



Aalborg Universitet

AALBORG UNIVERSITY  
DENMARK

## Comparing contracting performance

*Culture, competition, contracts, capabilities and collaboration in UK and Scandinavia*

Lindholst, Andrej Christian

*Publication date:*  
2017

[Link to publication from Aalborg University](#)

*Citation for published version (APA):*

Lindholst, A. C. (2017). *Comparing contracting performance: Culture, competition, contracts, capabilities and collaboration in UK and Scandinavia*. Paper presented at 21th Annual International Research Society for Public Management (IRSPM) Conference, Budapest, Hungary.

### General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- ? Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- ? You may not further distribute the material or use it for any profit-making activity or commercial gain
- ? You may freely distribute the URL identifying the publication in the public portal ?

### Take down policy

If you believe that this document breaches copyright please contact us at [vbn@aub.aau.dk](mailto:vbn@aub.aau.dk) providing details, and we will remove access to the work immediately and investigate your claim.

# **Comparing contracting performance: Culture, competition, contracts, capabilities and collaboration in UK and Scandinavia**

*Paper for IRSPM 2017, Panel H1 Management and organizational performance  
Budapest, Hungary*

## **Author**

Andrej Christian Lindholst, Associate Professor  
[acl@dps.aau.dk](mailto:acl@dps.aau.dk)  
+ 45 2872 8520  
Department of Political Science, Aalborg University  
Fibigerstræde 1, DK-9220, Aalborg, Denmark

## **Paper abstract**

One key question for the last three decades of public management reforms is whether country differences between radical and modest reform approaches make a difference for performance in a reformed public sector. This paper compares differences in the implementation and performance of contracting out, i.e. private delivery of public services, between the United Kingdom and Scandinavia. The reform approach in the United Kingdom are argued to have generated a ‘deep’ contracting culture embedding stronger contracting institutions than the more ‘shallow’ contracting cultures in Scandinavia. Hypotheses are suggested for the role of culture, competition, contracts, capabilities and collaboration for contracting performance between and across the countries. Arguments are tested against data from on four comparable national surveys of private delivery of park and road maintenance services in local governments. Contracting performance is found to be higher in the United Kingdom due to a stronger overall combination of institutional features while performance differences across the countries in ranked order are found to be collaboration, contracts, capabilities and competition. Overall, the study support arguments in favor of a leading edge in radical reforming countries for some types of service settings but also prospective paths for catching up by more modest reforming countries.

## **Key words:**

Comparative analysis, contracting out, performance, reform, survey

## **Introduction**

A handful of countries have become international standards for measuring the scale and scope as well as evaluating the merits and perils of new public management (NPM) reforms and what have followed of post-NPM reform initiatives. In international comparisons these countries have become renowned as ‘vanguards’ (Hood & Dixon, 2015), ‘trailblazers’ (Christensen & Lægreid, 2007) or ‘benchmarks’ (Barzelay, 2001) by harboring public management reforms earlier and through more profound and radical approaches than what can be found in other and more modest reform countries. One key question is whether country differences between ‘radical’ and ‘modest’ reform approaches make a difference for performance in a reformed public sector. However, as noted by Greve et al. (2016), there is a genuine lack of systematic evaluations based on quantitative comparative research on reform outcomes going beyond a focus on a single country, reform or service. This paper adds a piece to the puzzle by addressing whether one of the global cornerstones in the last three decades’ public management reforms, the reliance on private delivery of public services through contracting out, differs between radical and modest reform countries in terms of the strength of contracting institutions and level of contracting performance.

From a broad institutionalist perspective on organization (Scott, 2001) it is argued that some contracting ‘cultures’ through radical reform approaches have become deeply institutionalized by embedding strong regulative, normative and cultural-cognitive rule systems while more modest reform countries embeds more shallow institutional supports for guiding contracting behaviors. The institutionalist perspective on contracting is refined in a set of hypotheses which link differences in contracting culture with differences in contracting performance.

The research calls on comparative survey data on the organization and performance of private delivery of park and road maintenance collected in 2014-16 from local governments in the United Kingdom (UK) and the three Scandinavian countries: Denmark, Sweden and Norway. Park and road services are mostly organized within technical departments in local governments and often contracted out in some degree (Lindholm, 2009). From a theoretical perspective park and road services furthermore represent a ‘most-likely’ setting for successful use of contracting out due to compliant transactional characteristics such as relatively ease of service specification and measurement and relatively low requirements for specialized investments (Brown & Potoski, 2005). In international comparisons of public reforms, the United Kingdom is often regarded as a more

radical reform country with profound emphasis on marketization. In contrast, the three Scandinavian countries represent contexts with more incremental and modest reform approaches with less emphasis on marketization and greater emphasis on modernization (Greve et al., 2016). The research's combination of a focus on contracting out of the same types of services across dissimilar national contexts should allow explorations of the importance of institutional differences for contracting performance.

The remainder of the paper is organized in the following four sections. A section on 'Theories of contracting performance' provides arguments on the importance of contracting culture and institutions for contracting performance. A section on 'Methods' described data sources, operationalizations, analytic tools as well as some limitations/reservations. Findings from the analysis is presented and discussed in a section on 'Results and Discussions'. Finally, in the section 'Conclusions' key findings are summarized and broader implications for policy and research are highlighted.

### **Theories of contracting performance**

This section review mainstay arguments in the literature on the performance determinants of contractual governance in the private delivery of public services and provides theoretical arguments on the importance of contracting culture based on an institutional perspective. The section delivers four main hypotheses on links between contracting institutions and contracting performance.

#### *Contracting culture*

Public management reforms within different national contexts have over the years emphasized and implemented contracting out differently. Differences in the institutionalization of key features such as markets for public services, contract management capabilities, contract standards and legal frameworks emerges as well as a more implicit 'contracting culture' emerges when contracting practices becomes everyday routine and 'hands-on' tacit knowledge becomes ever more refined (Greve & Ejersbo, 2005). From the perspective of institutional theory (Scott, 2001) such 'contracting culture' can be viewed as a multilayered institutional framework in terms of stable rule systems (guiding behavior) constituted by regulative, normative and cultural-cognitive systems ranging from the formal to the informal. In the institutional perspective a 'deep' contracting culture can be denoted as a context where the institutional framework for contractual governance not only

are staged as ‘artifacts’ in formal regulative systems, e.g. public policies and laws, but also woven into deeper layers of behavioral norms and core values as well as shared cognitive and interpretive schemes.

### *UK as a deep contracting culture*

Overall, successive policies and policy learning have shaped the use of contracting out within the local government park and road sectors in the UK since the 1980s (Dempsey et al., 2016). In the 1980s and 1990s, the policy framework of compulsory competitive tendering (CCT) promoted the use of contractual governance through a set of cohesive policy instruments and a clear cost-cutting focus. Park and road services formed part of several core services in the early implementation of CCT. In particular, local governments were only allowed to provide services through in-house arrangements after a tightly regulated market test of service delivery. In the late 1990s, the central government replaced the CCT regime with the Best Value (BV) regime – a less strict policy based on the four principles of competition, comparison, challenge and consulting (Boyne, 1999). In addition, a scheme for attracting private finance for investment in public services, the private finance initiative (PFI), was promoted as an alternative contractual arrangement as well as new forms for ‘partnerships with the market’ were invented (Bovaird, 2006). Overall, a shift in focus from competition to collaboration in public service delivery with a variety of service providers were visible in the overall policy framework (Entwistle & Martin, 2005). In the late 2000s, ‘strategic commissioning’ came to denote an even more encompassing policy approach to public service delivery including a focus on delivering ‘public value’ in local contexts as well as on trusts, social enterprises and non-profit organizations as alternatives to provision by local governments and private for-profit contractors (Bovaird et al., 2014).

Given the long history of successive and encompassing policies related to development of contractual governance as well as reliance on contracts and competition in public service delivery in the UK context it is argued that the country context represents a ‘deep contracting culture’ with a strong institutionalization of contracting in regulative as well as normative and cultural-cognitive systems.

### *Shallow contracting cultures in Scandinavia*

In general, Scandinavia, as part of the Nordic group of countries, leans toward modernization rather than marketization in their reform initiatives (Greve & Ejersbo, 2016). The three Scandinavian countries, Sweden, Norway and Denmark differs in important aspects in the degree contracting out has been pursued and regulated within various sectors, but in comparison to the UK they all represent national contexts with shorter histories, less systematic and more modest policy approaches to contracting out including less institutionalization of formal rules and regulations as well as a traditional reliance on the public sector for service provision. In Denmark, for example, the overall policy approach and implementation of contracting out within park and road services have been characterized as one of pragmatism, decentralization and incrementalism rather than ideology, centralization and radical reform (Lindholst et al., 2016). Similar approaches within park and road sector are found in Norway (Leiren et al., 2016) and Sweden (Bretzer et al., 2016). Greve & Ejersbo (2005) found that the contracting culture in Scandinavia has been one of ‘development’ compared to a ‘mature’ contracting culture in the US (another vanguard country in relation to marketization and privatization reforms). Learning processes are ongoing in Scandinavia where public purchasers, private providers and costumers learn to become more like ‘contractual men’. Given the policy approach the contracting culture in Scandinavia is argued to be more ‘shallow’ with less institutionalization of regulative, normative and cultural-cognitive systems. Given the relatively shallow contracting culture in Scandinavia it is argued that contracting performance is relatively inferior in the three Scandinavian countries compared to the UK.

*Hypothesis (H1): Contracting performance will be higher (lower) in countries with relatively deeper (shallower) contracting cultures.*

#### *Four contracting institutions*

This section decomposes further the contracting framework into four key institutional components which have been highlighted in the contracting literature and linked to contracting performance.

#### *Competition*

In the contracting literature, well-functioning and competitive markets have long been argued to be a key determinant for high contracting performance of both private and public delivery of public services (Boyne, 1998; Dehoog, 1990; Donahue, 1989; Kettl, 1993; Savas, 2000). The basic argument in the literature on the role of competition for contracting performance has been delivered

by neoclassic economics and public choice theory. In the classic public choice argument, competitive contexts, where governments can choose between alternative service providers based on prices offered through competitive markets, are believed to supply the incentives for providers to ensure cost-efficient service provision and reduce service expenditures. In an extended argument based on transaction cost theory, a shift from bureaucratic supply to markets may incur additional transaction costs for managing and safeguarding contracts with external service providers due to the risk of opportunistic behaviors. However, a situation with limited competition, e.g. dependence on a single supplier or a small number of suppliers, is still a source of more contracting hazards and higher transactions cost than a situation with higher levels of competition.

### *Contracts*

Formal contracts are commonly perceived as mechanisms for governing exchanges between two or more contracting parties. Contracts specify rights and obligations in an exchange relation for contracting parties. Successful contract design contributes to maximizing the received benefits from an exchange relation for the contracting parties. Contracting theory argues that formal contracts exhibit dual functions for contracting parties as a mechanism for mitigating contractual risks as well as a mechanism for enabling planning and coordination (Schepker et al., 2014). The importance of formal contracts for contracting performance of private delivery of public services has been highlighted by a number of studies over the years (Amirkhanyan et al., 2010; Donahue, 1989; Kim & Brown, 2012; Romzek & Johnston, 2005). The degree of contract completeness, i.e. how far contingencies in an exchange relation can be specified and written into a formal document in advance, has been associated with higher contracting performance.

### *Capabilities*

Several researchers have emphasized the pivotal role of governments' management capabilities or capacity for providing services by contracts for contracting performance (Amirkhanyan et al., 2007; Brown & Potoski, 2003; Choi & Heinrich, 2006; Romzek & Johnston, 2002). Contracting capabilities are seen as necessary for overcoming various problems stemming from, for example, lack of competition or incomplete contracting as well as developing and managing contracts successfully. In the literature contracting capabilities has broadly become a reference to the capability or capacities to manage different phases or components of the contracting process. Kettl wrote of the need for governments to learn the role as 'smart buyer' by building the capacity to

know ‘what to buy’, ‘who to buy from’ and finally ‘what is bought’. Later Brown & Potoski (2003) divided contracting capacities into the components assessing feasibility, implementing and evaluating service delivery by contract while the ability to align contracts with particular contracting objectives and the market situation can be added from the contracting framework proposed by Brown et al. (2006). Yang et al. (2009) have synthesized further earlier work of particular Brown & Potoski (2003) and Romzek & Johnston (2002) and propose a conceptualization based on the four components: ‘agenda setting’, ‘formulation’, ‘implementation’ and ‘evaluation’.

### *Collaboration*

It is a longstanding assumption within sociological theories of exchange relations that the collaborative orientation and the normative character of relations between exchange partners – or what is sometimes is discussed as ‘relational’ contracting or governance – have immense importance for the performance of exchange relations (Macaulay, 1963; Macneil, 1974; Schepker et al., 2014). The idea has gained empirical support in several studies within a private sector context (Cannon et al., 2000; Poppo & Zenger, 2002) as well as a public sector context (Amirkhanyan et al., 2010; Fernandez, 2009). In the argument institutionalization of shared collaborative behavioral norms allow contracting parties effectively to adapt an ongoing exchange relation (contract) to new needs or demands, deal with unforeseen contingencies arising from uncertainties or minimizing efforts to safeguard the contract from behavioral risks arising from exchange specific investments or incomplete contracting. Overall, stronger institutionalization of norms for collaborative behaviors is argued to improve contracting performance.

*Hypothesis two (H2): Higher levels of respectively competition (H2a), use of formal contracts in exchange relations (H2b), contract management capability (H2c) and collaboration (H2d) improve contracting performance.*

### *Contracting culture and institutions*

The paper’s key proposition is that the radical reforming countries’ development of deep contracting cultures results in relatively higher contracting performance than more modest reforming countries. In the more detailed account it can be argued that the development of a deep contracting culture to a large degree involves development of stronger contracting institutions related to higher levels of competition, contracts, capabilities and collaborative behaviors. In

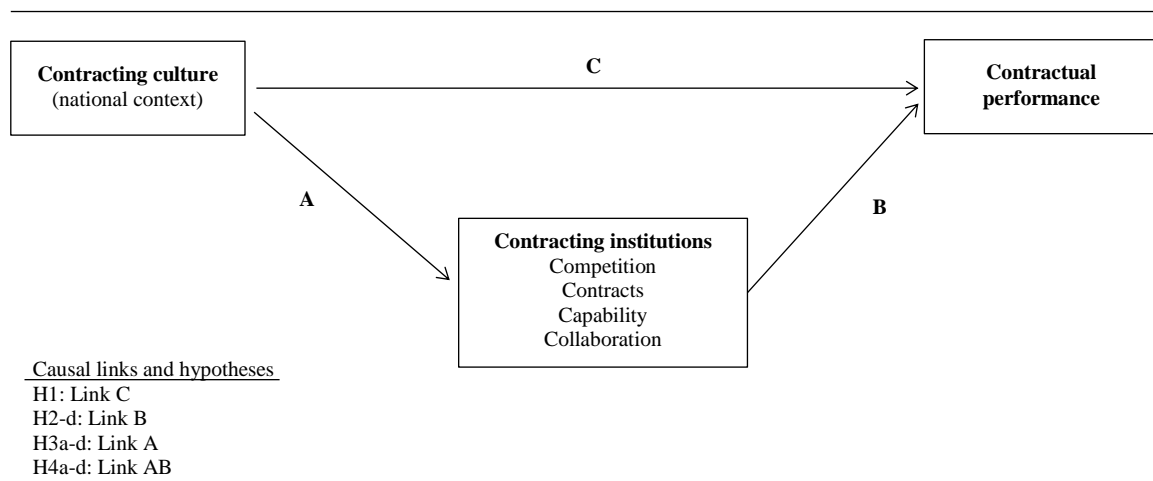


consequence, it is argued that the hypothesized association between contracting culture and contracting performance is partly mediated by the institutional strength of competition, formal contracts, contract management capabilities and collaborative behaviors.

*Hypothesis (H3): The contracting institutions, in terms of competition (H3a), formal contracts (H3b), contract management capabilities (H3c) and collaborative behaviors (H3d) will be more (less) developed in deeper (shallower) contracting cultures.*

*Hypothesis three (H4): A relatively higher (hypothesized) contracting performance in deep contracting cultures are (partly) mediated by higher levels of competition (H4a), use of formal contracts in exchange relations (H4b), contract management capability (H4c) and collaboration (H4d).*

Figure 1 summarizes the expected (hypothesized) main associations between contracting performance and contracting culture as well as the four contracting institutions related to competition, contracts, capabilities and collaboration.



**Figure 1.** Conceptual model: Contractual performance and three causal links

## Methods

### Data

The study relies on data from a web-based survey administered to local governments in the UK, Sweden, Denmark and Norway in the period 2014-2016. The survey took place as part of a larger research project and had a broader purpose of assessing why and how local governments use private and public service delivery for park and road maintenance services as well as assessing outcomes from the use of public and private service delivery. The survey data deliver unique comparable insights across four countries for phenomena where no other data sources exists.

The survey was firstly developed and carried out within a Danish context and subsequently translated and slightly adjusted for use in the three other countries. Pre-tests and adjustment were carried out in all country contexts. The setup of the survey and data collection in each country followed the same basic procedure. Small adjustments were made such as the number of reminders and the period for data collection as well as the number and formulation of survey items (see appendix for translated formulations). While translations of items where sought kept as comparable in their meaning the items also had to be formulated in order to be as clear and understandable by respondents within each country.

Targeted primary respondents were midlevel managers in local government organizations with responsibilities for roads and/or park services with expected insights in operational dimensions as well as strategic dimensions of park and/or road services. List of respondents and contact details was collected by local research partners in each country through a combination of contact with professional associations, use of phone books for professionals, as well as inspection of websites and direct phone contact. Due to variations in internal organization it was in some cases necessary to identify more than one respondent for some municipalities. Initial invitations were followed up by multiple reminders for partly and non-responding local governments. In case that data from more than one respondent were received for a local government the characteristics of respondents were inspected. Selection of data for the local government in these instances was based on assessment of years of employment, job title, and responsibilities and organizational position of respondents.

The final and full dataset includes data for 122 out of 391 local authorities in the UK (equal to 31 %), 115 out of 290 municipalities in Sweden (40 %), 75 out of 98 municipalities in Denmark (77 %) and 95 out of 490 municipalities in Norway (22 %). Drop-out analyses show that smaller local governments are underrepresented in all countries except Denmark. Non-response in Norway and

Sweden are likely to be due to high number of small municipalities (e.g. less than 10,000 inhabitants) with relatively few or no park and road responsibilities.

Not all local governments use private contractors for delivery of park and road maintenance services, thus the dataset for analysis builds on a subset of the full dataset. A total of 308 local governments indicated that they used private contractors to some degree for parks and/or road services including 70 UK, 77 Norwegian, 75 Danish and 86 Swedish local governments. One-hundred and fifty-five out of the 308 local governments contained data for private delivery of both park and road maintenance while 153 entries contained data for either park or road maintenance. The 155 local governments with unique data for private delivery of both park and road maintenance could therefore be split into a case for parks and a case for roads. The initial data set for analysis of private delivery contained a total of 463 cases. Exclusion of cases with missing values for survey items used for constructs in the main analysis reduced the number of cases to a total of 304 (before outlier diagnosis) of which 62 cases were from the UK 61 from Norway, 94 from Denmark and 87 from Sweden.

### *Operationalizations*

The dependent or outcome variable for contracting performance is operationalized through a four item composite construct. The variable is based on a multidimensional conceptualization of performance measured through respondents' (public managers) perceived satisfaction with key performance dimensions of private delivery of respectively park and road maintenance. The operationalization reflects similar approaches used in earlier studies of contracting performance in the public (e.g. Fernandez, 2009) and private sector (e.g. Cannon et al, 2000; Poppo and Zinger, 2002) where the number of items ranges from three to eight. The operationalization in the analysis includes four dimensions for evaluating contractual performance which are measured by an 11-point numeric scale with the two end anchors 'very unsatisfied' (0) and 'very satisfied' (10). The performance dimensions include items for: 1) service quality, 2) cost/pricing, 3) flexibility for change/improvements and 4) responsiveness to issues and deficiencies (see appendix for formulation for survey items). Reliability of the construct is high measured by Cronbach Alpha ( $\alpha = .898$ ) and by principal component analysis (Eigenvalue = 3.07 with 77 % of total inter-item variance explained). For scaling the composite construct for contracting performance the aggregated score for the items were divided by the number of items.

The analysis includes four constructs for respectively, competition, contracts, capabilities and collaboration as key explanatory variables. In addition four country variables are included together with three control variables.

The variable for competition is based by a single survey item. The item measures a respondent's perception of the degree in which the local government receives sufficient qualified bids (competition) when (park and/or road) services are tendered to the market. The measurement scale used an 11-point numeric scale with the two end anchors 'not at all' (0) and 'in very high degree' (10). The scale is treated as continuous in the analysis.

A composite construct for measurement of the level of formalized contracts within a local government's exchange relations with their private contractors were based on six survey items (see appendix for formulations). Earlier research has emphasized congruent contractual characteristics but operationalizations vary greatly in both content and the number of items across studies. (Cannon et al., 2000), for example, utilize three items to measure 'legal bonds', (Amirkhanyan et al., 2010) utilize weight items to measure 'completeness' while (Fernandez, 2009) utilizes one item for a measure of 'specificity'. With similar purposes to gauge the level of completeness/specificity/legal bounds in exchange relations, but also ensure a close match with the service context of the study, the survey items were primarily inspired from findings in a study of the contractual infrastructure used in maintenance contracts in public parks and green spaces (Lindholst, 2009). The items were measured by an 11-point numeric scale with the two end anchors 'not at all' (0) and 'in very high degree' (10). For scaling the composite construct the aggregated score for the items were divided by the number of items. Inter-item reliability of the construct is high measured by Cronbach Alpha ( $\alpha = .825$ ) and acceptable measured by principal component analysis (Eigenvalue = 3.24 with 54 % of total inter-item variance explained). A more complex operationalization of contracts based on two constructs derived from principal component analysis of altogether eight survey items was also possible. The one construct option was favored for keeping the analysis relatively simple and parsimonious with regard to the importance of contracts. The two construct option would mainly explore further which particular contract features relate to (higher) contracting performance.

The measure for contract management capacity is a composite construct based on four survey items. What is here understood as contract management capabilities has been labelled, conceptualized and operationalized fairly different in various survey based studies (Amirkhanyan et al., 2010; Fernandez, 2009). The concept may encompass very different and specific managerial activities such as drafting documents, evaluating tenders or monitoring contractors. The formulations in the survey have a more general orientation by referring to organizational concepts such as ‘systems’, ‘methods’, ‘procedures’, ‘expertise’ or ‘time’ for managing contracts. The items were measured by an 11-point numeric scale with the two end anchors ‘not at all’ (0) and ‘in very high degree’ (10). For scaling the composite construct the aggregated score for the items were divided by the number of items. Inter-item reliability of the construct is high measured by Cronbach Alpha ( $\alpha = .858$ ) and by principal component analysis (Eigenvalue = 2.85 with 71 % of total inter-item variance explained).

Collaboration within contractual relations has been conceptualized and operationalized slightly differently in earlier research and furthermore partly been associated with trust between contracting parties. Cannon et al (2000) used six items to measure ‘cooperative norms’, Poppo & Zinger (2002) used three items to measure ‘relational governance’ while Amirkhanyan et al. (2010) ended up with 20 items measuring ‘current relationship strength’. The survey included altogether six items primarily adapted from Cannon et al (2000). However, the final measure for collaboration is a composite construct based on three survey items. One item performed relatively poorly in a composite construct while two items related to trust were omitted for theoretical reasons. A measure of trust can be argued to be problematic in terms of reverse causality in relation to a measure of performance, i.e. trust may be a consequence of past performance as well as trust may enhance performance as suggested by for example Lamothe & Lamothe (2012). The items in the construct for collaboration were measured by an 11-point numeric scale with the two end anchors ‘not at all’ (0) and ‘in very high degree’ (10). For scaling the composite construct the aggregated score for the items were divided by the number of items. Inter-item reliability of the construct is high measured by Cronbach Alpha ( $\alpha = .882$ ) and by principal component analysis (Eigenvalue = 2.45 with 82 % of total inter-item variance explained).

*Controls:*

A dichotomous variable for sector is included as a simple control for differences in respondents' performance evaluation of private delivery of respectively park and road maintenance. The survey included separate items for measuring performance within the two sectors. Statistical tests showed that performance differences were negligible or non-existing for individual items as well as composite constructs.

Local government size was included as a second control. Earlier studies of contracting out have included measures for local government size for various purposes. In studies of differences in contracting levels it has been argued that relatively larger local governments should be expected to contract out a greater share of their services as they are likely to entertain larger markets and attract more competition – a precondition for well-working use of contracting out. Larger local governments also represent larger organizations which can devote more resources and entertain greater managerial and administrative specialization. Larger municipalities should therefore retain greater capability for managing various tasks including contracts. Theoretically speaking, local government size is expected to be correlated with higher levels of contract management capabilities and competition. The inclusion of a variable in the analysis for local government size should therefore moderate the size of the direct effect on performance from capabilities and competition. The control is also important as the size of local governments differs greatly among the four countries. The size of local government is measured by the number of inhabitants in 2014. Data for this variable were drawn from national statistical services in each of the four countries. The distribution of the data for the four countries combined is highly right-skewed (i.e. a small group of local governments account for disproportionate large numbers of inhabitants) and the data is transformed to a natural logarithmic scale for use in the analysis.

Finally, a control was included for potential influence of systematic differences or bias in performance evaluation by respondents with different organizational interests. The control is based on a survey item addressing whether a respondent's department had responsibilities for internal service provision (yes or no). The variable aims to control for any differences in performance evaluations from respondents with likely different interests or reference for evaluation arising from a status as responsible for internal service provision. Respondents employed in departments with direct (operational) responsibilities for internal provision of services may systematically evaluate

the performance of their direct competitors (i.e. private providers) differently than respondents in departments where no such competition exists.

### *Analysis*

The analysis is based on computation of several statistics. SPSS was used as main software package for computation of statistics. The statistics guide the evaluation of the main hypothesis (H1–H4) presented in the theory section. In an initial sequential outlier diagnosis based on a regression analysis with complete data for all included variables ( $n = 304$ ) identified altogether six ‘influential outliers’. The data for the six outliers were inspected and no (visible) erroneous data were found. The outliers were removed and the 298 cases are used as basis for all analysis and statistics in the result section. The six outliers are interpreted as a distinct group of cases with a ‘deviant’ pattern contrasted to the ‘normal’ pattern emerging in the analysis of the remaining 298 cases. Key differences between the two groups are presented and discussed as part of the result section.

In the result section the data is firstly explored by descriptive statistics and comparisons of country differences for all key variables. The UK is used as a reference country in comparisons with the three Scandinavian countries individually and combined. Statistical significance of differences is tested by one-way anova analysis with post hoc tests.

Secondly, the data is explored in a main analysis based on two ordinary least square (OLS) multiple regression analyses. In a first model (the ‘main analysis’), the UK is used as a reference country for comparison of country differences, i.e. the analysis includes dummy variables for each of the three Scandinavian countries. In a second model a single dummy variable for the UK replaced the three dummy variables for Scandinavia. The two models show corresponding results.

Thirdly, because a degree of correlation exists between the independent variables, additional statistics are reported to assist the interpretation of the estimates and findings in the OLS regressions (Nathans et al., 2012). Squared structure coefficients ( $r_s^2$ ) and relative weights (RW and RW %) were calculated for these purposes. In light of the multi-correlational character of most real life data, structure coefficients estimate effects from a predictor variable on the predicted values of the outcome variable, help assessing the degree of overall shared variance among predictor variables and identifying suppressor variables when compared to beta weights. According to

Tonidandel & LeBreton (2011), relative weights perform far better than regression weights when partitioning variance in the presence of correlated independent variables. Relative weights help identifying the contribution of predictor variables to the explained variance ( $R^2$ ) in a regression model by minimizing impacts from multicollinearity. The computation of additional statistics is based on a SPSS syntax provided by Lorenzo-Seva et al. (2010).

Fourthly, a mediation analysis is carried out in order to assist in the evaluation of hypothesis H1 and H3. The mediation analysis evaluate whether the hypothesized effect from the UK context on contracting performance are indirectly effected (mediated) by differences in competition, contracts, capabilities or collaboration or the effect is a genuine ‘direct’ effect. The analysis is based on a SPSS software package and guidelines for mediation analysis provided by Hayes (2013).

### *Reservations*

The study is based on a research design combining cross-sectorial survey data from four countries with correlational analysis (except data for one control variable) for explanatory purposes. Claims toward causality based on the present (nonexperimental) research design must be regarded as ‘weak’ in comparison with ‘stronger’ (quasi-)experimental research designs. For example, causal claims might be (partly) spurious due to omitted ‘third’ variables or causality may be (partly) reversed.

Research based on survey data may suffer from several methodological biases (Podsakoff et al., 2012). Potential biases were sought minimized through ex ante remedies as well as post hoc tests were carried out to gauge eventual influence of any substantial amount of common method variance. Ex ante remedies in the design phase included adoption of different response scales for items used for constructing the predictor and response variables, the design and exact wording of survey items included input from separate test pilots in all countries, respondents were selected due to their supposed insights in the survey’s main topic, and the survey was carried out with anonymity and confidentiality at both the level of respondents and municipalities.

Subjective measures may be biased by the identity and interest of the respondent. A small control is included in the analysis in order to gauge whether organizational interests, in terms of the respondents’ belonging to a department with direct operational responsibilities, systematically



influence the evaluation of contracting performance. Respondents belonging to a department with direct operational responsibilities are likely to be in a situation with direct or indirect competition with private delivery of similar services.

The influence from likely confounding ‘third’ variables was gauged by inclusion of additional variables in alternative OLS regression models (not shown). The inclusion did, however, not alter main findings in the analysis. However, some variables had explanatory power. In particular, the level of contracting out (self-reported percentage of budgets spend on private contractors) were found to be associated with higher performance in simple models with few variables but not when controlled for levels of competition and formal contracts.

Harman’s single factor test and a confirmatory factor analysis (CFA) were carried out as post hoc tests to diagnose for any severe influence from common method variance. In the Harman’s single factor test all survey items from the five main variables were loaded into an exploratory factor analysis with direct oblimin as rotation method. The analysis resulted in four factors with eigenvalues larger than 1 explaining 66 per cent of the total variance whereof the first factor explained 33 per cent (of total variance). In the rotated solution all items had high primary loadings into their respective constructs and low loadings on other constructs (indicative of good construct validity). According to criteria for acceptable model fit given by Hu and Bentler, (1999) the results from the CFA (SPSS AMOS) showed that the items provided a poor fit when loaded into a single factor model:  $\chi^2 = 1377.672$  ( $df = 135$ ,  $p < 0.000$ ,  $\chi^2 / df = 10.205$ ); SRMR=0.1325; RMSEA=0.176; PCLOSE=0.000; CFI=0.560; TLI=0.501. The post hoc statistics indicate that a substantial amount of common method variance is unlikely to be present in the data and confound interpretation of findings.

## **Results**

Table 1 shows simple descriptives for all variables at country level and at the aggregate level as well as results from significance test of country differences based on UK as the reference country.

In comparison, the differences in mean scores for contracting performance turns out as initially expected. The mean score is highest for the UK (7.7) and lowest for Sweden (6.9). While the mean scores for the three Scandinavian countries are close to similar, the performance differences

between UK and each of three Scandinavian countries are found to be statistically significant for Sweden and Norway while insignificant for Denmark. Furthermore a significance test of the difference for contracting performance for the UK and all three Scandinavian countries combined (n=238; mean score = 7.0) finds that the difference is significant ( $p < .001$ ). Overall, the direct comparisons build some support for the first hypothesis on relatively superior contracting performance in countries with deep contracting cultures versus countries with shallow contracting cultures. However, the direct comparison do not say much about whether the differences can be attributed to different contracting cultures or whether they can be accounted for by differences in micro-level institutions.

Findings from a direct comparison of country differences for each of the four micro-level determinants are mixed. The UK has the highest mean scores for contracts and capabilities and partly for collaboration. However, Denmark has the highest mean score for competition (7.2) while the UK has the second lowest mean score (5.6) very close to Norway with the lowest score (5.5). The direct comparisons indicate partial support for the assumption that a deeper contracting culture embeds stronger micro-level institutions. Apparently, all countries have almost similar levels of collaboration while ensuring competition might have become an issue in the UK. However, the estimated standard deviation for competition in the UK indicates that local governments experience very different competition levels, i.e. some local governments attracts high levels of competition while other experience highly insufficient levels. If the UK is compared to all Scandinavian countries as one group (n=238) the mean difference for competition (-0.9), contracts (0.6) and capabilities (1.1) are all significant ( $p < 0.05$ ) while the mean difference for collaboration (0.3) remains insignificant ( $p = .192$ ). The direct comparisons indicate that the four micro-level determinants (if they work as expected) have diverging effects on the performance differences between the UK and the three Scandinavian countries. For example, holding the competition level constant across the countries would result in higher performance differences.

**Table 1. Descriptive statistics and mean differences**

	N	Mean <sup>a</sup>	S.D
<b>Contracting performance (index, four items)</b>			
Measure: 11-point scale (0 = very unsatisfied, 10 = very satisfied).			
<i>UK</i>	60	7.7	1.9
<i>Norway</i>	60	7.0 *	1.4
<i>Denmark</i>	92	7.2	1.3
<i>Sweden</i>	86	6.9 **	1.5
<i>ALL</i>	298	7.2	1.4
<b>Competition level (single item)</b>			
Measure: 11-point scale (0 = not at all, 10 = In very high degree)			
<i>UK</i>	60	5.6	3.2
<i>Norway</i>	60	5.5	2.9
<i>Denmark</i>	92	7.2 ***	2.1
<i>Sweden</i>	86	6.6	2.6
<i>ALL</i>	298	6.4	2.7
<b>Contract level (index, six items)</b>			
Measure: 11-point scale (0 = not at all, 10 = In very high degree)			
<i>UK</i>	60	7.5	1.7
<i>Norway</i>	60	6.8	2.0
<i>Denmark</i>	92	7.1	1.9
<i>Sweden</i>	86	6.5 **	2.4
<i>ALL</i>	298	7.0	2.0
<b>Contract management capability (index, four items)</b>			
Measure: 11-point scale (0 = not at all, 10 = In very high degree)			
<i>UK</i>	60	7.4	1.4
<i>Norway</i>	60	5.6 ***	2.1
<i>Denmark</i>	92	6.9	1.6
<i>Sweden</i>	86	6.1 ***	2.1
<i>ALL</i>	298	6.5	1.9
<b>Collaborative relation (index, three items)</b>			
Measure: 11-point scale (0 = not at all, 10 = In very high degree)			
<i>UK</i>	60	7.9	1.8
<i>Norway</i>	60	7.9	1.5
<i>Denmark</i>	92	7.6	1.3
<i>Sweden</i>	86	7.3	1.8
<i>ALL<sup>b</sup></i>	298	7.7	1.6
<b>Internal provider</b>			
Measure: (Dummy) 1 = yes, 0 = no			
<i>UK</i>	60	0.7	0.5
<i>Norway</i>	60	0.8	0.4
<i>Denmark</i>	92	0.6	0.5
<i>Sweden</i>	86	0.8	0.3
<i>ALL<sup>c</sup></i>	298	0.7	0.5
<b>Sector</b>			
Measure: 1 = parks, 0 = roads			
<i>UK</i>	60	0.7	0.5
<i>Norway</i>	60	0.3 ***	0.5
<i>Denmark</i>	92	0.4 ***	0.5
<i>Sweden</i>	86	0.3 ***	0.4
<i>ALL</i>	298	0.4	0.5
<b>Municipal size</b>			
Measure: logarithmic scale (natural)			
<i>UK</i>	60	12.0	0.6
<i>Norway</i>	60	9.3 ***	1.4
<i>Denmark</i>	92	10.9 ***	0.7
<i>Sweden</i>	86	10.2 ***	1.0
<i>ALL</i>	298	10.6 ***	1.4

The table shows descriptive statistics and results from analysis of mean differences for variables included in the main analysis (N = 298).

<sup>a</sup> Significance levels reported for mean difference between the UK and the three Scandinavian countries. Differences tested by ONE-WAY ANOVA with post hoc tests (Turkey / Games-Howell). Significance levels: \* p < 0.1, \*\* p < 0.05, \*\*\* p < 0.01.

<sup>b</sup> While no significant mean differences are found between UK and each of the three Scandinavian countries there can be identified a significant difference between Norway and Sweden (p-level < .1).

<sup>c</sup> While no significant mean differences are found between UK and each of the three Scandinavian countries there can be identified a significant difference between Norway and Denmark (p-level < .05).

## Main analysis

Table 2 shows overall results from the analysis from two OLS multivariate regression analysis. The first model in regression analysis includes a comparison of performance differences between each of the three Scandinavian countries and the UK. The second model compares the performance

difference between the UK and the three Scandinavian countries combined. Overall no substantial difference in results or interpretations emerges from the two models. Table 3 reports auxiliary statistics on the contributions of each predictor variable to regression effects and variance. Table 4 shows results from the mediation analysis of contracting performance in UK. Table 5 provides a brief comparison of the characteristics of ‘deviant’ cases and ‘normal’ (groups based on outlier analysis).

Overall the results in the two OLS-regressions reported in Table 2 build further support for the hypothesis that contracting performance is higher in the deep contracting culture of the UK compared to the shallow contracting cultures in the three Scandinavian countries. All regression coefficients (*b*) for the three Scandinavian countries are negative in model 1, the coefficient for UK is positive in model 2 while all associations are significant at *p*-levels < 0.05. The findings provide support for hypothesis H1.

**Table 2. OLS-regressions: Determinants of contracting performance for road and park services in four countries**

Independent variables (scales)	Model 1			Model 2		
	<i>B</i>	SE	P	<i>B</i>	SE	P
UK, dummy (yes=1, no=0)				0.525	0.208	0.012
Norway, dummy (yes=1, no=0)	-0.671	0.292	0.022			
Denmark, dummy (yes=1, no=0)	-0.534	0.217	0.014			
Sweden, dummy (yes=1, no=0)	-0.506	0.241	0.037			
Competition level (scale: 0–10)	0.066	0.026	0.013	0.068	0.026	0.009
Formal contract level (scale: 0–10)	0.139	0.036	0.000	0.136	0.035	0.000
Contract mgmt. capability (scale: 0–10)	0.152	0.040	0.000	0.157	0.039	0.000
Collaborative relation (scale: 0–10)	0.424	0.046	0.000	0.418	0.045	0.000
<i>Controls</i>						
Sector (parks=1, roads=0)	-0.100	0.138	0.467	-0.092	0.137	0.502
Internal provider, dummy (yes=1, no=0)	-0.009	0.149	0.950	-0.019	0.148	0.899
Municipal size (LN)	-0.073	0.070	0.303	-0.055	0.064	0.389
<i>Model summary</i>						
Constant	2.790	0.921	0.003	2.076	0.725	0.005
Max VIF		3.165 <sup>a</sup>			1.670	
R <sup>2</sup> / Adj. R <sup>2</sup>		0.471 / 0.452			0.469 / 0.455	

N = 298. Six influential outliers removed in two steps with standardized residuals > -3.00 in final analysis. Analysis in model 1 without removal of influential outliers (n = 304) shows a significant negative beta-coefficient (-0.142) for municipal size (*p* < 0.1), slightly higher *p*-values for all Scandinavian countries and a slightly lower adj. R<sup>2</sup> = 0.402.

<sup>a</sup>The maximal value for variance inflation factor (VIF) are related to the relatively smaller number of inhabitants in local governments in Norway (and partly Sweden) compared to the United Kingdom.

The hypothesized association (H2a-d) between higher levels of respectively competition, contracts, capability and collaboration and higher level of contracting performance are also supported. All regression coefficients for the four variables for contracting institutions are positive and three of the

four associations are significant at p-levels  $< 0.000$  while the last association is significant at p-level  $= 0.013$  (in model 1). The respective results for four variables provide support for hypothesis H2a-d.

Analysis of the statistics in Table 3 provides further insights into the findings from the analysis of regression coefficients ( $B$ ) and significance of associations in model 1 in Table 2. Overall it was found that the predictor variables together account for a good amount of the variance in contracting performance ( $R^2 = 0.471$ ). However, the statistics for relative weights indicate that while country differences all are significant they only account for very little of the explained variance in contracting performance in the multivariate analysis. In total the three Scandinavian countries account for less than 5 % of the explained variance (RW %). In a separate calculation of relative weights (full stats not shown) based on the variables in model 2 (see table 2) it is congruently found that the UK compared to the three Scandinavian countries combined accounts for 4.3 per cent of the explained variance.

**Table 3.** Statistics determining contributions to regression effects and explained variance.

Variable	$\beta$	$r_s$	$r_s^2$	$r$	RW	RW %
Norway	-0.175	-0.091	0.008	-0.062	0.008	1.6
Denmark	-0.161	0.007	0.000	0.005	0.005	1.1
Sweden	-0.149	-0.169	0.029	-0.116	0.010	2.1
Competition	0.117	0.345	0.119	0.237	0.028	5.9
Contract	0.185	0.578	0.334	0.397	0.077	16.4
Capability	0.192	0.645	0.416	0.443	0.094	19.9
Collaboration	0.447	0.867	0.752	0.593	0.244	51.8
Sector	-0.032	0.010	0.000	0.007	0.000	0.1
Internal provider	-0.003	-0.067	0.004	0.046	0.001	0.2
Municipal size	-0.063	0.202	0.041	0.139	0.005	1.0

Statistics:  $\beta$  = standardized regression coefficient,  $r_s$  = structure coefficient,  $r_s^2$  = Squared structure coefficient,  $r$  = zero-order correlations (Pearson's), RW (%) = relative weight (% of explained variance).  $R^2 = 0.471$  (equal to sum of RWs).

The variables for competition, contracts, capability and collaboration account for substantially larger amounts of the explained variance in contracting performance. Overall, the analysis of the contributions of each predictor to the explained variance in contracting performance shows that the variance in some predictors is relatively more important than the variance in other variables. Evaluated by statistics for respectively  $\beta$ ,  $r_s^2$  and RW%, the most important predictors for explaining the observed variance in contracting performance in ranked order are collaboration, capability, contracts and competition. This finding can be further detailed by inspection of the variance for each predictor variable given by estimates for standard deviations in Table 1. For

example, the relatively high variation in competition (standard deviation = 2.7) matters less for explaining performance than the relatively lower variation in collaboration (standard deviation = 1.6). The variance in contracting performance is less sensitive to the relatively more extreme differences in the level of competition than the relatively smaller differences in the level of collaboration. In other words, good collaboration looks to be more important for contracting performance than good competition.

Comparison of the statistics for squared structural coefficients ( $r_s^2$ ) with regression weights ( $\beta$ ) for the three countries variables shows a degree of suppressor effect for all three variables. The suppressor effect is strongest for Denmark ( $r_s^2 = 0.000$  and  $\beta = -0.161$ ). The statistics tell that if one or more of the three dummy variables is dropped from the analysis the regression effects will be reduced for other variables in the analysis (but not which variables). However, in this case it is clear that if, for example, Denmark is dropped the dummy variables for Norway and Sweden will be comparing contracting performance in the two countries with a groups of cases from both Denmark and the UK (and similar if one of the other country variables are dropped).

The estimated effects for respectively competition, contracts, capabilities and collaboration are all significant (in the relevant models). The general patterns are that higher levels of competition, formalized contracts in exchange relations, contract management capabilities, and shared norms for collaborative behaviors are associated with higher contracting performance.

#### *Mediation analysis*

Mediation effects with UK as a predictor of contracting performance were calculated in one analysis with four mediators included in parallel together with three controls. Confidence intervals (CI) for evaluating significance of mediation effects were calculated at a 90 per cent level based on a bootstrap sample of 5,000. The main results from the mediation analysis are presented in Table 4.

Overall, the analysis of potential mediation effects finds that the total mediation effect from the four variables combined does not change the finding that the UK context contributes in a unique way to explaining variance in contracting performance in the UK and Scandinavia. The decrease in the coefficient for UK from .663 to .525 indicates a partial mediation effects from all four mediating

variables combined. The total indirect effect (.138) of all four mediating variables, however, is found insignificant by calculation of a 90 per cent CI (-.166 / .434).

**Table 4.** Mediation effect analysis with UK as predictor of contracting performance (CP)

Mediators (M)	a path (UK → M)	b path (M → CP)	Indirect effect (UK → M → CP)	CI90% lower / upper (indirect effect)	Effect size (ratio of total effect)	CI90% lower / upper (effect size)
Contracts	.165 (p= .661)	.136 (p= .001)	.022	-.044 / .108	.034	-.093 / .174
Competition	-2.047 (p< .000)	.068 (p= .024)	-.139	-.277 / -.050	-.210	-.848 / -.059
Capabilities	.432 (p= .112)	.157 (p= .001)	.068	.004 / .167	.102	.008 / .299
Collaboration	.449 (p= .123)	.418 (p< .000)	.188	.004 / .401	.238	.020 / .683
Total	c path .663 (p= .013)		.138	-.166 / .434	.209	-.414 / .603

N = 298. All reported regression coefficients are unstandardized. Direct effect of UK on contracting performance (c' path) = .525 (p. = .016). Confidence intervals (CI) calculated with bootstrapping (5,000 samples). A sensitivity analysis based on a 95% CI finds that only competition remains a significant mediator.

Going beyond the total indirect effect, a significant degree of partial mediation is found for three out of the four mediating variables. For capabilities, the analysis finds a significant indirect effect of UK context on contracting performance through capabilities, indirect effect = 0.068, CI90% (0.004 / 0.167). Capabilities mediate about one tenth of the total effect of UK on contracting performance (effect size = 0.102). For collaborative behaviors, the analysis finds a significant indirect effect of UK context on contracting performance through collaborative behaviors, indirect effect = 0.188, CI90% (0.004 / 0.401). Collaboration mediates about one quarter of the total effect of UK on contracting performance (Effect size = 0.238). For competition, the analysis finds a significant indirect effect of UK context on contracting performance through competition, indirect effect = -0.139, CI90% (-0.277 / -0.050). Competition accounts for one-fifth of the total effect of UK on contracting performance (effect size = -0.210). The mediation effect is negative (inconsistent mediation) and partly cancels out effects from other mediating variables. Finally, the analysis shows that contracts do not function as a mediating variable indirect effect = -0.022 CI90% (-0.044 / 0.108).

The findings equal to say that the relatively higher performance in the UK compared to Scandinavia partly is explained by higher levels of capabilities (support for H4c) and collaborative behaviors (support for H4d). However, the relatively higher performance in UK is also limited by a relatively lower degree of competition (no support for H4a) in the UK compared to Scandinavia while the level of contract (no support for H4b) don't make a difference for contracting performance in the comparison.

The relatively higher contracting performance in the UK compared to Scandinavia is partly accounted for by more robust institutionalization of collaborative behaviors (one-quarter of total effect) and contract management capabilities (one-tenth of the total effect). However, contradictory to initial expectations competition is relatively lower in the UK and the mediation effect of competition (one-fifth of the total effect) cancels out a part of the indirect effects of collaborative behaviors and contract management capabilities. Overall, the findings from the mediation analysis only bring partly support for H4.

*Interpreting outliers as ‘big failures’*

Six outliers were excluded in the main analysis by their highly deviant reported values for contracting performance compared to expected values in the model (all identified by standardized residuals larger than -3.00). By further inspection and qualitative assessment of data and background information the data for outlier cases are suggested to be, not erroneous, but representative for a qualitatively different type of case where contractual governance severely fails or underperform. Outliers were identified across all country contexts.

**Table 5.** Comparing key characteristics of ‘outlier’ and ‘normal’ case groups

Case	Performance		Competition		Contracts		Capability		Collaboration		Size (LN)	
Normal (298)	7.2	(1.4)	6.4	(2.7)	7.0	(2.0)	6.5	(1.9)	7.7	(1.6)	10.6	(1.4)
Outliers (6)	2.3	(1.5)	7.3	(3.1)	7.6	(2.0)	5.8	(2.8)	6.9	(1.9)	11.7	(1.0)
ALL (304)	7.1	(1.7)	6.4	(2.8)	7.0	(2.0)	6.5	(2.0)	7.6	(1.6)	10.6	(1.3)

Note: The table reports means and standard deviations (in brackets) between cases in the outlier group (std. residuals larger than -3.00) and ‘normal’ cases included in the main analysis. Mean differences are (only) statistically significant ( $p < 0.1$ ) for performance and size (LN) SPSS ONE-WAY ANOVA with post-hoc tests.

Table 5 compares key characteristics and performance evaluations in the group of ‘normal’ cases with the groups of ‘deviant’ outlier cases. The group of outliers consists mainly of cases of large municipalities with relatively favorable characteristics related to competition, contracts, capacity and collaboration but still reporting severe contract failures. From the average performance evaluation in the group of outlier cases (score = 2.0) it is indicated that these local governments have had exceptionally disappointing experiences with private service delivery.

The differentiation between a normal and deviant group of cases leads to the finding (by interpretation) that contractual governance has a ‘business-as-usual baseline’ where it works routinely as predicted (by theory), but can be disrupted by ‘big failures’ (like buying a ‘lemon car’). The ‘business-as-usual’ situation is reflected in the final main analysis ( $n = 298$ ).



The particular reasons for performance problems in the outlier cases cannot be empirically assessed further in the analysis. The problems might reflect severe ‘adverse selection’ or ‘moral hazards’ problems apparently hard to mitigate albeit having relatively good contracting institutions in place. Indication is given in Table 3 that possessing good contract management capabilities and embarking on a collaborative approach in exchange relations may be more important for higher contracting performance than being a large local government, having formal contracts in place and being within competitive markets.

#### *Summary of key findings*

Overall, the analysis finds that reformed service provision in a country with a deep contracting culture in comparison with more modest reform countries with more shallow produce superior results – at least within a most-likely setting for a reform instruments. The findings in the analysis support the first hypothesis (H1).

The findings for links between performance and respectively collaboration, capacity, contracts and competition across local governments in the UK and the three Scandinavian countries are all congruent with key notions in different parts of the contracting literature. Overall, the second main hypothesis including sub-hypotheses (H2a-d) is supported by findings in the analysis. Interestingly, collaborative behaviors are found to be far more important for explaining variations in contracting performance than competition.

The analysis of differences in the strength of contracting institutions shows that the UK has substantially stronger contracting institutions in comparison with the three Scandinavian countries combined. However, pairwise country comparisons show that the difference is neglectable for some contracting institutions as well as the UK have relatively weaker contracting institutions in some areas. Denmark, in particular, is very close to having similar strong contracting institutions related to contracts and capabilities, equal levels of collaboration as well as higher levels of competition. As a consequence the third main hypothesis (H3a-d) is only partly supported.

Findings on the links between the UK context, contracting institutions and contracting performance (mediation effects) are only partly supported. Initially, the differences in contracting performance

between the UK and Scandinavia cannot be explained in any substantially way by the total effects from differences in contracting institutions. Overall, the lack of insignificance for the total mediating effect from the four contracting institutions indicates a null finding. However, subsequent analysis finds that differences in the strength of some contracting institutions make differences for contracting performance between the UK and Scandinavia. In particular, weaker competition in the UK reduces the performance difference while stronger contracting capabilities and collaborative behaviors increase the difference. Overall, the findings support hypotheses H4c and H4d while no support is found for hypothesis H4a and H4b.

### **Discussions and conclusions**

The findings sustain the idea that performance in a reformed public sector differs between radical reform countries in comparison with more modest reforming countries. Performance are found to be higher in a radical reform country where a deep contracting culture has developed and embedded relatively strong contracting institutions – at least within a most-likely context. However, the narrative is mixed with both positive and negative experiences. Sustaining sufficient competition in terms of well-qualified bids is indicated to be an issue in the UK whereas contracting institutions related to capabilities and collaborative behaviors are stronger.

The relatively minor role for competition for explaining contract performance in comparison with the greater role of other contracting institutions is partly surprising. The shift from emphasis on competition toward emphasis on collaboration in UK public policies seems to be well-advised. Collaboration is found to be a strong driver of higher contract performance. The finding also indicates why Denmark while embedding very high competition doesn't benefit much from this advantage.

Contracting is found to perform relatively better in the deep contracting culture of the UK than in the three Scandinavian countries. This finding is congruent with the findings by Petersen et al. (2017) on the relative superior economic performance in terms of cost savings in Anglo-Saxon countries compared to other countries as well as the relatively poor economic performance of contracting out in the Scandinavia contexts of Norway found by Leiren et al. (2016) and Sweden found by Bretzer et al. (2016).

However, the analysis also shows that belonging to a particular national context turns out to account only for a very limited (albeit statistical significant) part of the overall variation in contracting performance. For practical purposes the magnitude of the contributions of national context for explaining the variation in contracting performance is negligible compared to the four contracting institutions. In ranked order the substantial determinants for explaining variations in contracting performance are collaboration, capabilities, contracts and competition.

The low ranking of competition for explaining differences in contracting performance compared to the three other contracting institutions are unexpected given the longstanding emphasis in the literature on the role of competition for contracting performance. One reason for the low ranking might be that the study gauges the importance of competition when a certain ‘threshold’ level of competitive pressures are already in place when local governments use private delivery. The emphasis on competition in the theoretical parts of the literature is mainly centered upon a contrast to a non-competitive situation when services are delivered by a (public) provider endowed with a monopoly, i.e. the ‘big’ gains from competition is mainly produced when service delivery shifts from a non-competitive to a competitive context.

It is also found that the strength of key contracting institutions varies across contracting culture and a deep contracting culture partly can be associated with well-developed institutions which in turn support high-performing contracting practices. The finding is slightly modified by two important observations. Firstly, the level of competition in the deep contracting culture of the UK is on the average lower than in the three Scandinavian countries combined and at level with Norway which entertain the lowest level of competition among the four countries. Secondly, collaboration is apparently equally well-developed in the three Scandinavian countries as in the UK albeit collaboration is observed to be slightly lower in Sweden. The deep contracting culture is mainly associated with higher levels of contract management capabilities as well as formal contracts.

### *Perspectives*

In perspective the paper provide several suggestions for further research.

In contrast to the most-likely service setting for contractual governance adopted in the presented research, future research might focus on ‘least-likely’ settings for positive reform outcomes. A

focus on less likely service settings, such as various welfare services and/or services with a high political profile, might yield complementary results on the relative importance of contracting culture.

Research may also adopt alternative and/or more robust research designs, measures and data were possible including quasi-experimental rather than non-experimental designs, longitudinal rather than cross-sectorial data, integrate measures of key constructs from multiple data sources as well as evaluating the same model with different performance measures such as register-based as well as perceptual measures.

## References

- Alonso, J. M., Clifton, J. & Díaz-Fuentes, D. (2017). The impact of government outsourcing on public spending: evidence from European Union countries. *Journal of Policy Modeling*,
- Amirkhanyan, A. A., Kim, H. J. & Lambright, K. T. (2007). Putting the Pieces Together: A Comprehensive Framework for Understanding the Decision to Contract Out and Contractor Performance. *International Journal of Public Administration*, 30(6), 699-725.
- (2010). Do relationships matter? Assessing the association between relationship design and contractor performance. *Public Performance and Management Review*, 34(2), 189-220.
- Barzelay, M. (2001). *The new public management: improving research and policy dialogue*. Berkeley: University of California Press.
- Blom-Hansen, J. (2003). Is private delivery of public services really cheaper? Evidence from public road maintenance in Denmark. *Public Choice*, 115(3), 419-438.
- Bovaird, T. (2006). Developing new forms of partnership with the 'market' in the procurement of public services. *Public Administration*, 84(1), 81-102.
- Bovaird, T., Briggs, I. & Willis, M. (2014). Strategic Commissioning in the UK: Service Improvement Cycle or Just Going Round in Circles?. *Local Government Studies*, 40(4), 533-559.
- Boyne, G. A. (1998). Competitive Tendering in Local Government: A Review of Theory and Evidence. *Public Administration*, 76(4), 695-712.
- (1999). *Managing Local Services: From CCT to Best Value*. Ilford: Frank Cass.

- Bretzer, Y. N., Persson, B. & Randrup, T. B. (2016). Is public procurement efficiency conditioned by market types? A critical test in park and road sectors in Sweden. *International Journal of Public Sector Management*, 29(5), 488-501.
- Brown, T. L. & Potoski, M. (2003). Contract-management capacity in municipal and county governments. *Public Administration Review*, 63(2), 153-164.
- (2005). Transaction Costs and Contracting: The Practitioner perspective. *Public Performance & Management Review*, 38(3), 326-351.
- Brown, T. L., Potoski, M. & Van Slyke, D. M. (2006). Managing public service contracts: Aligning values, institutions, and markets. *Public Administration Review*, 66(3), 323-331.
- Cannon, J. P., Achrol, R. S. & Gundlach, G. T. (2000). Contracts, Norms, and Plural Form Governance. *Journal of the Academy of Marketing Science*, 28(2), 180-194.
- Choi, Y. & Heinrich, C. J. (2006). Design and management of performance-based contracts for public welfare services. In G. A. Boyne, K. J. Meier, L. J. O'Toole Jr. & R. M. Walker (Eds.). *Public Service Performance: Perspectives on Measurement and Management*. (pp. 189-211). Cambridge: Cambridge University Press.
- Christensen, T. & Lægreid, P. (2007). *Transcending New Public Management: The Transformation of Public Sector Reforms*. Burlington: Ashgate.
- Dehoog, R. (1990). Competition, Negotiation, or Cooperation: " Three Models for Service Contracting". *Administration and Society*, 22(3), 317.
- Dempsey, N., Burton, M. & Selin, J. (2016). Contracting out parks and roads maintenance in England. *International Journal of Public Sector Management*, 29(5)
- Donahue, J. D. (1989). *The privatization decision: Public ends, private means*. New York: Basic Books.
- Entwistle, T. & Martin, S. (2005). From competition to collaboration in public service delivery: A new agenda for research. *Public Administration*, 83(1), 233-242.
- Fernandez, S. (2009). Understanding Contracting Performance: An Empirical Analysis. *Administration & Society*, 41(1), 67-100.
- Greve, C. & Ejersbo, N. (2005). *Contracts as Reinvented Institutions in the Public Sector: A Cross-Cultural Comparison*. Westport: Praeger.
- (2016). Reform Context and Status. In C. Greve, P. Lægreid & L. H. Rykkja (Eds.). *Nordic Administrative Reforms: Lessons for Public Management*. (pp. 37-56) Palgrave.
- Greve, C., Lægreid, P. & Rykkja, L. H. (2016). Introduction: The Nordic Model in Transition. In C. Greve, P. Lægreid & L. H. Rykkja (Eds.). *Nordic Administrative Reforms: Lessons for Public Management*. (pp. 1-21). London: Palgrave Macmillian.

- Hayes, A. F. (2013). *Introduction to Mediation, Moderation, and Conditional Process Analysis: A Regression-Based Approach*. New York: The Guilford Press.
- Hood, C. & Dixon, R. (2015). *A Government that Worked Better and Cost Less? Evaluating Three Decades of Reform and Change in UK Central Government*. Oxford: Oxford University Press.
- Hu, L. & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, 6(1), 1-55.
- Kettl, D. F. (1993). *Sharing power: public governance and private markets*. Washington: The Brookings Institution.
- Kim, Y. W. & Brown, T. L. (2012). The Importance of Contract Design. *Public Administration Review*, 72(5), 687-687-696.
- Lamothe, M. & Lamothe, S. (2012). To Trust or Not to Trust? What Matters in Local Government-Vendor Relationships?. *Journal of Public Administration Research and Theory*, 22(4), 867-892.
- Leiren, M. D., Lindholst, A., Ingjerd, S. & Randrup, T., Barfoed. (2016). Capability versus efficiency: contracting out park and road services in Norway. *International Journal of Public Sector Management*, 29(5), 474-487.
- Lindholst, A. C. (2009). Contracting-Out in Urban Green-Space Management: Instruments, approaches and arrangements. *Urban Forestry & Urban Greening*, 8(4), 257-268.
- Lindholst, A. C., Hansen, M. B. & Petersen, O. H. (2016). Marketization trajectories in the Danish road and park sectors: A story of incremental institutional change. *International Journal of Public Sector Management*, 29(5), 457-473.
- Lorenzo-Seva, U., Ferrando, P. J. & Chico, E. (2010). Two SPSS programs for interpreting multiple regression results. *Behavior Research Methods*, 42(1), 29-35.
- Macaulay, S. (1963). Non-contractual Relations in Business: A Preliminary Study. *American Sociological Review*, 28(1), 1-15.
- Macneil, I. R. (1974). The Many Futures of Contracts. *Southern California Law Review*, 47(3), 691-816.
- Nathans, L. L., Oswald, F. L. & Nimon, K. (2012). Interpreting Multiple Linear Regression: A Guidebook of Variable Importance. *Practical Assessment, Research and Evaluation*, 17(9), 1-19.
- Podsakoff, P. M., MacKenzie, S. B. & Podsakoff, N. P. (2012). Sources of Method Bias in Social Science Research and Recommendations on How to Control It. *Annual Review of Psychology*, 63(539-569)

- Pollitt, C. & Bouckaert, G. (2011). Comparative public management reform: An introduction to the key debates. In C. Pollitt, & G. Bouckaert (Eds.). *Public Management Reform. Comparative Analysis of New Public Management, Governance and the Neo-Weberian State*. (pp. 1-25). Oxford: Oxford University Press.
- Poppo, L. & Zenger, T. (2002). Do formal contracts and relational governance function as substitutes or complements?. *Strategic Management Journal*, 23(8), 707-725.
- Romzek, B. S. & Johnston, J. M. (2002). Effective contract implementation and management: A preliminary model. *Journal of Public Administration Research and Theory*, 12(3), 423-453.
- Romzek, B. S. & Johnston, J. M. (2005). State Social Services Contracting: Exploring the Determinants of Effective Contract Accountability. *Public Administration Review*, 65(4), 436-449.
- Savas, E. S. (2000). *Privatization and Public-Private Partnerships*. New York: Chatham House Publishers.
- Schepker, D. J., Oh, W., Martynov, A. & Poppo, L. (2014). The Many Futures of Contracts: Moving beyond structure and safeguarding to coordination and adaptation. *Journal of Management*, 40(1), 193-225.
- Scott, W. R. (2001). *Institutions and Organizations - Second Edition*. London: Sage Publications Ltd.
- Tonidandel, S. & LeBreton, J. M. (2011). Relative Importance Analysis: A Useful Supplement to Regression Analysis. *Journal of Business and Psychology*, 26, 1-26.
- Yang, K., Hsieh, J. Y. & Li, T. S. (2009). Contracting Capacity and Perceived Contracting Performance: Nonlinear Effects and the Role of Time. *Public Administration Review*, 69(4), 681-696.

Appendix I. Survey items

Variables	Question and items <sup>a</sup>	Response scale
<b>Performance</b> <sup>d</sup>	Specify on a scale of 0 to 10 how satisfied or unsatisfied you are with the job (the) private contractor(s) carry out for your department in relation to: <ul style="list-style-type: none"> <li>• Overall quality of maintenance operations.</li> <li>• General pricing and cost of the service provided.</li> <li>• Flexibility to change and/ or develop services if required.</li> <li>• Addressing issues and deficiencies in the service provided.</li> </ul>	Bipolar 11-point scale with end anchors <sup>c</sup>
<b>Competition</b> <sup>e</sup>	Do you usually get sufficient with qualified bids (competition) for tasks you tender?	Unipolar 11-point scale with end anchors <sup>b</sup>
<b>Formal contracts</b> <sup>e</sup>	Specify on a scale of 0 to 10 the degree to which the following content is central in your relation with private contractor(s) for maintaining parks and green spaces and/ or roads: <ul style="list-style-type: none"> <li>• Formalized and written legal clauses (e.g. a signed contract).</li> <li>• Performance specifications – describing overall goals, functionality and guidelines for operation and development.</li> <li>• Prescriptive specifications – based on quantities, instructions and performance measures.</li> <li>• Formal sanctions (e.g. financial penalties) for noncompliance.</li> <li>• Agreement on close collaboration and joint planning of operations and development.</li> <li>• Competence requirements (e.g. professional affiliation or qualification).</li> </ul>	Unipolar 11-point scale with end anchors <sup>b</sup>
<b>Contract management capacity</b> <sup>e</sup>	Specify on a scale of 0 to 10 the degree to which you think that the following statements describe your department's capacity to manage private contractor(s): <ul style="list-style-type: none"> <li>• We have sufficient organisational resources (e.g. time and staff)</li> <li>• We have sufficient experience and expertise</li> <li>• We have sufficient methods and systems (e.g. quality standards, GIS and ICT systems)</li> <li>• Our management practices and procedures are sufficient.</li> </ul>	Unipolar 11-point scale with end anchors <sup>b</sup>
<b>Collaborative relation</b>	On a scale of 0 to 10 to what degree do you think the following statements characterize the relationship between your department and private contractor(s): <ul style="list-style-type: none"> <li>• We are both of the opinion that it is necessary to co-operate in order for each of us to attain our goals</li> <li>• We are both prepared to make operational changes if it makes the work easier for one of the parties</li> <li>• We are both concerned with the other party attaining their goals</li> </ul>	Unipolar 11-point scale with end anchors <sup>b</sup>
<b>Internal provider</b> <sup>d</sup>	Which functions are your department responsible for? <ul style="list-style-type: none"> <li>• Practical delivery (e.g. contractor role / day to day maintenance)</li> </ul>	Yes / No

<sup>a</sup> All items translated for the purpose of this article. Original survey languages: Danish, English, Norwegian and Swedish.

<sup>b</sup> Response categories are based on an 11 point scale with end anchors, where 0 = not at all, 10 = in very high degree.

<sup>c</sup> Response categories are based on an 11 point scale with end anchors, where 0 = very unsatisfied, 10 = very satisfied.

<sup>d</sup> Question is formulated individually for respectively parks and roads.

<sup>e</sup> Question encompasses both parks and roads.