

Does performance information affect politicians' attitudes to spending and organizational reform?

Poul Aaes Nielsen
Aarhus University, Denmark

Martin Baekgaard
Aarhus University, Denmark

Work in progress – please do not cite without the authors' permission

Abstract:

Performance information systems have been widely adopted by national and local governments. This development has been argued to assist elected politicians in deciding on important questions of resource allocation and organizational reform, potentially creating a system of performance-based budgeting where decision-makers reward high-performers and punish low-performers. Others have argued, however, that politics will continue to undermine these efforts but also that inherent ambiguity of performance information—for instance that low performance signals a need for more resources—makes this unrealistic (Moynihan 2008). Empirically based knowledge of this is limited, however, and existing work relies on case studies or cross-sectional survey data unable to fully control for contextual factors.

The aim of this study is to examine the impact of performance information on politicians' attitudes toward spending and organizational reform. We did this by conducting a randomized survey experiment among all Danish city councillors, in which the treatment group was exposed to a weak information treatment concerning whether their municipality performed above, below, or around the national average within a highly salient policy area. We found, first, that performance information does not have one general effect on politicians' attitudes but that it depends on the level of performance reported. Second, information on performance both above and below—but not around—the national average leads to attitudes of increased spending, which indicates that lessons from performance information are ambiguous. Third, only high performance makes politicians less favorable to undertaking organizational reform. We discuss the theoretical and practical implications.

Introduction

Performance measurement schemes have been widely adopted by both national and local governments in many different countries (Bouckaert and Halligan 2008). One of the primary motivations behind this development was the perceived need to assist elected politicians in holding government agencies accountable and making more informed decisions on resource allocation and institutional design (Joyce 1993; Halachmi 2002). By shifting decision-makers focus from input and process regulation to government outputs and outcomes, elected representatives as well as agency managers were to pay greater attention to performance and creating public value (Kettl 1997).

Linking performance information to budget decisions, this has been argued to create a system of performance-based budgeting whereby decision-makers reward high-performers and punish low-performers, ultimately allocating resources according to the value-added of different programs and organizations (Breul 2007). Others have argued, however, that politics will continue to undermine these efforts but also that performance information itself, and the lessons to be learned from it, are inherently ambiguous (Moynihan 2006).

Empirically based knowledge of how performance information affects the decision-making of elected politicians is limited, however. Previous quantitative work has focused mainly on different types of performance information use or perceptions of its usefulness in decision-making (Poister and Streib 1999; Bogt 2004; Melkers and Willoughby 2005; A. T. K. Ho 2006; Pollitt 2006; Askim 2007; Askim 2009). While these are important questions, they can only tell us whether information was used but not if or how decisions were actually affected by performance information. Other studies have sought to trace the impact of performance scores on budget appropriations but so far results are mixed (Gilmour and Lewis 2006a; Gilmour and Lewis 2006b; Zaltsman 2009; Joyce 2011).

The aim of this study is to examine how performance information affects politicians' attitudes to resource allocation and organizational reform. These are primary concerns of a system of performance-based budgeting, and if performance information truly affects political decision-making, we should expect to find an impact here. When we study the impact of performance information, however, it is important to take into account what this information actually tells decision-makers. Particularly, lessons from performance information should differ depending on whether performance is portrayed as high or low. So far, this has only received limited attention in existing studies of how not only elected politicians but also public managers respond to performance information (two exceptions are Salge 2011; Nielsen 2012).

We identify three different theoretical perspectives on how performance information might influence political decision makers' attitudes to spending and reform. While traditional prescriptions of performance-based budgeting suggest that high performance will be rewarded, others argue that this is an open question subject to, among other aspects, existing preferences and the success of policy advocates in framing the information (Moynihan 2006; Moynihan 2008). Third, from a behaviorally based learning perspective we might expect systematic reactions depending on performance relative to aspirations, but not necessarily in the same manner as traditional prescriptions of performance-based budgeting. A further question is whether we should expect the effects of performance information on politicians' attitudes to be the same for attitudes to resource allocation and organizational reform. We argue that this is not the case.

An important feature of existing work on performance information is that it, like most work in public administration, relies on case studies or cross-sectional survey data that make it difficult to isolate its causal effects from different contextual factors. In order to address this challenge, we employ data from a randomized experiment embedded in a survey. Such experiments are often criticized for their limits in external validity. While not studying actual decision making, we move

some way toward accommodating this critique by surveying the population of Danish city councilors. But more importantly, our central concern here relates to establishing a causal link between performance information, its content, and how it is processed by decision-makers.

In the following section, we combine existing theoretical work in developing hypotheses concerning the impact of performance information on political attitudes. We then present the research design and empirical setting followed by a presentation and discussion of our findings. We conclude with a discussion of the theoretical and practical implications of this study and point to areas of future research.

Integrating Budgeting and Performance Information

The budget is the central tool in prioritizing how public resources are allocated across policies, programs, and societal groups. As Schumpeter phrased it, "...the budget is the skeleton of the state stripped of all misleading ideologies..." (1991, x). Understanding and improving the budgeting process has attracted considerable attention among scholars and practitioners alike, and numerous approaches to budgeting have been proposed and adopted over time (Rubin and Willoughby 2011). These approaches have been targeted at introducing a stronger element of rationality in the budgeting process (Willoughby 2011).

In recent decades, efforts to introduce performance-based budgeting systems have been widespread among national and local governments in different countries (Joyce and Tompkins 2002; Diamond 2003; Bouckaert and Halligan 2008). Performance-based budgeting is an attempt to develop and integrate non-financial performance indicators in the different stages of budgeting in order to establish a stronger link between how government allocates resources and the actual outputs or

outcomes produced by service providers (Joyce and Tompkins 2002). At the extreme, this has been described as a movement toward replacing the politics of budgeting with a scientific approach to budgeting by performance, where government spending follows program performance, i.e. its ability to create public value. More cautious observers have argued that this is unrealistic and that performance information will only be one variable among many other budget determinants that are likely to dominate, or at least limit, its impact (Joyce 2011).

A growing number of studies deal with the impact of performance information on budget processes. Some of these examine whether performance information is used or perceived as useful in budget decision making and they identify some general antecedents of use (Poister and Streib 1999; Bogt 2004; Melkers and Willoughby 2005; A. T. K. Ho 2006; Pollitt 2006; Askim 2007; Askim 2009; Zaltsman 2009; Hou et al. 2011; Raudla Forthcoming). While dissemination of performance information is widespread, these studies also point to barriers against its integration into budgeting. One of the primary obstacles to actual use has been a lack of appropriate data for measuring results or outcomes rather than activity or output, and for some services it may even be difficult to measure output (Joyce 2011; Heinrich 2012). This problem is further exacerbated when attempting to compare performance across agencies performing different tasks.

Other studies have sought to trace the impact of performance scores on budget appropriations but so far results are mixed, and other factors appear to be more important (Gilmour and Lewis 2006a; Gilmour and Lewis 2006b; Moynihan 2006; Zaltsman 2009; Heinrich 2012). There appears to be a tendency that performance information might have greater value at more disaggregated levels of government, such as at the municipal or county level (Poister and Streib 1999; Berman and Wang 2002; Melkers and Willoughby 2005) or for intradepartmental budget changes (A. T.-K. Ho 2011). Still, a widespread sentiment is that performance-based budgeting so far has largely failed to deliver

the fundamental changes in budgeting processes that its proponents intended (Moynihan 2006; Joyce 2011; Heinrich 2012).

Theory and Hypotheses

The lack of a major change in budgeting is less than surprising given the experiences from similar previous reform efforts (Joyce 1993). There are also compelling theoretical arguments for why we might be skeptic about any such changes. A general argument has been made that politics will tend to thwart attempts at basing budgeting on performance data (Downs and Larkey 1986; Joyce 2011). Decision makers often have strong interests in the size of budget appropriations, particularly if these affect important constituents or concern issues that are placed highly on the public agenda (Jones and Baumgartner 2005). Many policy decisions and reforms also have important symbolic effects that politicians are often well aware of (March and Olsen 1983).

Apart from politics, the complexity of the budget process itself may also undermine the systematic incorporation of performance information. Particularly, incrementalism posits that budgetary decision makers, rather than undertaking the heroic effort of drawing up an entirely new budget each year, will look to last year's base and existing compromises with conflict and negotiation limited to only few items and with incremental changes from year to year (Wildavsky and Caiden 1988). Due to bounds on human cognitive capacities, decision makers will thus resort to simplifying heuristics over comprehensive processing of performance information.

In contrast to this image, however, several authors note that although performance information does not determine budget appropriations, it is not necessarily ignored either as the previously cited studies of performance information use for budgeting has shown. Also, in an attempt to avoid

information overload, attention has been directed at simplifying performance data by collapsing multiple measures into fewer overall and performance assessments and making them more comparable (Moynihan 2006).

Taking into account the political nature of budgeting, Moynihan (2006; 2008) argues that rather than being ignored, performance information has become part of politics, with decision makers and policy advocates actively using the information to strategically pursue their interests. Proposing a dialogue theory of performance information use, he argues that all performance information is inherently ambiguous. Apart from differing opinions on what good performance is or should be, performance data typically does not convey why performance occurred or how to decide on which of multiple measures should be given priority (Moynihan 2006; Moynihan 2012; Nutley et al. 2012). Perhaps more important, at least compared to the traditional understanding of performance-based budgeting, is the choice ambiguity in relating performance assessments to budget appropriations. As Moynihan notes, “If a program is consistently performing well, does that indicate that it should receive greater resources or that it is already amply provided for? Is the poorer-performing program a candidate for elimination or just in need of additional resources?” (2006, 159). Thus, while the inherent ambiguity of performance information excludes any automatic impact on budgetary decision making, the same ambiguity makes different political and institutional actors engage in a game of interpreting the meaning and lessons of performance information in a favorable way, and it is how this might affect budget processes that we should seek to understand (Moynihan 2008; Moynihan and Hawes 2010; Moynihan 2012).

A separate strand of research that so far appears to have been absent in studies of public budgeting is the application of behaviorally based models of learning from performance feedback (Greve 2003). Building on the work of March & Simon (1958) and Cyert & March (1963), these models posit that because high and low performers face different adaptive pressures and learning potentials,

their responses should be sensitive to how well they perform. Responses to performance information are further argued to follow a satisficing logic (March and Simon 1958) by depending on performance relative to behaviorally based aspiration levels or targets that reflect information cues on what might be expected of an organization's performance. Thus, when performance data shows that performance falls short of aspirations, this is expected to provide a signal to an organization that some sort of change is required (Salge 2011).

Accordingly, a large number of private sector studies have shown that performance feedback systematically affects resource allocation and organizational change (for a review of this literature, see Greve 2003). Recently this model has also been applied successfully to how public organizations and managers respond to performance feedback (Salge 2011; Nielsen 2012).

The diagnosis that a performance problem exists has been found to increase resources spent on problemistic search for potential solutions to the problem (Greve 2003). Conversely, performing above the aspiration level is argued to reduce the need for search. Similarly, performance problems indicate that organizational change may be needed. Assuming that aspiration levels are fundamental to how choice problems are represented and understood (Kahneman and Tversky 1979), performing below the aspiration level should increase actors' risk tolerance as they have little to lose, and consequently they are expected to be more willing to pursue risky organizational change (Greve 1998). If decision makers are sensitive to performance levels, it is essentially an instance of learning behaviour in which decision makers seek to adapt to environmental signals (Greve 2003). Such adaptive behavior might arise from learning what the organization's real performance potential is, or it might reflect political or social norm pressure to meet performance aspirations (Greve 2003).

Whether also politically elected officials, like many public and private managers, follow this pattern of response to performance information is an open question. Concerns specific to politicians may

moderate or undermine their impact. In the following section, we present our hypotheses concerning the impact of performance information on political attitudes to spending and reform.

The Impact on Attitudes to Spending and Reform

Performance information serves as an instrument for elected officials to making more informed decisions and to better hold agencies accountable (Joyce 1993; Halachmi 2002). When studying the impact of introducing performance information into the political decision making process, focus tends to be on its impact on resource allocation.

Another important aspect, however, concerns institutional design. The organization of public services is a key part of securing satisfactory performance (Meier and O'Toole 2009), and undertaking organizational reform is a common political response when there is a perceived need for change (March and Olsen 1983; Hood 2011). In times of austerity it may also be a more accessible option for pursuing performance improvements. Just as resource allocation can be used as reward or punishment of managers and their agencies, so may organizational reform (Hood 2011). Organizational reform comes in many guises, including new rules, restructuring of responsibilities, or departmental amalgamations, but often reforms will have consequences for the tasks and autonomy of agencies (Wilson 1989; Carpenter and Krause 2012). In many instances, impacts of performance information on organizational reform may therefore be as realistic and consequential as similar changes in resource allocation.

A first question to ask is how we might expect the functional form of the relationship between performance information and politicians' attitudes to look. Most studies of performance information examines whether performance information is used, irrespective of the content of the information,

i.e. whether performance is presented as high, low, or medium (Melkers and Willoughby 2005; A. T. K. Ho 2006; Askim 2009). Other studies seek to test whether the level of performance has an impact on resource allocation (Gilmour and Lewis 2006a; Gilmour and Lewis 2006b; Zaltsman 2009; Heinrich 2012). Considering the theoretical approaches presented above, the latter approach appears more useful for this study.

The theory of performance-based budgeting as espoused by its proponents suggests that decision-makers reward high-performers and punish low-performers. Thus, their responses to performance information should clearly differ depending in the content of the information. The same is true for a learning perspective on performance feedback, as it suggests that decision makers adapt their responses to the level of performance. The dialogue theory of performance budgeting gives a somewhat more blurry account. It perceives the interpretation of performance information as a more open process where advocacy and negotiation between different actors play an important role due to the ambiguity of performance information (Moynihan 2006). When limiting this ambiguity perspective to the content of the information (and not subsequent choices), however, the importance of ambiguity is likely to differ depending on the quality and source of the performance data. Particularly concerning high quality performance information, we might expect that while the information will be open for interpretation, some interpretations will often be easier to support than others. Generally, we might therefore expect that:

H1. The impact of performance information on politicians' attitudes to spending depends on the level of performance.

H2. The impact of performance information on politicians' attitudes to organizational reform depends on the level of performance.

Hypotheses 1 and 2 state that the impact of performance information is expected to be conditioned by the content of the information, i.e. by the measured level of performance. We now turn to more exact predictions of how this conditioning relationship behaves for attitudes to spending and reform.

Concerning attitudes to organizational reform, it is worth noting that organizational reform may have different rationales. The espoused theory of performance-based budgeting would generally suggest that decision makers seeking to spend tax-payers' money responsibly should support well-functioning programs and show little interest in maintaining poor performers. When decision makers consider undertaking organizational reform, low-performing organizations should therefore be the primary candidates. This can also be perceived as an incentive mechanism. Thus, given the potential consequences of reform to organizational and managerial autonomy in deciding over task and organization, introducing reform can be seen as a threat to low-performing organizations and, conversely, being given discretion to decide on internal matters might be a benefit of performing well (Carpenter and Krause 2012).

Turning to the learning perspective, several studies of private sector businesses have shown that performance below aspirations increases the rate of organizational change (Greve 1998; Greve 2003). Performance problems signal that an organization is unable to perform at the level that can be expected of it compared to its own past performance and the performance of others (Greve 2003) and hence suggests a need for reform. Because consequences of organizational change typically contain an element of uncertainty, decision makers' level of risk tolerance is also of importance. With aspiration levels serving as central thresholds of evaluation, decision makers' assessments of their performance is expected to depend primarily on how performance relates to the threshold rather than on gradations in how far performance is over or under the threshold (March and Simon

1958). Thus, performing below aspirations will tend to increase decision makers' risk tolerance. This is also the mechanism that is present in Kahneman and Tversky's (1979) prospect theory according to which being in the domain of losses makes actors more risk seeking. As a consequence, their willingness to adopt measures of organizational reform should also increase.

For the ambiguity perspective of dialogue theory, predictions are less clear due to the strategic responses of actors engaged in interpretation and political advocacy. As already mentioned, the existence of high quality performance information may limit the importance of information ambiguity. Some interpretation frames of performance levels may be easier to support and justify than others. But when discussing the impact of the information on particular attitudes, the ambiguity of choice also comes into play. Choice ambiguity entails that agreement on how to interpret the level of performance does not warrant agreement on what action should be taken as a consequence of this knowledge (Moynihan 2006).

As the previous quote illustrates, the importance of choice ambiguity has primarily been noted in relation to budget allocations. We argue, however, that this will not be the case for attitudes to organizational reform. Whereas both high and low performance can and often is construed to warrant increased funding, it is less obvious how high performance could make decision makers more inclined to pursue organizational change. High performance would rather suggest that the right formula has been found and should be sustained. High performance arguably also makes it easier for managers to defend their organizations against pressures for reform (Carpenter and Krause 2012). Similarly, low performance would likely make it easier to justify reform measures, and likewise more difficult for managers and other stakeholders to fight reform. One possible exception might be if managers or stakeholders succeed in attributing blame for poor performance to their continuous struggles with implementing past reforms, but we do not expect this to be the usual outcome. We therefore propose that the significance of choice ambiguity identified in

dialogue theory is of little importance to attitudes to organizational reform. Concerning organizational reform we are thus able to present one common hypothesis for the different theoretical perspectives discussed here.

H3. Performance information has a negative impact on politicians' attitudes to organizational reform when performance is above and a positive impact when performance is below the aspiration level.

In contrast to the attitudes to organizational reform, the theoretical expectations regarding the impact of performance information on spending preferences are not in agreement. According to dialogue theory, the choice ambiguity in relating performance levels to attitudes and decision making makes it difficult to predict consistent response patterns. Rather, actors' responses are likely to differ depending on decision makers' existing preferences and the different institutional roles they assume (Moynihan 2006) as well as, for instance, the nature of the program in question. We therefore refrain from offer hypotheses based on this perspective.

When we turn to the espoused theory of performance-based budgeting, however, the prediction is clearer. The suggestion here is that resources should be allocated according to how well programs and organizations perform, so that public resources are spent more efficiently. While this may by now have turned into somewhat of a straw man, the defense and rhetorical use of this argument has not seized (Joyce 2011). This perspective also sets out a clear rationale for how decision makers do or should respond to performance information that is of academic interest. This results in the following hypothesis.

H4. Performance information has a positive impact on politicians' attitudes to spending when performance is above and a negative impact when performance is below the aspiration level.

Finally, turning to the behavioral learning perspective we argue that a different logic is at play. To our knowledge, this perspective has not previously been applied to how public resources are allocated. While not directly transferable, studies of resource allocation primarily within private business firms have shown that low performance relative to an organization's aspiration level leads to increases in research and development budgets (Greve 2003) and spending on innovative search (Salge 2011). The logic behind this pattern is that because all performance points below the aspiration level are perceived as unsatisfactory, resources should be spent locating the cause of the performance problem and delivering potential solutions that may help increase future performance. In this sense, learning from performance data enables decision makers to adjust their investments in organizational search to ensure organizational adaptation to environmental pressures.

There are some limitations, however, in adapting this model to political budget processes. In some cases, an alternative to increasing the funding for poorly performing programs is to simply shut them down.¹ Often, however, public programs are perceived as too important or too politically costly to shut down. If this is the case, decision makers must either accept poor performance or take measures to change it. These measures are not necessarily to increase funding but this is often one way of dealing with the problem. Conversely, cutting budgets in response to poor performance will typically do little to improve future performance. This therefore points to the opposite prediction of that in Hypothesis 4, though with the restriction that it suggests an effect of performance

¹ This may also happen in business firms, though, if performance is so poor that owners prefer to take their losses up front rather than keep running deficits.

information when performance is below but not necessarily when performance is either around or above the aspiration level.

H5. Performance information has a positive impact on politicians' attitudes to spending when performance is below the aspiration level.

Which of the different logics political decision makers actually follow is an empirical question that also requires elaboration of the context surrounding the particular budgetary process under study. In the following section we present our empirical case, and we detail the research design and data that we have used to test our hypotheses.

Testing the Impact of Performance Information

In order to test our hypotheses, we used data from an email-based survey sent to the population of politicians in Danish city councils. Data was gathered in May and June 2012. 1,295 responded to the survey corresponding to a response rate of 52.9 per cent. Denmark is characterized by a high degree of decentralization and the municipalities are responsible for a large number of policy areas; for instance public schools, day care, roads, and elderly care. In total their functions amount to approximately half of all public expenditures (Blom-Hansen and Heeager 2011, 227). The research question is thus examined in a setting in which performance is highly important to citizens. We checked whether the politicians who responded to the survey differed from the population of Danish city council politicians as a whole with regard to their gender, party affiliation, political posts, and the population size of their municipalities. Response rates were considerably higher (57.0 per cent)

among members of socialist parties than among members of other parties (47.4 per cent). Moreover, response rates were also somewhat higher for committee chairmen (56.6 per cent) than for other respondents (50.4 per cent). This is only a minor problem in this particular case, since our main concern is to establish the link between one particular information cue and politician preferences. Moreover, our analyses show that there are no significant differences in the effect of performance information and its content on preferences between the over- and underrepresented groups (these results are not shown in the paper but are available upon request). Hence, we feel confident that our results can indeed be generalized to all Danish city councilors.

It is associated with several challenges to detect the effect of performance information on preferences. Selection bias is a likely problem in cross-sectional analyses because the use of performance information may be systematically correlated with known or unknown characteristics (James 2011). For instance, the ideological standpoint differs between politicians, and ideology may be correlated with the knowledge of performance as well as politician preferences. Moreover, we may experience endogeneity bias if policy preferences affect the extent to which performance information is gathered or used. For instance, politicians with more extreme attitudes may be more eager to gather information on performance.

In order to deal with these challenges we used a survey experimental design in which information was randomly assigned to some but not all politicians. The randomization creates groups that are probabilistically similar to each other on average entailing that differences in outcomes can be used to estimate the causal effect of the treatment. While the internal validity is generally high in randomized experiments, it often also comes at a price, as external validity suffers due to the very specific experimental treatment that is being used (Morton and Williams 2010, 253–265). We will return to this topic in the concluding section.

Using simple randomization across municipalities, we created three groups of politicians of which two are relevant for the purpose of this paper. For this reason, only two thirds of the 1,295 respondents were included in the analysis. Since the respondents were assigned randomly to the three groups, the exclusion of one of the groups does not affect the representativity of the remaining two groups that were included. One group received an information cue about how the public schools in their municipality perform compared to public schools in all other Danish municipalities, while the other group received no such information. The cue was based on real performance data and contained information that their schools performed either among the best, the middle, or the worst third of Danish schools with regard to their grades and when accounting for student composition. Compared to information cues in previous research (James 2011) this is arguably a rather weak cue as the performance categories are very broad and hence encompass large differences in performance. The exact wording of the information cue is presented in the appendix. The policy area of public schools is considered one of the major municipal tasks in Danish municipalities and is highly salient compared to other policy areas (Bækgaard 2011). The policy area is thus very important to politicians and citizens alike. This means that our analysis is in a policy area where performance actually matters and, hence, where we should expect that performance information affects preferences to a higher extent than on less salient policy areas. As can be seen in Table 1, the treatment was indeed randomly assigned among individuals of different gender, age, length of education, city council seniority, party affiliation, and for individuals coming from municipalities of different population sizes entailing that we do not need to control for such variables in our analysis.

[Table 1 about here]

Existing studies have shown that the source of information may have a substantial impact on responses to frames and information (Cohen 2003, 811–812; James 2011; Slothuus and de Vreese 2010; Stubager and Slothuus 2012). These findings may also apply to this case. Potentially, the credibility of the source may be evaluated differently by politicians from different parties and this may in turn affect their responses to performance information. If there are also systematic differences between how municipalities governed by, for instance, leftist and rightist parties perform, the findings from our survey experiment may be biased. For our purposes, we therefore need a source of information that is conceived of as being fairly neutral in political terms. We use information from the independent government agency, KREVI. The purpose of this organization is exactly to provide neutral and reliable information on how municipalities perform relative to one another.

In the survey, the information cue was placed prior to the two dependent variables measuring politicians' preferences. Between the cue and the preference questions, we placed a number of questions about politicians' attitudes to performance information and about the extent to which they use performance information. The idea was not to make the link between the information cue and the preference questions too obvious. The dependent variables measure two different kinds of preferences. Spending preferences were measured in a five-point scale battery in which the respondents were asked to indicate if they would prefer much less, less, the same, more or much more spending on a number of different policy areas governed by the municipalities. A similar battery has previously been used in other studies on politician spending preferences (Sørensen 1995; Baekgaard 2010). For our purposes, the item about preferences for spending on public schools was the only one of interest in the battery. This item was placed as number four out of eleven in the battery. Again, the idea was to make the link between the information cue and the

response variable less obvious. The variation on this item is limited since by far the most politicians (86 per cent) responded that they prefer ‘the same’ or ‘more’ spending. This is a pattern that corresponds closely with findings in earlier investigations of Danish politicians’ preferences for spending on public schools. Apparently, local politicians tend to agree to a very high extent on this particular topic (Serritzlew 2003; Baekgaard 2010). The lack of variation may entail that an effect of information is underestimated compared to effects on other preference variables.

Preferences for organizational change were measured by a single-item question in which the respondents were asked to consider a hypothetical case: ‘Imagine a situation in which, according to anticipations, 4-8 per cent of the budget for two schools can be saved each year if all educational services are placed on one of the schools, while the other one is closed. The teaching quality is not expected to be affected by the school merger. To what extent do you agree or disagree that the two schools ought to be merged in this case?’ The answers were given on a scale ranging from 0-10 with 0 meaning totally disagree and 10 meaning totally agree. As can be seen, the formulation of the question entails that there are benefits associated with merging the schools thus making this a rather conservative test since politicians arguably will tend to agree on the question regardless of whether performance information was received or not. Table 2 provides descriptive statistics on the dependent variables:

[Table 2 about here]

Content of Performance Information and Preferences

While the information cue is randomly assigned to individuals, for ethical reasons we have chosen not to randomize what kind of information is given to individuals in the treatment group. Respondents from municipalities with well-performing schools have to receive a treatment indicating that their schools are performing well, in case they are assigned to the treatment group. Since larger municipalities generally perform better than smaller municipalities, and since larger municipalities also have larger city councils, more respondents were given a positive information cue than a neutral or a negative cue. Among the respondents who received information, 176 (41 per cent) received positive, 117 (27 per cent) neutral, and 140 (32 per cent) negative information.

This also entails that unobservables at the municipal level (for instance, actual knowledge of performance, recent school mergers etc.) may render the effects spurious, in case such unobservables are both correlated with the kind of information that is provided and the dependent variables. In order to account for omitted variable bias on the municipal level, we employ fixed effects regressions, entailing that all municipal differences are effectively leveled out and treatment effects are compared only between individuals from the same municipality (Allison 2009). This means that the preference effect of being given, for instance, positive information is estimated only by comparison to politicians from the same high-performing municipality who did not receive performance information. Any information effects can therefore not be caused by politicians knowing or sensing in advance how well their municipality performs. The fixed effects specification therefore effectively creates a real experimental design even though actual municipal performance is not randomly assigned across municipalities.

The hypotheses state that the effect of performance information depends on whether the information provided is positive, neutral, or negative. We test this by interacting the information treatment with

dummies describing the content of the information (good, neutral, bad). As municipality fixed effects are used to control for all factors that are constant within municipalities, the constitutive terms describing municipality school performance are not included in the regression analysis (Allison 2009). Table 3 provides the main findings:²

[Table 3 about here]

Model 1 concerns the effect of performance information on spending preferences, while Model 3 shows the effect on preferences for organizational change. In neither of the models do we account for the content of the performance information. As predicted by Hypotheses 1 and 2, both effects are insignificant. Performance information does not in itself affect preferences in this case. Models 2 and 4 take the content of information into account. In these analyses, the effect of neutral performance can be found by looking at the effect of the information variable, while the effects of information about good and bad performance, respectively, can be deduced by interpreting the effect of the information variable in conjunction with the effect of the interaction terms.

Looking first at Model 2, we see that information on good performance as well as information on bad performance significantly affects politicians' preferences for spending on public schools but in a way which neither accords with Hypothesis 4, nor with hypothesis 5. In both cases information is associated with a positive effect on spending preferences. In contrast, politicians exposed to neutral

² We also tested the hypotheses in a model in which random effects was used and, hence, the interaction term between the information cue and information content as well as both constitutive terms were included. In addition, we tested the hypotheses in an analysis in which we divided the sample into three different experiments entailing that we analyzed the respondents from well-performing municipalities in a separate regression, the respondents from neutral-performing in another, and the respondents from worse-performing in a third. Such alternative specifications produce results similar to the ones in the fixed-effects specification. We have chosen only to show the findings from the fixed-effects specification here since this specification is advantageous compared to other specifications due to the fact that unobservables do not affect our results here.

information to a significantly lesser extent than the control group prefer increased spending on public schools.

Turning to Model 4, we see that neither neutral information nor information about bad performance affect politicians' preferences for organizational change. On the other hand, information about good performance has a negative and significant effect on preferences entailing that politicians, who are exposed to this kind of information, to a lesser extent agree that organizational change should take place than do their counterparts in the control group. Hence, the findings partly corroborate Hypothesis 3.

While the findings in Models 2 and 4 differ in terms of the significance of the findings, a common pattern also appears; information about good performance seem to increase the likelihood of preferring more spending and to decrease the likelihood of preferring organizational change. This pattern accords very well with the prescriptive theory on performance based budgeting which suggests that good performance will be rewarded. On the other hand, this perspective has less to offer in terms of explaining the findings regarding the poorly performing municipalities. It seems as if punishment mechanisms matter less in this particular case. One explanation might be that it is not politically feasible to punish bad performers on a highly salient policy area. As the service (public schools) is one which is highly demanded by citizens, a natural reaction to bad performance might be to allocate further resources to the policy area in order to strengthen future performance.

Effect Size of the Performance Information Cue

To get a better grasp of effect sizes, we calculate mean values on both dependent variables for each mode of information, and for the treatment as well as the control group. Table 4 provides additional information on several aspects of the nature of the effect:

[Table 4 about here]

The average difference in spending preferences between those who receive and those who do not receive the information treatment in municipalities that are performing well is 0.14 point, corresponding to approximately 18 per cent of the standard deviation of the dependent variable. Similar effect sizes are found when looking at the effect of neutral and bad information. Regarding preferences for organizational change, the difference between the treatment and control group is 0.44 point in well performing municipalities, corresponding to approximately 13 per cent of the standard deviation of the dependent variable. In other words, the effect of the treatment is rather weak in substantial terms. We should, however, keep in mind that the treatment presented to the politicians is also rather weak thus suggesting that we might find stronger effects if a stronger treatment was used.

The table also adds additional information that is useful when interpreting the results in Table 3. In Table 3, Model 2 we found that information has a negative effect on spending preferences for politicians from municipalities performing neither good nor bad. However, as shown in Table 4, politicians from such municipalities on average are slightly positive towards increasing spending regardless of whether they have received the information cue or not. A plausible explanation of the

negative effect would be that neutral information is interpreted by the politician as a sign that the municipality is performing as it is supposed to and that no further resource allocation is needed. Finally, the findings in Table 4 indicate that actual performance matters for politician preferences and that politicians have some knowledge of actual performance. This is particularly evident when looking at the preferences for school mergers where politicians from well performing municipalities to a much lesser extent than politicians from worse performing municipalities prefer school mergers regardless of whether they have received the information treatment or not. This puts the effect of the information treatment in perspective because knowledge is likely to confound the effect of the information cue. The fact that the treatment has an effect despite politicians having some knowledge of performance highlights the potential of performance information. Performance information has an effect on preferences even when politicians are presented to a rather weak cue.

Discussion and Conclusion

This paper examined the impact of performance information on politicians' attitudes to spending and organizational reform. The aim was not to test different theories but instead to use such theories to form and subsequently test specific predictions about the impact of performance information on politicians' preferences. Overall, our analyses present evidence that performance information matters in the sense that it affects the preferences of central decision-makers. This is novel information as empirically based knowledge about the effects of performance information is limited. The internal validity of this finding is high, because information on real performance was used in the experiment, and because the evidence was based on a randomized experiment in which, moreover, municipal differences, were effectively leveled out by fixed effects estimation. We found that performance information does not in itself affect preferences. Contrary to most other studies on

the effects of performance information, we furthermore investigated whether the effect of performance information depends on the performance level reported. The findings are non-trivial and complex to interpret.

We found that the effects depend on what kind of information is provided. Information about good performance is associated with more positive attitudes towards the policy area in question, as politicians in the treatment group prefer a higher level of spending and to a lesser extent prefer organizational reform than the control group. These findings accord very well with traditional prescriptions of performance-based budgeting that good performance should be rewarded while bad performance should be punished. However, turning to the impact of information in poorly performing municipalities, we see that low performance is not punished, as information about low performance is actually associated with preferences for higher spending and about the same preferences for organizational reform as in the control group. In other words, reward- and punishment mechanisms only to a limited extent explain the impact of different modes of performance information.

The question is how we should make sense of these findings? For one thing, we can note that extant theories are insufficient in explaining these findings except perhaps for the dialogue theory which, however, is hard to test, because performance information is not associated with any specific outcome. In line with dialogue theory (Moynihan 2006; Moynihan 2012), one interpretation might be that different rationales are linked to different modes of information. While the reward mechanism may be significant for high performers, for low-performing municipalities, however, the information about performance may signal that more resources are needed in order to increase the level of performance to an acceptable level. The latter part of the interpretation may be particularly important in highly politically salient policy areas as the one examined in this paper, because the pressure for enhancing performance in such areas is particularly strong. The question obviously

calls for more systematic research into the rationales linked to different modes of performance information. Such research could, for instance compare the effects of performance information on highly salient and less salient policy areas with the purpose of examining whether the punishment mechanism matters more in less salient policy areas. As for now, the only certain conclusion that we can draw on basis of our findings is that the effects of performance information are ambiguous.

Moreover, our theoretical discussion as well as our findings indicate that the effects of performance information depend on what kind of preferences we focus on. While this might sound trivial it does, in fact, have important implications. From a theoretical perspective it entails that some theories are better able to explain the impact on some kinds of preferences while other theories are better able to explain the impact on others. What is needed here is a more nuanced theoretical framework which takes into account that performance information may have a different impact on different outcome measures.

A common critique of experiments remains to be touched upon; do our findings travel to other cases as well (McDermott 2002, 334)? While the internal validity is high, this does not automatically entail that external validity is low. It might be argued that information is often not presented in quite as clear a version as in this case. Yet, the information presented bears much in resemblance with actual benchmarking activities, in which public organizations are compared to their peers (Ammons 2000). Moreover, in contrast to laboratory studies, where college students typically are used as respondents, our respondents are actual politicians. Also worth mentioning is the fact that the examination was made on a highly salient topic which politicians arguably care about. This may entail that our findings only travel to highly salient topics. What is needed here is a similar examination of less salient topics where, for instance, information on low performance may be associated with much stronger effects on preferences for organizational reform.

In conclusion, the main lesson to be learned from this paper is that the effects of performance information on decision makers' preferences are complex; that no single theory is able to fully explain the patterns found in this paper; and that different rationales may apply to different kinds of information. We therefore strongly encourage new empirical studies to reproduce and expand on the findings in this paper and to study the mechanisms that link different modes of performance information to preferences. Such studies are needed in order to build theories that more accurately predict the effects of performance information than the theories existing today.

Appendix

The information cue was included in the wording placed prior to a number of likert-scaled questions about the importance of performance information for politicians when they are to decide on important political matters.

The treatment group received the wording presented below. Respondents in the control group received exactly the same wording apart from the information presented in italics.¹

'The quality of public services is greatly debated. For instance, there has been a marked focus on measuring teaching quality in Danish public schools.

The last 3-year report about the grades at the final public school exam showed that the average grade in Aarhus² municipality was placed in the middle³ third among the Danish municipalities when taking the social composition of pupils into account (KREVI report 2011).

To what extent do you agree or disagree in the following:

When I am to decide on important political matters about the municipal service, I use...'

¹: This information was of course not presented in italics in the original survey.

²: The name was inserted on the respondent's municipality.

³: Best, middle or worst was inserted here depending on the performance of the respondent's municipality.

Tables

Table 1: Testing the randomization

	Model 1: Experimental treatment and background information (logistic regression)
Gender (male = 1)	0.171 (0.154)
Age (years)	0.002 (0.007)
Length of education (university degree or alike = 1)	-0.063 (0.156)
City council seniority (years)	-0.006 (0.010)
Party affiliation (member of socialist party = 1)	0.199 (0.140)
Municipal population size (ln)	-0.100 (0.100)
N	837
Nagelkerke's R ²	0.01

Entries are logistic regression coefficients. Standard errors in parentheses. **: $p < 0.01$; *: $p < 0.05$.

Table 2: Descriptive statistics on the dependent variable

	Min	Max	Mean	Std. dev.	N
Spending preferences	1	5	3.42	0.75	844
Preferences for school mergers	0	10	6.03	3.24	806

Table 3: Content of performance information and politicians' preferences.

	Model 1	Model 2	Model 3	Model 4
	Spending preferences	Spending preferences	Preferences for school mergers	Preferences for school mergers
	Ordered logit; fixed effects	Ordered logit; fixed effects	Ordinary least squares; fixed effects	Ordinary least squares; fixed effects
Information	0.159 (0.145)	-0.643 (0.285) *	-0.198 (0.235)	0.484 (0.457)
Information x good performance		1.186 (0.367)**		-1.172 (0.589) *
Information x bad performance		0.971 (0.381) *		-0.641 (0.608)
N	844	844	806	806
R ²	-	-	0.19	0.19
Nagelkerke's R ²	0.09	0.10	-	-

Entries are regression coefficients (OLS and ordered logit). Standard errors in parentheses; **: p < 0.01; *: p < 0.05

Table 4: Effect sizes

	Spending preferences		Preferences for school mergers	
	Treatment group	Control group	Treatment group	Control group
Bad performance	3.47 (0.71)	3.34 (0.78)	6.05 (3.33)	6.26 (3.41)
Neutral performance	3.38 (0.78)	3.48 (0.82)	6.80 (2.92)	6.23 (3.10)
Good performance	3.49 (0.68)	3.35 (0.74)	5.38 (3.26)	5.82 (3.22)

Entries are mean values. Standard deviations in parentheses.

References

- Allison, Paul D. 2009. *Fixed Effects Regression Models*. Thousand Oaks, CA: Sage Publications.
- Ammons, David N. 2000. 'Benchmarking as a Performance Management Tool: Experiences Among Municipalities in North Carolina'. *Journal of Public Budgeting, Accounting and Financial Management* 12 (1): 106–124.
- Askim, J. 2007. 'How Do Politicians Use Performance Information? An Analysis of the Norwegian Local Government Experience'. *International Review of Administrative Sciences* 73 (3): 453–472.
- . 2009. 'The Demand Side of Performance Measurement: Explaining Councillors' Utilization of Performance Information in Policymaking'. *International Public Management Journal* 12 (1): 24–47.
- Baekgaard, Martin. 2010. 'Self-selection or Socialization? A Dynamic Analysis of Committee Member Preferences'. *Legislative Studies Quarterly* 35 (3): 337–359.
- Berman, Evan, and XiaoHu Wang. 2002. 'Performance Measurement in US Counties: Capacity for Reform'. *Public Administration Review* 60 (5): 409–420.
- Blom-Hansen, Jens, and Anne Heeager. 2011. 'Denmark: Between Local Democracy and Implementing Agency of the Welfare State'. In *The Oxford Handbook of Local and Regional Democracy in Europe*, ed. John Loughlin, Frank Hendriks, and Anders Lidström, 221–240. Oxford: Oxford University Press.
- Bogt, Henk J. ter. 2004. 'Politicians in Search of Performance Information? Survey Research on Dutch Aldermen's Use of Performance Information'. *Financial Accountability & Management* 20 (3): 221–252.
- Bouckaert, Geert, and John Halligan. 2008. *Managing Performance: International Comparisons*. New York: Routledge.
- Breul, Jonathan. 2007. 'GPRA—A Foundation for Performance Budgeting'. *Public Performance & Management Review* 30 (3): 312–331.
- Bækgaard, Martin. 2011. 'Committee Bias in Legislatures with a High Degree of Party Cohesion: Evidence from Danish Municipalities'. *European Journal of Political Research* 50 (3): 315–335.
- Carpenter, Daniel P., and George A. Krause. 2012. 'Reputation and Public Administration'. *Public Administration Review* 72 (1): 26–32.
- Cohen, Geoffrey L. 2003. 'Party Over Policy: The Dominating Impact of Group Influence on Political Beliefs'. *Journal of Personality and Social Psychology* 85 (5): 808–822.
- Cyert, Richard M., and James G. March. 1963. *A Behavioral Theory of the Firm*. Englewood Cliffs, NJ: Prentice-Hall.
- Diamond, J. 2003. *Performance Budgeting: Managing the Reform Process (EPub)*. International Monetary Fund.
- Downs, George W., and Patrick D. Larkey. 1986. *The Search for Government Efficiency*. New York: Random House.

- Gilmour, John B., and David E. Lewis. 2006a. 'Assessing Performance Budgeting at OMB: The Influence of Politics, Performance, and Program Size'. *Journal of Public Administration Research and Theory* 16 (2): 169–186.
- . 2006b. 'Does Performance Budgeting Work? An Examination of the Office of Management and Budget's PART Scores'. *Public Administration Review* 66 (5): 742–752.
- Greve, Henrich R. 1998. 'Performance, Aspirations, and Risky Organizational Change'. *Administrative Science Quarterly* 43 (1): 58–86.
- . 2003. *Organizational Learning from Performance Feedback: A Behavioral Perspective on Innovation and Change*. Cambridge: Cambridge University Press.
- Halachmi, Arie. 2002. 'Performance Measurement, Accountability, and Improved Performance'. *Public Performance & Management Review* 25 (4): 370–374.
- Heinrich, Carolyn J. 2012. 'How Credible Is the Evidence, and Does It Matter? An Analysis of the Program Assessment Rating Tool'. *Public Administration Review* 72 (1): 123–134.
- Ho, A. T. K. 2006. 'Accounting for the Value of Performance Measurement from the Perspective of Midwestern Mayors'. *Journal of Public Administration Research and Theory* 16 (2): 217–237.
- Ho, Alfred Tat-Kei. 2011. 'PBB in American Local Governments: It's More Than a Management Tool'. *Public Administration Review* 71 (3): 391–401.
- Hood, Christopher. 2011. *The Blame Game: Spin, Bureaucracy, and Self-Preservation in Government*. Princeton University Press.
- Hou, Yilin, Robin S. Lunsford, Katy C. Sides, and Kelsey A. Jones. 2011. 'State Performance-Based Budgeting in Boom and Bust Years: An Analytical Framework and Survey of the States'. Wiley-Blackwell.
- James, Oliver. 2011. 'Performance Measures and Democracy: Information Effects on Citizens in Field and Laboratory Experiments'. *Journal of Public Administration Research and Theory* 21 (3): 399–418.
- Jones, Bryan D., and Frank R. Baumgartner. 2005. *The Politics of Attention: How Government Prioritizes Problems*. Chicago, IL: University of Chicago Press.
- Joyce, Philip G. 1993. 'Using Performance Measures for Federal Budgeting: Proposals and Prospects'. *Public Budgeting & Finance* 13 (4): 3–17.
- . 2011. 'The Obama Administration and PBB: Building on the Legacy of Federal Performance-Informed Budgeting?'. *Public Administration Review* 71 (3): 356–367.
- Joyce, Philip G., and Susan S. Tompkins. 2002. 'Using Performance Information for Budgeting: Clarifying the Framework and Investigating Recent State Experience'. In *Meeting the Challenges of Performance-oriented Government*, ed. Kathryn Newcomer, Edward T. Jennings, Cheryl Broom, and Allen Lomax, 61–96. Washington, DC: American Society for Public Administration.
- Kahneman, Daniel, and Amos Tversky. 1979. 'Prospect Theory: An Analysis of Decision Under Risk'. *Econometrica* 47 (2): 263.
- Kettl, Donald F. 1997. 'The Global Revolution in Public Management: Driving Themes, Missing Links'. *Journal of Policy Analysis and Management* 16 (3): 446–462.

- March, James G., and Johan P. Olsen. 1983. 'Organizing Political Life: What Administrative Reorganization Tells Us About Government'. *The American Political Science Review*: 281–296.
- March, James G., and Herbert A. Simon. 1958. *Organizations*. Cambridge, MA: John Wiley.
- McDermott, Rose. 2002. 'Experimental Methodology in Political Science'. *Political Analysis* 10 (4): 325–342.
- Meier, Kenneth J., and Laurence J. O'Toole. 2009. 'The Proverbs of New Public Management Lessons From an Evidence-Based Research Agenda'. *The American Review of Public Administration* 39 (1): 4–22.
- Melkers, Julia, and Katherine Willoughby. 2005. 'Models of Performance-Measurement Use in Local Governments: Understanding Budgeting, Communication, and Lasting Effects'. *Public Administration Review* 65 (2): 180–190.
- Morton, Rebecca B., and Kenneth C. Williams. 2010. *Experimental Political Science and the Study of Causality: From Nature to the Lab*. New York: Cambridge University Press.
- Moynihan, Donald P. 2006. 'What Do We Talk About When We Talk About Performance? Dialogue Theory and Performance Budgeting'. *Journal of Public Administration Research and Theory* 16 (2): 151–168.
- . 2008. *The Dynamics of Performance Management: Constructing Information and Reform*. Washington, DC: Georgetown University Press.
- . 2012. 'Does Performance Data Lead to Performance Budgeting? Evidence from an Experiment'. In New Orleans.
- Moynihan, Donald P., and Daniel P. Hawes. 2010. 'The Organizational Environment and Performance Information Use'. In *Annual Conference of the European Group of Public Administration*. Toulouse, France.
- Nielsen, Poul Aaes. 2012. 'Performance Information, Aspiration Levels, and Managerial Priorities: How Managers Learn from Performance Feedback'. In Chicago, IL.
- Nutley, S., R. LEVITT, W. SOLESBURY, and S. MARTIN. 2012. 'SCRUTINIZING PERFORMANCE: HOW ASSESSORS REACH JUDGEMENTS ABOUT PUBLIC SERVICES'. *Public Administration*.
- Poister, T. H., and G. Streib. 1999. 'Performance Measurement in Municipal Government: Assessing the State of the Practice'. *Public Administration Review*: 325–335.
- Pollitt, C. 2006. 'Performance Information for Democracy: The Missing Link?' *Evaluation* 12 (1): 38–55.
- Raudla, Ringa. Forthcoming. 'THE USE OF PERFORMANCE INFORMATION IN BUDGETARY DECISION-MAKING BY LEGISLATORS: IS ESTONIA ANY DIFFERENT?' *Public Administration*
- Rubin, M., and K. Willoughby. 2011. 'Measuring Government Performance: The Intersection of Strategic Planning and Performance Budgeting'. In *7th Transatlantic Dialogue Conference, Dialogue on Strategic Management of Public Organizations, ASPA-EGPA, Newark, New Jersey, June, 23–25*.
- Salge, Torsten O. 2011. 'A Behavioral Model of Innovative Search: Evidence from Public Hospital Services'. *Journal of Public Administration Research and Theory* 21 (1): 181–210.

- Schumpeter, J. A. 1991. 'The Crisis of the Tax State'. In *Joseph A. Schumpeter: The Economics and Sociology of Capitalism*, ed. R. A. Swedberg. Princeton, NJ: Princeton University Press.
- Serritzlew, Søren. 2003. 'Shaping Local Councillor Preferences: Party Politics, Committee Structure and Social Background'. *Scandinavian Political Studies* 26 (4): 327–348.
- Slothuus, Rune, and Claes H. de Vreese. 2010. 'Political Parties, Motivated Reasoning, and Issue Framing Effects'. *The Journal of Politics* 72 (3): 630–645.
- Stubager, Rune, and Rune Slothuus. 2012. 'What Are the Sources of Political Parties' Issue Ownership? Testing Four Explanations at the Individual Level'. *Political Behavior*: 1–22.
- Sørensen, Rune J. 1995. 'The Demand for Local Government Goods: The Impact of Parties, Committees, and Public Sector Politicians'. *European Journal of Political Research* 27 (1): 119–141.
- Wildavsky, Aaron A., and Naomi Caiden. 1988. *The New Politics of the Budgetary Process*. Glenview, Ill.: Scott, Foresman & Co.
- Willoughby, Katherine. 2011. 'Introduction to the Symposium: PBB-Works Like the BCS?' *Public Administration Review* 71 (3): 352–355.
- Wilson, James Q. 1989. *Bureaucracy: What Governments Do and Why They Do It*. New York: Basic Books.
- Zaltsman, A. 2009. 'The Effects of Performance Information on Public Resource Allocations: A Study of Chile's Performance-Based Budgeting System'. *International Public Management Journal* 12 (4): 450–483.