies performing regulatory activities and their likelihood of being criticized on the performative dimension. This may be problematic. Threats towards the other dimensions (e.g. on corruption, professional competences) may well spill over to questions on the performance of agencies with a service delivery task (Salomonsen, Boye, & Boon, 2021). Also, as we reasoned earlier, the very ambiguity of regulatory performance may lead to debates on the performative dimension in the media. In a longitudinal study on four agencies' communication strategies, Christensen and Gornitzka (2019) find that the performative dimension is the most prominently stressed dimension on regulators' websites, yet they also find it is increasingly balanced with symbols related to the other dimensions. However, this latter study does not compare the reputation management activities between regulators and non-regulators. Rimkuté (2020), also restricting her focus on communication by agencies, shows that regulators tend to stress their technical, moral and procedural reputation more than non-regulators. We choose to formulate our hypothesis in line with this limited empirical attention that is closest to the interests of this study, yet remain aware of potential contrasting mechanisms in interpreting results. We formulate the following hypothesis:

> H3: Agencies that perform regulatory tasks will be more likely to face reputational threats in the media towards the technical, moral and procedural dimensions compared to agencies which do not perform regulatory tasks.

No hypothesis is formulated for agencies with a redistributive task, who as mentioned occupy a position somewhere between agencies with a regulatory task and agencies with a service delivery task. Agencies with redistributive tasks, collecting from and allocating financial resources to their target groups, share some of the abovementioned characteristics with agencies with a regulatory task. They also have

administrative actions spelled out in protocols and regulations in order to avoid misjudgment, and are also expected to make allocative decisions according to ethical standards.

Task-related control variables

In our study we included three kinds of control variables for which we obviously do not formulate hypotheses, being task-related, organization-related and country-related control variables. However, as to the first type of task-related control variables, being task multiplicity and policy issue multiplicity, we will justify the inclusion in our analyses in this section. BRT poses that agencies will seek to establish and convey to its audiences a unique reputation, which entails the prioritization of certain aspects, tasks and goals. Yet agencies have multiple audiences which each have their own expectations in terms of what the agency needs to do and what values it needs to uphold. These expectations of these different audiences are often not aligned, implying that agencies within the constraints posed by their resources, need to make choices about which audience to please at what moment. Carpenter and Krause (2012, p. 27) indeed assert that "the audiences are multiple and diverse, so satisfying one audience often means perturbing another".

Task multiplicity and policy issue multiplicity are important variables to control for, because these variables may affect the extent of reputational threats agencies face in different ways. First, the balancing or prioritizing of demands can be argued to be much more difficult for agencies with multiple tasks and which are active on multiple policy issues.. These agencies find it difficult to portray a unique image to the outside world, because each task and policy issue the agency has to deal with will be defended by different organizational units, staff profiles, and professions within the agency (Carpenter & Krause, 2012; Wæraas & Byrkjeflot, 2012). Busuioc and Lodge (2017) make a similar point when they state that "even secondary tasks carry a potential for reputational risk when they catch public and/or media

attention" (p. 256). Carpenter and Krause (2012, p. 29) refer to the "bad apple-effect" and "least common denominator effect", meaning that a certain organizational unit with a specific role or task which accumulates a bad reputation might negatively influence the reputation of other units in the organization or of the organization in general.

Task multiplicity and policy issue multiplicity might relate to both cognitive and affective reputational judgment by audiences. A particular audience may observe that an agency handles many different tasks or that it is handling different policy issues, and hence fails to reach the quality, performance or behavioral norms this audience expects (cognitive judgment). But the involved audience might also be emotionally upset because it feels that her priorities and demands are neglected, while other audiences seems to get what they want (affective judgment). A second and alternative perspective would state that task multiplicity and policy issue multiplicity actually help agencies in balancing or prioritizing demands. The reputation management repertoire of single task agencies is restricted to relatively few audiences and arenas. Multiple tasks or policy issues which they handle may give agencies more flexibility to use a wider variety of symbols in promoting their reputation, or to engage a wider array of defensive tactics (Christensen & Gornitzka, 2019).

Data and methods

To test the hypotheses, we have selected 20 public agencies in both Denmark and Flanders (40 in total). Flanders is one of the autonomous regions of the Belgian federal system with its own parliament, cabinet and public administration, consisting of departments and agencies. Flemish public agencies exercise their competences exclusively in the Flemish region. Hence, the Flemish region can be considered a full-fledged state for the competences under its remit, due to the 'dual federalism' in the Belgian state.

Denmark and Flanders both have a parliamentary system, with similar degrees of agencification (Verhoest et al., 2012), and their population is of a similar magnitude. However, they differ in terms of their politico-administrative culture and state tradition (Denmark–Nordic and Flanders–Napoleonic). Nevertheless, since the two research contexts have news media positioned in similar media systems—the democratic corporatist system (Hallin & Mancini, 2017)—we do not expect significant differences between these research contexts.

We analyze the media audience, which serves as an intermediary between different audiences. As pointed out by Maor "…the mass media plays a key role in channeling, and sometimes even structuring, interactions between agencies and audiences" (2020:1045). Quality media outlets, as the ones examined here, are expected to discuss stories from several perspectives and to give the floor to several sources. This is expected to lead to the inclusion of both cognition-based and affect-based interpretations of agencies' functioning. The empirical focus on the traditional (written) news media is not uncommon in empirical studies on bureaucratic reputation (Maor et al., 2013; Gilad et al., 2013). Traditional news media are a critical source of mediated information about agencies, and serve as an important informal accountability forum in which agencies provide accounts of their behavior to the general public (Jacobs and Schillemans 2016), with important repercussions for agency reputations (Busuioc & Lodge, 2017). The media, therefore, have an active influence on public opinions, including their judgment about organizations reputations. As a result, an organization's media reputation is largely reflected in the reputation held by other audiences, such as the general public. Media attention has proven to be of vital relevance to agency survival due to the (positive) relationship between media attention and organizational survival (Bertelli and Sinclair 2015).

The specific 40 agencies were selected from the population of Danish and Flemish agencies on the basis of their salience in the media, their type and task both to ensure variation in the primary independent variable, agency task, and to have agencies that are representative for the population of agencies in Denmark and Flanders. Furthermore, to ensure functional equivalence between the countries , Danish agencies and Flemish agencies were paired so that each pair is to a considerable degree alike in terms of task and type (e.g., both the Danish and Flemish roads and traffic agency were selected).

Our strive to ensure functional equivalence relates to the recent concern raised by Carpenter in terms of the need for being cautious when categorizing agencies, as similar or different, not least for the purpose of conducting quantitative research. Carpenter points to the relevance of ensuring that agencies are compared which "…perform a common or similar function or inhabit what sociologists call an organizational field…" (Carpenter 2020:92).

For the 40 agencies we have read all articles in two national-wide quality newspapers, 'Berlingske' in Denmark and 'De Standaard' in Flanders in a 10-year-period from 1/01/ 2006 till 31/12/2015. Both newspaper are comparable in terms of political orientation (center-right), and considered as leading broadsheet for national, political, foreign, economic, and social news. Being somehow more to the right than to the middle of the political ideological spectrum, one potential bias could be that those newspapers are compared to more center-left newspapers more critical in their coverage of agencies with a regulatory task. Apart from two agencies in Flanders where sampling occurred due to the high number of articles, the full population of articles mentioning one or more of these agencies was coded.

All articles for each agency were found by using the full and official name of the agency, part of the full and official name of the agency, full name of agency and extra key words. In a first step, each article was read by one of the authors (N = 13,387), and coded according to whether the article involved an explicit

negative opinion of the agency (i.e. *reputational threat*). In a second step, we have only coded the explicit threats or praises on several variables, including the reputational dimensions.

Interpreting and coding Carpenter's four-dimensional framework has been a challenging endeavor over the past years. As the original conceptualization – developed for usage on a single regulatory agency – leaves some room for interpretation when applied to other organizational contexts, it is paramount for reputation scholars to be transparent in their coding decisions. The following operationalization of the reputational dimensions was used in this study:

[Please include Table 1 here]

The coding of a reputational threat is based on all of the parts of an article in which the agency is discussed negatively by one or more of its audiences. Each article may contain multiple threats. "Negativity" includes both cognitive references to e.g. lack of efficiency or effectiveness by the agency, but also affective references through negatively charged words such as "bad" or "terrible". The coding is binary; if the article contains one or more threats, a value of 1 is assigned (and 0 otherwise). If a reputational threat applied to more than one dimension (for instance: an agency that breaches formal rules about ethical behavior), then each dimension to which the threat applied was coded (in the example here: both the procedural and the moral dimension).

To ensure reliable coding across the two research contexts, a combination of best practice strategies was made (Lacy et al., 2015); see supplementary material for a detailed overview. Overall, 9.1% (N = 1,217) of the 13,291 articles in the sample contain a reputational threat. There are considerable differences in the relative amount of threats the agencies are subjected to, ranging from 0.0% (minimum relative share of threats) to 35.5% (maximum relative share of threats). Likewise, there is a country difference as 8.5%

of the articles contain a threat in the Danish sample, while 10.0% of the articles contain a threat in the Flemish sample.

Our first dependent variable - *likelihood of reputational threats* - is operationalized as the relative share of articles in which a reputational threat is voiced towards the involved agency in relation to the total number of articles in which the involved agency was mentioned in given year.[2] So this variable take into account that some agencies might receive much media coverage while others a low coverage. The second dependent variables – *likelihood of reputational threats towards specific dimensions* – are measured as the 'share of reputational threats towards a specific dimension (be it performative, technical, moral, procedural) on the total of reputational threats voiced towards the involved agency' in a given year.

Task-related independent variables

First, we constructed three binary variables that indicate whether an agency has a *service delivery task in its task portfolio* (yes/no), a *regulatory task in its task portfolio* (yes/no), and/or a *redistributive task in its task portfolio* (yes/no). While an agency may have several of these tasks, multicollinearity checks show no issues of multicollinearity and indicate sufficient independence of these variables. Second, in order to shed light on the impact of particular combinations of task categories, we created a categorical variable with all potential combinations of task categories ('solely regulatory task'; 'regulatory task + service delivery task'; 'service delivery task + redistributive task' or 'service delivery task + regulatory task + redistributive task') as values.

Control variables

We include two task-related control variables in the analyses. *Multiplicity of tasks* is operationalized in a dummy, measuring whether the agency have a mixed task portfolio. This means whether or not the agency combines two or more kinds of tasks (e.g. an agency with both a regulatory and service delivery

task). *Multiplicity of policy issues* is measured by how many policy issues the agency handles. We used the Classification of the Functions of Government (COFOG) published by the UN Statistical Division. The coding is done by using the most detailed level of COFOG, being the third level of Classes. We have constructed a dummy variable taking the values 0 (single policy issue) or 1 (multiple policy issues). The coding of tasks and policy issues was done based on task descriptions on the agencies' websites and annual reports.

We also include organizational control variables which are common in agency studies and for the country context. First, political salience is a well-recognized control variable in the agency literature (Pollitt et al., 2004; Van Thiel & Yesilkagit, 2014). Scholars argue that agencies that are politically salient will face more accountability demands (Wilson, 1989). Whether the increased attention for salient agencies translates into more reputational threats, however, is not straightforward. Salient agencies, after all, may have more power and can use diverse instruments of reputation that place them in a good position to impress the media and the public (Boon et al., 2019; Fredriksson et al., 2015). This study uses *parliamentary attention* as a measurement of political salience. Parliamentary attention is measured by the number of mentions of the agency in written questions in parliament on an annual basis. The log value is used in the analyses as the distribution of the variable is highly skewed.

Second, we rely on Van Thiel's (2012) classification of agencies' *legal type*, focusing on Type 1 agencies (departmental agencies without own legal identity) and Type 2 agencies with their own legal identity. Third, the analyses control for *agency size* as a measurement of agency capacity, measured as full time equivalents (FTE) and collected as time-variant data on FTE in the agencies' annual reports. The data on *size* was log transformed because of skewness. Fourth, *agency budget* is measured by the agency's grand total budget expenditure measured in millions euro. Being a measurement of the financial resources an

agency has expended in that year, it is a measurement of salience in terms of the amount of resources a government wants to spend on the tasks deal with by the agency (Pollitt et al. 2004). Time-variant data on budget were collected in the agencies' annual reports and were log transformed because of skewed distributions. We also include a *country dummy* to control for country differences. Finally, for the models of reputational threats, we have added year-dummies to take account for aggregate time trends.

Table 2 and Table A1 (online annex) show the descriptives and correlation matrix of the dependent variables, independent variables and control variables for the full sample. The data is structured in agency-year.

[Please include Table 2 here]

Estimation methods

We investigate the percentage of threats relative to the number of articles for each agency on a yearly basis. In sum, we have 393 agency-year observations (two agencies were established during the investigated period). Most years, the share of threats on total number of articles is quite low: for 140 out of 393 agency-years (35.6%) in the data the share of threats is zero (meaning no article in that year included a threat towards the agency's reputation). To accommodate for this, we first transform the dependent variables by using the log + 1 of the distribution. Second, we apply Tobit models with specified upper and lower limits to take account of the censored dependent variable (Wooldridge, 2018, pp. 186-188, 574-575). As the dependent variable is log-transformed, (exp(x)-1)*100 is applied to get the change (in percentage) in the proportion of threats at a given year. The data has panel-structure and we present random effects Tobit models. Standard errors are clustered by agency to take account for autocorrelation. As seen in Table A1 (in online annex), the pairwise correlation between political

salience, FTE, and expenditures sometimes reaches significance, which could result in problems of multicollinearity. Thus, we have calculated a mean variance inflation factor of 3.55, with political salience = 4.14, FTE = 3.97, and expenditures = 3.97. We do not consider problematic (Wooldridge, 2018, p. 92). Lastly, analyses were run on reputational threats and praises combined as dependent (see online supplementary materials tables A4-A6-A7), which provided highly similar results. The online supplementary materials also hold a wide range of robustness checks with analyses that shed light on potential country differences (Tables A8-A13), model with cross-section data only with averages for the variables over time (Tables A14-A15), supplementary multilevel logistic regression analyses with articles as the level of analysis and standard errors clustered at agency-level (Tables A16-A17), analyses with year-agency standard deviation (instead of the mean) on the reputational threats as the dependent variable to investigate if the variance of reputational threats varies across agencies (Tables A19-A22).

Results

We will first present the results analyzing the proportion of reputational threats. Table 3 shows the results of the regressions on the full sample with the logged percentage of threats as the dependent variable. Table 3 shows the analysis with binary task variables.

[Please include Table 3 here]

H1a expected more reputational threats to be voiced towards agencies with a service delivery task. Instead, the results show that agencies with a service delivery task in their task portfolio do not have a higher likelihood of reputational threats compared to agencies without a service delivery task in their task portfolio. Having a service delivery task compared to not having the task portfolio in fact reduces the proportion of reputational threats by 10 percent (i.e., (exp(-0.107)-1)*100 = -10.07). Table 3 further

shows no effect of having a regulatory task, and a negative significant effect of having a redistributive task (albeit at $p < 0.10$).

While Table 3 sheds a first light on the relations between task categories and the likelihood of attracting reputational threats, further analyses are needed. To start, the reference categories used in Table 3 allow to compare 'having a certain task' (e.g. service delivery) vs. 'not having that task', but do not allow to compare task categories to each other (e.g. regulation vs. redistribution). Furthermore, a considerable share of the agencies included in our dataset does not perform one task but multiple tasks. Therefore, Table A2 (in online annex) sheds light on how specific combinations of tasks in the agencies' task portfolio affect the likelihood of attracting reputational threats. In these analyses, 'solely service delivery task' is used as a reference category.

Table A2 demonstrates that agencies with solely a regulatory task have a higher likelihood of reputational threats compared to agencies with solely a service delivery task. Similarly, agencies with solely a redistributive task have a higher likelihood of reputational threats compared to agencies with solely a service delivery task. The difference amounts to 9% increase in the proportion of threats in a given year. This suggests that the theoretical effect of 'authoritative and coercive nature' of task (cf. H1b) primarily holds when comparing agencies with solely a regulatory or redistributive tasks with agencies with solely service delivery tasks. Additionally, agencies that combine a regulatory task with either a service delivery task or with a redistributive task have a higher likelihood of facing reputational threats compared to agencies with solely a service delivery task, while agencies combining service delivery and redistributive tasks, or a combination of all three tasks, do not have a higher likelihood of reputational threats.

Our analyses control for a series of variables. We return to Table 3 to discuss these effects (while not presented in Table A2 (in online annex) in order to ease readability, the variables were also run and led

to highly similar effect sizes and significance levels). Agencies with task multiplicity or policy issue multiplicity were not found to attract more reputational threats (in fact, the opposite holds for policy issue multiplicity). Parliamentary attention has a strong and positive effect on the likelihood of reputational threats. None of the other control variables (country dummy, agency type, agency size, agency budget) have an effect.

In a second round of analyses, we tested the formulated hypotheses on the likelihood that agencies would face reputational threats geared towards a specific reputational dimension (performative, technical, moral and procedural). Table 4 and Table A3 (in online annex) show the results for each of the different reputational dimensions. Again, Table 4 (presented in the paper) shows the analysis with binary task variables. Table A3 (in online annex) gives insight into the effect of particular task combinations, with 'solely service delivery task' as reference category.

[Please include Table 4 here]

To discuss the hypotheses on the relations between task categories and reputational threats to specific dimensions, we rely on Tables 4 and A3 which together give a rather complete picture. H2 formulated a positive relation between agencies with a service delivery task and reputational threats towards the performative dimension. Tables 4 (negative coefficient 'service delivery task') and Table A3 (positive coefficients of other task categories vis-à-vis 'solely service delivery task' as reference category) refute this claim, which leads us to reject H2.

Next, H3 expected a positive effect of agencies with a regulatory task on reputational threats towards the technical, moral and procedural dimension. Table A3 (in online annex) supports this expectations regarding the technical dimension. Both pure regulators (i.e. without other tasks) and agencies that

combine regulatory with redistributive tasks are more likely to get criticized on the technical dimension compared to agencies which solely perform service delivery tasks. Table 4, however, shows no significant effect between agencies with a regulatory task and agencies without a regulatory task, which probably has to do with the absence of a significant difference between agencies with a regulatory task and agencies with a redistributive task (this difference was also not hypothesized). Agencies with redistributive tasks have no increase in threats to the technical dimension compared to agencies with solely a service delivery task. Concerning the moral dimension, both Table 4 and Table A3 support the expectation that regulatory tasks attract reputational threats to the moral dimension (again no such effect is found for redistributive tasks). Concerning the procedural dimension, Table A3 (in online annex) supports the notion that agencies with regulatory tasks are more criticized on their procedural correctness (particularly when compared to agencies with solely a service delivery task). This holds both agencies with solely regulatory tasks and when regulatory tasks are combined with either a service delivery or redistributive task (but not both). Taken together, these observations lead us to largely accept H3.

Concerning control variables, Table 4 shows that task multiplicity has little significant effects (it is only positively related to threats to the procedural dimension). Policy issue multiplicity leads to less reputational threats across all dimensions. Parliamentary attention leads to more reputational threats across all dimensions. Budget only has a significant positive effect on threats to the moral dimension. The remaining control variables show no significant results, except in the model on threats to the technical dimension where Flemish agencies (negative effect) and type-2 agencies (positive effect) reach significance.

Discussion and conclusions

This article sought to contribute to the literature by shedding light on the relations between a series of task-related variables and agencies' likelihood to attract reputational threats (both in general and towards reputational dimensions). After conceptually distinguishing task from related concepts, we borrowed insights from the agency literature to discuss different features related to the activities agencies perform, and developed hypotheses based on the extent to which these features are likely to trigger an interplay of cognitive (based on observability and attributability of behavior and results) and affective (based on the coercive/authoritative nature of performed activities) mechanisms.

Our findings are suggestive of some main conclusions. First, concerning the task categories (service delivery, regulation, redistribution), agencies performing regulatory and redistributive tasks were more likely to be confronted with reputational threats compared to agencies that perform service delivery tasks. This finding rejects H1a and supports H1b. Turning to the reputational dimensions, the expected effect concerning service delivery tasks – being positively related to threats towards the performative dimension (H2) – was again rejected. Supporting H3, agencies with regulatory tasks were more likely to be criticized on the technical, moral and procedural dimension. Furthermore in terms of task-related control variables, the results show virtually no effect of task multiplicity and a consistent negative effect of policy issue multiplicity on attracting reputational threats. Political saliences measured by parliamentary attention leads to more threats across dimensions and agency budget led to more threats in relation to the moral dimensions. The other organizational and country controls did not have any effects.

Some of our observations run counter existing studies that found agencies with a service delivery task to be more likely to attract media attention (Schillemans, 2012), including negative media attention (Boon, Salomonsen & Verhoest, 2018). We see different potential explanations. First, studies differ in terms of

the research context (country, government level and media systems and outlets). For instance, if the focus is on the national government level of a country that has decentralized substantially to lower government tiers, leaving primarily regulatory activities at the national government level (such as the Danish context that forms a substantial part of this study), this could signal to audiences that regulatory activities should be the main target of scrutiny. Second, there might be methodological explanations. The descriptive statistics (cf. Tables 2 and 3) in the present study show that service delivery in absolute numbers actually attract most negative media attention, yet given that our operationalization uses a relative measure of threats (as a portion of *all* media attention) this high number is offset by the even higher number of neutral articles among agencies with a service delivery task. Studies that take different approaches to dealing with this distinction between absolute vs. relative amount of threats might find different results. Furthermore, the analytical strategy is important. For instance, Qualitative Comparative Analysis studies (e.g. Boon, Salomonsen, & Verhoest, 2018) examine how explanatory factors *combine* to produce certain outcomes, whereas the present study uses statistical methods to analyze the *independent* effect of each explanatory factor. Lastly, compared to previous studies, the present study controls for a variety of factors which might offset certain bivariate effects. For instance, the correlation matrix (Table A1 – in online annex) shows that some bivariate correlation disappears when controlling for additional task elements. To summarize, therefore, this study speaks to the idea that agency task is a relevant and intriguing factor to look upon, and calls upon future studies to continue research into task using more refined groupings and categorizations in different institutional contexts.

These findings contribute to the literature in several ways. First, this study used a large dataset that includes longitudinal data (10 years) on the media reputation of 40 agencies that differ in a series of characteristics (incl. task) in two European countries; and hence brings a complementary perspective to

the focus towards regulatory, non-European and single case studies in contemporary reputation studies (Boon et al., 2021; Waeraas & Maor, 2015). Our case (and country) selection strategy sought to (a) create sufficient variation in terms of organizational characteristics, while (b) minimizing potential country effects by selecting countries of similar size, degree of agencification and media system. In terms of the generalizability of our findings, we would expect that our findings are generalizable to contexts which similar degrees of agencification as Denmark and Flanders covered by news media positioned in a democratic corporatist system.

This study points at the relevance of what an agency does (task) as a variable to explain reputational threats and hence adds to BRT. It is among the first to examine – and confirm – that agencies with a regulatory task are more likely to attract (all types of) reputational threats by the media, a finding that provides food for thought not only for the reputation literature and agency heads of agencies with a regulatory task interested in cultivating their reputation, but also for scholars interested in the behavior, governance and control of regulators. Future research may pursue the question of whether this is also the case for the likelihood of attracting reputational praises by the media. For example, agencies with service delivery tasks might get a relatively higher share of reputational praises, compared to agencies without such tasks, just because such agencies invest more in good performance due to the observability and attributability of their task (Bertelli, 2016). Due to the scarcity of reputational praises, we were not able to run any analyses with solely reputational praises as a dependent, yet analyses where we combined reputational threats and praises (see supplementary materials tables A4-A5-A6-A7) provided highly similar results.

In addition our empirical focus was on agencies' *media* reputation. The media acts both as an audience in its own right and as an intermediary between agencies and their other audiences, and is hence an

important audience to study. However, future studies may further explore whether and how the media indeed affects the reputational perceptions of other audiences. For instance, it could be that audiences with direct contact with the agency may give less importance to mediated perceptions and judgments of the agency relative to their own direct experience.

Second, this study is the first to apply Carpenter's four-dimensional framework to examine the content of reputational threats towards different types of agencies (with different tasks). Empirical contributions on the dimensionality of reputation have been scarce and mostly focused on agencies' own external communication in annual reports (Rimkuté, 2020) or on their websites (Christensen et al., 2020; Christensen & Lodge, 2018). One interpretation of our findings is that most of the expected relations related to regulatory reputations – i.e. more oriented towards the technical, moral and procedural – are confirmed (Overman et al., 2020; Rimkutė, 2020). Another interpretation is that there is limited variation in how task-related factors affect reputation and the reputational dimensions across the models. For instance, regulatory activities lead to an increase in threats on *all* dimensions, including the performative dimension which was unexpected given the challenges of observing and attributing regulatory outcomes, but which might be exactly related to the ambiguities surrounding the results of regulatory agencies (Coglianese, 2012; Jordana & Sancho, 2004; Mathieu et al., 2016). This finding may also corroborate insights from Overman, Busuioc and Wood's (2020) recent study, arguing that audiences further away from an organization are less likely to make dimension-specific judgments, but will resort to general agency assessments. Our findings may indicate that this is true for the media – as an intermediary audience – as well, since reputational judgment towards similar agencies (in terms of task) are largely consistent across reputational dimensions.

Third, our results should inspire future research on the underlying mechanisms of the observed effects. Our theoretical framework formulated a series of cognitive processes – e.g. the ambiguity of regulatory performance that may expose these agencies to (undeserved?) criticism – and affective processes – e.g. the coercive and authoritative nature of regulatory behavior that evokes negative emotions – that may be at the basis of the higher likelihood to of regulators to face threats. These mechanisms, however, could not be directly tested with our research design. Future studies could measure directly what explains the observed negativity bias towards regulators, focusing on the interplay between cognitive and affective processes. Experimental designs, focus groups or more refined content analyses offer interesting approaches to delve into the role of emotions more directly. Moreover, future studies should also focus on what the consequences might be of certain agency tasks being more prone to reputational threats. For example, our findings might speak to the literature on the survival/non-survival of agencies, which is at least sometimes an extreme consequence of reputational damage. For example, Greasley and Hanretty (2016) point at the impact of task on agency termination, and find that in the UK agencies created with a credible commitment rationale, like regulators, are less likely to be terminated compared to other agencies. Combined with the findings of our study, this might raise the interesting question why agencies with a regulatory task which face comparatively more reputational threats are nevertheless able to survive longer than other agencies.

Lastly, for reputation scholars, our findings on the effect of the task-related control variables, being task and policy issue multiplicity, offer food for thought, given that the notion of how to deal with multiple task environments forms an important part of BRT, yet has been scarcely empirically examined using large N quantitative data. Contrary to expectations, no effect was found for task multiplicity (though we do observe interesting variations in the models with distinct combinations of task categories). In addition,

the observations of a consistent negative relation between policy issue multiplicity and reputational threats further strengthens the claim that multiplicity - at least in terms of the policy issues agencies handle - assists rather than obstructs agencies in their reputation management. This might suggest that dealing with multiple policy issues gives agencies the flexibility to stress certain strengths towards specific audiences; however future studies should look further into the exact nature of the relation between task/policy multiplicity and reputational dynamics. For instance, studies could account for, on the one hand, the extent to which the different policy issue an agency has to attend to are fundamentally different in terms of audiences and, on the other hand, for potential interaction effects of agency capacity (e.g. agency size). One might expect that the extent to which agencies with multiple tasks or policy issues will face higher chances of reputational threats also depends on their staff capacity to invest sufficient efforts in their different tasks.

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

The research was funded by the Danish Research Council (grant 1182-00110B), and by the Flemish Research Council (grant 1244720N). The research is part of the GOVTRUST Centre of Excellence at the University of Antwerp.

References

Baldwin, R., Cave, M., & Lodge, M. (2013). Understanding Regulation: Theory, Strategy, and Practice, 2nd Edition (2nd edition). Oxford University Press.

Bertelli, A. M. (2016). Who Are the Policy Workers, and What Are They Doing? Citizen's Heuristics and Democratic Accountability in Complex Governance. Public Performance & Management Review, 40(2), 208–234.

Boon, J., Salomonsen, H. H., & Verhoest, K. (2019). The Effects of Organisational Features on Media Attention for Public Organisations. Policy & Politics, 47(2), 245–264.

Boon, J., Salomonsen, H. H., & Verhoest, K. (2021). A reputation for what, to whom, and in which task environment: A commentary. Regulation & Governance, 15(2), 428–441.

Busuioc, E. M., & Lodge, M. (2017). Reputation and Accountability Relationships: Managing Accountability Expectations through Reputation. Public Administration Review, 77(1), 91–100.

Capelos, T., Provost, C., Parouti, M., Barnett, J., Chenoweth, J., Fife-Schaw, C., & Kelay, T. (2016). Ingredients of institutional reputations and citizen engagement with regulators. Regulation & Governance, 10(4), 350–367.

Carpenter, D. (2001). The Forging of Bureaucratic Autonomy: Reputations, Networks, and Policy Innovation in Executive Agencies 1862-1928. Princeton University Press.

Carpenter, D. (2004). Protection without Capture: Product Approval by a Politically Responsive, Learning Regulator. The American Political Science Review, 98(4), 613–631.

Carpenter, D. (2010). Reputation and power: Organizational image and pharmaceutical regulation at the FDA. Princeton Univ. Press.

Carpenter, D., & Krause, G. A. (2012). Reputation and public administration. Public Administration Review, 72(1), 26–32.

Christensen, T., & Gornitzka, Å. (2019). Reputation Management in Public Agencies: The Relevance of Time, Sector, Audience, and Tasks. Administration & Society, 51(6), 885–914.

Christensen, T., Lægreid, P., & Røvik, K. A. (2020). Organization Theory and the Public Sector: Instrument, Culture and Myth (2nd edition). Routledge.

Christensen, T., & Lodge, M. (2018). Reputation Management in Societal Security: A Comparative Study .The American Review of Public Administration, 48(2), 119–132.

Coglianese, C. (2012). Evaluating the Impact of Regulation and Regulatory Policy. OECD. https://www.oecd.org/gov/regulatory-policy/1_coglianese%20web.pdf

Fredriksson, M., Schillemans, T., & Pallas, J. (2015). Determinants of Organizational Mediatization: An Analysis of the Adaptation of Swedish Government Agencies to News Media. Public Administration, 93(4), 1049–1067.

Greasley, S., & Hanretty, C. (2016). Credibility and Agency Termination Under Parliamentarism. Journal of Public Administration Research and Theory, 26(1), 159–173.

Hood, C. C. (1983). The Tools of Government. Macmillan Education UK. https://doi.org/10.1007/978-1-349-17169-9

Jordana, J., & Sancho, D. (2004). Regulatory Designs, Institutional Constellations and the Study of the Regulatory State. In In: Jordana J,Levi-Faur D (eds) The Politics of Regulation: Institutions and Regulatory Reforms for the Age of Governance (pp. 296–319). Edward Elgar.

Lacy, S., Watson, B. R., Riffe, D., & Lovejoy, J. (2015). Issues and Best Practices in Content Analysis. Journalism & Mass Communication Quarterly, 92(4), 791–811. https://doi.org/10.1177/1077699015607338

Laegreid, P., Roness, P. G., & Verhoest, K. (2011). Explaining the innovative culture and activities of state agencies. Organization Studies, 32, 1321–1347.

Lange, D., Lee, P. M., & Dai, Y. (2011). Organizational Reputation: A Review. Journal of Management, 37(1), 153–184.

Lazarus, R. S., & Folkman, S. (1984). Stress, Appraisal, and Coping (1st edition). Springer Publishing Company.

Lowi, T. J. (1972). Four Systems of Policy, Politics, and Choice. Public Administration Review, 32(4), 298–310.

Luoma-aho, V. (2008). Sector reputation and public organisations. International Journal of Public Sector Management, 21(5), 446–467.

Maggetti, M. (2012). The media accountability of independent regulatory agencies. European Political Science Review, 4(3), 385–408.

Maor, M. (2016). Missing Areas in the Bureaucratic Reputation Framework. Politics and Governance, 4(2), 80–90.

Maor, M., Gilad, S., & Ben-Nun Bloom, P. (2013). Organizational Reputation, Regulatory Talk, and Strategic Silence. Journal of Public Administration Research and Theory, 23(3), 581–608.

Maor, M., & Sulitzeanu-Kenan, R. (2016). Responsive Change: Agency Output Response to Reputational Threats. Journal of Public Administration Research and Theory, 26(1), 31–44.

Mathieu, E., Verhoest, K., & Matthys, J. (2016). Measuring multi-level regulatory governance: Organizational proliferation, coordination, and concentration of influence. Regulation & Governance, https://doi.org/10.1111/rego.12127

Overman, S., Busuioc, M., & Wood, M. (2020). A Multidimensional Reputation Barometer for Public Agencies: A Validated Instrument. Public Administration Review, https://doi.org/10.1111/puar.13158

Pollitt, C., Talbot, C., Caulfield, J., & Smullen, A. (2004). Agencies: How Governments Do Things through Semi-autonomous Organizations. Palgrave Macmillan.

Rimkutė, D. (2018). Organizational reputation and risk regulation: The effect of reputational threats on agency scientific outputs. Public Administration, 96(1), 70–83.

Rimkutė, D. (2020). Building organizational reputation in the European regulatory state: An analysis of EU agencies' communications. Governance, 33(2), 385–406.

Rolland, V. W., & Roness, P. G. (2010). Mapping Organizational Units in the State: Challenges and Classifications. International Journal of Public Administration, 33(10), 463–473.

Schillemans, T. (2012). Mediatization of Public Services: How Organizations Adapt to News Media. Peter Lang.

Sternberg, R. J. (2011). Cognitive Psychology (6th edition). Cengage Learning. Belmont, CA.

Strömbäck, J., & Esser, F. (2014). Meditization of Politics: Towards a Theoretical Framework. In Mediatization of Politics: Understanding the Transformation of Western Democracies (pp. 3–28).

Thorbjørnsrud, K. (2015). Mediatization of Public Bureaucracies: Administrative versus Political Loyalty. Scandinavian Political Studies, 38(2), 179–197.

Van Dooren, W., Bouckaert, G., & Halligan, J. (2015). Performance Management in the Public Sector (2nd edition). Routledge.

Van Thiel, S., & Yesilkagit, K. (2014). Does task matter? The effect of task on the establishment, autonomy and control of semi-autonomous agencies. International Review of Administrative Sciences, 80(2), 318–340.

Verhoest, K., Roness, P., Verschuere, B., Rubecksen, K., & MacCarthaigh, M. (2010). Autonomy and Control of State Agencies: Comparing States and Agencies (2010 edition). Palgrave Macmillan.

Verhoest, K., Van Thiel, S., Bouckaert, G., & Laegreid, P. (2012). Government agencies: Practices and lessons in 30 countries. Palgrave Macmillan.

Verhoest, K., Rommel, J., & Boon, J. (2014). How organizational reputation and trust may affect the autonomy of independent regulators: The case of the Flemish energy regulator. In Organizational reputation in the public sector (eds. A. Waeraas & M. Maor) (pp. 118–139). Routledge.

Waeraas, A., & Byrkjeflot, H. (2012). Public Sector Organizations and Reputation Management: Five Problems. International Public Management Journal, 15(2), 186–206.

Waeraas, A., & Maor, M. (2015). Organizational Reputation in the Public Sector. Routledge.

Wilson, J. Q. (1989). Bureaucracy: What Government Agencies do and Why they do it. Basic Books.

Endnotes

[1] We also performed the analysis with the absolute number of reputational threats as dependent variable. Results which are not included in this article, are very similar to the results reported in this article.

List of legends

Tables to include

Figure 1: Conceptual model

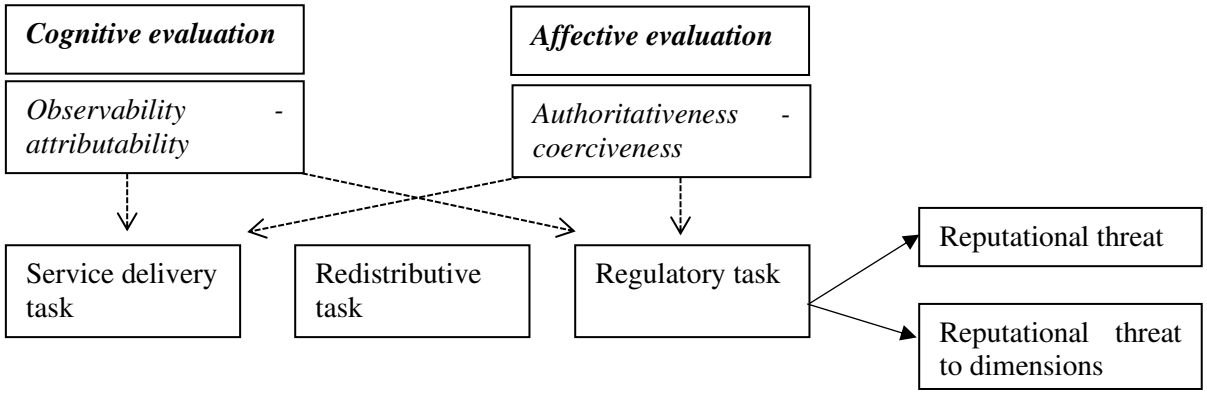

Table 3: Operationalization reputational dimensions

Dimension	Operationalization
Performative dimension	Output/outcome-oriented; addresses judgements on the quality, efficiency and/or effectiveness of the services that are considered an agency' outputs and outcomes. These services can be a set of activities, policy instruments, but also an initiative, a program, a report, that is the final manifestation of the agency's core task that is delivered to society, to politicians or to other public actors; It also includes references to the agency pursuing the right objectives "performing the right outputs" and the agency having a good/ambitious vision, or correct prioritization of task and goals.
Technical dimension	Input-oriented: refers to the adequacy and sufficiency of the inputs of the agency such as: staff, equipment, resources (e.g. money, organizational structure, or basic ICT systems), qualified personnel, leadership in terms of leaders competencies, management structure, internal capacity and structures, in-house expertise, knowledge, skills. It also captures the efficacy and efficiency of internal processes and steps which are not directly reflecting the final manifestation of the agency's task delivered to society, but which are 'merely' serving as means (or preliminary steps) to realize those ends.
Moral dimension	refers to agencies' well-meaning (regardless of actual output); attentiveness and compassion to different client groups; honesty; integrity; fairness; ethical behavior; openness and transparency; ability to prevent inequity, bias, and abuse of office; turf protection; trustworthiness, scandals etc.; indecent

	behavior. It also includes references to the agency's attentiveness to democratic values (e.g. transparency, equal rights, neglect of legal rights or responsiveness to citizens in terms of citizens being heard in processes and procedures, equal access to service delivery for different groups).
Procedural dimension	references to the justness of the processes by which its behavior is generated; procedures, standards, norms and rules (which can be internal but also external to the organization, e.g. the constitution) that were followed or not followed.

Table 4: Descriptives

	Variable	N	Mean	Std. dev.	Min	Max
1	Relative threats (logged)	393	0.075	0.096	0.000	0.511
2	Relative performative threats (logged)	393	0.058	0.084	0.000	0.511
3	Relative technical threats (logged)	393	0.012	0.034	0.000	0.373
4	Relative moral threats (logged)	393	0.016	0.036	0.000	0.336
5	Relative procedural threats (logged)	393	0.012	0.038	0.000	0.333
6	Task: Regulation	40	0.575	0.501	0.000	1.000
7	Task: Service delivery	40	0.700	0.464	0.000	1.000
8	Task: Redistributive	40	0.475	0.506	0.000	1.000
9	Multiple tasks	40	0.575	0.501	0.000	1.000
10	Type 2 agency	40	1.400	0.496	0.000	1.000
11	Multiple policy issues	40	0.350	0.483	0.000	1.000
12	Parliamentary attention	393	54.987	97.341	0.000	788
12	Parliamentary attention (logged)	393	2.713	1.804	0.000	6.671
13	Agency size: Full time equivalent	393	650.708	856.099	11.000	5054.000
13	Agency size: Full time equivalent (logged)	393	5.862	1.174	2.398	8.528
14	Budget: Expenditures in Euro	393	169.273	305.809	2.212	3500.137
14	Budget: Expenditures in Euro (logged)	393	4.125	1.483	0.794	8.161

Table 3: Reputational threats, with binary task variables

	(1)
Service delivery task	-0.106**
	(-2.894)
Regulatory task	0.046
	(1.475)
Redistributive task	-0.059+
	(-1.878)
Multiple tasks	0.021
	(0.499)
Multiple policy issues	-0.052*
	(-2.341)
Parliamentary attention (logged)	0.032***
	(3.931)
Legal type (Type 2)	-0.003
	(-0.148)
Size: Full time equivalent (logged)	-0.004
	(-0.269)
Budget: Expenditures in Euro (logged)	0.016
	(1.459)
Country: Flemish agency	-0.055
	(-1.633)
Constant	0.022
	(0.350)
/	
sigma_u	0.035***
	(3.501)
sigma_e	0.119***
	(20.312)
Observations	393

z statistics in parentheses. Year-dummies applied but coefficients now shown.
+ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 4: Reputational threats (dimensions), with binary task variables

	(1) Performative	(2) Technical	(3) Moral	(4) Procedural
Service delivery task	-0.120***	-0.056	-0.054+	-0.108*
	(-3.491)	(-1.520)	(-1.797)	(-2.500)
Regulatory task	0.028	0.027	0.044+	0.025
	(0.971)	(0.890)	(1.730)	(0.718)
Redistributive task	-0.064*	-0.045	-0.043+	-0.132***
	(-2.197)	(-1.485)	(-1.692)	(-3.631)
Multiple tasks	0.029	0.042	0.007	0.116*
	(0.738)	(0.981)	(0.202)	(2.331)
Multiple policy issues	-0.051*	-0.055*	-0.039*	-0.040+
	(-2.488)	(-2.429)	(-2.249)	(-1.783)
Parliamentary attention (logged)	0.026**	0.027***	0.017**	0.020*
	(3.269)	(3.634)	(2.692)	(2.051)
Legal type (Type 2)	-0.007	0.032+	0.007	-0.006
	(-0.416)	(1.654)	(0.453)	(-0.302)
Size: Full time equivalent (logged)	0.002	0.012	0.006	0.002
	(0.188)	(0.867)	(0.516)	(0.140)
Budget: Expenditures in Euro (logged)	0.016	-0.008	0.020*	0.006
	(1.560)	(-0.722)	(2.295)	(0.511)
Country: Flemish agency	-0.019	-0.077*	-0.018	0.007
	(-0.601)	(-2.380)	(-0.667)	(0.167)
Constant	-0.016	-0.150*	-0.172***	-0.132*
	(-0.282)	(-2.368)	(-3.441)	(-2.047)
/				
sigma_u	0.028**	0.038***	0.026**	0.014
	(2.686)	(3.834)	(3.049)	(0.626)
sigma_e	0.115***	0.076***	0.075***	0.104***
	(18.925)	(11.863)	(13.170)	(9.810)
Observations	393	393	393	393

z statistics in parentheses. Year-dummies not applied due to non-convergence of models.
$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Supplementary materials (online annex)

This document contains the following supplementary materials to the paper "How does Organizational Task matter for the Reputation of Public Agencies?":

- Coding reliability procedure
- Table A1: correlation matrix
- Table A2: sheds light on how specific combinations of task categories affect the likelihood of attracting reputational threats.
- Table A3: sheds light on how specific combinations of task categories affect the likelihood of attracting reputational threats towards different reputational dimensions.
- Tables A4 to A7: analyses with as dependent variable reputational threats + reputational praises
- In Tables A8 to A13, we have split the analyses for Flemish and Danish agencies to uncover potential different effects. In general, the results are robust to these analyses as well.
- We have included a robustness check, in which we have run the model with cross-section data only with averages for the variables over time (N=40). The regressions are to be found in table A14 and table A15.
- We have added a supplementary analyses (table A16 and A17) with articles as the level of analysis (n = 13,290). The analyses are multilevel logistic regression with standard errors clustered at agency-level so the comparison of the size of the coefficients in our main models in table 3 and A2 are not possible. However, from the direction of the variables of interest and the level of significance, the analyses are robust to this specification as well.
- We have compared the period before the financial crisis (2006 and 2007) with the period after the crisis (2008 to 2015) in appendix table A18. We find no significant difference in the relative share of articles containing a reputational threat. For the dimensions, we with some indication of the technical dimension being more exposed before the crisis, and the moral dimension less.
- To investigate if the variance of reputational threats varies across agencies, in table A19 to A22 we have calculated the year-agency standard deviation (instead of the mean) on the reputational threats as the dependent variable. The variable is skewed, so we have log + 1-transformed it following the same argumentation in the main models in the manuscript. As the reviewer correctly notes, the variance of the threats (operationalized as the standard deviation in an organization-year) is affected by agency type.

Coding reliability

Reliability had to be ensured at two levels: within-country (between student coders in each country) and between-country (between master coders of the research team). First, a protocol was developed by senior members of the research team, based on iterative rounds of coding and discussion. Second, We have calculated the intercoder reliability coefficient (Krippendorff's alpha, KA) between countries and within-country coders. The KA's for intra-country reliability were on average 0.69 for reputational threats (on a total of 169 articles, and eight different coders). The mean inter-country reliability calculated on a set of English articles (n = 40, three different coders) was 0.67 for reputational threats. These alphas, in particular, for the intra-country reliability tests, are relatively low. Two mitigating factors exist, being (a) variable complexity: As Lacy et al. (2015, 797) discuss, "coding news story topic is easier than coding the valence (positive or negative leaning toward an object, person, or issue)."; and (b) difficulties of the KA measure to deal with skewed variables, in which case KA is known to yield lower coefficients even when the levels of simple agreement are high (Lacy et al., 2015). In our case, newspaper coverage towards agencies is typically overwhelmingly neutral (Boon et al., 2019; Schillemans, 2012). When calculating simple agreements, we observe an agreement of 88% for reputational threats (intra-country); and 87% for reputational threats (inter-country). Furthermore, we took additional measures to maximize reliability: (a) in Microsoft Access (which was used as an interface for coding), a "doubt box" was added which coders were strongly urged to use in case they experienced doubt. All articles that were ticked as "doubtful" were later discussed and decided on by the master coders in both countries; (b) the Danish master coder visited the Belgian team to finalize the data. In preparation of this visit, all master coders again went through *all* articles, and again listed potential problem cases which were discussed and solved together.

Table A1: Correlation matrix

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1	1.000													
2	0.933*	1.000												
3	0.465*	0.276*	1.000											
4	0.558*	0.409*	0.280*	1.000										
5	0.512*	0.431*	0.250*	0.284*	1.000									
6	0.117*	0.059	0.159*	0.032	0.167*	1.000								
7	-0.253*	-0.248*	-0.152*	-0.123*	-0.168*	-0.329*	1.000							
8	-0.061	-0.063	0.030	0.038	-0.075	-0.084	0.062	1.000						
9	-0.112*	-0.158*	0.066	-0.056	-0.017	0.302*	0.422*	0.608*	1.000					
10	0.032	0.021	0.084	0.119*	-0.013	-0.054	-0.018	0.073	-0.103*	1.000				
11	0.004	0.005	-0.060	-0.024	0.073	0.300*	-0.297*	-0.270*	-0.204*	-0.075	1.000			
12	0.134*	0.158*	-0.016	0.181*	0.031	-0.162*	0.186*	0.098	-0.052	0.277*	-0.003	1.000		
13	-0.022	-0.036	-0.053	0.080	-0.082	-0.007	0.379*	-0.045	0.144*	-0.165*	0.273*	0.211*	1.000	
14	0.007	0.009	-0.090	0.185*	-0.106*	-0.256*	0.383*	0.266*	0.245*	0.078	0.055	0.432*	0.707*	1.000

* shows significance at the 0.05 level. Correlations with 12, 13, and 14 are for logged variables.

Table A2. Reputational threats, with combinations of task categories (i.e. combinations of tasks in the agencies' task portfolios)

	(1)
Solely service delivery task	Ref.
Solely regulatory task	0.106**
	(3.119)
Solely redistributive task	0.087+
	(1.906)
Service delivery task + regulatory task	0.084**
	(2.705)
Service delivery task + redistributive task	-0.071*
	(-2.050)
Regulatory task + redistributive task	0.124**
	(3.139)
Service delivery task + regulatory task + redistributive task	0.001
	(0.044)
Multiple policy issues	-0.039+
	(-1.757)
Parliamentary attention (logged)	0.032***
	(4.007)
Legal type (Type 2 agency)	-0.006
	(-0.305)
Size: Full time equivalents (logged)	-0.011
	(-0.751)
Budget: Expenditures in Euro (logged)	0.014
	(1.223)
Country: Flemish agency	-0.039
	(-1.177)
Constant	-0.033
	(-0.426)
/	
sigma_u	0.031**
	(3.052)
sigma_e	0.119***
	(20.385)
Observations	393

z statistics in parentheses
+ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A3. Reputational threats (dimensions), with combinations of task categories (i.e. combinations of tasks in the agencies' task portfolios)

	(1) Performative	(2) Technical	(3) Moral	(4) Procedural
Solely service delivery task	Ref.	Ref.	Ref.	Ref.
Solely regulatory task	0.112***	0.053+	0.074**	0.095*
	(3.509)	(1.664)	(2.667)	(2.495)
Solely redistributive task	0.094*	-0.057	-0.003	-0.080
	(2.256)	(-1.081)	(-0.073)	(-1.162)
Service delivery task + regulatory task	0.073*	0.045	0.046+	0.127***
	(2.542)	(1.594)	(1.806)	(3.612)
Service delivery task + redistributive task	-0.057+	-0.062+	-0.070*	-0.088+
	(-1.785)	(-1.772)	(-2.234)	(-1.928)
Regulatory task + redistributive task	0.116**	0.104**	0.072*	0.133**
	(3.163)	(2.904)	(2.218)	(3.016)
Service delivery task + regulatory task + redistributive task	-0.010	0.001	-0.002	-0.010
	(-0.414)	(0.037)	(-0.077)	(-0.353)
Multiple policy issues	-0.039+	-0.060**	-0.037*	-0.037+
	(-1.888)	(-2.860)	(-2.080)	(-1.704)
Legal type (Type 2 agency)	-0.009	0.023	0.003	-0.015
	(-0.534)	(1.318)	(0.183)	(-0.758)
Parliamentary attention (logged)	0.026***	0.025***	0.017**	0.019*
	(3.381)	(3.599)	(2.642)	(2.059)
Size: Full time equivalents (logged)	-0.003	-0.000	-0.002	-0.014
	(-0.189)	(-0.001)	(-0.179)	(-0.848)
Budget: Expenditures in Euro (logged)	0.013	0.003	0.024*	0.017
	(1.225)	(0.274)	(2.424)	(1.191)
Country: Flemish agency	-0.007	-0.070*	-0.012	0.018
	(-0.225)	(-2.314)	(-0.443)	(0.473)
Constant	-0.097	-0.140+	-0.180**	-0.155+
	(-1.332)	(-1.928)	(-2.753)	(-1.757)
/ sigma_u	0.024*	0.029**	0.024**	0.000

	(2.208)	(3.014)	(2.863)	(0.000)
sigma_e	0.114***	0.076***	0.075***	0.104***
	(19.005)	(11.818)	(13.226)	(10.177)
Observations	393	393	393	393

z statistics in parentheses. Year-dummies not applied due to non-convergence of models.
$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A4. Reputational threats and praises combined, with binary task variables

	(1)
Service delivery task	-0.154***
	(-3.458)
Regulatory task	0.017
	(0.455)
Redistributive task	-0.097*
	(-2.540)
Multiple tasks	0.073
	(1.405)
Multiple policy issues	-0.061*
	(-2.256)
Parliamentary attention (logged)	0.026**
	(2.592)
Legal type (Type 2 agency)	-0.025
	(-1.031)
Size: Full time equivalents (logged)	0.004
	(0.205)
Budget: Expenditures in Euro (logged)	0.014
	(1.093)
Country: Flemish agency	0.019
	(0.469)
Constant	0.081
	(1.066)
/	
sigma_u	0.040**
	(3.091)
sigma_e	0.153***
	(21.469)
Observations	393

z statistics in parentheses. Year fixed effects
+ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A5. Reputational threats and praises combined, with combinations of task categories (i.e. combinations of tasks in the agencies' task portfolios)

	(1)
Solely service delivery task	Ref.
Solely regulatory task	0.133**
	(3.144)
Solely redistributive task	0.110+
	(1.929)
Service delivery task+ regulatory task	0.112**
	(2.891)
Service delivery task+ redistributive task	-0.044
	(-1.034)
Regulatory task + redistributive task	0.151**
	(3.052)
Service delivery task + regulatory task + redistributive task	-0.009
	(-0.275)
Multiple policy issues	-0.046+
	(-1.682)
Parliamentary attention (logged)	0.026**
	(2.668)
Legal type (Type 2 agency)	-0.025
	(-1.059)
Size: Full time equivalents (logged)	-0.001
	(-0.050)
Budget: Expenditures in Euro (logged)	0.010
	(0.740)
Country: Flemish agency	0.032
	(0.798)
Constant	-0.039
	(-0.394)
/	
sigma_u	0.038**
	(2.867)
sigma_e	0.152***
	(21.500)
Observations	393

z statistics in parentheses. Year fixed effects
+ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A6. Reputational threats and praises combined (dimensions), with binary task variables

	(1) Performative	(2) Technical	(3) Moral	(4) Procedural
Service delivery task	-0.127***	-0.055+	-0.043	-0.088*
	(-3.387)	(-1.827)	(-1.499)	(-2.300)
Regulatory task	0.010	0.003	0.048+	0.030
	(0.322)	(0.112)	(1.938)	(0.969)
Redistributive task	-0.084**	-0.037	-0.041+	-0.113***
	(-2.622)	(-1.432)	(-1.677)	(-3.510)
Multiple tasks	0.060	0.044	-0.001	0.098*
	(1.376)	(1.267)	(-0.043)	(2.212)
Multiple policy issues	-0.054*	-0.034+	-0.027+	-0.030
	(-2.398)	(-1.901)	(-1.648)	(-1.496)
Parliamentary attention (logged)	0.017*	0.023***	0.021***	0.017+
	(2.018)	(3.568)	(3.357)	(1.899)
Legal type (Type 2 agency)	-0.033	0.033*	0.012	-0.002
	(-1.640)	(2.051)	(0.823)	(-0.123)
Size: Full time equivalents (logged)	0.005	0.014	-0.001	0.005
	(0.362)	(1.208)	(-0.108)	(0.342)
Budget: Expenditures in Euro (logged)	0.015	-0.011	0.020*	0.004
	(1.368)	(-1.148)	(2.429)	(0.364)
Country: Flemish agency	0.045	-0.073**	-0.030	0.014
	(1.340)	(-2.677)	(-1.149)	(0.400)
Constant	0.036	-0.122*	-0.151**	-0.140*
	(0.577)	(-2.317)	(-3.185)	(-2.423)
/				
sigma_u	0.032**	0.029***	0.025**	0.013
	(2.822)	(3.446)	(3.004)	(0.631)
sigma_e	0.128***	0.077***	0.075***	0.099***
	(20.238)	(13.236)	(13.664)	(10.477)
Observations	393	393	393	393

z statistics in parentheses. + $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A7. Reputational threats and praises combined (dimensions), with combinations of task categories (i.e. combinations of tasks in the agencies' task portfolios)

	(1) Performative	(2) Technical	(3) Moral	(4) Procedural
Solely service delivery task	Ref.	Ref.	Ref.	Ref.
Solely regulatory task	0.121***	0.032	0.067*	0.089**
	(3.331)	(1.130)	(2.523)	(2.626)
Solely redistributive task	0.074	0.013	-0.012	-0.070
	(1.535)	(0.339)	(-0.296)	(-1.124)
Service delivery task + regulatory task	0.083*	0.044+	0.042+	0.115***
	(2.503)	(1.756)	(1.733)	(3.697)
Service delivery task + redistributive task	-0.030	-0.023	-0.078*	-0.064+
	(-0.827)	(-0.792)	(-2.539)	(-1.679)
Regulatory task + redistributive task	0.112**	0.078*	0.060+	0.116**
	(2.661)	(2.408)	(1.930)	(2.957)
Service delivery task + regulatory task + redistributive task	-0.013	0.000	-0.004	0.001
	(-0.465)	(0.009)	(-0.216)	(0.051)
Multiple policy issues	-0.046+	-0.032+	-0.025	-0.028
	(-1.953)	(-1.738)	(-1.490)	(-1.461)
Parliamentary attention (logged)	0.017*	0.023***	0.021***	0.015+
	(2.066)	(3.555)	(3.325)	(1.887)
Legal type (Type 2 agency)	-0.033	0.030+	0.008	-0.009
	(-1.624)	(1.870)	(0.542)	(-0.514)
Size: Full time equivalents (logged)	0.004	0.009	-0.009	-0.007
	(0.245)	(0.716)	(-0.790)	(-0.480)
Budget: Expenditures in Euro (logged)	0.012	-0.009	0.024*	0.012
	(1.046)	(-0.882)	(2.549)	(0.941)
Country: Flemish agency	0.052	-0.065*	-0.023	0.022
	(1.500)	(-2.397)	(-0.893)	(0.647)
Constant	-0.078	-0.140*	-0.148*	-0.163*
	(-0.936)	(-2.126)	(-2.402)	(-2.075)
/				
sigma_u	0.031**	0.028**	0.023**	0.000
	(2.766)	(3.185)	(2.792)	(0.000)
sigma_e	0.128***	0.076***	0.074***	0.098***

	(20.263)	(13.232)	(13.725)	(10.847)
Observations	393	393	393	393

z statistics in parentheses. + $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A8. Reputational threats, with binary task variables, country

	DK	FL
Service delivery task	-0.134**	-0.197+
	(-2.787)	(-1.865)
Regulatory task	0.025	0.031
	(0.517)	(0.697)
Redistributive task	-0.084+	-0.132
	(-1.842)	(-1.273)
Multiple tasks	0.065	0.094
	(1.196)	(0.713)
Legal type (Type 2 agency)	-0.019	0.013
	(-0.564)	(0.498)
Multiple policy issues	-0.046	-0.077*
	(-1.484)	(-1.993)
Parliamentary attention (logged)	0.032**	0.025+
	(3.183)	(1.869)
Size: Full time equivalents (logged)	-0.080+	0.029
	(-1.933)	(1.140)
Budget: Expenditures in Euro (logged)	0.066	0.017
	(1.558)	(1.302)
Constant	0.323*	-0.139+
	(2.264)	(-1.928)
/		
sigma_u	0.035*	0.016
	(2.516)	(0.746)
sigma_e	0.107***	0.130***
	(14.611)	(14.097)
Observations	199	194

z statistics in parentheses. Year fixed effects
+ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A9. Reputational threats, with combinations of task categories (i.e. combinations of tasks in the agencies' task portfolios), country

	DK	FL
Solely service delivery task	Ref.	Ref.
Solely regulatory task	0.149***	0.228**
	(4.145)	(2.799)
Solely redistributive task	0.156**	0.066
	(2.786)	(0.972)
Service delivery task + regulatory task	0.115**	
	(3.278)	
Service delivery task + redistributive task	-0.034	-0.038
	(-0.637)	(-0.782)
Regulatory task + redistributive task	0.146***	
	(4.355)	
Servce + regulatory task + redistributive task	0.038	-0.007
	(0.773)	(-0.227)
Legal type (Type 2 agency)	0.009	0.013
	(0.289)	(0.498)
Multiple policy issues	-0.027	-0.077*
	(-0.936)	(-1.993)
Parliamentary attention (logged)	0.029**	0.025+
	(3.258)	(1.869)
Size: Full time equivalents (logged)	-0.092*	0.029
	(-2.319)	(1.140)
Budget: Expenditures in Euro (logged)	0.083*	0.017
	(2.021)	(1.302)
Constant	0.142	-0.336*
	(1.280)	(-2.284)
/		
sigma_u	0.000**	0.016
	(2.828)	(0.746)
sigma_e	0.109***	0.130***
	(15.159)	(14.097)
Observations	199	194

z statistics in parentheses. Year fixed effects
+ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A10. Reputational threats (dimensions), with binary task variables, Denmark

	(1) Performative	(2) Technical	(3) Moral	(4) Procedural
Service delivery task	-0.138**	-0.072	-0.049	-0.117***
	(-3.225)	(-1.418)	(-1.344)	(-3.292)
Regulatory task	0.002	0.064	0.053	0.002
	(0.050)	(1.228)	(1.307)	(0.067)
Redistributive task	-0.073+	-0.056	-0.052	-0.081*
	(-1.799)	(-1.279)	(-1.618)	(-2.561)
Multiple tasks	0.053	0.070	0.021	0.061+
	(1.131)	(1.288)	(0.534)	(1.650)
Legal type (Type 2 agency)	-0.015	0.029	0.006	-0.024
	(-0.547)	(0.846)	(0.250)	(-1.017)
Multiple policy issues	-0.062*	-0.045	-0.005	-0.050*
	(-2.291)	(-1.272)	(-0.245)	(-2.244)
Parliamentary attention (logged)	0.024**	0.021*	0.018**	0.015*
	(2.625)	(2.401)	(2.683)	(1.981)
Size: Full time equivalents (logged)	-0.061	-0.053	-0.014	-0.116**
	(-1.568)	(-1.214)	(-0.455)	(-3.097)
Budget: Expenditures in Euro (logged)	0.070+	0.029	0.017	0.115**
	(1.730)	(0.651)	(0.519)	(2.967)
Constant	0.199	0.075	-0.060	0.323**
	(1.600)	(0.521)	(-0.607)	(2.983)
/				
sigma_u	0.024+	0.037*	0.018+	0.016
	(1.705)	(2.470)	(1.776)	(1.211)
sigma_e	0.100***	0.074***	0.062***	0.077***
	(13.335)	(9.298)	(8.855)	(10.656)
Observations	199	199	199	199

z statistics in parentheses. Year fixed effects
+ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A11. Reputational threats (dimensions), with binary task variables, Flanders

	(1) Performative	(2) Technical	(3) Moral	(4) Procedural
Service delivery task	-0.242*	-0.134	-0.016	-0.110
	(-2.376)	(-1.477)	(-0.193)	(-1.249)
Regulatory task	0.015	0.005	0.043	0.030
	(0.368)	(0.151)	(1.282)	(0.849)
Redistributive task	-0.203*	-0.128	0.030	-0.077
	(-2.065)	(-1.459)	(0.400)	(-0.913)
Multiple tasks	0.175	0.130	-0.071	0.055
	(1.389)	(1.141)	(-0.726)	(0.505)
Legal type (Type 2 agency)	0.008	0.024	0.011	0.022
	(0.325)	(1.167)	(0.585)	(1.057)
Multiple policy issues	-0.078*	-0.065*	-0.046	-0.048
	(-2.140)	(-2.110)	(-1.595)	(-1.555)
Parliamentary attention (logged)	0.021	0.029*	0.010	0.030**
	(1.614)	(2.418)	(0.956)	(2.659)
Size: Full time equivalents (logged)	0.030	0.036+	0.014	0.020
	(1.218)	(1.792)	(0.690)	(0.973)
Budget: Expenditures in Euro (logged)	0.022+	-0.001	0.015	0.010
	(1.761)	(-0.053)	(1.435)	(0.935)
Constant	-0.120+	-0.315***	-0.223***	-0.260***
	(-1.738)	(-4.126)	(-3.806)	(-3.896)
/				
sigma_u	0.000*	0.000	0.000**	0.013
	(2.449)	(0.000)	(2.828)	(0.850)
sigma_e	0.127***	0.076***	0.086***	0.088***
	(13.827)	(7.609)	(9.801)	(10.190)
Observations	194	194	194	194

z statistics in parentheses. Year fixed effects
+ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A12. Reputational threats (dimensions), with combinations of task categories (i.e. combinations of tasks in the agencies' task portfolios), Denmark

	(1) Performative	(2) Technical	(3) Moral	(4) Procedural
Solely service delivery task	Ref.	Ref.	Ref.	Ref.
Solely regulatory task	0.122***	0.111*	0.081**	0.130***
	(3.509)	(2.164)	(2.694)	(3.472)
Solely redistributive task	0.168**	-0.371	-0.279	0.109*
	(3.212)	(-0.004)	(-0.010)	(2.146)
Service delivery task + regulatory task	0.081*	0.093+	0.044	0.094**
	(2.372)	(1.773)	(1.446)	(2.750)
Service delivery task + redistributive task	-0.047	-0.034	-0.304	0.001
	(-0.912)	(-0.451)	(-0.016)	(0.019)
Regulatory task + redistributive task	0.124***	0.136**	0.061*	0.109**
	(3.878)	(2.781)	(2.111)	(3.272)
Service delivery task + regulatory task + redistributive task	0.013	0.032	-0.014	0.018
	(0.288)	(0.486)	(-0.329)	(0.400)
Legal type (Type 2 agency)	0.006	0.006	-0.010	-0.003
	(0.224)	(0.163)	(-0.407)	(-0.094)
Multiple policy issues	-0.040	-0.067+	-0.024	-0.029
	(-1.435)	(-1.722)	(-1.002)	(-1.181)
Parliamentary attention (logged)	0.024**	0.020*	0.017*	0.016*
	(2.893)	(2.307)	(2.468)	(2.070)
Size: Full time equivalents (logged)	-0.075*	-0.046	-0.016	-0.127***
	(-2.020)	(-1.075)	(-0.502)	(-3.336)
Budget: Expenditures in Euro (logged)	0.082*	0.026	0.021	0.125**
	(2.121)	(0.610)	(0.631)	(3.155)
Constant	0.051	0.040	-0.060	0.181+
	(0.500)	(0.295)	(-0.691)	(1.849)
/				
sigma_u	0.000	0.033*	0.015	0.010
	(0.000)	(2.209)	(1.366)	(0.548)
sigma_e	0.099***	0.074***	0.063***	0.077***
	(13.867)	(9.249)	(8.835)	(10.707)
Observations	199	199	199	199

z statistics in parentheses. Year fixed effects
+ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A13. Reputational threats (dimensions), with combinations of task categories (i.e. combinations of tasks in the agencies' task portfolios), Flanders

	(1) Performative	(2) Technical	(3) Moral	(4) Procedural
Solely service delivery task	Ref.	Ref.	Ref.	Ref.
Solely regulatory task	0.257**	0.140+	0.075	0.141*
	(3.247)	(1.933)	(0.977)	(2.021)
Solely redistributive task	0.039	0.006	0.054	0.033
	(0.623)	(0.117)	(0.883)	(0.629)
Service delivery task + regulatory task				
Service delivery task + redistributive task	-0.028	0.001	-0.039	-0.022
	(-0.603)	(0.034)	(-0.866)	(-0.553)
Regulatory task + redistributive task				
Service delivery task + regulatory task + redistributive task	-0.013	0.007	0.004	0.008
	(-0.423)	(0.284)	(0.151)	(0.328)
Legal type (Type 2 agency)	0.008	0.024	0.010	0.022
	(0.321)	(1.167)	(0.445)	(1.057)
Multiple policy issues	-0.078*	-0.065*	-0.054	-0.048
	(-2.129)	(-2.110)	(-1.456)	(-1.555)
Parliamentary attention (logged)	0.021	0.029*	0.009	0.030**
	(1.610)	(2.418)	(0.809)	(2.659)
Size: Full time equivalents (logged)	0.030	0.036+	0.020	0.020
	(1.214)	(1.792)	(0.811)	(0.973)
Budget: Expenditures in Euro (logged)	0.022+	-0.001	0.014	0.010
	(1.757)	(-0.053)	(1.106)	(0.935)
Constant	-0.362*	-0.449**	-0.269+	-0.371**
	(-2.520)	(-3.257)	(-1.911)	(-2.898)
/				
sigma_u	0.003	0.000	0.026+	0.013
	(0.030)	(0.000)	(1.852)	(0.850)
sigma_e	0.127***	0.076***	0.083***	0.088***
	(13.407)	(7.609)	(9.682)	(10.190)
Observations	194	194	194	194

z statistics in parentheses. Year fixed effects
+ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A14. Cross-sectional robustness check, with binary task variables

	(1)
Service delivery task	-0.086***
	(-4.248)
Regulatory task	0.007
	(0.418)
Redistributive task	-0.038*
	(-2.138)
Multiple tasks	0.024
	(1.025)
Multiple policy issues	-0.036**
	(-2.970)
Legal type (Type 2 agency)	0.004
	(0.325)
Parliamentary attention (logged)	0.001
	(0.108)
Size: Full time equivalents (logged)	0.010
	(1.193)
Budget: Expenditures in Euro (logged)	0.006
	(0.881)
Country: Flemish agency	0.020
	(0.720)
Constant	0.040
	(1.247)
Observations	40

z statistics in parentheses
+ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A15. Robustness check, Cross-sectional at agency-level, with combinations of task categories (i.e. combinations of tasks in the agencies' task portfolios)

	(1)
Solely service delivery task	Ref.
Solely regulatory task	0.066**
	(3.612)
Solely redistributive task	0.064*
	(2.553)
Service delivery task + regulatory task	0.039*
	(2.312)
Service delivery task + redistributive task	-0.033+
	(-1.758)
Regulatory task + redistributive task	0.085***
	(4.092)
Service delivery task + regulatory task + redistributive task	-0.010
	(-0.718)
Multiple policy issues	-0.029*
	(-2.428)
Legal type (Type 2 agency)	-0.000
	(-0.026)
Parliamentary attention (logged)	0.002
	(0.211)
Size: Full time equivalents (logged)	0.004
	(0.499)
Budget: Expenditures in Euro (logged)	0.006
	(0.840)
Country: Flemish agency	0.026
	(1.010)
Constant	-0.010
	(-0.235)
Observations	40

z statistics in parentheses
+ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A16. Robustness check, articles as level of observations, with binary task variables (logistic regression)

	(1)
Service delivery task	-1.572**
	(-3.259)
Regulatory task	-0.079
	(-0.262)
Redistributive task	-0.717*
	(-2.206)
Multiple tasks	0.728
	(1.588)
Multiple policy issues	-0.673*
	(-2.198)
Legal type (Type 2 agency)	-0.039
	(-0.206)
Parliamentary attention (logged)	0.108+
	(1.756)
Size: Full time equivalents (logged)	0.108
	(0.785)
Budget: Expenditures in Euro (logged)	0.103
	(0.889)
Country: Flemish agency	0.352
	(1.059)
Constant	0.022
	(0.350)
Observations	13290

z statistics in parentheses. Standard errors are clustered by agency. Year fixed effects.
+ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A17. Robustness check, articles as level of observations, with combinations of task categories (i.e. combinations of tasks in the agencies' task portfolios) (logistic regression)

	(1)
Solely service delivery task	Ref.
Solely regulatory task	1.352**
	(3.191)
Solely redistributive task	0.581
	(0.819)
Service delivery task + regulatory task	0.589+
	(1.780)
Service delivery task + redistributive task	-0.363
	(-0.864)
Regulatory task + redistributive task	1.675**
	(2.977)
Service delivery task + regulatory task + redistributive task	-0.138
	(-0.524)
Multiple policy issues	-0.680*
	(-2.024)
Legal type (Type 2 agency)	-0.056
	(-0.317)
Parliamentary attention (logged)	0.097
	(1.585)
Size: Full time equivalents (logged)	0.032
	(0.216)
Budget: Expenditures in Euro (logged)	0.167
	(1.112)
Country: Flemish agency	0.370
	(1.089)
Constant	-3.489***
	(-3.940)
Observations	13290

z statistics in parentheses. Standard errors are clustered by agency. Year fixed effects.
+ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A18. Percentage and type of reputational threats, before and after the financial crisis (start 2008)

	Before crisis (2006-2007)	After crisis (2008-2015)	Difference
Reputational threats	0.069	0.077	0.016 (n.s)
Performative threats	0.047	0.060	0.013 (n.s.)
Technical	0.019	0.011	0.009 (p = 0,04)
Moral	0.010	0.018	0.008 (p = 0,09)
Procedural	0.010	0.015	0.006 (n.s.)
n	77	316	77

Note: Test of difference is two-sample t test with equal variances.

Table A19. Variance (standard deviation) in reputational threats, with binary task variables

	(1)
Service delivery task	-0.102+
	(-1.811)
Regulatory task	0.113*
	(2.368)
Redistributive task	-0.101*
	(-2.106)
Multiple tasks	0.016
	(0.237)
Multiple policy issues	-0.075*
	(-2.213)
Legal type (Type 2 agency)	0.002
	(0.055)
Parliamentary attention (logged)	0.059***
	(4.642)
Size: Full time equivalents (logged)	-0.018
	(-0.813)
Budget: Expenditures in Euro (logged)	0.038*
	(2.257)
Country: Flemish agency	-0.126*
	(-2.452)
Constant	0.054
	(0.547)
/	
sigma_u	0.056***
	(3.807)
sigma_e	0.178***
	(19.847)
Observations	393

z statistics in parentheses. Year fixed effects
+ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A20. Variance (standard deviation) in reputational threats, with combinations of task categories (i.e. combinations of tasks in the agencies' task portfolios)

	(1)
Solely service delivery task	Ref.
Solely regulatory task	0.167**
	(3.079)
Solely redistributive task	0.053
	(0.723)
Service delivery task + regulatory task	0.151**
	(3.062)
Service delivery task + redistributive task	-0.118*
	(-2.156)
Regulatory task + redistributive task	0.137*
	(2.187)
Service delivery task + regulatory task + redistributive task	0.022
	(0.514)
Multiple policy issues	-0.060+
	(-1.694)
Legal type (Type 2 agency)	-0.001
	(-0.037)
Parliamentary attention (logged)	0.060***
	(4.706)
Size: Full time equivalents (logged)	-0.025
	(-1.078)
Budget: Expenditures in Euro (logged)	0.034+
	(1.929)
Country: Flemish agency	-0.109*
	(-2.111)
Constant	0.002
	(0.019)
/	
sigma_u	0.054***
	(3.683)
sigma_e	0.177***
	(19.871)
Observations	393

z statistics in parentheses. Year fixed effects
+ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A21. Variance (standard deviation) in reputational threats (dimensions), with binary task variables

	(1) Performative	(2) Technical	(3) Moral	(4) Procedural
Service delivery task	-0.120***	-0.056	-0.054+	-0.108*
	(-3.491)	(-1.520)	(-1.797)	(-2.500)
Regulatory task	0.028	0.027	0.044+	0.025
	(0.971)	(0.890)	(1.730)	(0.718)
Redistributive task	-0.064*	-0.045	-0.043+	-0.132***
	(-2.197)	(-1.485)	(-1.692)	(-3.631)
Multiple tasks	0.029	0.042	0.007	0.116*
	(0.738)	(0.981)	(0.202)	(2.331)
Multiple policy issues	-0.051*	-0.055*	-0.039*	-0.040+
	(-2.488)	(-2.429)	(-2.249)	(-1.783)
Legal type (Type 2 agency)	-0.007	0.032+	0.007	-0.006
	(-0.416)	(1.654)	(0.453)	(-0.302)
Parliamentary attention (logged)	0.026**	0.027***	0.017**	0.020*
	(3.269)	(3.634)	(2.692)	(2.051)
Size: Full time equivalents (logged)	0.002	0.012	0.006	0.002
	(0.188)	(0.867)	(0.516)	(0.140)
Budget: Expenditures in Euro (logged)	0.016	-0.008	0.020*	0.006
	(1.560)	(-0.722)	(2.295)	(0.511)
Country: Flemish agency	-0.019	-0.077*	-0.018	0.007
	(-0.601)	(-2.380)	(-0.667)	(0.167)
Constant	-0.016	-0.150*	-0.172***	-0.132*
	(-0.282)	(-2.368)	(-3.441)	(-2.047)
/				
sigma_u	0.028**	0.038***	0.026**	0.014
	(2.686)	(3.834)	(3.049)	(0.626)
sigma_e	0.115***	0.076***	0.075***	0.104***
	(18.925)	(11.863)	(13.170)	(9.810)
Observations	393	393	393	393

z statistics in parentheses
+ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A22. Variance (standard deviation) in reputational threats (dimensions), with combinations of task categories (i.e. combinations of tasks in the agencies' task portfolios)

	(1) Performative	(2) Technical	(3) Moral	(4) Procedural
Solely service delivery task	Ref.	Ref.	Ref.	Ref.
Solely regulatory task	0.112***	0.053+	0.074**	0.095*
	(3.509)	(1.664)	(2.667)	(2.495)
Solely redistributive task	0.094*	-0.057	-0.003	-0.080
	(2.256)	(-1.081)	(-0.073)	(-1.162)
Service delivery task + regulatory task	0.073*	0.045	0.046+	0.127***
	(2.542)	(1.594)	(1.806)	(3.612)
Service delivery task + redistributive task	-0.057+	-0.062+	-0.070*	-0.088+
	(-1.785)	(-1.772)	(-2.234)	(-1.928)
Regulatory task + redistributive task	0.116**	0.104**	0.072*	0.133**
	(3.163)	(2.904)	(2.218)	(3.016)
Service delivery task + regulatory task + redistributive task	-0.010	0.001	-0.002	-0.010
	(-0.414)	(0.037)	(-0.077)	(-0.353)
Multiple policy issues	-0.039+	-0.060**	-0.037*	-0.037+
	(-1.888)	(-2.860)	(-2.080)	(-1.704)
Legal type (Type 2 agency)	-0.009	0.023	0.003	-0.015
	(-0.534)	(1.318)	(0.183)	(-0.758)
Parliamentary attention (logged)	0.026***	0.025***	0.017**	0.019*
	(3.381)	(3.599)	(2.642)	(2.059)
Size: Full time equivalents (logged)	-0.003	-0.000	-0.002	-0.014
	(-0.189)	(-0.001)	(-0.179)	(-0.848)
Budget: Expenditures in Euro (logged)	0.013	0.003	0.024*	0.017
	(1.225)	(0.274)	(2.424)	(1.191)
Country: Flemish agency	-0.007	-0.070*	-0.012	0.018
	(-0.225)	(-2.314)	(-0.443)	(0.473)
Constant	-0.097	-0.140+	-0.180**	-0.155+
	(-1.332)	(-1.928)	(-2.753)	(-1.757)
/				
sigma_u	0.024*	0.029**	0.024**	0.000
	(2.208)	(3.014)	(2.863)	(0.000)
sigma_e	0.114***	0.076***	0.075***	0.104***
	(19.005)	(11.818)	(13.226)	(10.177)
Observations	393	393	393	393

z statistics in parentheses
+ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$