Quasi-contractual ministerial steering of state agencies: Its intensity, modes, and how agency characteristics matter

Jostein Askim, Karl Hagen Bjurstrøm & Jonas Kjærvik (University of Oslo)

Accepted for publication in *International Public Management Journal*

ABSTRACT

The article demonstrates the value of conceptualizing four ideal types when studying ministries' contract steering of state agencies: relational, double-whammy, performance and behavioral steering – each defined by its combination of input- and output-oriented steering. In the system under study – Norway – about half of all agencies are subjected to steering with a clear profile; the other half is not. The two profiles often dismissed or overlooked in existing research – relational and double-whammy steering – are most common. Thus, introducing a contract regime has not meant a clear shift from input to output control, as posited by some. Most agency characteristics under study – size, age, political salience, and tasks – have significant effects on the likelihood that an agency is subjected to one type of steering or another. Still, the analysis suggests that agency-level characteristics only to a limited extent constrain ministerial choice about which steering to practice.

INTRODUCTION

From a system that relies on contracting with private and third-sector providers to deliver public services, the "contract state" has grown to a system that formalizes even relationships between entities within central government as "deliverables" in return for funding (Greve 2007, 4). A contract regime tends to reduce informal and ad hoc superior-subordinate interaction, and to limit interaction instead to documents and meetings with high degrees of formality and predictability. Furthermore, a contract regime tends to "convert diffuse, non-specific expectations into more specific and concrete obligations" (Freedland and King 2003, 468) and to concentrate the focus of superior-subordinate interaction toward performance vis-à-vis contract obligations (Freedland and King 2003, 466; Greve 2007, 4).

Two discussions dominate scholarship on the use of contracts within central government. One is concerned with input control and output control. Some have claimed that a shift from input to output control is an integral part of introducing contracts to regulate hierarchical relationships (e.g., Faraneti, Padovani, and Young 2010, 258; Osborne 2006, 383; Schrijvers 1993, 598;). Empirical research across sectors and countries shows, however, that traditional regulation, detailing how government entities should work, often continues alongside “modern” output-oriented performance management, with goals and performance indicators that define which deliverables and ends to achieve (Binderkrantz and Christensen 2009; Hood 2004; Moynihan and Pandey 2006; Van Thiel et al. 2012; Verhoest et al. 2004).A second discussion sees steering as a one-dimensional variable going from low to high intensy. Here, the difference between input and output control is not essential; it is assumed that an equal level of oversight can be achieved by compensating for a reduction in input control by increasing output control, and vice versa (Askim 2015; Bouckaert 1998; Roness et al. 2008). In the first discussion a key question is why some agencies are controlled predominantly on inputs and others predominantly on outputs; in the second a key question is why some agencies are tightly controlled and others are loosely controlled by their ministries.

Majone (2001, 117) has called for students of hierarchical relations, and agency theorists in particular, to improve the understanding of governance by recognizing its complexity. To answer this call, this article bridges the two discussions just mentioned and introduces a typology with four ideal typical ministry-agency steering practices: performance steering, behavioral steering, double-whammy steering, and relational steering. The four are distinguished by their relative application of input-oriented and output-oriented steering. We also include a fifth category, that is, a category for steering that lacks a clear profile. In constructing the typology, the article sees a contract as predated by the hierarchical relation it regulates and assumes that contracts reflect established inter-organizational relations and professional cultures (Ferlie 1992, 86; Freedland and King 2003, 467). Performance and behavioral steering, we argue, reflect underlying relationships where the subordinate is seen as the superior's agent, while relational and double-whammy steering reflect underlying relationships that are best understood when using perspectives other than principal-agency theory.

The empirical relevance of the typology is explored by a study of ministry-agency relationships in Norway. We demonstrate, first, how ministry-agency relationships can be categorized by analyzing the contents of the equivalent of contracts in this context, which is annual letters of appropriation to state agencies. We ask: How many steering relationships do and how many do not have a clear profile, and which of the four ideal types of steering are more and less common in practice? A preview of the findings is that the profiles relational and double-whammy steering are more common than are performance and behavioral steering.

Beyond improving description, and offering the opportunity systematically to compare practices across space and time (Kristiansen et al. 2017; Pollitt 2013), a typology of steering practices can be empirically useful in efforts to understand how variation in steering affects performance. Due in part to data limitations and endogeneity problems, this article does not pursue this question empirically, but suggestions for future research along this line are made in the conclusion. In this article the typology is used instead to contribute to scholarly exploration of how steering practices are context dependent (Pollitt 2006). Among the many potentially relevant contextual factors – some external and some at the ministry side and some at the agency side – this article focuses on three agency characteristics and asks: How are ministry-agency steering practices affected by variation in the size, age, political salience, and tasks of agencies?

The next section develops the typology and expectations about the relationship between agency characteristics and steering practices. Then, the data and methods are presented. The article's dependent variables are measured using data derived from content analysis of 142 letters of appropriation (quasi-contracts) between ministries and agencies in Norway. Independent variables are measured using register data. Thus, we avoid common method-bias problems often associated with public management research (Jakobsen and Jensen 2015). Next, the results of the empirical study are presented. The final section discusses findings vis-à-vis expectations, answers the research questions, and offers suggestions for future research, building on insights from this article.

THEORY

A typology of steering practices

This section extends existing theoretical and empirical work (Bouckaert 1998; Eisenhardt 1989; Moynihan and Pandey 2006; Roness et al. 2008; Van Dooren, Bouckaert, and Halligan 2010; Verhoest et al. 2004) to develop a typology that distinguishes four ideal-type steering practices. In *performance steering*, the superior steers the subordinate primarily on performance against objectives referring to organizational outputs and outcomes (output-oriented steering); emphasis on steering the subordinate's inputs and work processes is low. This is the so-called managerial version of steering.

Behavioral steering is the opposite: It is characterized by high reliance on steering subordinates with what-to-do instructions concerning organizational inputs and work processes (input-oriented steering), and by low emphasis on output-oriented steering. The label behavioral steering is borrowed from agency theorist Kathleen Eisenhardt's (1989, 58) conceptualization of "behavior-oriented contracts," which she contrasts with "outcome-oriented contracts." Van Dooren, Bouckaert, and Halligan call practices that correspond with behavioral steering "the administrative default mode" and see such practices as the "most obvious (…) alternative" to performance-based steering (2010, 185). Others portray behavioral steering as outmoded – a form of steering made superfluous and/or inappropriate by ideas and reforms associated with New Public Management – while performance steering is seen as the emerging dominant form (Moynihan and Pandey 2006, 121; Verhoest et al. 2004, 33).

Double whammy is a term used by Hood (2004, 16) to describe situations where attempts at more managerial approaches to public service provision in reality add new, *ex post* controls without reducing the old, *ex ante* ones. Institutional theory uses the term layering to describe such phenomena (Askim 2015; Thelen 2003). We borrow Hood's term to characterize steering relationships with high emphasis on steering subordinates on both inputs and outputs.

Relational steering is the antithesis; it characterizes a steering practice with low emphasis on steering subordinates on both inputs and outputs. Figure 1 shows a graphical representation of the two-by-two typology of steering modes. In the middle of the figure is a field called "No profile," meant to capture steering practices where the use of input-oriented and/or output-oriented steering is of medium intensity.

<<<Figure 1 about here>>>

We study steering practices within a quasi-contract regime. Following organizational sociology, and as mentioned in the introduction, a contract can be perceived as predated by the hierarchical relation it regulates – "a social relation with its own history and norms" (Ferlie 1992, 86).

Individual contracts can be assumed to be adjusted to and to reflect established inter-organizational relations and professional cultures (Freedland and King 2003, 467). Steering practices, as expressed in (quasi-)contracts, should therefore be understood in light of underlying superior-subordinate relations. This sociological perspective adds a layer to the proposed typology. Performance and behavioral steering reflect underlying relationships where the subordinate is seen as the superior's agent. Such relationships are associated with agency problems: information asymmetry regarding production of the contracted task, diverging interests, and bureaucratic drift, that is, the ability of the agent to act in ways that differ from the superior's preferences. Given diverging interests, the agent may use his information advantage to hide bureaucratic drift from the principal (Eisenhardt 1989; Pollack 1997, 108–109; Schillemans and Busuioc 2015; Waterman and Meier 1998). The difference between performance and behavioral steering lies in what sort of control the ministry relies on to contain bureaucratic drift by the state agency. What sort of control the ministry relies on may be defined by the underlying superior-subordinate relation. If this relation is characterized by a high degree of trust, steering may be regarded as what Pierre and Peters refer to as a trust-based regime: "Trust … refers to the relative absence of performance control or measurement" (Pierre and Peters 2017, 163). However, trust does not necessarily imply absence of control, but rather that control is exercised ex ante with an expectation that the subordinate will act according to the preferences of the principal (Pierre and Peters 2017, 163).

Double-whammy and relational steering can reflect a variety of superior-subordinate relationships. Neither of them reflects a principal-agent relationship, though. Our perspective on double-whammy steering is that it represents a relationship in which the superior (e.g., a ministry) sees its subordinate (e.g., a state agency) as its extended arm, not as an entity to hold and manage at arm's length. With strong constraints on both which ends to achieve and which means to employ in the process, the ministry has in a sense colonized the agency – permanently or temporally. Relational steering represents underlying inter-organizational relationships where control is unnecessary, inappropriate, or ineffective. In some cases, the superior does not care very much what the subordinate does. Since control is costly, the rational thing for the superior to do is to control the subordinate as little as possible. In other cases, considerable constraints on ends or means would prevent the subordinate from fulfilling its *raison d'être*. The effectiveness and legitimacy of several types of government functions require a credible commitment to policy and agency autonomy (Majone 2001, 103) and hence a hands-off approach from the minister and the ministry. Examples include ombudsman and central banking functions, and financial and market regulation. In such cases, so-called loose or incomplete contracts are most appropriate and effective, since what lies underneath is what Majone (2001, 116–117) calls a fiduciary relationship and not a principal-agent relationship between the ministry and the agency. Incomplete contracts "do not determine all terms of the agreement in advance of the execution" (Amirkhanyan, Kim, and Lambright 2010, 192). Instead, "parties agree not on detailed plans of action but on (…) what to do when unforeseen contingencies arise, on who has the power to act and the range of actions that can be taken, and on dispute resolution mechanisms to be used if disagreements do occur" (Majone 2001, 117). Others have used the term "relational contracts" – thereby inspiring our label "relational steering" – to capture the same phenomenon, for example in organizational sociology (e.g., Ferlie 1992, 86) and contract law (e.g., Amirkhanyan, Kim, and Lambright 2010; Macneil 1978; Macneil and Campbell 2001).

Students of inter-firm relationships have long acknowledged that trust is a potent mechanism for making behavior predictable, and as such a viable and, in the long run, economical alternative to detailed, formalized steering. Trust increases superior entities' willingness to take calculated risks – and to avoid detailed contracts – "because of their confident expectation that [subordinates] will act responsibly" (Gulati 1995, 94). Public administration scholars Van Dooren, Bouckaert, and Halligan (2010, 185) discuss trust-based control as an alternative to performance steering, albeit a less obvious alternative than behavioral steering.

Agency characteristics

How should we expect steering practices to vary across agencies? Size, age, political salience, and tasks are agency-level characteristics identified in the existing literature as important explanations for how ministries steer their agencies (Pollitt 2006; Verhoest et al. 2010). The role of *agency size* is hard to study, or at least to interpret, since size can be seen as a proxy for several agency phenomena that can affect steering practices in different ways. Knowledge about how size and steering are connected is therefore inconclusive (Askim 2015; Chenhall 2003; Lægreid, Roness, and Rubecksen 2006; Verbeeten 2008). Most phenomena for which size can be seen as a proxy– agency complexity, societal importance, political salience, influence – do however point in the direction of low autonomy. We therefore expect greater agency size to decrease the likelihood of relational steering.

One perspective on the possible influence of *agency age* is that with the passage of time, the likelihood of relational steering increases. Repeated interactions and personal relationships across organizational entities are factors that enable using loose contracts in situations that would otherwise have demanded detailed contracts (Gulati 1995, 92–95). Both factors are common within the central government apparatus. Relationships between ministries and agencies are normally long-term, in principle eternal, relationships, with numerous interactions. An alternative perspective on time and organizational practices – the birthmark (Guillén et al. 2002,

24) or imprinting perspective (Stinchcombe 1965) – is that the agency's age as such matters less than in which period it was established. Before the institutionalization of the contract state and ministry-agency performance management – that is before 1997 in the Norwegian case[1] – ministry-agency relations were characterized by "traditional regulation" with controls predominantly targeting agency inputs (Van Dooren, Bouckaert, and Halligan 2010, 185). Since governance relations can attain "characteristics that reflect prominent features of the environment (…) during a (…) period of susceptibility" (Marquis and Tilcsik 2013, 199) , we can expect that pre-performance-age agencies to this day are subjected to behavioral steering while performance-age agencies are subjected to performance steering.

Agency tasks: An assumption well established in research on state agencies is that higher *political salience* is associated with lower autonomy and with being held politically accountable (Askim 2015; Koop 2011; Kristiansen 2016; Pollitt 2006). The political salience of an agency and its tasks is affected by interrelated characteristics such as the size of its budget, the agency's importance to society, and the risk of electoral retribution for politicians in case of low agency performance and legitimacy. We do not expect to see relational steering practiced vis-à-vis high-salience agencies. Double-whammy steering and behavioral steering are the most likely observations. Performance steering is a less likely observation because issuing what-to-do instructions to the agency can be considered a more unambiguous proof of ministerial involvement than is specifying performance objectives.

Table 1 summarizes our expectations for how steering practices vary depending on agency size, age, and political salience.

<<<Table 1 about here>>>

Finally, certain tasks can be expected to be associated with a hands-off approach by the ministry – in our conceptualization: relational steering – to avoid threatening the effectiveness and legitimacy of the agencies that perform them. Examples, in addition to those mentioned above, include research and higher education, compiling national statistics, mediation, arbitration, reviews, and appeals. We expect to see relational steering practiced vis-à-vis agencies that are responsible for such hands-off tasks. Double-whammy steering is the least likely observation. Among the remaining two types, performance steering is a more likely observation than gbehavioral steering is. Constraints on ends can be considered less intrusive of agency autonomy than are constraints on means. Our empirical assessment of this expectation is arrived at by subjective interpretation rather than by statistical analysis.

DATA AND METHODS

The units of analysis here were Norwegian ministries' annual letters of appropriation ("tildelingsbrev") to subordinate state agencies – organizational entities that are part of central government in legal terms. Formally, these documents are contract-like operational arrangements rather than contracts in a strict sense; they do not operate via a horizontal approach, using voluntarism and negotiation, as contracts usually do. The letters of appropriation can be characterized as quasi-contracts operating on a vertical approach, based on a foundation of authority (Greve 2007, 4). Norwegian state agencies cannot decline to be bound by ministerial instruction or demand external resolution of disagreements However, most agencies are allowed a say by their ministries on the specification of performance demands. A recent survey shows that 34 percent of ministry employees and 27 percent of agency employees in Norway fully agree that ministries and agencies "collaborate in the formulation of performance objectives and indicators." Only 1–2 percent fully disagree with this statement (5-point scale) (Survey of State Administration 2016).[2] The Norwegian performance management system can therefore be regarded as dialogue oriented (Lægreid et al. 2006),. Case studies show that the contents of these documents are very important in steering relations between Norwegian ministries and agencies (Askim and Kjærvik 2015; Eltun 2013; Fremstad 2013; Helle 2016; Kaasin 2016). Aspects of formal steering not included in the empirical study are reporting routines and formal meetings. The empirical study also ignores informal ministry-agency steering; implications are discussed in the article's final section.

Among items coded from the letters of appropriation, this article used data on the number of performance objectives and performance indicators and on the number of what-to-do instructions. For illustrative purposes, Box 1 shows coding examples from two state agencies. The sum of performance objectives and indicators is the variable "performance demands," which operationalizes output-oriented steering. The number of direct what-to-do instructions is the variable "behavioral demands," which operationalizes input-oriented steering.

<<<Box 1 about here>>>

Letters of appropriation are voluminous documents (jokingly referred to as "appropriation novels" by practitioners) with elements of professional phraseology. Several steps were taken to avoid misunderstandings and errors in the coding. Background interviews were conducted in 2015 with key personnel in the Ministry of Finance and the Government Agency for Financial Management, which share responsibility for guiding ministry-agency steering practices. Guidelines and regulations from these two organizations on how to design letters of appropriation were also consulted. To ensure consistency, all letters of appropriation were coded by two authors, and inconsistencies (there were practically none on the data items used in this article) were resolved together with a third author.

The data set on the contents of letters of appropriation was initially established with data from 2012, used first in Askim (2015). The present article uses an extension of the data set to cover 2015 in addition to 2012, and to cover appendices to letters of appropriation in addition to the letters themselves (appendices sometimes contain some performance indicators). Any supplementary letters of appropriation have not been analyzed.(3) The material includes data coded from 142 letters of appropriation from 72 agencies (two agencies in the sample were merged with other agencies between 2012 and 2015). The sample represents a balance between the need for a number of observations sufficient for statistical analysis, research economy (reliable coding), and of having a sample representative of the relevant universe of agencies. The 72 agencies were sampled to represent analytically relevant variety among the universe of about 200 state agencies in Norway (parent ministry, strength of formal affiliation to the ministry, main tasks, and organizational size). For a complete list of agencies in the sample, see Table A1 in the appendix. Table 2 shows descriptive statistics for the variables performance and behavioral demands.

<<<Table 2 about here>>>

We analyze the distribution of ministry-agency steering practices among the four categories of steering developed in the analytical section. Empirically to assign letters of appropriation to one category or another, we used threshold values to identify polar observations on each of the two variables performance and behavioral demands. Thresholds were defined as the values that delineate the bottom and top third of the observations on each variable, using observations for the year 2012 (below 7 and above 26 for behavioral demands and below 11 and above 26 for performance demands). Then, the two variables were combined to produce four polar categories and a zone in between for steering practices without a clear profile. Having this fifth category means we can be confident about the validity of our claim that all observations we assign to one ideal-type category do in fact represent a group of observations sharing a quality that differentiates them from observations in the other three ideal-type categories. We want to avoid having very small differences between two observations, say, one more performance demand in contract A than in contract B, mean that A and B are assigned to different ideal-type categories.

We analyze the distribution of ministry-agency steering practices in light of the independent variable, agency tasks, using subjective interpretation, and in light of agency size, age, and political salience, using bivariate analysis and multinomial regression analysis .

Multinomial regression analysis is used for two reasons. First, it is a suitable method of regression when the dependent variable is nominal (i.e., cannot be ordered).Second, it allows us to explore each independent variable's effect on the outcome category of the typology relative to the same reference point (no profile). In other words, "the effect of the independent variables is allowed to differ for each outcome" (Long 1997, 149), enabling comparison of the conditional effect of each independent variable.

The independent variables are measured as follows: Agency size is measured as the number of full-time equivalents (FTEs). Agency age is measured as the number of years between the agency's initial formation and the observation year. Political salience is measured as the number of times an agency was mentioned in national printed or online mass media the year before the contract. Data on agency size and age were collected from the Norwegian Centre for Research Data's (NSD) State Administration Database (2016). Data on media hits were collected from the A-press database of the Nordic media monitoring company Retriever. Descriptive data for the independent variables are shown in Table 3.

<<<Table 3 about here>>>

In the regression analyses, one unit of agency size is 100 FTEs, and one unit of political salience is 100 media hits. Agency age is operationalized in two ways: one unit of age as 10 years (in model 1 in Table A.3) and as a generational dummy – a dichotomous variable distinguishing between agencies established before and after the year 1997 (in models 2–3 in Table A.3). Twenty-six of the 72 agencies were established in 1997 or later. Robustness checks with different thresholds have been performed, with no substantial differences in results, with one exception as discussed below (see also note to Table A.3).

RESULTS

Figure 2 shows a graphical representation of the distribution of observations on the variables performance demands (left panel) and behavioral demands (right). When Figure 2 is combined with the descriptive statistics in Table 2, we see that the distribution is skewed right for both variables; means are higher than the median values and tails are longer on the right- than on the left-hand side. One agency faces 250 steering demands (the Directorate of Health in 2012) and a handful face between 170 and 200, but the median is as low as 32. The most common range, when distinguished by tens of steering demands, is the zero-to-ten range.

<<<Figure 2 about here>>>

Since Table 2 and Figure 2 contain observations from both 2012 and 2015, they mask the fact that the 2015 cohort of letters of appropriation contains fewer performance and behavioral demands than the 2012 cohort does. The average number of demands combined was reduced by 24 percent, from 51 to 39 (the reduction is fairly even across the two types of demands). We include in the modelling a dummy variable that measures whether a performance contract was written in 2012 or 2015; it allows us to control for differences between the two cohorts of letters.

Given the method used, that is, using top and bottom thirds on two variables and combining them to construct four categories, there might be no empirical observations in any of the categories we have defined. In principle, all agencies steered with either few or many performance demands may be steered with an intermediate number of behavioral demands; and vice versa, all those steered with either few or many behavioral demands may be steered with an intermediate number of performance demands. Thus, all observations would fall in the zone in between the categories (no profile). In reality, though, a considerable proportion of our empirical observations fall in one of the categories.

<<<Figure 3 about here>>>

Figure 3 shows a scatter plot of all observations, 2012 and 2015 combined. Almost half (45 percent; 66 of 142 observations) fall in one of the four categories. Twenty-one observations illustrate relational ministry-agency governance, 13 illustrate behavioral governance, 10 illustrate performance governance, and 22 illustrate the double-whammy mode. Detailed observations by steering category are shown in appended Table A.2.

<<<Figure 4 about here>>>

The bivariate analysis compares mean values and variation within each of our four categories of steering. We see that agency characteristics vary significantly among agencies steered similarly (same ideal-type category). For example, there are quite a few small agencies in the double-whammy category, not only large ones, and there are quite a few old agencies in the performance-steering category, not only agencies established in the age of performance measurement. For a more comprehensive picture of the relationship between our three explanatory variables and steering, we now move on to the multivariate analysis. Then we will discuss separately the results from the bivariate and multivariate analyses for each variable.

<<<Figure 5 about here>>>

The coefficient plot in Figure 5 shows a graphical representation of the results of a multinomial regression analysis (detailed results are in the appended Table A.3).[4] Coefficients are displayed as odds ratios. They should be interpreted as the factor change in odds for being in one of the four ideal-type categories of steering vs. being in the no-profile category for one unit change in the explanatory variable, holding all other variables constant. In the discussion of the results below, the most notable effects are also interpreted as probabilities. Unlike odds ratios, probabilities are dependent on the level of the explanatory variable and on the level of the other covariates in the analysis (Long 1997, 169).

Agency size

A bivariate analysis shows that agency size varies considerably between the four categories of steering. The box plot in Figure 4 (top panel) shows that agencies subjected to double-whammy steering tend to be larger and that agencies subjected to relational steering tend to be smaller than are agencies subjected to other types of steering. In the full sample, the average agency size is

1001 FTEs (see Table 3); in the double-whammy category it is 2835 and in the relational steering category it is 99.

Agencies in the two remaining categories, performance and behavioral steering, are typically in between the two former ones in size. Agencies in the behavioral steering category are slightly larger than are those in the performance steering category. Figure 4 also shows that agencies in the relational and performance steering categories clearly cluster in the low range of agency size. In the other two categories, and especially in the double-whammy category, there is large internal variety in agency size.

As shown by the first row in Figure 5, the effect of size on steering holds when controlling for variation in the other independents. Increased agency size has significant ($p < 0.1$) negative effect on being in the relational steering category and a small but significant positive effect on being in the double-whammy category. When agency size increases by 100 employees, we expect a reduction in the odds of being in the relational category by a factor of 0.65 (a 35 percent reduction in odds) and an increase in the odds for being in the double-whammy category by a factor of 1.025 (a 2.5 percent increase in odds), compared to the reference category. The predicted probability of being in the double-whammy category increases from 9.5 to 14 percent if an agency has about 7000 employees (3rd quartile) as compared with 640 employees (1st quartile).[(5)] We should note that with only 142 observations, the results are vulnerable to outliers. As an illustration: Robustness checks (see note to Table A.3) show that two observations strongly influence size's estimated effect on relational steering. If we operationalize the ideal-type categories more narrowly, for example, by using quartiles as cut-offs instead of thirds, two relatively large agencies change category from relational to no-profile steering (the Directorate of Public Construction and Property and the Norwegian Immigration Appeals Board). As a result, the effect of a one-unit change in agency size on being in the relational-steering category increases from -0.4 to -2.6.

Agency age

Agency *age*, measured as the number of years since the agency was established, does not vary systematically between categories of steering practice. The bivariate analysis (Figure 4, middle panel) shows that the median age for agencies is slightly higher in the categories behavioral and double-whammy steering than in the other two categories. Variation across categories is smaller than variation within each category is, though. This null finding is confirmed by the regression analysis, irrespectively of whether agency age is measured in years (not reported) or as a generational dummy (Figure 5, second row). Although the effects are not statistically significant (as illustrated by confidence interval lines crossing the value 1), variation in age has relatively large effects on the estimated likelihood of agencies' being in the different categories of steering (as illustrated by the point estimate's distance from the value 1). When transforming regression coefficients to probabilities, it emerges, for example, that compared to the older (pre-1997) ones, the younger (post-1997) agencies have twice the probability of being subjected to performance steering and half the probability of being subjected to behavioral steering.[(6)] The lack of statistical significance may be due to the relatively low number of observations.

Political salience

The bivariate analysis indicates that which kind of steering that is practiced varies with the level of agency-level *political salience,* measured as media attention. Figure 4 (bottom panel) shows that political salience is higher among agencies in the categories double-whammy and behavioral steering than among agencies in the categories relational and performance steering. Figure 4 also shows that agencies in the relational steering category clearly cluster in the low range of political salience; in the other categories there is large internal variety in political salience. The regression analysis confirms that there is a statistically significant relationship between political salience and the probability that relational steering is used (Figure 5, third row). If an agency's media exposure increases from about every fourth day (1st quartile) to about twice per day (3rd quartile), the predicted probability of being subjected to relational steering is reduced from 11 to 0.3 percent.

These findings are corroborated by subjective interpretation of observations (Table A.2). Virtually all agencies that handle issues of lasting high political salience in the Norwegian context (Druckman and Warwick 2005) are subjected to double-whammy steering (e.g., police, health, environment, employment, social welfare, railways, and petroleum). Furthermore, having directorate status – an alternative proxy for high political salience (Grønlie and Flo 2009) – increases the chance that an agency is found in the double-whammy category (82 percent of agencies with directorate status is in this category; the share is 54 percent in the full sample) and reduces the chance that it is found in the relational steering category (31 percent).

Agency tasks

Using subjective interpretation of the task profiles of agencies in the four categories (Table A.2), we make two observations. First, virtually all agencies responsible for tasks such as mediation, arbitration, reviews, and appeals are in the relational steering category. Examples include the Accident Investigation Board, the Criminal Cases Review Commission, and the Immigration Appeals Board. These tasks concern issues that may well be important to the ministry, but where hands-on steering is still avoided, possibly because it is perceived as damaging to the legitimacy and reputation of the ministry, the agency, and the agency's task performance. Second, regulatory agencies and research and higher education institutions are found in several categories of steering.

CONCLUDING DISCUSSION

Variance in steering practices

The article has explored and demonstrated the value of applying a two-dimensional understanding of steering in the study of ministry-agency steering relationships. Distinguishing between input- and output-oriented steering is conventional in public management research, but our conceptualization and operationalization of four ideal types or profiles are new. We asked how many steering relationships have a clear profile and how many do not, and which of the four profiles are more and less common in practice. In the present empirical context, and according to content analysis of letters of appropriation, about half of all ministry-agency relationships have a clear steering profile; the other half do not. The results show, moreover, that the profiles relational and double-whammy steering are more common than are performance and behavioral steering.

This finding speaks to Majone's (2001, 117) call for agency theorists, especially, to recognize the complexity of governance. Many students of steering in hierarchical relationships tend either to overlook practices corresponding to double-whammy and relational steering (e.g., Eisenhardt 1989; Moynihan and Pandey 2006) or to dismiss them as transitional or ineffective. As an illustration, Verhoest et al. (2004) argue that practices corresponding to double-whammy steering are a recipe for a "surplus" of control and that practices corresponding to relational steering are a recipe for control "deficiency" (see also Bach, Fleischer, and Hustedt 2010, 55). The present study does not support this assertion. Since double-whammy and relational steering are indeed common and pervasive practices, scholars should not overlook or dismiss them, but rather should aim to understand their rationales, contents, and dynamics.

For example, it appears that relational steering can occur in any functional context, as long as the underlying relationship is of the trustee kind, that is, one where the superior sees the subordinate as a trusted steward. According to stewardship theory, subordinates are loyal servants whose goals are aligned with those of their principals. Fulfilling the tasks defined by their superiors is all the motivation they need to behave effectively and responsibly (Davis, Schoorman, and Donaldson 1997; Le Grand 2010; Schillemans 2013).

The article furthermore confirms that a mixture of "traditional" behavioral steering in ministry-agency relationships can continue alongside "modern" performance steering even after the introduction of a contract regime within government (Binderkrantz and Christensen 2009). Contract regimes may well have several constitutive effects on the relationships they regulate (Ferlie 1992, 87), but a clear shift from input to output control, as posited by some, is not one of them, at least not in the Norwegian case. Rather than necessarily signaling such a shift, the introduction of a contract regime means that government has decided "to clothe" governance in the form of contracts (Freedland and King 2003, 466). When hierarchical steering outside contract-related channels loses leverage and legitimacy, ministries use flexibility in contract formats to incorporate as much as possible of their steering into contracts, behavioral steering included. Sometimes behavioral instructions are repackaged in targets and objectives language, and other times they are communicated in plain "the agency shall" language. In the words of Freedland and King (2003, 466), studies of contractual relations can "illustrate the ironical way in which [strict control practices are] couched in avowedly liberal forms."

The solidity of the data used here adds weight to our findings concerning variance in steering practices. Letters of appropriation (quasi-contracts) are important elements in operational, year-to-year ministry-agency governance, closely tied to the budgeting process and hence to politics, and to the continuous dialogue between ministries and agencies. Using data coded from such policy documents is not unique (see e.g., Binderkrantz and Christensen 2009; Binderkrantz, Holm, and Korsager 2011) but it is rare; the orthodoxy is rather to study ministry-agency relationships with data obtained from surveys of public sector employees. Given that they are instrumental and not merely symbolic artifacts, letters of appropriation are valid expressions of actual steering practices, more so than survey data, which offer perceptions of steering practices, sometimes laced with views on how steering should be. Systematic document studies are time-consuming when subjective interpretation is required, as is the case here, and set certain limits for how many observations one can obtain. Nonetheless, we believe this method's advantages in validity and reliability more than outweigh its disadvantages relative to survey research when the objective is to understand ministry-agency steering relationships.

Note, though, that the typology concerns steering practices in general; the use here of the contents of formal documents to measure steering is just a design choice. Future research could incorporate data, for example from surveys, on important aspects of informal ministry-agency steering. Incorporating data from a recent survey of central government employees in Norway suggests, however, that adding data on informal steering practices does not necessarily change the location of steering practices within the space defined by the typology (Survey of State Administration 2016).(7)

How agency characteristics matter

Our second research question is how steering practices are affected by the agency characteristics size, age, political salience, and tasks. Answers here can inform reasoning about a related question: To what extent is a ministry's choice about how to steer its underlying agencies conditioned by exogenous factors at agency level? One conclusion that can be drawn from the analysis is that the scope for ministerial choice about which steering to practice is, perhaps unsurprisingly, large. All variants of steering are observed in conjunction with practically all agency characteristics. A higher number of observations might strengthen causal effects, but we acknowledge that none of the agency characteristics studied have a decisive influence on steering practices. We assume that steering, to a large extent, is influenced by factors unobserved here – for example by situational contingency and by endogenous factors such as political programs and personal preferences and interpersonal relationships among those involved in steering.

We can also conclude that agency size and political salience do matter – especially for whether relational or double-whammy steering is practiced. With increased agency size and higher political salience, it is more likely that an agency is subjected to double-whammy steering and less likely that it is subjected to relational steering. The results are inconclusive concerning whether higher political salience also increases the likelihood that an agency will be subjected to behavioral steering. Surprisingly, higher age does not increase the likelihood that an agency will be subjected to relational steering. As for tasks, relational steering is strongly associated with tasks such as arbitration and appeals. Agencies performing other "hands-off" tasks, such as regulation, are subjected to a variety of steering practices. Agencies usually perform a mix of different tasks. Having agencies as the unit of observation has therefore prevented us from pursuing other important questions related to tasks, for example whether so-called wicked issues are associated with a particular style of steering (Head and Alford 2013, 732) argue that such issues ought to be subjected to performance steering to allow "flexibility concerning the processes by which outcomes are achieved."). Finally, agency age does influence steering, but the effects are not statistically significant. Compared to those established later, agencies established before the institutionalization of quasi-contracting are more likely to be subjected to behavioral steering and less likely to be subjected to performance steering.

Limitations and suggestions for future research

This article has used observations from the years 2012 and 2015 to increase the number of observations, not to study change and stability in steering practices over time. The steering did decrease in intensity between the two observation points. We are reluctant, however, to make claims on this basis about possible trends in Norwegian ministry-agency relationships. As part of a modernizing government initiative launched by the government that came into office in 2013, the ministers of finance and modernization wrote letters to all ministries, strongly encouraging them to issue less detailed and voluminous letters of appropriation to subordinate agencies (Ministry of Finance and Ministry of Local Government and Modernization 2013, 2014). It is no surprise, therefore, to find a general "slimming" of the contracts from 2012 to 2015. If and when political attention to ministry-agency relationships returns to its normal, lower, level, steering practices may well change again, possibly back to a less "liberal" version. To uncover trends, future research can incorporate observations further back in time in addition to observing practices in years to come. That would enable a discontinuity design not achievable in this study.

Longer time series (and ideally more agencies) would furthermore enable an examination of the importance of ministry characteristics on the steering relation between ministry and agency. With only 16 ministries (in 2015; 18 in 2012) and a limited number of agencies per ministry in the data set, it was not possible in the present study to estimate effects of ministry characteristics. With more data, future research can focus on ministries and connect, for example, to research on the potential "drifting" not of agents but of principals in hierarchical relationships (Schillemans and Busuioc 2015). If an agency is subjected to behavioral steering at t_0, relational at t_1, and performance steering at t_2, we have an indication of a drifting ministry. Such drift can be compared for variation in, for example, the capacity of ministries and the stability of key ministry personnel.

Furthermore, the present analysis cannot answer whether it actually matters, for example for substantive agency performance or legitimacy, whether agencies are steered in one format or the other. Future research could use steering profiles as independent variables rather than as dependent variables. Endogeneity would be a problem, though, since the object of research is steering relations with mutual interactions, where it is difficult to distinguish cause and effect. To cope with endogeneity one can, for example, introduce a time lag between steering and the dependent variable (e.g., performance), thus increasing the chance of capturing how steering affects performance and not vice versa.

A third suggestion for future research is to use the typology developed here to study other hierarchical relationships in the public sector. The theorizing was done with ministry-agency relationships in mind, but could just as well be applied to relationships between entities that are not part of the same legal entity, for example, central government vs. government enterprises, central vs. local government, and local government vs. local government enterprises. Combinations of behavioral and performance steering exist in all these relationships, and a classification scheme is helpful for describing practices and relating steering practices to causes and consequences.

Finally, it might be that ministry-agency steering practices are dependent upon yet another unobserved variable: how agencies conduct their internal steering (Verhoest and Wynen 2016). As pointed out by Hood (2004, 16), there is limited knowledge about "the conditions under which formal and external oversight can be linked to internal or immanent controls in organizations." This article does not address that issue empirically. Hood (2004, 16–17) suggests that deliberate linkages between external and internal controls are less likely to be prevalent in the public sector than in the private sector. Still, future research should analyze whether cases of relational-type ministry-agency steering are matched with extraordinarily tight (or well-functioning) internal agency steering, and, vice versa, whether cases of double-whammy type ministry-agency steering are matched with extraordinarily loose (or poorly functioning) internal agency steering.

NOTES

(1) Norway institutionalized ministry-agency relationships performance management through a budget reform in 1986, new regulations for agency plans in 1990, and, importantly, new regulations for economic steering in 1997 (more emphasis on measuring performance), 2003 and 2013 (Lægreid, Roness, and Rubecksen 2006; Grønlie and Flo 2009; Askim 2015).

(2) The survey data do not cover enough agencies to examine whether there are significant differences between agencies subjected to different steering practices.

(3) Some agencies receive many supplementary letters of appropriation per year and others receive none. Those that receive the most supplementary letters are agencies whose incomes are unpredictable, for example, due to incomes from fees and sales. Supplementary letters of appropriation rarely contain performance demands; their primary purpose is to authorize the use of funds beyond those covered by that year's original letter of appropriation, which is tied to the state budget. This authorization normally contains one or more behavioral demands but they can often be a repetition of demands from the original letter. In many cases, counting them – without also doing a detailed comparison with the original letter – would therefore mean to double-count the same demands for one year.

(4) The coefficients in table A.3 are displayed as log odds. They should be interpreted as the change in the log odds for being in one of the four ideal-type categories of steering vs. being in the no-profile category, holding all other variables constant. The odds ratios displayed in figure 5 were calculated by applying the exponential function to the coefficients in table A.3.

(5) Given that the agency is established after 1997, the contract year is 2015, and political salience is held at sample mean.

(6) With agency size and political salience fixed at sample mean, the predicted probability of being subject to performance steering is 2.7 percent for younger agencies and 1.3 percent for older ones. The probability of being subject to behavioral steering is 4.8 percent for younger agencies and 9.1 percent for older ones.

(7) The survey (see also note 2) contains data from employees in 30 agencies covered by the present study. Two questions can be considered as proxies for informal ministry-agency steering. One asked about the frequency of contact with top civil servants in the ministry; the other asked about the emphasis they put on signals from political leaders in the ministry. We grouped the answers by agency and by category of steering (as used in the present study). Both questions had a 1 to 5 range for answers. On the first question, the maximum difference in mean scores between categories was 0.1, on the second it was 0.3.

REFERENCES

Amirkhanyan, A. A., H. J. Kim, and K. T. Lambright. 2010. "Do Relationships Matter? Assessing the Association Between Relationship Design and Contractor Performance." *Public Performance & Management Review* 34(2): 189–220.

Askim, J. and J. Kjærvik. 2015. *Etatsstyring i forsvarssektoren: En analyse av mål- og resultatstyringen av Forsvarsdepartementets virksomheter*. Technical report. Oslo: University of Oslo.

Askim, J. 2015. "The Role of Performance Management in the Steering of Executive Agencies: Layered, Imbedded, or Disjointed?" *Public Performance & Management Review* 38(3): 365–394.

Bach, T., J. Fleischer, and T. Hustedt. 2010. *Organisation und Steurung zentralstaatlicher Behörden. Agenturen im westeuropäischen Vergleich*. Berlin: Sigma.

Binderkrantz, A. S. and J. G. Christensen. 2009. "Governing Danish agencies by contract: From negotiated freedom to the shadow of hierarchy." *Journal of Public Policy* 29(1): 55–78.

Binderkrantz, A. S., M. Holm, and K. Korsager. 2011. "Performance Contracts and Goal Attainment in Government Agencies." *International Public Management Journal* 14(4): 445–463.

Bouckaert, G. 1998. "Public sector performance management in a principal/agent context: Some reflections." Pp. 137–145 in P. Boorsma and A. Halachmi, eds., *Inter and Intra government arrangements for productivity: An agency approach*, Dordrecht: Kluwer Academic Publishers.

Chenhall, R. H. 2003. "Management control systems design within its organizational context: Findings from contingency-based research and directions for the future." *Accounting, Organizations and Society* 28(2/3): 127–168.

Davis, J. H., F. D. Schoorman, and L. Donaldson. 1997. "Towards a stewardship theory of management." *The Academy of Management Review* 22(1): 20–47.

Druckman, J. N. and P. V. Warwick. 2005. "The missing piece: Measuring portfolio salience in Western European parliamentary democracies." *European Journal of Political Research* 44(1): 17–42.

Eisenhardt, K. M. 1989. "Agency Theory: An Assessment and Review." *Academy of Management Review* 14(1): 57–74.

Eltun, R. 2013. *Mål- og resultatstyring av Norges forskningsråd 2004–2012*. MA thesis. Oslo: University of Oslo.

Faraneti, F., E. Padovani, and D. W. Young. 2010. "Governance of outsourcing and contractual relationships." Pp. 255–269 in S. P. Osborne, ed., *The new public governance? Emerging perspectives on the theory and practice of public governance*, Abingdon: Routledge.

Ferlie, E. 1992. "The creation and evolution of quasi markets in the public sector: A problem for strategic management." *Strategic Management Journal* 13(S2): 79–97.

Freedland, M. and D. King. 2003. “Contractual governance and illiberal contracts: Some problems of contractualism as an instrument of behavior management by agencies of government.” *Cambridge Journal of Economics* 27(3): 465–477.

Fremstad, E. 2013. *En målstyrt etat i en regelstyrt verden? En studie av styringspraksisen i Politi- og lensmannsetaten*. MA thesis. Oslo: University of Oslo.

Greve, C. 2007. *Contracting for Public Services*. London: Routledge.

Grønlie, T. and Y. Flo. 2009. *Sentraladministrasjonens historie etter 1945. Den nye staten? Tiden etter 1980* [The history of the state administration since 1945. The new state? The period after 1980]. Bergen: Fagbokforlaget.

Guillén, M. F., R. Collins, P. England, and M. Meyer. 2002. “The revival of economic sociology.” Pp. 1–34 in M. F. Guillén, R. Collins, P. England, and M. Meyer, eds., *The new economic sociology: Developments in an emerging field*, New York, NY: Sage.

Gulati, R. 1995. “Does Familiarity Breed Trust? The Implications of Repeated Ties for Contractual Choice in Alliances.” *The Academy of Management Journal* 38(1): 85–112.

Head, B. W., and J. Alford. 2015. “Wicked Problems: Implications for Public Policy and Management.” *Administration & Society* 47(6): 711–739.

Helle, P. 2016. *Etatsstyring med flere prinsipaler. En studie av styringen av fylkesmannen.* MA thesis. Oslo: University of Oslo.

Hood, C. 2004. “Controlling public services and government: Towards a cross-national perspective.” Pp. 3–24 in C. Hood, O. James, B. G. Peters, and C. Scott, eds., *Controlling Modern Government. Variety, Commonality and Change*, Cheltenham: Edward Elgar.

Jakobsen, M., and R. Jensen. 2015. “Common Method Bias in Public Management Studies.” *International Public Management Journal* 18(1): 3–30.

Kaasin, H. 2016. *Etatsstyringen av Innovasjon Norge 2011–2015. Omlegging til koordinert styring mot alle odss?* MA thesis. Oslo: University of Oslo.

Koop, C. 2011. “Accountability of Independent Agencies: The Importance of Political Salience.” *Journal of Public Policies* 31(2): 209–234.

Kristiansen, M. B. 2016. “One Scandinavian approach to management by objectives and results?” *Scandinavian Journal of Public Administration* 20(1): 45–70.

Kristiansen, M. B., P. Dahler-Larsen, and E. M. Ghin. 2017. “On the Dynamic Nature of Performance Management Regimes.” *Administration & Society*. Published online July 2, 2017 https://doi.org/10.1177/0095399717716709

Le Grand, J. 2010. “Knights and Knaves Return: Public Service Motivation and the Delivery of Public Services.” *International Public Management Journal* 13(1): 56-71.

Long, J. S. 1997. *Regression Models for Categorical and Limited Dependent Variables.* Thousand Oaks: SAGE Publications.

Lægreid, P., P. G. Roness, and K. Rubecksen. 2006. “Performance management in practice: The Norwegian way.” *Financial Accountability and Management* 22(3): 251–270.

Macneil, I. R. 1978. “Contracts: Adjustments of long-term economic relations under classical, neoclassical and relational contract law.” *Northwestern University Law Review* 72(6): 854–90.

Macneil, I. R. and D. Campbell. eds. 2001. *The Relational Theory of Contract: Selected Works of Ian Macneil,* London: Sweet & Maxwell.

Majone. G. 2001. “Two logics of delegation: Agency and fiduciary relations in EU governance.” *European Union Politics* 2(1): 103–122.

Marquis C. and A. Tilcsik. 2013. “Imprinting: Toward a multilevel theory.” *Academy of Management Annals* 7(1): 195–245.

Ministry of Finance and Ministry of Local Government and Modernization. 2013. *Færre mål og rapporteringskrav i tildelingsbrevene for 2014* [Fewer performance objectives and reporting requirements in letters of appropriation for 2014]. Note to all ministers of state and all ministry secretaries general (7.11.2013).

Ministry of Finance and Ministry of Local Government and Modernization. 2014. *Videreutvikle mål og rapportering i etatsstyringen – arbeidet med Prop 1S og tildelingsbrev for 2015* [Improving performance objectives and reporting requirements in ministries’ steering of subordinate agencies – working on the national budget and letters of appropriation for 2015]. Letter to all ministers of state (19.6.2014).

Moynihan, D. P. and S. K. Pandey. 2006. “Creating desirable organizational characteristics.” *Public Management Review* 8(1): 119–140.

Norwegian Centre for Research Data (NSD). 2016. *The Norwegian State Administration Database*. Available at http://www.nsd.uib.no/polsys/en/civilservice/

Osborne, S. P. 2006. “The New Public Governance?” *Public Management Review* 8(3): 377–387.

Pierre, J. P. and B. Guy Peters. 2017. “The Shirking Bureaucrat: a Theory in Search for Evidence?” *Policy & Politics* 45(2): 157–172.

Pollack, M. A. 1997. “Delegation, agency, and agenda setting in the European Community.” *International Organization* 51(1): 99–134.

Pollitt, C. 2006. “Performance Management in Practice: A Comparative Study of Executive Agencies.” *Journal of Public Administration Research and Theory* 16(1): 25–44.

Pollitt, C. 2013. “The logics of performance management.” *Evaluation* 19(4): 346–363.

Roness, Paul G., K. Verhoest, K. Rubecksen, and K. MacCarthaigh. 2008. “Autonomy and Regulation of State Agencies: Reinforcement, Indifference or Compensation?” *Public Organization Review* 8(2): 155–174.

Schillemans, T. 2013. “Moving beyond the clash of interests: On stewardship theory and the relationship between central government departments and public agencies.” *Public Management Review* 15(4): 541–562.

Schillemans, T. and M. Busuioc. 2015. “Predicting Public Sector Accountability: From Agency Drift to Forum Drift.” *Journal of Public Administration Research and Theory* 25(1): 191–215.

Schrijvers, A. 1993. “The Management of a Larger Town.” *Public Administration* 71(4): 595–603.

Stinchcombe, A. L. 1965. “Social Structure and Organizations” Pp. 142-193 in J. G. March ed., *Handbook of Organizations,* Chicago: Rand McNailly & Company.

Survey of State Administration 2016. 2016. Norwegian Centre for Research Data (NSD). http://www.nsd.uib.no/polsys/forvaltning/sentraladm2016.html

Thelen, K. 2003. “How institutions evolve: Insights from comparative historical analysis.” Pp. 208–240 in J. Mahoney and D. Rueschemeyer, eds., *Comparative Historical Analysis in the Social Sciences*, Cambridge: Cambridge University Press.

Van Dooren, W., G. Bouckaert, and J. Halligan. 2010. *Performance Management in the Public Sector.* Abingdon: Routledge.

Van Thiel, S., K. Verhoest, G. Bouckaert, and P. Lægreid. 2012. “Lessons and recommendations for the practice of agencification.” Pp. 413–439 in K. Verhoest, S. Van Thiel, G. Bouckaert, and P. Lægreid, eds., *Government agencies: Practices and lessons from 30 countries,* Basingstoke, UK: Palgrave Macmillan.

Verbeeten, F. H. M. 2008. “Performance management practices in public sector organizations: Impact on performance.” *Accounting, Auditing & Accountability Journal* 21(3): 427–454.

Verhoest, K. and J. Wynen. 2016. “Why Do Autonomous Public Agencies Use Performance Management Techniques? Revisiting the Role of Basic Organizational Characteristics.” *International Public Management Journal* (advance access): 1–31. DOI: 10.1080/10967494.2016.1199448

Verhoest, K., P. G. Roness, B. Verschuere, K. Rubecksen, and M. MacCarthaigh. 2010. *Autonomy and control of state agencies: Comparing states and agencies*. Basingstoke, UK: Palgrave Macmillan.

Verhoest, K., B. Vershuere, B. G. Peters, and G. Bouckaert. 2004. “Controlling Autonomous Public Agencies as an Indicator of New Public Management.” *Management international* 9(1): 25–35.

Waterman, R. W. and K. J. Meier. 1998. “Principal-Agent Models: An Expansion?” *Journal of Public Administration Research and Theory* 8(2): 173–202.

Tables and figures

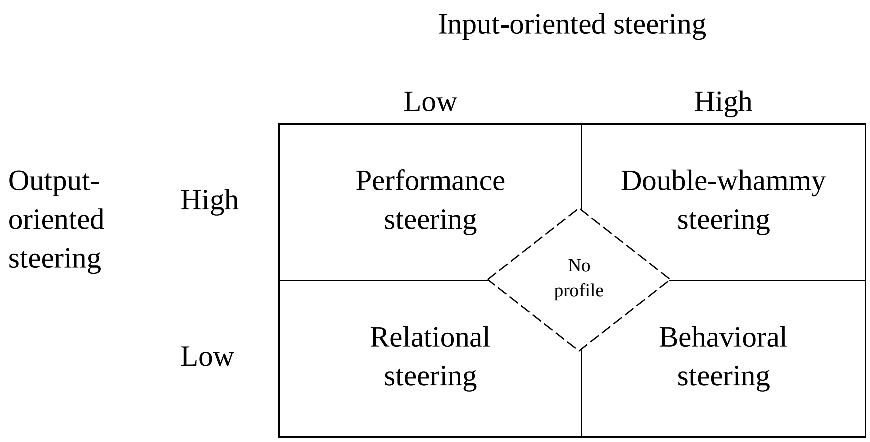


Figure 1: Four modes of steering in ministry-agency relationships

Table 1: Summary of expectations: Agency characteristics and steering practices

	Performance steering	Behavioral steering	Double-w. steering	Relational steering
Size				-
Age	-	+		+
Political salience	-	+	+	-

Box 1: Coding illustrations from two agencies

From the Directorate of Health's 2015 letter of appropriation the following was coded as a performance objective: "Reduced social inequalities in health"; the following as a performance indicator: "Decrease in the proportion of children and young people with poor dental health"; and the following as a behavioral demand: "The Directorate of Health shall collaborate in the establishment of internet-assisted treatment for people with addictions and mental illness."

From the Directorate of Integration and Diversity's 2015 letter of appropriation the following was coded as a performance objective: "Immigrants participate in democracy and society"; the following as a performance indicator: "An increased proportion of members with immigration background in voluntary organizations"; and the following as a behavioral demand: "[The directorate] shall plan and implement measures to encourage greater voter turnout among people with immigration background in the 2015 local government election."

Table 2: Descriptive statistics. Performance demands, behavioral demands, and steering demands (irrespective of format) in letters of appropriation, years 2012 and 2015 (n = 142)

	Mean	St. dev.	Min	Median	Max
Performance demands	23.0	23.2	0	17.5	127
Behavioral demands	22.2	28.4	0	12	152
Steering demands (irrespective of format)	45.2	42.5	0	32	250

Table 3: Descriptive statistics. Agency size, age, and political salience (n = 142)

	Min	Median	Max	Mean	St. dev.
Agency size	1	197.8	14630	1001	2485.7
Agency age (years)	2	27	157	44.1	42.8
Political salience	5	314	5023	549.9	783.9

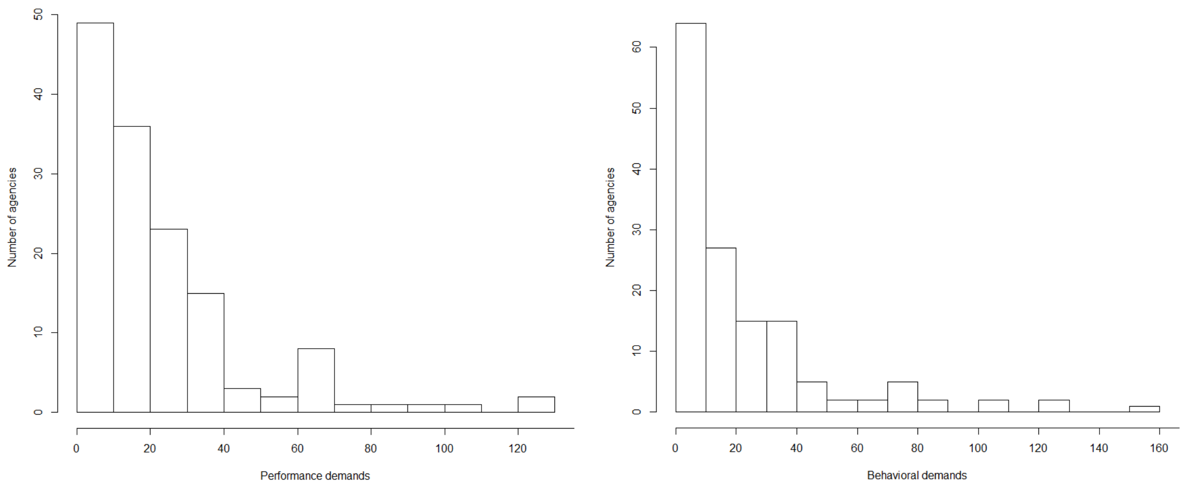


Figure 2. Distribution of agencies according to the number of performance demands (left) and behavioral demands (right) in letters of appropriation, years 2012 and 2015 combined (n = 142)

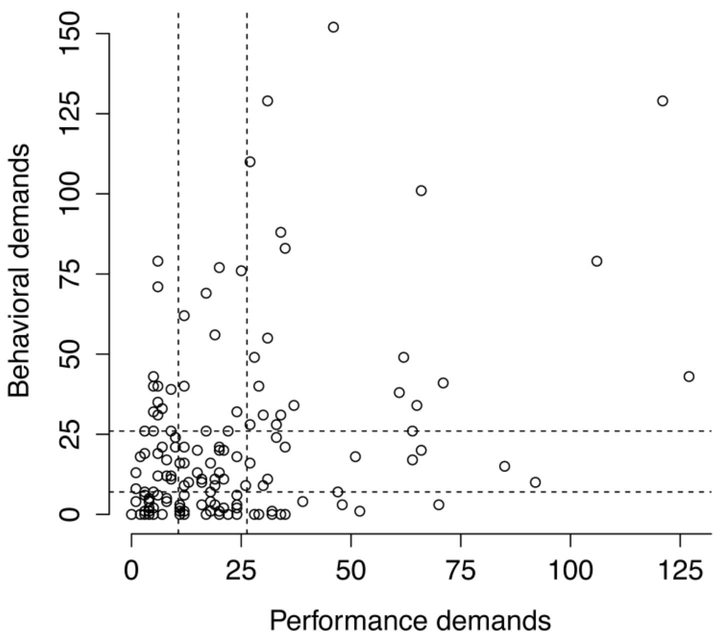


Figure 3: Scatter plot of behavioral demands vs. performance demands in letters of appropriation. Dotted lines mark 33 and 66 pct. of observations for each variable (n = 142).

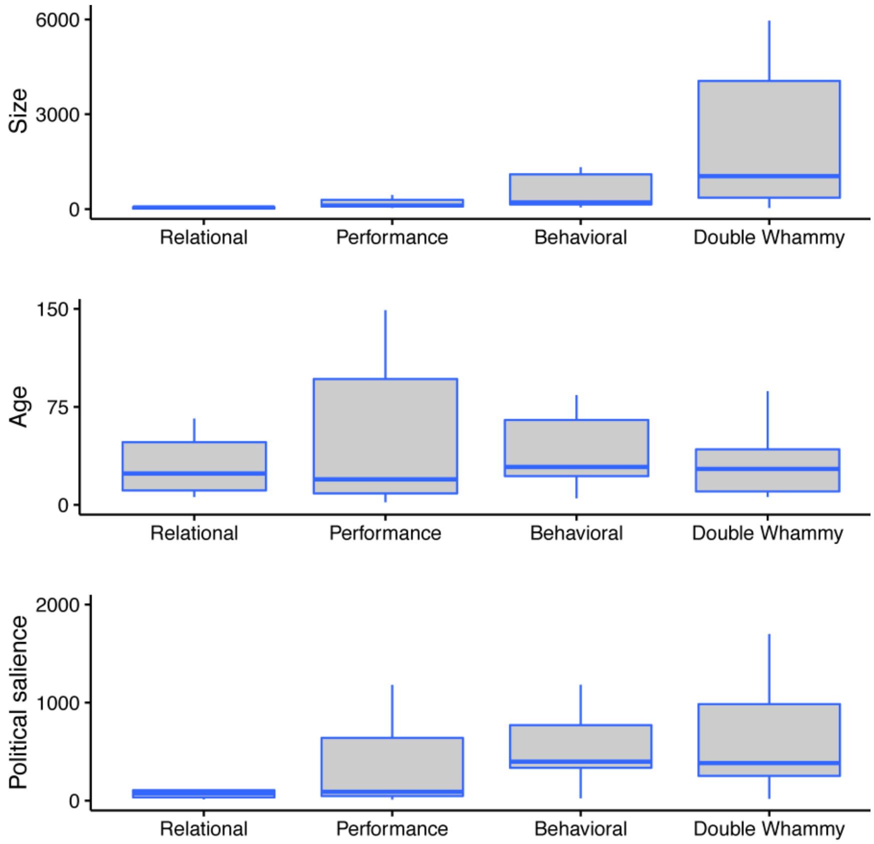
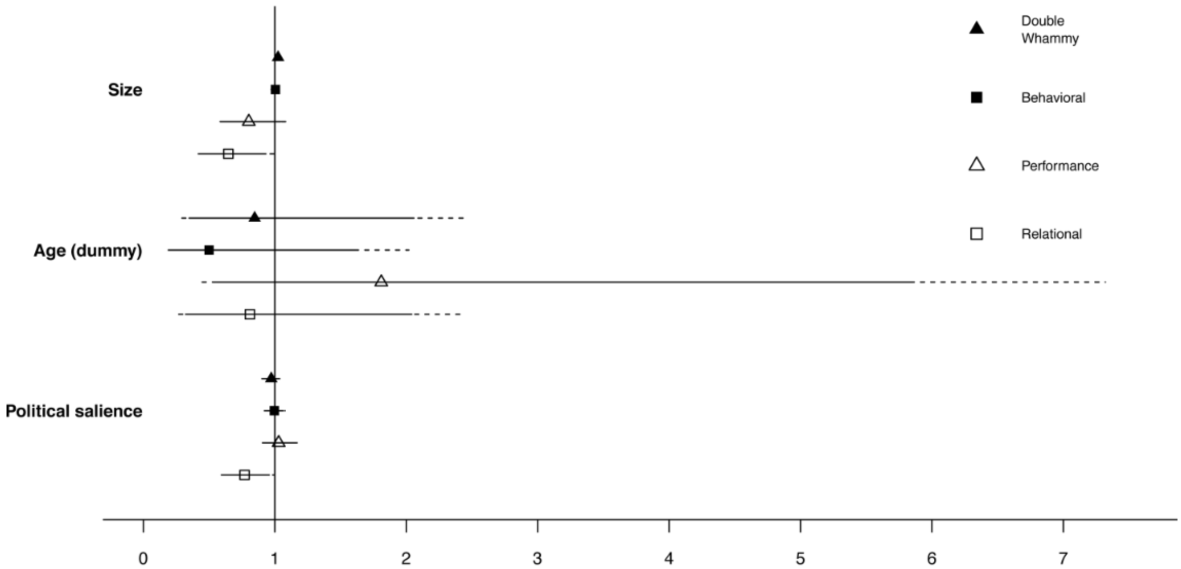


Figure 4: Box plot of agency size, age, and political salience by category of steering. Median values, upper and lower quartiles, without upper and lower extremes.



Note: Points mark estimated coefficients; full lines mark 90 percent confidence intervals; dotted lines mark 95 percent confidence intervals. See A.3., Model 3, for full results.

Figure 5: Coefficient plot from multinomial regression analysis (odds ratios). Marks indicate odds ratios of agencies' being subjected to each type of steering vs. being in the reference category (no clear steering profile). Results sorted by predictors (n = 142).

Table A.1: List of agencies included in the sample

English	Norwegian
Norwegian (Norw.) Labour and Welfare Administration	Arbeids- og velferdsdirektoratet
Norw. Directorate for Children, Youth and Family Affairs	Barne-, ungdoms- og familiedirektoratet
Ombudsman for Children	Barneombudet
Norw. Institute for Agricultural and Environmental Research	Bioforsk
Norw. Biotechnology Advisory Board	Bioteknologinemnda
Brønnøysund Register Centre	Brønnøysundregistrene
Government Administration Services	Departementenes servicesenter
Norw. Labour Inspection Authority	Direktoratet for arbeidstilsynet
Directorate for Building Quality	Direktoratet for byggkvalitet
Agency for Public Management and e-Government	Direktoratet for forvaltning og IKT
Directorate for Emergency Communication	Direktoratet for nødkommunikasjon
Norw. Government Agency for Financial Management	Direktoratet for økonomistyring
Directorate for Civil Protection and Emergency Planning	Direktoratet for samfunnssikkerhet og beredskap
National Courts Administration	Domstoladministrasjonen
Financial Supervisory Authority	Finanstilsynet
Directorate of Fisheries	Fiskeridirektoratet
Consumer Council of Norway	Forbrukerrådet
Norw. Defence Research Establishment	Forsvarets forskningsinstitutt
Norw. Guarantee Institute for Export Credits	Garantiinstituttet for eksportkreditt
Institute of Marine Research	Havforskningsinstituttet
Norw. Directorate of Health	Helsedirektoratet
Rent Disputes Tribunal	Husleietvistutvalget
Directorate of Integration and Diversity	Integrerings- og mangfoldsdirektoratet
Norw. National Rail Administration	Jernbaneverket
Church Council	Kirkerådet
Norw. Criminal Cases Review Commission	Kommisjonen for gjenopptakelse av straffesaker
Norw. Competition Authority	Konkurransetilsynet
Norw. Coastal Administration	Kystverket
Equality and Anti-discrimination Board of Appeals	Likestillings- og diskrimineringsnemnda
Norw. Gaming Board	Lotteri- og stiftelsestilsynet
Norw. Food Safety Authority	Mattilsynet
Norw. Media Authority	Medietilsynet
Norw. Meteorological Institute	Meteorologisk institutt
Norw. Environment Agency (Climate and Pollution Agency in 2012)	Miljødirektoratet (Klima- og forurensningsdirektoratet i 2012)
Office of the National Librarian	Nasjonalbibliotekaren
Norw. Institute of Public Health	Nasjonalt folkehelseinstitutt
Restoration Workshop of Nidaros Cathedral	Nidaros Domkirkes Restaureringsarbeider
Geological Survey of Norway	Norges geologiske undersøkelse
Norw. Film Institute	Norsk fiminstitutt
Norw. Social Research	Norsk institutt for forskning om oppvekst, velferd og aldring
Norw. Forest and Landscape Institute	Norsk institutt for skog og landskap
Arts Council Norway	Norsk kulturråd
Norw. Polar Institute	Norsk polarinstitutt
Norw. Space Centre	Norsk romsenter
Norw. Institute of International Affairs	Norsk utenrikspolitisk institutt
Norw. Petroleum Directorate	Oljedirektoratet
Norw. Church Endowment	Opplysningsvesenets fond
Patients' Injury Compensation Board	Pasientskadenemnda
National Police Directorate	Politidirektoratet
Norw. Post and Telecommunication Authority	Post- og teletilsynet
Norw. Reindeer Husbandry Administration	Reindriftsforvaltningen
Directorate for Cultural Heritage	Riksantikvaren - direktoratet for kulturminneforvaltning
National Mediation Service	Sekretariatet for konfliktrådene
Norw. Centre for ICT in Education	Senter for IKT i utdanningen
Norw. Maritime Directorate	Sjøfartsdirektoratet
Norw. Tax Administration	Skattedirektoratet
Norw. Language Council	Språkradet
National Institute of Occupational Health	Statens arbeidsmiljøinstitutt
Accident Investigation Board Norway	Statens havarikommisjon for transport
Norw. Board of Health Supervision	Statens helsetilsyn
National Institute for Alcohol and Drug Research	Statens institutt for rusmiddelforskning
State Educational Loan Fund	Statens lånekasse for utdanning
Statistics Norway	Statistisk sentralbyrå
Directorate of Public Construction and Property	Statsbygg
Governor of Svalbard	Sysselmannen på Svalbard
Directorate of Customs and Excise	Toll- og avgiftsdirektoratet
National Insurance Court	Trygderetten
Norw. University of Life Sciences	Universitetet for miljø - og biovitenskap
Norw. Directorate of Immigration	Utlendingsdirektoratet
Norw. Immigration Appeals Board	Utlendingsnemnda
Directorate of Public Roads	Vegdirektoratet
Norw. Agency for Lifelong Learning	VOX, Nasjonalt fagorgan for kompetansepolitikk

Table A.2: Observations by type of steering practice, 2012 and 2015

Performance Steering Arts Council Norway (2015); Directorate (Dir.) of Integration and Diversity (2012); National Mediation Service (2012); Norw. Agency for Lifelong Learning (2012, 2015); Norw. Centre for ICT in Education (2012); Norw. Maritime Dir. (2015); Norw. Meteorological Institute (2012, 2015); Norw. Space Centre (2012)	*Double whammy* Dir. for Building Quality (2015); Dir. for Civil Protection and Emergency Planning (2012); Dir. of Customs and Excise (2012, 2015); Governor of Svalbard (2012); Institute of Marine Research (2015); National Police Dir. (2012); Norw. Dir. for Children, Youth and Family Affairs (2012, 2015); Norw. Dir. of Health (2012, 2015); Norw. Environment Agency (2012, 2015); Norw. Forest and Landscape Institute (2012); Norw. Institute for Agricultural and Environmental Research (2012); Norw. Labour and Welfare Administration (2015); Norw. Labour Inspection Authority (2012); Norw. Maritime Dir. (2012); Norw. National Rail Administration (2015); Norw. Petroleum Dir. (2015); Norw. Tax Administration (2012, 2015)
Relational steering Accident Investigation Board Norway (2012, 2015); Dir. of Public Construction and Property (2015); Equality and Anti-discrimination Board of Appeals (2012, 2015); National Insurance Court (2012, 2015); National Mediation Service (2015); Norw. Biotechnology Advisory Board (2015); Norw. Church Endowment (2015); Norw. Criminal Cases Review Commission (2012, 2015); Norw. Guarantee Institute for Export Credits (2012, 2015); Norw. Immigration Appeals Board (2015); Norw. Social Research (2012); Ombudsman for Children (2015); Patients' Injury Compensation Board (2012, 2015); Restoration Workshop of Nidaros Cathedral (2012, 2015)	*Behavioral steering* Dir. of Public Roads (2012); Financial Supervisory Authority (2012); Geological Survey of Norway (2012); Institute of Marine Research (2012); Norwegian (Norw.) Centre for ICT in Education (2015); Norw. Coastal Administration (2015); Norw. Food Safety Authority (2012, 2015); Norw. Petroleum Dir. (2012); Norw. Polar Institute (2012); Norw. Post and Telecommunication Authority (2012, 2015); Norw. Reindeer Husbandry Administration (2012)

Table A.3: Multinomial regression analysis. Coefficients as log odds of agencies' being subjected to each of the four types of steering vs. reference category (no clear steering profile).

	Model 1				*Model 2*				*Model 3*			
	Relational	Performance	Behavioral	Double whammy	Relational	Performance	Behavioral	Double whammy	Relational	Performance	Behavioral	Double whammy
Size	-0.429^{*}	-0.225	0.002	0.024^{**}	-0.422^{*}	-0.225	0.004	0.025^{**}	-0.436^{*}	-0.221	0.005	0.025^{**}
	(0.221)	(0.164)	(0.017)	(0.010)	(0.218)	(0.163)	(0.017)	(0.010)	(0.222)	(0.162)	(0.017)	(0.010)
Age	0.032	0.062	0.027	-0.038								
	(0.070)	(0.083)	(0.069)	(0.063)								
Age (dummy)					-0.202	0.597	-0.687	-0.163	-0.210	0.594	-0.693	-0.166
					(0.557)	(0.712)	(0.708)	(0.539)	(0.562)	(0.713)	(0.712)	(0.540)
Political salience	-0.266^{**}	0.001	-0.004	-0.024	-0.267^{**}	0.031	-0.003	-0.027	-0.261^{**}	0.029	-0.004	-0.027
	(0.129)	(0.070)	(0.040)	(0.040)	(0.132)	(0.065)	(0.040)	(0.039)	(0.133)	(0.065)	(0.040)	(0.039)
Contract year ('15)									0.525	-0.361	-0.871	-0.290
									(0.548)	(0.699)	(0.646)	(0.510)
Constant	0.152	-1.648^{***}	-1.892^{***}	-1.290^{***}	0.344	-1.739^{***}	-1.580^{***}	-1.385^{***}	0.063	-1.577^{**}	-1.227^{**}	-1.247^{***}
	(0.453)	(0.565)	(0.474)	(0.401)	(0.499)	(0.593)	(0.422)	(0.365)	(0.586)	(0.657)	(0.478)	(0.433)
Akaike Inf. Crit.	364.301	364.301	364.301	364.301	363.505	363.505	363.505	363.505	367.574	367.574	367.574	367.574

Notes: $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$. The robustness of the results has been tested by running the models with thresholds defined by using observations from 2012 and 2015 combined instead of from 2012 alone and with thresholds defined by using quartile values instead of thirds (see methods section). The results are not sensitive to these alternative specifications, with one exception: With thresholds defined by using quartile values (i.e., more narrow empirical categories) the two agencies mentioned in the text – the Directorate of Public Construction and Property and the Norwegian Immigration Appeals Board – are no longer categorized as subject to relational steering. As a result, this model specification returns a stronger negative relationship between increased size and the chance of being in the relational category than does the specification reported in Model 3 (whose results are illustrated also in Figure 5).