

This is an Open Access document downloaded from ORCA, Cardiff University's institutional repository:<https://orca.cardiff.ac.uk/id/eprint/156421/>

This is the author's version of a work that was submitted to / accepted for publication.

Citation for final published version:

Walker, Richard, Andrews, Rhys , George, Bert and Tu, Xuan 2024. Organizational size and public service performance: a meta-analysis and an agenda for future research. *Asia Pacific Journal of Public Administration* 46 (1) , pp. 32-65. 10.1080/23276665.2023.2176333

Publishers page: <https://doi.org/10.1080/23276665.2023.2176333>

Please note:

Changes made as a result of publishing processes such as copy-editing, formatting and page numbers may not be reflected in this version. For the definitive version of this publication, please refer to the published source. You are advised to consult the publisher's version if you wish to cite this paper.

This version is being made available in accordance with publisher policies. See <http://orca.cf.ac.uk/policies.html> for usage policies. Copyright and moral rights for publications made available in ORCA are retained by the copyright holders.



Organizational Size and Public Service Performance: A Meta-Analysis and Agenda for Future Research

Richard M. Walker, Rhys Andrews, Bert George and Vivian Tu

Abstract

The question whether small or large organizations are associated with the best public service performance has long been a subject of debate in public administration research, and has had profound ramifications for practice. This article seeks to bring clarity to this debate by conducting a meta-analysis of studies scrutinizing the relationship between organizational size and public service performance (45 articles, 122 effect sizes). Meta-analytical and meta-regression results show mostly null findings. We discuss the circumstances in which organizational size matters for public service performance, and propose rekindling venerable research agendas about nonlinearity and contingency in the size-performance relationship.

Key words

Organizational size, public service performance, meta-analysis, null findings

Forthcoming in: Asia Pacific Journal of Public Administration

Introduction

The question whether small or large organizations are associated with the best public service performance has long been debated by public administration scholars and policy-makers (e.g. Dahl and Tufte 1973; Faulk and Hicks 2011; Kuhlmann and Bouckaert 2016; OECD 2005; Ostrom 1976). These debates have largely been shaped by economic theories that focus on the pursuit of economies of scale and scope hypothesizing larger organizations deliver superior performance, and political theories that regard closeness and responsiveness to citizens as the *raison d'être* of government promoting the performance benefits of smaller organizations (Newton 1982; Sharpe 1970). These economic and political perspectives on organizational size rest on very strong assumptions about principles of organizing that can sit uneasily with the practical real-world orientation of public administration. Indeed, prescriptions for being big or for being small have been supplemented within the public administration literature with arguments from public choice theory and from organization theory that attempt to address real-world complexity (Boyne 1998, Kimberley 1976). In particular, public choice arguments regarding the likelihood of an inverted u-shaped relationship between size and performance (see Niskanen 1971; Tullock 1965), and organization theorists' assertions about the importance of organization structure (Child 1972; Hall 1987), and the reach of an organization's network of partners (Gulati et al., 2011).

Explicit decisions to make substantial changes to organizational size in the public sector are typically political ones, and while such decisions are often outside the control of public managers a clear understanding of their consequences can better inform the practice of public administration. Existing theory and practice present alternative remedies to resolve the problem of whether being big or small is best for a public organization. Despite the on-going theoretical and practical importance of this debate and a large number of scholarly studies, an integrative analysis of the public service performance consequences of organizational size has not yet been

conducted. In this article, we ask: does organizational size matter for public service performance? We answer this question through an integration of prior empirical studies published in the field of public administration from 2000 to 2013. A meta-analysis is undertaken on 45 empirical studies (122 effect sizes) identified in the public administration section of Clarivate Analytics ISI Web of Science. This article contributes towards calls for more systematic knowledge in public administration and stronger linkages between theory and data (Abner, Kim and Perry 2017; Perry 2012).

In the next section, we discuss the nature of public service performance. We move on to explore how and why organizational size might matter for public service performance, and then discuss substantive, methodological and contextual moderators. The methods that we employ for our study are then presented. Meta-analysis and meta-regression findings mostly show a null and/or trivial relationship between organizational size and public service performance, and also among the moderators. The overall implication of these findings for public administration research is that organizational size does not appear to have a direct linear effect on the attainment of public service performance. Being big or small is not associated with better, nor, indeed, worse performance. Practically, these findings suggest that whatever other benefits they may bring, amalgamation and de-amalgamation of public organizations are unlikely to result in performance improvement. Nevertheless, as long averred by organizational theorists, organizational size may still be an important explanatory variable influencing the connection between other internal and external contingencies, and public service performance. In the discussion, we therefore call upon public management scholars to return to venerable theory and conceptualization on the organizational size-public service performance relationship when conducting empirical studies.

Organizational size and public service performance

Organizational size is commonly regarded as an important structural characteristic that influences the physical milieu in which the members of an organization interact (see Dalton et al. 1980). Although Aldrich (1972, 3) indicated that organizational size refers to “the scale of an organization's operations and not necessarily to the size of the labor force”, size has often been conceptualized as the number of employees because this captures the sheer number of people who are being organized (Kimberley, 1976). Nevertheless, the scale of operations for many public organizations is often defined less by the number of people being organized than by the size of the budget the organization oversees, the number of clients they serve or the number of contractors that they manage (Jung 2013; Light 1999). As Kimberley (1976) emphasizes, there are several important aspects of organizational size, which vary in their importance depending upon the type and function of the organization. For that reason, we follow Blau (1972) in defining organizational size broadly as “the scope of an organization and its responsibilities” (page 3), though as part of our meta-analysis below, we investigate whether measuring size in terms of employees, clients, or budget has a bearing on the likelihood of performance effects.

Gooding and Wagner (1985) suggest that research on organization size typically operationalizes performance in terms of productivity or efficiency. While this approach may facilitate comparisons of the economic performance of private firms, public service performance is a complex and multi-faceted construct (Andersen, Boesen and Pederson 2016; Heinrich and Lynn 2000; Walker, Boyne, and Brewer 2010). Complexity arises from the range of possible conceptual dimensions of public service performance, and how these are evaluated by stakeholders, together with the different types and sources of data available to measure performance (Andersen, Boesen and Pederson 2016; Walker and Andrews 2015; Walker, Boyne, and Brewer 2010). The complexity of the public service performance construct suggests

an ideal analysis of the relationship between organizational size and public service performance would comprehensively describe the achievements of an organization across a range of performance dimensions (including outputs such as quantity and quality and outcomes like efficiency, effectiveness, responsiveness and equity) using perceptual and archival data collected from internal and external stakeholders. However, practical considerations often limit the range of performance indicators that can be included in public service performance studies, and, in the case of organizational size and public service performance, policy debates are often focused only on cost input measures and indicators of efficiency (Keating 1995). As a result, the connection between the size and the performance of public organizations has conventionally been approached from the perspective of economies of scale. This theory generally contends that bigger organizations are more efficient and effective than smaller ones because of the scale economies that they are able to generate (Fox and Gurley 2006; Stigler 1958). Such economies are likely to take three main forms.

Firstly, in larger organizations the fixed costs of service production can be spread across more units of output. Examples of such fixed costs within public organizations include the common administrative overheads, technical knowledge and equipment that are utilized to achieve organizational aims (Boyne 1996; Dollery and Crase 2004). Secondly, large organizations are assumed to be able to deploy their bigger budget in ways that are conducive to better performance (Boyne 2003). In addition to having more financial resources to invest in service development, large organizations have the purchasing power to drive down suppliers' prices and obtain more favorable rates on credit and loans for new investments (Beck, Demirgüç-Kunt, and Maksimovic 2005; Snyder 1998) – though public sector organizations are more likely than private firms to be constrained by rules and regulations regarding responsible procurement (Flynn and Davis 2016). Thirdly, large organizations offer a wider array of opportunities for professional development than their smaller counterparts, which can attract

more talented employees (Black, Noel and Wang 1999; Kalleberg and Van Buren 1996). In particular, bigger organizations may be able to hire better and more experienced managers who are drawn to the opportunity to take on prestigious and influential leadership roles within such organizations (Jung 2013).

In contrast with economic perspectives, political theory tells us that large public organizations may be disconnected from the citizens that they serve and smaller organizations therefore have stronger public service performance (e.g. Boyne 1996; Newton 1982). Firstly, large public organizations confront greater collective actions problems than smaller ones (Olson 1971), and have little incentive to reconfigure and tailor service provision to the needs of specific niche groups who are unable to exert influence over organizational decision-making (Dahl and Tufte 1973). Secondly, large bureaucracies are generally perceived to be more remote from and less interested in the concerns of the public as a whole (Denters et al. 2014). Again, this makes it more difficult for service users to hold organizations accountable for the quality of the public services that they provide, and implies that small organizations are more responsive to users' needs and demands than large ones (Lowndes and Sullivan 2008). Thirdly, due to all of the problems identified above, large public organizations are less likely to be trusted by citizens, which, in turn, may limit their capacity to get user "buy-in" for the development of new and innovative ways of providing services (Pestoff 2014).

Despite the impact of economic and political arguments regarding the benefits of being big or small, it remains conceivable that the organizational size-public service performance relationship follows a non-linear pattern. Public choice theorists, in particular, claim that scale economies are likely to be exhausted by bureaucratic congestion in large organizations as the sheer number of activities requiring co-ordination overwhelm the central administrative function (Boyne and Meier 2013; Niskanen 1971). Hence, medium-sized organizations may be able to capitalize on being large enough to capture some scale economies and small enough to

retain a connection with service users. For example, Bohte and Meier's (2001) study that operationalizes size as the number of clients provides empirical evidence that bigger organizations are able to offer benefits, such as higher salaries to attract better caliber personnel but that the performance gains associated with these benefits eventually diminish. Building explicitly on that research, Theobald and Nicolson-Crotty (2005) show that larger organizational size gradually impedes communication, coordination and supervision, and past a certain point has a detrimental effect on effectiveness and equity. At the same time, medium-sized organizations may be less likely to suffer from the dysfunctions of large or small size. For instance, a number of studies indicate that it is only the largest public organizations that suffer the effects of bureaucratic congestion (e.g. Andrews, Boyne and Mostafa 2017; Rutherford 2016).

From a sociological point of view, organizational theorists have long pointed towards the task complexity associated with coordinating large organizations (Marsden, Cook, and Kalleberg 1994; Rushing 1967). In particular, because the number of possible social relationships increases as an exponential function of the organization's size (Caplow 1957), large organizations experience serious communication problems between and within the different levels of the organizational hierarchy (Maguire 2003). The relationship between organizational size and complexity may though be curvilinear, with levels of complexity growing at a decreasing rate as organizations grow to a size where they have sufficient administrative capacity to efficiently accommodate further expansion (Blau 1970; Beyer and Trice 1979). Moreover, research on private sector firms finds that organizational size is often a critical intervening variable in a complex set of inter-relationships between the environment, management, organization and outcomes (e.g. Ling et al. 2007; Real et al. 2014; Vaccro et al. 2012) – an approach that has been especially evident in the development of contingency theory (Donaldson 2001). The notion of fit found within contingency theory suggests that

organizations need to align their strategy, structure and other key contingencies, such as organizational size, to generate high performance (Van de Ven et al. 2013). Despite the strength of these arguments, contingency theorists have struggled to conclusively demonstrate measurable performance effects from strategic fit, partly due to a focus on one contingency factor at a time and the challenge of fully capturing contextual variations (Shenkar and Ellis, 2022). All of which implies that the conventional expectations of a strong unidirectional relationship between organizational size and public service performance may be unrealistic and inappropriate, especially given the goal ambiguity and task complexity that confront public organizations (Rainey 2009).

A meta-analysis of the available evidence on the organizational size-public service performance relationship can therefore cast invaluable light on whether arguments in favor of being big or small *per se* are associated with good organizational performance or whether a more nuanced understanding of the relationship that considers the practical nature of public administration is required.

Moderators

The main purpose of a quantitative integration of research results is to identify whether there are generalizable relationships between key variables. To shed light on the potential boundary conditions affecting aggregate effect size analysis it is necessary to explore sources of inconsistencies in those relationships and understand if they are contingent on certain conditions (Hunter and Schmidt 1990). We did this by examining potential moderators of the main effects under consideration. We therefore assess the likely effects of *substantive* (theoretical perspective and the direction hypothesized in the article, organizational goals, level of government, size measurement, performance dimension), *methodological* (performance measurement and research design) and *contextual* (country of study) moderators on the

organizational size -public service performance relationship.

Hypothesized direction

Conventional theoretical perspectives on the effect of organizational size on public service performance offer competing hypotheses. As we discussed in the conceptual section, amalgamation into larger organizations and disaggregation into smaller units are both argued to have a positive effect on public service performance. The meta-analysis technique may drown out these conflicting hypotheses because it examines the aggregate data on organizational size-performance. To overcome this problem, we therefore include as the first moderator the principal arguments made by authors in their hypotheses or their statement of expectations about the direct effect of organizational size on public service performance – smaller or bigger organizational size is related to higher public service performance.

Level of government

The level of government at which a public organization is located (i.e. federal or central versus local) may have an influence on the likelihood that being big or small is beneficial for public service performance. Arguments in favor of small public organizations, in particular, tend to focus on the local level of government, which is assumed to need a close physical connection to citizens (Dahl and Tufte 1973; Denters et al. 2014). By contrast, large federal or central government agencies, departments and ministries, are rarely required to be accessible to citizens, and are often orientated towards the internal production of policy outcomes and so may be better able to capture both production and administration scale economies than their local level counterparts (Jung 2013). Federal or central government organizations may also confront particularly challenging levels of task complexity, which require considerable capacity to address effectively (Beyer and Trice 1978).

Organizational type

The tasks and authorities that are allocated to public organizations vary considerably within and across different policy fields. In particular, many organizations are tasked with a single public purpose (e.g. border control, welfare payments), whereas others are multi-purpose organizations responsible for providing a wide range of public services (e.g. city and county governments). Proponents of management reforms, including New Public Management contend that single-purpose organizations will perform better than multi-purpose ones because they are smaller, have clearer lines of managerial and political accountability and are less susceptible to “producerist” capture by powerful professional interest groups or trade unions (Hood 1991; Talbot 2004). Indeed, many governments across the world embraced these ideas in pursuit of performance improvements and have disaggregated larger organizations into smaller single-purpose entities through processes of agencification (Pollitt and Bouckaert 2017; Verhoest et al. 2016). Nevertheless, there are still sound economic reasons for anticipating that multi-purpose organizations might out-perform their single-purpose counterparts. In particular, public organizations providing many different services may be able to generate more internal economies of scope than single-purpose organizations by spreading fixed costs, such as administration, ICT, buildings, estates and equipment, across multiple production units (Grosskopf and Yaisawamg 1990). Moreover, empirical studies analyzing agency performance suggest that they may in fact be inefficient and unproductive organizational forms (Overman and van Thiel 2016).

Size measurement

The next substantive moderator is the operationalization of organizational size, which is variously measured as the number of employees, the client population, or the budget of an organization. Attempts to capture overall size effects are most easily achieved by measuring

the number of employees. However, data on the number of employees is not always readily available (Light 1999). When these data are not available researchers turn to proxy measures, of which one of the most frequently used is the number of citizens that the organization serves (referred to as population in the analysis below). The number of clients has also been used as a surrogate measure of organizational size (Kimberly 1976). An alternative measure of organizational size is budget. When budget is used to operationalize the concept of organizational size competing arguments have been presented (Boyne 2003). The first is that more resources produce better services, albeit that they need to be managed effectively. Second, this is countered by public choice prescriptions that bigger budgets results in bloated and inefficient organizations due to rent-seeking by bureaucrats.

Organizational size is typically measured as a static concept. However, it also has an important dynamic aspect. The rate at which organizations increase or decrease in size may be as important an influence on performance as whether or not they are large or small in the first place – for example, the empirical literature on downsizing points to a decline in public service quality and equity as the slack resources needed to respond to environmental change are “hollowed out” (Raudla, Savi and Randma-Liiv 2015). Dynamic size effects are reflected in changes in employees and budgets, and have often be captured using measures of change in the client population (e.g. Ladd 1992). Typically, studies drawing on dynamic size measures suggest that increases in the size of an organization’s client population generate more resources for the organization and are therefore associated with improved performance (e.g. Andrews et al. 2005; though see Ladd1992).

Performance dimension

Public service performance is a multi-dimensional construct. Public organizations produce a range of outputs and outcomes and a number of models have been advanced for classifying

those outputs and outcomes (Boyne 2002; Selden and Sowa 2004; Walker, Boyne, and Brewer 2010). The 3Es model, for example, emphasizes economy, efficiency, and effectiveness, while the IOO model, stresses relationships between inputs, outputs and outcomes. Within the IOO model, key outputs include: speed, quality, and quantity, and key outcomes focus on efficiency, effectiveness, equity, and responsiveness (Boyne 2002). Walker, Boyne, and Brewer (2010) extend this logic to include broader concerns about the governance of public organizations, arguing that several additional dimensions should be included in models of public service performance, such as accountability, probity and respect for human rights.

Organizational size may influence these different dimensions of public service performance in different ways. Due to varying levels of resource capacity, human capital and political legitimacy, big and small organizations are likely to encounter contrasting challenges in pursuit of improved efficiency, effectiveness and responsiveness to service users. For example, much of the economic argument reviewed above suggests that larger organizations are inherently more efficient (e.g. Stigler 1958) and effective (e.g. Jung 2013). Furthermore, organizational size is often assumed to have a positive relationship with outputs more generally (e.g. quantity and quality), due to the quality of the personnel that big organizations attract (Boyne 1996).

Performance measurement

The measurement of performance is our first methodological moderator. Previous quantitative studies of management and public service performance outcomes find that differences in the measurement of the performance construct influence statistical findings (Jakobsen and Jensen 2015; Meier and O'Toole 2013b). Typically, PSP indicators are drawn from stakeholders' perceptions or from archival sources (Selden and Sowa 2004; Walker, Boyne, and Brewer 2010). Perceptual measures are useful, as they cover a wide range of non-economic dimensions

of performance, such as responsiveness and service satisfaction. However, they are problematic when independent variables (in this case organizational characteristics) and the dependent variable (public service performance) are drawn from a single survey source (Podsakoff et al. 2003; George and Pandey 2017). Organizational size is typically measured with secondary data sources, thus common source bias may not affect the coefficients in the studies we examine. Nevertheless, a critical element of debates about performance measurement focuses on the likelihood that survey respondents will overestimate the performance of their organization (Meier and O'Toole 2013a) resulting in a greater prospect of rejecting the null hypothesis and making a Type I error. We therefore undertake moderator analysis on whether performance is measured using archival or perceptual data.

Research design

Public administration researchers increasingly employ robust research designs with the power to detect causal relationships. Central to the identification of causal relationships is time. In observational studies, this entails the measurement of the independent variables before the dependent variable, and is accommodated by longitudinal data that also offer the opportunity to lag the measurement of the dependent variable. Such robust longitudinal designs are expected to identify organizational size-public service performance relationships. However, not all observational research designs implement lagged structures. Cross-sectional research designs are only able to identify associations between variables, and are more susceptible to Type I and Type II errors. While improvements to research design have reduced the publication of cross-sectional studies in recent years in public administration, the use of cross-sectional data was widespread during the time period of this study. For example, Walker and Andrews (2015) reported that around a third of studies in their integration study of management and local government performance adopted this practice.

Country

Context is an important variable to consider in any study of organization, management and performance (Meier, Rutherford and Avellaneda 2017; O'Toole and Meier 2015; Thompson 1967). Public organizations vary in scope, size and purpose in ways that reflect national institutional governance arrangements. In particular, organizational size in the public sector is likely to be affected by the regulatory environment. The majority of the studies examined in this article are from Europe or the U.S.A. The regulatory environment in general is more intense in Europe than in the U.S.A. (Löfstedt and Vogel 2001). It is argued that public organizations in Europe tend to have less autonomy, more political oversight and are more constrained by labour market and other legal regulations than in the US (Meier, Rutherford and Avellaneda 2017). Consequently, regulators in Europe may prefer closely managing and supporting the performance of big organizations due to the greater risks associated with the failure of such organizations (Pollitt 2006). By contrast, public choice arguments about the relationship between small organizational size and public service performance may be more likely to hold in the American context, whereby public organizations generally have more room for maneuver to compete with one another for service users (Boyne 1998).

Methods

This study utilizes a meta-analysis, or an “analysis of analyzes”, to aggregate quantitative results from different studies into a single integrated literature review. A meta-analysis is defined by Glass (1976, 3) as “the statistical analysis of a large collection of analysis results from individual studies for the purpose of integrating the findings.” The research design follows the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines (Liberati et al. 2009).

Data

The empirical literature on organizational size and public service performance was located in the public administration journals listed in the Thompson Reuters Web of Science Social Sciences Citation Index (SSCI). Search terms included: perform*, output*, outcome*, responsive*, consequenc*, efficienc*, effective*, satisfac* and equit* to capture different dimensions of public service performance. This first stage of our search procedure resulted in the identification of 7,363 papers dealing with issues of public service performance for the period 2000-2013. These studies were examined in greater depth, and the following decision rules for inclusion in the analysis were implemented: the unit of analysis was an organization, or part thereof (thus excluding studies of individuals and groups of individuals)¹; studies included measures of management and organization; public service performance was operationalized as the dependent variable (see search terms above); and articles presented statistical results in the form of correlation coefficients or regressions coefficients that were suitable for the application of the meta-analytic technique used in this study (see below). This procedure resulted in a sample of 236 articles (see figure 1). Organization size is a widely used in statistical models of public service performance, as an independent variable and is often used as an important control variable (O'Toole and Meier 1999). The second stage of our search involved examining each of the 236 articles that included measures of public service performance and organization and management by hand to identify those that included a measure of organizational size.² The implementation of this final decision rule resulted in 45 studies representing just over 19% of all the articles identified in stage one, containing 122 effect sizes of organizational size and public service performance.³

[Figure 1 here]

The review strategy that we adopted benefits from focusing on peer-reviewed journal articles that were judged to be of suitable quality for publication by editors following a blind

review process, and were therefore expected to meet the basic requirements of theoretical and methodological rigor. It does, however, exclude unpublished papers on management and public service performance and work sponsored by government, national and global organizations, such as the OECD, with an interest in the achievements of public organizations, along with books and book chapters. This approach may lead to bias, by overstating the relationship between organizational size and public service performance if articles that contain statistically significant results are more likely to be published in journals. Hence, although estimates from other fields suggest that the magnitude of publication bias is likely to be small (Rosenthal 1991), we test for these problems in our analysis (see below).

The studies we examine were published between 2001 and 2013. The highest number of studies (8) were published in 2013 with no studies in 2002 and 2007. All of the studies contained at least one sample with an average sample size or number of organizations of 274 ranging from 22 to 1,748 organizations.⁴ Table 1 contains an overview of all included studies and their characteristics. Table 2 contains a summary of the distribution of effect sizes across our moderators.

[Insert Table 1 and 2 about here]

The majority of effect sizes offered no stated direction or hypothesis (53%). About 28% of effect sizes hypothesized that larger organizations would have a positive relationship with performance, while 5 percent stated smaller was better, and a null hypothesis was offered for 14% of effect sizes. Of the effect sizes in this analysis 55 percent examined single purpose organizations including fire brigades, police departments and school districts, and the balance investigated multi-purpose organizations that deliver several types of public service. Size was operationalized using the proxy measure of the population of the jurisdiction in 33 effect sizes (27%), the number of clients in 23 (19%)⁵, employees in 16 (13%), the budget of the organization was used in 39 effect sizes (32%), and dynamically by population change in 11

(9%).⁶ The studies drew on five dimensions of performance as well as an aggregate of multiple dimensions, efficiency and equity were used as a dependent variable the least (8% and 6%), effectiveness the most (36%).

The majority of effect sizes used secondary data to operationalize performance: archival data were used in 90 (74%) effect sizes and 59% of effect sizes (72) implemented research designs that sought to tease out causality, largely by introducing some semblance of time into the regression model. Over half of the effect sizes (66, 54%) were from the U.S.A, with 40% (49 effect sizes) coming from European studies.

Meta-analytical procedure

The meta-analytical procedure proposed specifically for public management and policy studies by Ringquist (2013) was used to integrate the findings of the empirical studies on the relationship between organizational size and public service performance. This procedure allowed us to derive a population effect size across effect sizes, and identify potential issues with publication bias. All analyses were conducted using Stata software. This approach has been used and validated by recent public administration meta-analyses (George, Walker and Monster 2019; Gerrish 2016). Below we describe the three-step approach implemented in this article.

Identifying effect sizes

The effect size for each relationship identified in a study was calculated. Correlation coefficients are a well-known and easily interpreted effect size across the social sciences (Field and Gillett 2010; Rosenthal and DiMatteo 2001; Ringquist 2013). Eleven studies reported correlation coefficients. Those that did not, did typically report regression coefficients and related information. It was therefore necessary to calculate correlation coefficients based on the

t-statistic and the degrees of freedom, thus:

$$r = \sqrt{[t^2 / (t^2 + df)]}$$

with t = the t-statistic and df = degrees of freedom

The t-statistic can also be calculated by dividing the unstandardized regression coefficient and their standard errors. If studies did not report unstandardized coefficients or t-values but rather z-values or standardized coefficients, the formulas proposed by Ringquist (2013 105-110) were used to calculate the correlation coefficients. To maximize the number of effect sizes, we did not calculate an average effect size per study at first but rather included all effect sizes between organizational size and public service performance reported in each study. However, to avoid that studies containing many effect sizes would skew our findings we did, as a robustness check, test whether results changed when only one average effect size per study was included.

Conducting the meta-analysis

A main population effect size was derived for all the identified effect sizes on organizational size and public service performance. A random-effects meta-analysis model using Fisher's r-to-Z transformation was selected for our analysis (Gerrish 2016; Hedges and Vevea 1998; Homberg, McCarthy and Tabvuma 2015). A random-effects model is preferred because we are working with "real-world data ... [that] are likely to have variable population parameters" (Field and Gillet 2010, 673). A fixed-effects model would not be appropriate given that sample characteristics vary considerably. In order to be able to generalize study findings to a broader population (i.e. not just the studies included in a meta-analysis), the unconditional inferences presented by random-effects models are required (Hedges and Vevea 1998).

Moreover, Fisher's r-to-Z transformation was used because this is a convention in meta-analysis in public management and policy as well as the broader social sciences (Field and Gillet 2010; Ringquist 2013). Following this, the overall variation in effect sizes is calculated

as well as the main population effect size, its significance and the accompanying 95% confidence interval. In addition to calculating an overall population effect size across all effect sizes and studies, we also conduct a subgroup analysis as a first exploration of potential moderators. Such an analysis centers on calculating the population effect size for specific groups in the dataset thus enabling one to compare this population effect size across groups and identify potential differences. Subsequent to this we perform a meta-regression.

Conducting the meta-regression analysis

A meta-regression analysis was performed to identify whether the moderators were significant sources of variation among effect sizes using a random-effects model with clustered standard errors at the study level (i.e. to account for the fact that studies report multiple effect sizes) based on the procedure of Ringquist (2013). All analyses were conducted in Stata. All moderators were dummy-coded and the dummies were included as independent variables into the model (based on the n-1 principle). Dummies with a significant coefficient are thus a source of variation in effect sizes overall.

Identifying publication bias

Visual and statistical tests to detect potential issues with publication bias were conducted. Publication bias in the scholarly community – also called the “file drawer problem” – is assumed to mean that more statistically significant research findings are being published than null results (Franco, Malhortra and Simonovits 2014; Ringquist 2013). In other words, studies with null results end up in a “file drawer” as opposed to being published because reviewers and editors favor articles with significant findings (van Witteloostuijn 2016). Hence, publication bias implies that a meta-analysis based on articles published in peer-reviewed journals is likely to overestimate population effects because these do not include unpublished articles with

statistically insignificant findings (Field and Gillett 2010). There is no “definitive” test which can precisely estimate the presence of publication bias (Ringquist 2013). Nonetheless, we report the Egger test and Begg test to statistically test for the presence of publication bias and generate a funnel plot to indicate potential asymmetry of effect sizes as a visual test for publication bias.

Results

Meta-analysis

Table 3 contains the results of the meta-analysis between organizational size and public service performance. The first analysis contains all effect sizes identified in the studies (i.e. 122). The population effect size of the association between organizational size and public service performance is positive and statistically significant but the strength of the effect size is trivial at best (i.e. far below .10). The second analysis contains one average effect size per study (i.e. 45) and paints an even grimmer picture, the population effect size becomes even smaller as well as statistically insignificant. The findings suggest that organizational size has not had much influence, on average, on public service performance. Figure 2 offers a forest plot visualizing the average effect size per study, its 95% confidence interval and its weight in the overall analysis.

[Figure 2 about here]

Importantly, in both analyses we find that the overall variation in the effect sizes is large and significant (i.e. as indicated by the Cochran's Q values). This implies that there may be important moderators influencing the effect sizes – such as how organizational size and public service performance are measured. We therefore conduct several moderator analyses in what follows including potential substantive, methodological and contextual moderators on the organizational size-public service performance relationship.

[Table 3 about here]

As a first moderator analysis, Table 4 presents the same analysis but this time per different group based on our proposed moderators. This table thus demonstrates whether the population effect size is different between groups in the dataset. As Table 4 shows, few differences pop up. The majority of moderators have insignificant population effect sizes, and when a population effect size is significant it is mostly trivial (i.e. below .10) with the exception of output and perceptual measures where population effect sizes are still considered small, however.

[Table 4 about here]

Meta-regression analysis findings

Table 5 presents the results of the meta-regression analysis. Importantly, few of the included moderators were a significant source of variation in effect sizes overall. In other words, the hypothesized direction, the organizational goals, the level of government, how size is measured, which methods were used and the country context do not – based on our data – matter much for variation in the organizational size-public service performance relationship. One moderator that did demonstrate some significance is the different dimensions of performance, with size having a significantly smaller impact on effectiveness, responsiveness and equity in comparison to output. But, again, effect sizes remain small at best. The model only explains 29% of variation, thus indicating that there are other conditions that could influence the size-performance relationship which we did not include. These findings offer additional evidence on the overall findings of a null relationship in the main meta-analysis reported in the main body of the article.

[Table 5 about here]

Publication bias

The weak meta-analysis findings are not suggestive of strong publication bias. Nevertheless, we report the Egger test and the Begg test results. Both tests provide significant results (Begg test significant at $p = .001$ level; Egger shows a bias coefficient of 1.32 significant at the $p < .01$

level). These tests further validate the null findings, as the positive, trivial population effect size could even be slightly inflated. The funnel plot illustrated in figure 3 further demonstrate some asymmetry observed in the sample.

[Figure 3 here]

Discussion

Implications for theory

Public administration is a design science developing theoretically informed solutions to complex and human-related real-world problems (Simon 1996). Recommendations from public administration scholarship can potentially have tangible and profound consequences for citizens. This is the case for organizational size: based on arguments from the fields of economics and political science on the benefits of large and small organizational size governments have undertaken major organizational restructuring. For example, Blom-Hansen et al. (2016, 813) note the “dramatic scope” of amalgamations in local governments that have affected millions of citizens. Despite the strength of the contrasting perspectives on the role of large and small organizations in achieving higher levels of public service performance, few studies have developed organizational samples that can adequately test the normative arguments about size that are proposed by economists and political scientists. As a result, a straightforward definitive point of view is not supported in the effects presented in this article. Furthermore, our meta-regression did not uncover any relationships with alternative operationalizations of organizational size or public service performance, apart from the more positive impact of size on output (although still small at best in terms of strength).

Given the mostly null findings on the linear relationship between organizational size and public service performance, we argue that the theory-practice focus of the applied social science discipline of public administration suggests the need to identify the conditions under

which size matters for performance. We discuss two perspectives that move beyond the blunt measurement of organizational size and performance alone, suggesting that future research should rekindle and reengage with venerable debates about curvilinear and contingency perspectives on the relationship between organizational size and public service performance that have been neglected in the empirical research reviewed in this article.

First, the organizational size-performance relationship could be conceptualized as nonlinear, especially in an inverted U-shape indicating that beyond an optimal organizational size the benefits of being big are lost – referred to as the Too-Much-of-a-Good-Thing effect by Pierce and Aguinis (2013) – though the precise dynamics of such a relationship would be difficult to identify in the interorganizational networks increasingly responsible for providing public services. Likewise, investigation of the moderating and mediating effect of organizational size on other key determinants of public service performance could contribute to the growing body of empirical studies and meta-analyses in public administration that seek to capture the complexity of the relationship between management, organization and outcomes (Avellaneda and Gomes 2015; Borst et al, 2020; Ding et al., 2021; George et al., 2021; Jugl 2019; O’Toole and Meier 1999; 2015; Provan and Kenis 2008; Walker, Avellaneda and Berry 2015). Although it is clearly impossible to draw lessons about the right organizational size across the whole of the public sector, research on public schools highlights the conditions where size matter and how size different effects on different organizational goals (e.g., Theobald and Nicholson-Crotty 2005). Once sufficient studies are conducted meta-analysis techniques can be applied to identify the optimal school size (see Colegrave and Giles 2008). Hence, public administration research could be advanced by revisiting, exploring and empirically testing theory conceptualizing the conditions that influence the organizational size-public service performance relationship to identify potential tipping-points.

The second part of this agenda to is to reconnect the organizational size-public service

performance relationship with longstanding perspectives. Management researchers are increasingly interested in the relationship between organizational size and a range of management and organizational variables, especially those relating to innovation, management tools and organizational learning (e.g. Lee and Kim 2016; Petruzzelli et al. 2018; Real et al. 2014 – for rare public administration examples in this area, see George, Van de Walle and Hammerschmid 2019; and Jung and Lee 2016). A number of the studies examined in this meta-analysis do locate organizational size within conceptual frameworks that examine the inter-relationship between management and organization variables, for example that of O'Toole and Meier (1999). However, investigation of joint relationships between these variables and organizational size is still rare. Conceptual and theoretical understanding of the organizational size-public service performance relationship could usefully be advanced by drawing upon theories from public administration and management that place organizational size within a wider organizational ecosystem, such as organizational ecology (e.g., Bidwell and Kasarda 1985), contingency theory (e.g., Donaldson 2001) or configurational theory (e.g., Meyer, Tsui, and Hinings 1993). A starting place for this investigation could be to revisit theories of bureaucracy. For example, Blau (1970 1972) found that some dimensions of bureaucratization, such as formalization, vertical and horizontal complexity and differentiation, grow alongside increasing organizational size, albeit at a diminishing marginal rate. Dimensions of bureaucratization could be theorized as mediators in future studies of organizational size-public service performance. Similarly, theories of organizational fit, such as Donaldson's (2001) Structural Adaptation to Regain Fit (SARFIT) model, indicate that the relationship between structure and organizational size moderates the achievement of public service performance. Fruitful lines of enquiry for public administration researchers would include examining moderators such as organizational strategy, governance structures and environmental uncertainty (Badri et al. 2000; Palmon and Wald 2002; Smith et al. 2016). This is suggestive of

future research conceptualizing and empirically testing management and organizational moderators and mediators when examining the organizational size-public service performance relationship.

In short, theoretical expectations on nonlinear and moderated or mediated relationships need to be empirically tested in numerous organizational settings in order to advance theory on the relationship between organizational size and public service performance and to provide robust recommendations to practice.

Implications for research

We focus on three implications for research drawing on our earlier discussion of the measurement of performance and organizational size and raise some research implications about the widespread use of size as a control variable in the studies included in our meta-analysis.

The moderator analysis suggests that when scholars analyze the organizational size-public service performance relationship they do not take a comprehensive view of the performance construct, in terms of its dimensions and measurement, and perhaps do not fully consider appropriate performance indicators. While the studies reviewed in this article drew upon a range of measures of public service performance the majority of studies examined effectiveness as the dependent variable. There was only limited inclusion of measures of efficiency, equity, outputs (quality and quantity), responsiveness and indexes. It is possible that the slight preponderance of evidence in favor of positive but statistically insignificant size effects reflects the impact of scale economies on the types of performance measures that are examined in this article. In particular, improvements in effectiveness generally imply increased output levels. Nevertheless, given the focus of attention on efficiency in many of the theoretical and policy debates on organizational size (Gendźwił 2020; Tavares 2018), the comparatively

limited consideration of this measure of performance is surprising. Scholars should seek to compare the effects of organizational size on efficiency with a range of other performance indicators simultaneously to fully tease out the potential for scale economies to be traded-off against other important aspects of public service performance, such as effectiveness and equity.

Organizational size is a complex construct that can be operationalized in a number of ways, none of which may separately capture the full dimensionality of organizational size and its effects. Moreover, changes in the management models adopted for public service delivery often obscure the scale of production. In addition to the growth in regional agencies coordinating public services across wide areas (Fesler, 1949), in recent times, the task of government has evolved from one involving direct service provision to the co-ordination of assorted contractors and networks of producers within the “hollow state” (Milward and Provan 2000). According to some observers, the hollowing out of the state in this way has paradoxically required an expansion of organizational capacity to cope with the complexity of the new public administration environment (Skelcher, 2000). As a result of these developments, it is increasingly difficult to estimate the “true size of government.” (Light 1999). Research on organizational size that is able to capture the contribution of all the staff employed by the organizations involved in networks of public service production, may reveal a very different relationship with public service performance to that presented here. All of which underlines the need to better understand the inter-connections between multiple management, organizational and network variables and public service performance before coming to any firm conclusions about the importance of organizational size.

Our last implication for research arises from an artefact of the studies that were reviewed. Organizational size was extensively used as a control in multiple regression models of management and public service performance. We were able to identify these controls by searching for studies on public service performance and then interrogating these to identify

measures of organizational size, rather than searching titles and abstracts alone. Organizational size is clearly identified as an important control variable in public service performance studies to rule out spurious relationships. When scholars include size as a control variable the expected relationship should be clearly articulated. Authors often failed to hypothesize the direction of the effect when including organizational size as a control. Public administration scholars are therefore encouraged to clearly hypothesize the anticipated impact of organizational size on public service performance in future studies. For example, Bernerth and Aguinis (2016) recommend developing theoretical arguments to justify the inclusion of control variables, using citations to show how the control has been used in other studies (including reporting results), and anticipating the direction of the relationships between the control and dependent variable. If possible, the validity and reliability of the control measure should also be explored. Given the widespread use of organizational size as a control, the inclusion of such information would provide clearer evidence on the organizational size-public service performance relationship.⁹

Implications for practice

It seems that, on average, neither large nor small organizations perform better than other organizations by simple virtue of their size. Hence, in general, bigger is not better and smaller is not more beautiful. Politicians and policy-makers concerned with improving performance in the public sector should therefore consider alternative approaches to large-scale restructuring for promoting improvements in public service performance. For example, performance management techniques, such as the use of strategic planning and benchmarking, have been shown to improve the effectiveness of public services and have a positive impact on outcomes (George et al. 2019; Gerrish 2016; Walker and Andrews 2015).

Governments that make changes to organizational size often evoke the prospect of performance improvement, however the wide research suggests that changes in organizational

size are motivated as much by political as technical considerations (Drew, Razin and Andrews 2019; Feiock, Carr and Johnson 2006). Thus, it cannot be assumed that performance changes will be automatically gained. Rather than seeking to change the size of organizations to demonstrate their political authority, governments should pay closer attention to the role that organizational size may play in determining the success of other approaches to driving public service performance. For instance, bigger public organizations may be better able to introduce the process innovations that could hold the key to successful organizational change (Walker 2014). Alternatively, in seeking to generate improved effectiveness and efficiency, time and resources spent exploring alternatives to a strategy of organizational growth or downsizing may therefore pay dividends. Government policies and initiatives could be formulated to support the development of the capabilities needed to lead and manage small and large public organizations drawing on the contingencies and configurations that result in successful performance outcomes.

Limitations

There are a number of limitations to this study which can inform future research on this topic. We have already touched on questions of nonlinearity, performance dimensions and the operationalization of organizational size and public service performance above. In addition to this, further research could be conducted in associated fields to ascertain if these findings apply to cognate disciplines such as economics, political science and sociology. For example, political science research suggests organizational size has varying effects on different areas of policy expenditure within government (Blom-Hansen et al. 2016). Subsequent studies could build on those findings to evaluate performance effects within and across governments, in different policy domains and in other public service organizations.

Second, the organizational size-public service performance relationship may suffer from

reversed causality, where changes in performance positively or negatively influence organizational size. In this article, we sought to control for some aspects of this by including moderators on research design and different measures of size. Nonetheless, studies measuring size as the number of clients typically examined schools which are susceptible to changes in organizational size as a result of performance changes, with ineffective schools shrinking in size and successful ones growing. Although most of the studies in our meta-analysis focused on public organizations whose size is unlikely to be determined by their performance, the potential for this to affect the findings presented here persists.

Third, as with many meta-analyses in public administration, correlation coefficients are extracted from regression or correlation tables based on observational data. This raises the possibility of potentially spurious relationships influenced by omitted variable bias. Finally, while we uncovered a trivial relationship between organizational size and public service performance, further interrogation of the studies in our analysis is suggestive of publication bias. As the evidence on this topic continues to grow, replication of our study would reveal whether the conclusions that we draw about the organizational size-public service performance relationship here are upheld or require revision.

Conclusion

The null meta-analysis findings suggest that when studied as a linear relationship, organizational size is not a major determinant of public service performance, yet governments around the globe invest substantial resources in amalgamating and disaggregating public service organizations (Askim et al. 2017; Blom-Hansen et al. 2016; Dollery et al. 2020; Lago-Peñas and Martinez-Vazquez 2013). Furthermore, and unlike prior meta-analyses (Gooding and Wanger 1985), we did not identify many moderating effects, indicating that the varying conceptualizations and measurement in the primary studies did not influence our meta-analysis

results much. This finding strikes at the heart of longstanding debates about government efficiency and effectiveness. These debates have typically suggested that “bigger is better” or that “small is beautiful”. We noted that these competing perspectives fail to recognize the real-world complexity of public administration, and the findings reported here could support a new proposition: organizational size alone is not a critical antecedent of public service performance.

Our findings suggest that sufficient research has been directed at the question of the linear relationship between organizational size and public service performance, and that better theorization of the complex interplay between organizational size and other internal and external contingencies in studies of public service performance published in public administration journals may be a more fruitful avenue to explore. Researchers could then empirically test nonlinear, mediated and moderated relationships, to verify or refute these venerable arguments. This meta-analysis lays down a research agenda on the relationship between organizational size and public service performance and we encourage others to contribute further empirical evidence to this debate to advance the theory and practice of public administration.

Notes

¹ Studies that included class size as a measure of organizational size were excluded from our analysis because a class is a group of individuals within an organization, rather than a sub-unit or department (see Meier and O’Toole 2013a 2013b).

² For example, Andersen, Boesen and Pedersen (2016) searched Scopus using the terms “performance” and “public” and returned 794 articles, which after close reading revealed that only 61 were suitable for the purposes of analysis.

³ We search for the use of organizational size as an independent variable, a control and for non-linear and moderated relationships. Non-linear, mediated and moderated relationships are

not examined in enough empirical studies of organizational size and public service performance to undertake a meta-analysis.

⁴ Identified studies were coded by two of the authors. They coded each study independently and then shared their coding results. Disagreements were dealt with through discussion until agreement was reached. This approach was taken because of the need to identify studies that included measures of organizational size as a control.

⁵ For example, the organizational size-public service performance relationship was studied in the education arena. These studies used clients to operationalize organizational size, and label this school size.

⁶ Integration studies, including those on questions of organizational size, frequently tackle concepts that are operationalized in a variety of ways, see for example Boyne (2003), Damanpour (1993), Gerrish (2016), Walker and Andrews (2015) and Yeung (2009). Indeed, if the studies reviewed in a meta-analysis were homogeneous on questions of population, measures and settings etc. little evidence would be generated to support casual generalizations, and might introduce method bias (Matt 2003).

References

* = Included in the meta-analysis

- Amirkhanyan, A. A., H. J. Kim, and K. T. Lambright. 2010. Do relationships matter? Assessing the association between relationship design and contractor performance. *Public Performance & Management Review* 34:189-220.*
- Abner, G. B., S. Y. and J. L. Perry. 2017. Building evidence from public human resource management: Using middle range theory to link theory and data. *Review of Public Personnel Administration* 37, 139-159.
- Aldrich, H. E. 1972. Technology and organizational structure: A reexamination of the findings of the Aston Group. *Administrative Science Quarterly* 17:26-43.
- Andersen, L. B., A. Boesen, and L. H. Pedersen. 2016. Performance in public organizations: Clarifying the conceptual space. *Public Administration Review* 76, 852-862.
- Andrews, R. 2009. Organizational task environments and performance: an empirical analysis. *International Public Management Journal* 12:1-23.*
- Andrews, R., and G. A. Boyne. 2010. Capacity, leadership and organizational performance: Testing the black box model of public management. *Public Administration Review* 70: 443-454.*
- Andrews, R., and G. A. Boyne. 2011. Corporate capacity and public service performance. *Public Administration* 89:894-908.*
- Andrews, R., G. A. Boyne, J. Law, and R. M. Walker. 2005. External constraints on local service standards: the case of comprehensive performance assessment in English local government. *Public Administration* 83, 639-656.*
- Andrews, R., G.A. Boyne, K.J. Meier, L.J. O'Toole Jr. and R.M. Walker. 2009. EU accession and public service performance, *Policy & Politics* 37:19-37.*
- Andrews, R, G.A. Boyne, K. J. Meier, L. J. O'Toole Jr. and R.M. Walker. 2012. Vertical

- strategic alignment and public service performance. *Public Administration* 90:77-98.*
- Andrews, R., G.A. Boyne, J. Law, and R.M. Walker. 2009a. Centralization, organizational strategy, and public service performance. *Journal of Public Administration Research and Theory* 19:57-80.*
- Andrews, R., G. A. Boyne, J. Law and R. M Walker. 2009b. Strategy formulation, strategy content and performance: An empirical analysis', *Public Management Review* 11: 1-22.*
- Andrews, R, G.A. Boyne, J. Law and R.M. Walker. 2011. Strategy implementation and public service performance. *Administration & Society* 43:643-671.*
- Andrews, R., G. Boyne, and A. S. M. Mostafa. 2017. When bureaucracy matters for organizational performance: The benefits of administrative intensity in big and complex organizations. *Public Administration* 95, 115-139.
- Andrews, R., G.A. Boyne, L.J. O'Toole, K.J. Meier, and R.M. Walker. 2013. Managing migration? EU enlargement, local government capacity and performance in England. *Public Administration* 91:174-194.*
- Andrews, R., R.J. Cowell and J. Downe. 2008. Support for active citizenship and public service performance: An empirical analysis of English local authorities. *Policy & Politics* 36:225-243.*
- Andrews, R., and T. Entwistle. 2010. Does cross-sectional partnership deliver? An exploration of public service effectiveness, efficiency and equity. *Journal of Public Administration Research and Theory* 20:679-701.*
- Andrews, R., and S. Van de Walle. 2013. New public management and citizens' perceptions of local service efficiency, responsiveness, equity and effectiveness. *Public Management Review* 15:762-783.*

- Askim, J., J. E. Klausen, S. I. Vabo and K. Bjurstrøm. 2017. Territorial upscaling of local governments: a variable-oriented approach to explaining variance among Western European countries. *Local Government Studies*, 43, 555-576.
- Avellaneda, C. N. 2008. Municipal performance: Does mayoral quality matter? *Journal of Public Administration Research and Theory* 19:285-312.*
- Avellaneda, C. N., and R. C. Gomes. 2015. Is small beautiful? Testing the direct and nonlinear effects of size on municipal performance. *Public Administration Review* 75, 137-149.
- Badri, M. A., D. Davis and D. Davis. 2000. Operations strategy, environmental uncertainty and performance: a path analytic model of industries in developing countries. *Omega*, 28 155-173.
- Beck, T., A. S. L. I. Demirgüç-Kunt, and V. Maksimovic. 2005. Financial and legal constraints to growth: does firm size matter? *The Journal of Finance* 60, 137-177.
- Bernerth, J. B., and H. Aguinis. 2016. A critical review and best-practice recommendations for control variable usage. *Personnel Psychology* 69, 229-283.
- Berry, C. R. and M. R. West. 2008. Growing pains: The school consolidation movement and student outcomes. *Journal of Law, Economics, & Organization* 26, 1-29.
- Beyer, J. M. and H. M. Trice. 1979. A reexamination of the relations between size and various components of organizational complexity. *Administrative Science Quarterly* 24, 48-64.
- Bidwell, C. E., and J. D. Kasarda. 1985. *The Organization and its ecosystem: A theory of structuring in organizations*. Greenwich, CT: JAI press.
- Black, D. A., B. G. Noel, and Z. Wang. 1999. On-the-job training, establishment size and firm size: Evidence for economies of scale in the production of human capital. *Southern Economic Journal* 1, 82-100.

- Blau, P. M. 1970. A formal theory of differentiation in organizations. *American Sociological Review* 35, 201-218.
- Blau, P. M. 1972. Interdependence and hierarchy in organizations. *Social Science Research* 1, 1-24.
- Blom-Hansen, J., K. Houlberg, S. Serritzlew, and D. Treisman. 2016. Jurisdiction size and local government policy expenditure: Assessing the effect of municipal amalgamation. *American Political Science Review* 110, 812-831.
- Bohte, J. 2004. Examining the impact of charter schools on performance in traditional public schools. *Policy Studies Journal* 32:501-520.*
- Bohte, J., and K. J. Meier. 2001. Structure and the performance of public organizations: Task difficulty and span of control. *Public Organization Review* 1, 341-354.*
- Borst, R.T., P. M. Kruiyen, C. J. Lako and M. S. de Vries. 2020. The attitudinal, behavioral, and performance outcomes of work engagement: A comparative meta-analysis across the public, semipublic, and private sector. *Review of Public Personnel Administration* 40, 613-640.
- Boyne, G. A. 1996 Scale, performance and the New Public Management: an empirical analysis of local authority services. *Journal of Management Studies* 33, 809-826.
- Boyne, G. A. 1998. *Public choice theory and local government: A comparative analysis of the UK and the USA*. Houndmills, UK: Macmillan.
- Boyne, G. A. 2002. Theme: Local Government: concepts and indicators of local authority performance: an evaluation of the statutory frameworks in England and Wales. *Public Money and Management* 22, 17-24.
- Boyne, G. A. 2003. Sources of public service improvement: A critical review and research agenda. *Journal of Public Administration Research and Theory* 13, 367-394.

- Boyne, G. A., and K. J. Meier. 2013. Burdened by bureaucracy? Determinants of administrative intensity in public organizations. *International Public Management Journal* 16, 307-327.
- Brewer, G. A. 2005. In the eye of the storm: Frontline supervisors and federal agency performance. *Journal of Public Administration Research and Theory* 15:505-527.*
- Brewer, G.A., and R.M. Walker. 2013. Personnel constraints in public organizations: The impact of reward and punishment on organizational performance. *Public Administration Review* 73:121-131.*
- Caplow, T. 1957. Organizational size. *Administrative Science Quarterly* 1, 484-505.
- Carmeli, A. 2006. The managerial skills of the top management team and the performance of municipal organisations. *Local Government Studies* 32:153-176.*
- Charbonneau, É., and G.G. Van Ryzin. 2012. Performance measures and parental satisfaction with New York City schools. *The American Review of Public Administration* 42:54-65.*
- Child, J. 1972. Organizational structure, environment and performance: The role of strategic choice. *Sociology*, 6, 1-22.
- Cho, C. L., C.A. Kelleher, D.S. Wright, and S.W. Yackee. 2005. Translating national policy objectives into local achievements across planes of governance and among multiple actors: Second-order devolution and welfare reform implementation. *Journal of Public Administration Research and Theory* 15:31-54.*
- Chun, Y. H., and H.G. Rainey. 2005. Goal ambiguity and organizational performance in US federal agencies. *Journal of Public Administration Research and Theory* 15:529-557*
- Colegrave, A. D. and M. J. Giles. 2008. School cost functions: A meta-regression analysis. *Economics of Education Review* 27, 688-696.
- Dahl, R. A. and E. R. Tuft. 1973. *Size and democracy*. California: Stanford University Press.

- Dalton, D. R., Todor, W. D., Spendoloni, M. J., Fielding, G. J. and L. W. Porter. 1980. Organizational Structure & Performance: A Critical Review. *Academy of Management Review* 5, 49—64.
- Damanpour, Faribroz. 1993. Organizational size and innovation. *Organization Studies* 13, 375-402.
- Davis, T. J., and G.T. Gabris. 2008. Strategic compensation utilizing efficiency wages in the public sector to achieve desirable organizational outcomes. *Review of Public Personnel Administration* 28:327-348.*
- Denters, B., M. Goldsmith, A. Ladner, P. E. Mouritzen, and L. Rose, Lawrence E. 2014. *Size and Local Democracy*. Edward Elgar Publishing.
- Ding, F., J. Lu and N. M. Riccucci. 2021. How bureaucratic representation affects public organizational performance: A meta-analysis. *Public Administration Review* 81, 1003-1018.
- Dollery, B. and L. Crase. 2004. Is bigger local government better? An evaluation of the case for Australian municipal amalgamation programs. *Urban Policy and Research* 22, 265-275.
- Dollery, B., H. Kitchen, M. McMillan and A. Shah. 2020. Structural reform: Municipal mergers. In: Dollery, B., H. Kitchen, M. McMillan and A. Shah. (eds) *Local public, fiscal and financial governance* (pp. 231-255). Palgrave Macmillan, Cham.
- Donaldson, L. 2001. *The contingency theory of organizations*. London, Sage.
- Drew, J., E Razin, and R. Andrews. 2019. Rhetoric in municipal amalgamations: a comparative analysis. *Local Government Studies* 45, 748-767.
- Faulk, Dagney G. and Michael J. Hicks. 2011. *Local government consolidation in the United States*. Amherst, NY, Cambria Press.
- Favero, N., and K.J. Meier. 2013. Evaluating urban public schools: Parents, teachers, and state

- assessments. *Public Administration Review* 73:401-412.*
- Feiock, R. C., J. B. Carr, and L. S. Johnson. 2006. Structuring the debate on consolidation: A response to Leland and Thurmaier. *Public Administration Review* 66, 274-278.
- Feiock, R. C., and L.G. Kalan. 2001. Assessing the performance of solid waste recycling programs over time. *American Review of Public Administration* 31:22-32.*
- Fesler, J. 1949. *Area and administration*. University of Alabama Press.
- Field, A. P. and R Gillett. 2010. How to do a meta-analysis. *British Journal of Mathematical and Statistical Psychology* 63, 665-694.
- Flynn, A. and P. Davis. 2016. The policy–practice divide and SME-friendly public procurement. *Environment and Planning C: Government and Policy*, 34, 559-578.
- Fox, W. F., and T. Gurley. 2006. Will consolidation improve sub-national governments? World Bank Working Paper 3913.
- Fournier, J-N and Å Johansson. 2016. The effect of the size and the mix of public spending on growth and inequality. *OECD Economic Department Working Papers*, No. 1344. Paris: OECD Publishing.
- Franco, A., N. Malhotra, and G. Simonovits. 2014. Publication bias in the social sciences: Unlocking the file drawer. *Science* 345, 1502-1505.
- Gazley, B., W.K. Chang, and L.B. Bingham. 2010. Board diversity, stakeholder representation, and collaborative performance in community mediation centers. *Public Administration Review* 70:610-620.*
- Gendźwiłł, A., A. Kurniewicz and P. Swianiewicz. 2020. The impact of municipal territorial reforms on the economic performance of local governments. A systematic review of quasi-experimental studies. *Space and Polity*, pp.1-20.

- George, B. and S. K. Pandey. 2017. We know the yin—But where is the yang? Toward a balanced approach on common source bias in public administration scholarship. *Review of Public Personnel Administration* 37, 245-270.
- George, B., R. M. Walker, and J. Monster. 2019. Does strategic planning improve organizational performance? A meta-analysis. *Public Administration Review* 79, 810-819.
- George, B., S. Van de Walle, and G. Hammerschmid. 2019. Institutions or contingencies? A cross-country analysis of management tool use by public sector executives. *Public Administration Review* 79, 330-342.
- George, B., S.K. Pandey, B. Steijn, A. Decramer, and M. Audenaert. 2021. Red tape, organizational performance, and employee outcomes: Meta-analysis, meta-regression, and research agenda. *Public Administration Review* 81, 638-651.
- Gerrish, E. 2016. The impact of performance management on performance in public organizations: A meta-analysis. *Public Administration Review* 76, 48-66.
- Glass, G. V. 1976. Primary, secondary, and meta-analysis of research. *Educational researcher* 5, 3-8.
- Goerdel, H. T. 2006. Taking initiative: Proactive management and organizational performance in networked environments. *Journal of Public Administration Research and Theory* 16: 351-367.*
- Gooding, R. Z. and J. A. Wagner III. 1985. A meta-analytic review of the relationship between size and performance: The productivity and efficiency of organizations and their subunits. *Administrative Science Quarterly* 30, 462-481.
- Grosskopf, S. and S. Yaisawamg, S. 1990. Economies of scope in the provision of local public services. *National Tax Journal* 43, 61-74.

- Gulati, R., D. Lavie and R. R. Madhavan. 2011. How do networks matter? The performance effects of interorganizational networks. *Research in Organizational Behavior* 31, pp.207-224.
- Gutierrez-Romero, R., D. Haubrich., and I. McLean. 2008. The limits of performance assessments of public bodies: external constraints in English local government. *Environment and Planning C-Government and Policy* 26:767-787.*
- Hall, R. H. 1987 *Organizations: Structure, process and outcomes*. (3rd ed.) Englewood Cliffs, NJ: Prentice-Hall.
- Heinrich, C. J., and L. E. Lynn Jr. 2000. *Governance and performance: New perspectives*. Washington, DC: Georgetown University Press.
- Hedges, L. V., and J. L. Vevea. 1998. Fixed-and random-effects models in meta-analysis. *Psychological Methods* 3, 486-504.
- Homberg, F, D. McCarthy, and V. Tabvuma. 2015. A meta-analysis of the relationship between public service motivation and job satisfaction. *Public Administration Review* 75, 711-722.
- Hood C.. 1991. A public management for all seasons. *Public Administration* 69, 3-19.
- Hunter, J. E., and F. L. Schmidt. 1990. *Methods of meta-analysis*. Beverly Hills, CA: Sage.
- Hur, Y. 2013. Racial diversity, is it a blessing to an organization? Examining its organizational consequences in municipal police departments. *International Review of Administrative Sciences* 79:149-164.*
- Jakobsen, Morten, and Rasmus Jensen. 2015. Common method bias in public management studies. *International Public Management Journal* 18, 3-30.
- Jugl, M.. 2019. Finding the golden mean: Country size and the performance of national bureaucracies. *Journal of Public Administration Research and Theory* 29, 118-132.

- Jung, C. S. 2013. Navigating a rough terrain of public management: Examining the relationship between organizational size and effectiveness. *Journal of Public Administration Research and Theory* 23, 663-686.
- Jung, C. S. 2011. Organizational goal ambiguity and performance: conceptualization, measurement, and relationships. *International Public Management Journal* 14:193-217.*
- Jung, C. S. and Lee, Geo. 2016. Organizational climate, leadership, organization size, and aspiration for innovation in government agencies. *Public Performance & Management Review* 39, 757-782.
- Kalleberg, A. L. and M. E. Van Buren. 1996. Is bigger better? Explaining the relationship between organization size and job rewards. *American Sociological Review* 61, 47-66.
- Keating, M. 1995. Size, efficiency and democracy: consolidation, fragmentation and public choice. In: David Judge, Gerry Stoker and Harold Wolman (eds) *Theories of Urban Politics*, London: Sage 117-134.
- Kimberly, J. 1976. Organizational size and the structuralist perspective: A review, critique and proposal. *Administrative Science Quarterly* 21, 571-597.
- Knott, J. H., and A.A. Payne. 2004. The impact of state governance structures on management and performance of public organizations: A study of higher education institutions. *Journal of Policy Analysis and Management* 23:13-30.*
- Kuhlmann, Sabine and Geert Bouckaert. (eds). 2016. *Local public sector reforms in times of crisis*. London, UK: Palgrave Macmillan
- Lago-Peñas, S. and J. Martinez-Vazquez (eds). 2013. *The challenge of local government size: theoretical perspectives, international experience and policy reform*. Cheltenham: Edward Elgar Publishing.
- Langbein, L., and E.C. Stazyk. 2013. Vive la différence? The impact of diversity on the turnover

- intention of public employees and performance of public agencies. *International Public Management Journal* 16:465-503.*
- Ladd, H. F. 1992. Population growth, density and the costs of providing public services. *Urban Studies* 29, 273-295.
- Lee, J and M. Kim. 2016. Market-driven technological innovation through acquisitions: The moderating effect of firm size. *Journal of Management* 42, 1934-1963.
- Lee, S. Y., and A.B. Whitford. 2013. Assessing the effects of organizational resources on public agency performance: Evidence from the US federal government. *Journal of Public Administration Research and Theory* 23:687-712.*
- Liberati, A., D. G. Altman, J. Tetzlaff, C. Mulrow, P. C. Gotzsche, J. P. A. Ioannidis, M. Clarke, P. J. Deveraux, J. Kleijnen, D. Moher. 2009. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: Explanation and elaboration. *PLoS Med* 6(7):e1000100.
- Light, P. C. 1999. *The true size of government*. Washington D.C.: The Brookings Institute.
- Ling, Y., H. Zhao, and R. A. Baron. 2007. Influence of founder—CEOs' personal values on firm performance: Moderating effects of firm age and size. *Journal of Management* 33, 673-696.
- Löfstedt, R. E. and D. Vogel. 2001. The changing character of regulation: A comparison of Europe and the United States. *Risk Analysis* 21, 399-405.
- Lowndes, V. and H. Sullivan. 2008. How low can you go? Rationales and challenges for neighborhood governance. *Public Administration* 86, 53-74.
- Maguire, E. R. 2003. *Organizational structure in American police agencies: Context, complexity, and control*. New York: SUNY Press.
- Marsden, P. V., C. R. Cook, and A. L. Kalleberg. 1994. Organizational structures: Coordination and control. *American Behavioral Scientist* 37, 911-929.

- Matt, G. E. 2003. Will it work in Münster? Meta-analysis and the empirical generalization of causal relationships.” In: R. Schulze, H. Holling, and D. Böhning, eds. *Meta-analysis: New developments and applications in medical and social sciences*. Cambridge, MA: Hogrefe and Huber 113-139.
- Meier, K. J., and L. J. O'Toole Jr. 2013a. I think (I am doing well), therefore I am: assessing the validity of administrators' self-assessments of performance. *International Public Management Journal* 16, 1-27.*
- Meier, K. J., and L. J. O'Toole Jr. 2013b. Subjective organizational performance and measurement error: Common source bias and spurious relationships. *Journal of Public Administration Research and Theory* 23, 429-456.
- Meier, K. J., A. Rutherford, and C. N. Avellaneda. Eds. 2017. *Comparative public management: Why national, environmental, and organizational context matters*. Washington DC: Georgetown University Press.
- Melkers, J., and K. Willoughby. 2005. Models of performance-measurement use in local governments: Understanding budgeting, communication, and lasting effects. *Public Administration Review* 65 180-190.*
- Meyer, A. D., A. S. Tsui, and C. R. Hinings. 1993. Configurational approaches to organizational analysis. *Academy of Management Journal* 36, 1175-1195.
- Milward, H. B. and K. G. Provan. 2000. Governing the hollow state. *Journal of Public Administration Research and Theory* 10, 359-380.
- Newton, K. 1982. Is small really so beautiful. Is big really so ugly? Size, effectiveness and democracy in local government. *Political Studies* 30, 190-206.
- Nicholson-Crotty, S., J.A. Grisom., and Nicholson-Crotty, J. 2012. Governance and the impact of public employee unions on organizational performance. *Public Performance & Management Review* 35:422-448.*

- Nicholson-Crotty, S., and L.J. O'Toole. 2004. Public management and organizational performance: The case of law enforcement agencies. *Journal of Public Administration Research and Theory* 14:1-18.*
- Nicholson-Crotty, S., N.A. Theobald., and Nicholson-Crotty, J. 2006. Disparate measures: Public managers and performance-measurement strategies. *Public Administration Review* 66:101-113.*
- Niskanen, W. 1971. *Bureaucracy and representative government*. Chicago IL: Aldine.
- OECD. 2005. *Modernizing government*. Paris: OECD
- Olson, M. 1971. Increasing the incentives for international cooperation. *International Organization*, 25, 866-874.
- Osborne, D. and T. Gaebler. 1993. *Reinventing government: How the entrepreneurial spirit is transforming the public sector*. New York: Plume.
- O'Toole, L. J., Jr., and K. J. Meier. 1999. Modeling the impact of public management: Implications and structural content. *Journal of Public Administration Research and Theory* 9, 505–526
- O'Toole, L. J., Jr. and K. J. Meier. 2015. Public management, context, and performance: In quest of a more general theory. *Journal of Public Administration Research and Theory* 25, 237-256.
- Ostrom, E. 1976. Size and performance in a federal system. *Publius*, 6, 33-73.
- Overman, S. and S. Van Thiel. 2016. Agencification and public sector performance: a systematic comparison in 20 countries. *Public Management Review* 18, 611-635.
- Palmon, O. and J. K. Wald. 2002. Are two heads better than one? The impact of changes in management structure on performance by firm size. *Journal of Corporate Finance*, 8, 213-226.

- Perry, J. L. 2012. How can we improve our science to generate more usable knowledge for public professionals? *Public Administration Review* 72, 492-482.
- Pestoff, V.. 2014. Collective action and the sustainability of co-production. *Public Management Review* 16, 383-401.
- Petruzzelli, A. M., L. Ardito and T. Savino. 2018. Maturity of knowledge inputs and innovation value: The moderating effect of firm age and size. *Journal of Business Research* 86, 190-201.
- Pierce, J. R. and H. Aguinis. 2013. The Too-Much-of-a-Good-Thing Effect in management. *Journal of Management* 39, 313-338.
- Podsakoff, P. M., S. B. MacKenzie, J-Y. Lee, and N. P. Podsakoff. 2003. Common method biases in behavioral research: a critical review of the literature and recommended remedies. *Journal of applied psychology* 88, 879-903.
- Pollitt, C. 2006. Performance management in practice: A comparative study of executive agencies. *Journal of Public Administration Research and Theory* 16, 25-44.
- Pollitt, C. and G. Bouckaert. 2017. *Public Management Reform: A Comparative Analysis-into the Age of Austerity*, 4th ed. Oxford, UK: Oxford University Press.
- Provan, K. G. and P. Kenis. 2008. Modes of network governance: Structure, management and effectiveness. *Journal of Public Administration Research and Theory* 18, 229-252.
- Rainey, H. G. 2009. *Understanding and managing public organizations*. 5th ed, San Francisco, CA: John Wiley & Sons.
- Real, J. C., J. L. Roldán and A. Leal. 2014. From entrepreneurial orientation and learning orientation to business performance: analysing the mediating role of organizational learning and the moderating effects of organizational size. *British Journal of Management* 25, 186-208.

- Raudla R. R. S. and T. Randma-Liiv. 2015. Cutback management literature in the 1970s and 1980s: Taking stock. *International Review of Administrative Sciences* 81, 433-456.
- Resh, W. G., and D.W. Pitts. 2013. No solutions, only trade-offs? Evidence about goal conflict in street-level bureaucracies. *Public Administration Review* 73:132-142.*
- Ringquist, E. 2013. *Meta-analysis for public management and policy*. San Francisco: Jossey-Bass.
- Rosenthal, R. 1991. *Meta-analytic procedures for social research*. London, UK: Sage.
- Rosenthal, R. and M. R. DiMatteo. 2001. Meta-analysis: Recent developments in quantitative methods for literature reviews. *Annual review of psychology* 52, 59-82.
- Rushing, W. A. 1967. The effects of industry size and division of labor on administration. *Administrative Science Quarterly* 12, 273-295.
- Rutherford, A. 2016. Re-examining causes and consequences: Does administrative intensity matter for organizational performance. *International Public Management Journal* 19, 342-69
- Salge, T. O., and A. Vera. 2012. Benefiting from public sector innovation: The moderating role of customer and learning orientation. *Public Administration Review* 72:550-559.*
- Selden, S. C. and J. E. Sowa. 2004. Testing a multi-dimensional model of organizational performance: Prospects and problems. *Journal of Public Administration Research and Theory* 14, 395-416.
- Sharpe, L. J. 1970. Theories and values of local government. *Political Studies* 18, 153-174.
- Shenkar, O. and S. Ellis. 2022. The rise and fall of structural contingency theory: A theory's 'autopsy'. *Journal of Management Studies* 59, 782-818.
- Simon, H. A. 1996. *Sciences of the artificial* (3rd ed.). Cambridge, MA: MIT Press.
- Skelcher, C. 2000. Changing images of the state: overloaded, hollowed-out, congested. *Public Policy and Administration* 15, 3-19.

- Smith, B. W. 2003. The impact of police officer diversity on police-caused homicides. *Policy Studies Journal* 31: 147-162.*
- Smith, K. G., J. P. Guthrie and M. J. Chen. 2016. Strategy, size and performance. *Organization Studies* 10, 63-81.
- Snyder, C. M. 1998. Why do larger buyers pay lower prices? Intense supplier competition. *Economics Letters* 58, 205-209.
- Stigler, G/ J. 1958. The economies of scale. *Journal of Law and Economics* 1, 54-71.
- Talbot, C. 2004. 'The Agency Idea', in C. Pollitt and C. Talbot (eds), *Unbundled Government: A Critical Analysis of the Global Trend to Agencies, Quangos and Contractualisation*. London: Routledge, pp . 1 – 21.
- Tavares, A.F. 2018. Municipal amalgamations and their effects: A literature review. *Miscellanea Geographica*, 22, 5-15.
- Theobald, N. A., and S. Nicholson-Crotty. 2005. The many faces of span of control: Organizational structure across multiple goals. *Administration & Society* 36, 648-660.*
- Thompson, James D. 1967. *Organizations in action: Social science bases of administrative theory*. New York: McGraw-Hill.
- Tullock, G. 1965. *The politics of bureaucracy*. Washington: Public Affairs Press.
- Vaccaro, I. G., J. J. P. Jansen, F. A. Van Den Bosch, and H. W. Volberda. 2012. Management innovation and leadership: The moderating role of organizational size. *Journal of Management Studies* 49, 28-51.
- Van de Ven, A. H., M. Ganco and C. R. Hinings. 2013. Returning to the frontier of contingency theory of organizational and institutional designs. *Academy of Management Annals* 7, 393-440.
- Verhoest, K., Van Thiel, S., Bouckaert, G., Lægreid, P. and Van Thiel, S. eds., 2016.

- Government agencies: Practices and lessons from 30 countries*. Berlin: Springer.
- Walker, R. M. 2014. Internal and external antecedents of process innovation: A review and extension. *Public Management Review* 16, 21-44.
- Walker, R. M., and R. Andrews. 2015. Local government management and performance: a review of evidence. *Journal of Public Administration Research and Theory* 25, 101-133.
- Walker, R. M., F. S. Berry, and C. N. Avellaneda. 2015. Limits on innovativeness in local government: Examining capacity, complexity and dynamism in organizational task environments. *Public Administration* 93, 663-683.
- Walker, R. M., G. A. Boyne, and G. A. Brewer, Eds. 2010. *Public management and performance: Research directions*. Cambridge: Cambridge University Press.
- Walker, R. M., and G. A. Brewer. 2009. Can management strategy minimize the impact of red tape on organizational performance? *Administration & Society* 41, 423-448.*
- Walker, R. M., C. S. Jung., and G. A. Boyne. 2013. Marching to different drummers? The performance effects of alignment between political and managerial perceptions of performance management. *Public Administration Review* 73: 833-844.*
- Weil, T. 2010. Hospital mergers: a panacea? *Journal of Health Services Research & Policy* 15, 251-253.
- van Witteloostuijn, A. 2016. What happened to Popperian falsification? Publishing neutral and negative findings: moving away from biased publication practices. *Cross Cultural & Strategic Management* 23, 481-508.
- Yeung, R. 2009. The effects of fiscal decentralization on the size of government: A meta-analysis. *Public Budgeting and Finance* 29, 1-23.

Figure 1. Search protocol and results

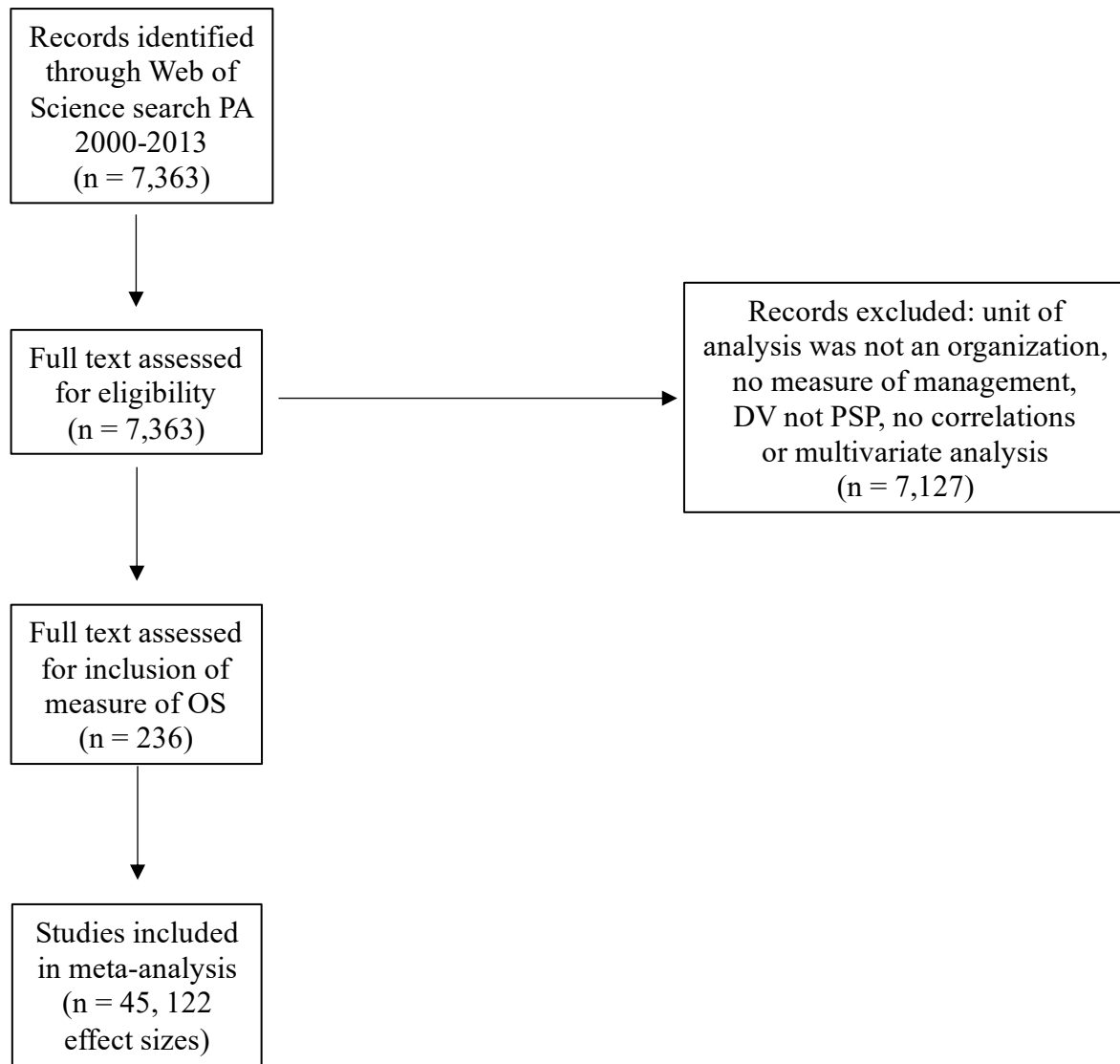


Figure 2. Forest plot of effect sizes (one average per study)

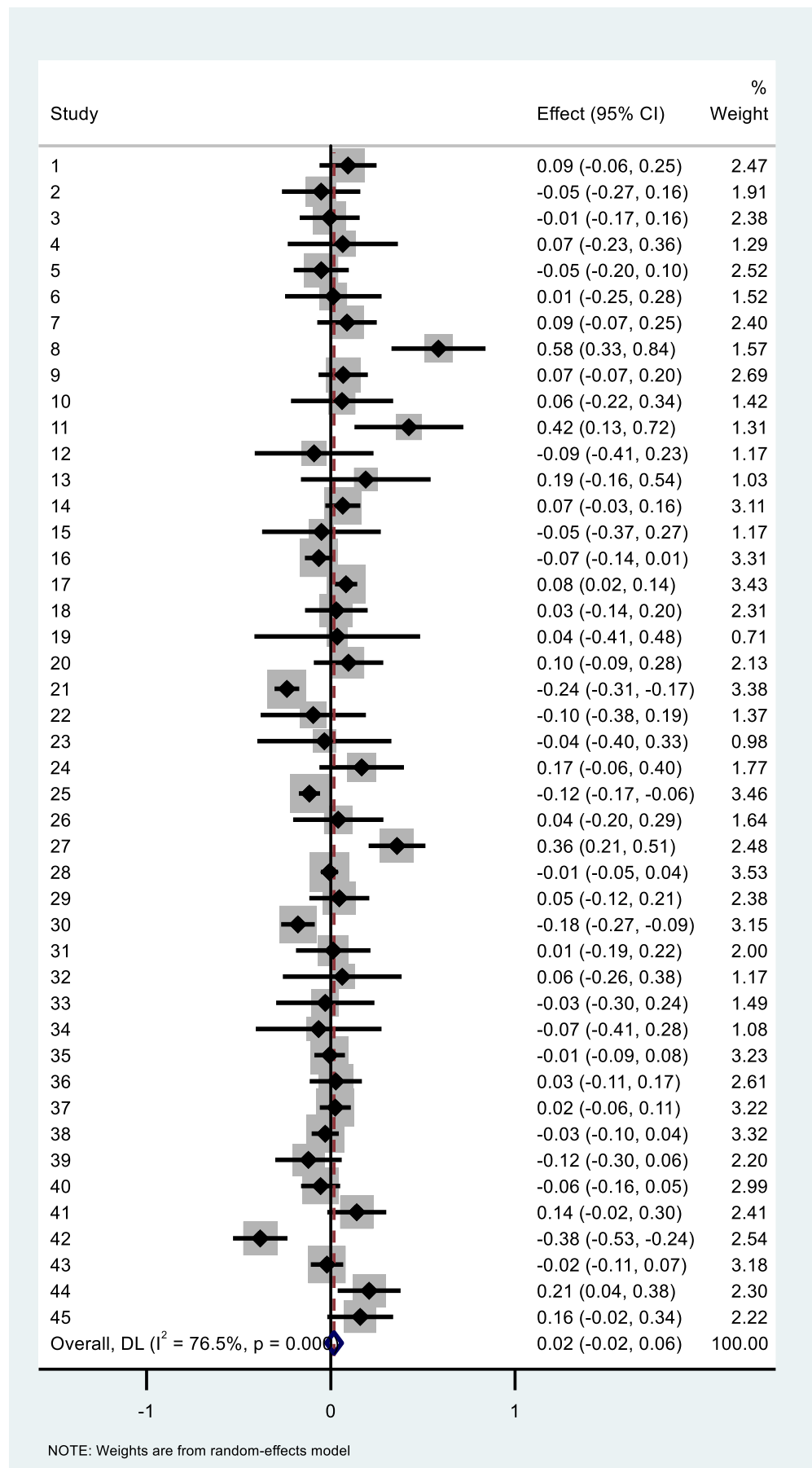


Figure 3. Funnel plot of effect sizes

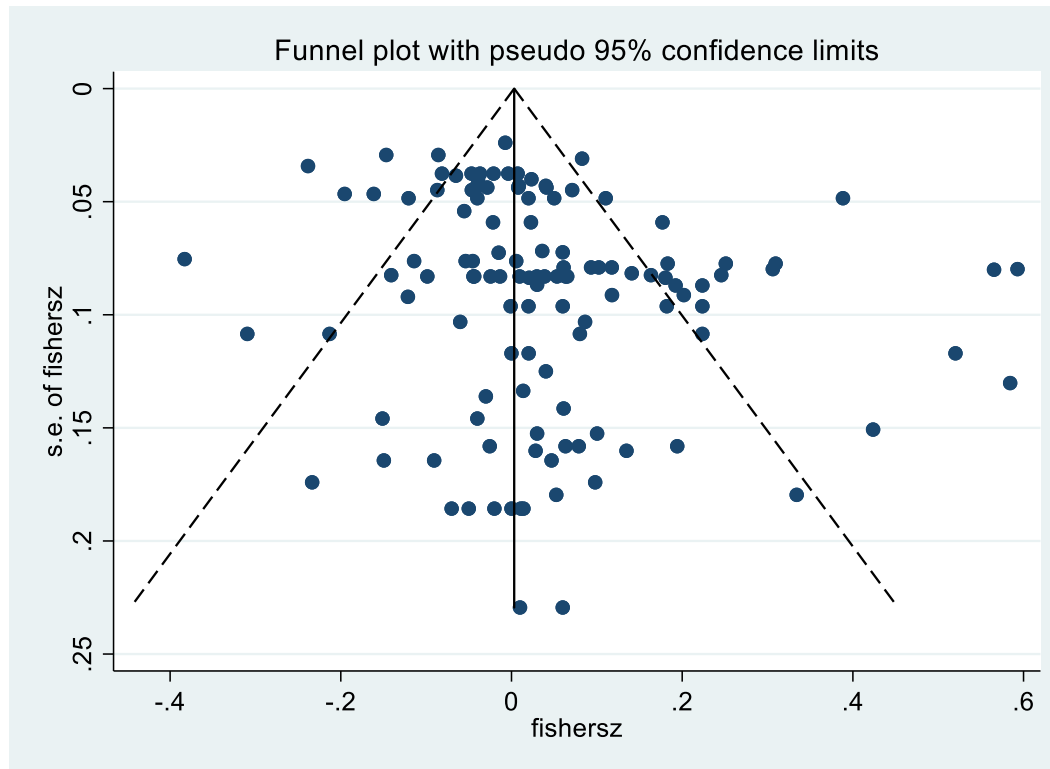


Table 1. Overview of included studies and their characteristics

Author(s)	# of effect sizes	N (average)	Hypothesized direction	Organizational goals	Level of government	Size measure	Performance dimension	Performance measurement	Research design	Country
Amirkhanyan et al. (2010)	4	163	Neutral	Single	Local	Budget, clients	Output, effectiveness	Perceptual, archival	Longitudinal	U.S.A.
Andrews & Boyne (2010)	4	88	Larger = higher P, no stated direction	Multiple	Local	Population, budget	Responsiveness, multiple	Archival	Longitudinal	Europe
Andrews & Entwistle (2010)	2	46	Larger = higher P	Single	Local	Budget	Effectiveness, equity	Archival	Longitudinal	Europe
Andrews & Boyne (2011)	9	148	Larger = higher P	Multiple	Local	Population, budget, population change	Efficiency, effectiveness, equity	Archival	Longitudinal	Europe
Andrews & Van de Walle (2013)	4	175	Neutral	Multiple	Local	Population	Efficiency, effectiveness, responsiveness, equity	Archival	Cross-section	Europe
Andrews (2009)	1	59	No stated direction	Single	Local	Budget	Multiple	Archival	Longitudinal	Europe
Andrews et al (2011)	1	40	Larger = higher P	Single	Local	Budget	Multiple	Archival	Longitudinal	Europe
Andrews et al (2012)	2	34	Larger = higher P, no stated direction	Multiple	Local	Population, budget	Multiple	Archival	Longitudinal	Europe
Andrews et al (2005)	3	150	Neutral, larger = higher P	Multiple	Local	Population, budget, population change	Multiple	Archival	Cross-section	Europe
Andrews et al (2008)	1	62	Larger = higher P	Multiple	Local	Budget	Multiple	Archival	Longitudinal	Europe
Andrews et al (2009)	6	218	Larger = higher P, no stated direction	Multiple	Local	Population, budget, population change	Responsiveness, multiple	Archival	Longitudinal	Europe

Andrews et al (2009a)	1	53	Larger = higher P	Multiple	Local	Budget	Multiple	Archival	Longitudinal	Europe
Andrews et al (2009b)	1	47	Larger = higher P	Single	Local	Budget	Multiple	Archival	Longitudinal	Europe
Andrews et al (2013)	6	429	Larger = higher P, no stated direction	Multiple	Local	Population, budget, population change	Efficiency, multiple	Archival	Longitudinal	Europe
Avellaneda (2008)	2	40	No stated direction	Multiple	Local	Population, budget	Equity	Archival	Longitudinal	Other
Bohte & Meier (2001)	1	678	Null	Single	Local	Clients	Effectiveness	Archival	Longitudinal	U.S.A.
Bohte (2004)	1	1046	No stated direction	Single	Local	Clients	Effectiveness	Archival	Longitudinal	U.S.A.
Brewer (2005)	2	22	No stated direction	Single	Federal	Budget, Employees	Multiple	Perceptual	Cross-section	U.S.A.
Brewer & Walker (2013)	1	136	No stated direction	Multiple	Local	Population change	Multiple	Archival	Longitudinal	Europe
Carmeli (2006)	5	111	No stated direction	Multiple	Local	Population	Efficiency, multiple	Perceptual, archival	Cross-section	Other
Gazley et al (2010)	6	160	No stated direction	Single	Local	Budget	Output, effectiveness	Perceptual	Cross-section	U.S.A.
Charbonneau & van Ryzin (2012)	1	856	No stated direction	Single	Local	Clients	Responsiveness	Archival	Cross-section	U.S.A.
Cho et al (2005)	2	50	No stated direction	Multiple	Local	Population	Effectiveness	Perceptual	Cross-section	U.S.A.
Chun & Rainey (2005)	4	32	No stated direction	Single	Federal	Employees	Output, effectiveness, responsiveness	Perceptual	Cross-section	U.S.A.
Davis & Gabris (2008)	3	76	No stated direction	Multiple	Local	Population, budget, employees	Output	Archival	Cross-section	U.S.A.
Favero & Meier (2013)	2	1164	No stated direction	Single	Local	Clients	Responsiveness	Archival	Longitudinal	U.S.A.
Feiock & Kalan (2001)	1	67	No stated direction	Multiple	Local	Population	Effectiveness	Archival	Longitudinal	U.S.A.

Goerdel (2006)	1	1748	Smaller = higher P	Single	Local	Clients	Effectiveness	Archival	Longitudinal	U.S.A.
Gutierrez-Romero et al (2008)	2	148	No stated direction	Multiple	Local	Population, budget	Multiple	Archival	Longitudinal	Europe
Hur (2013)	2	464	No stated direction	Single	Local	Population, employees	Effectiveness	Archival	Cross-section	U.S.A.
Jung (2011)	2	97	Smaller = higher P	Single	Federal	Budget, employees	Effectiveness	Archival	Cross-section	U.S.A.
Knott & Payne (2004)	8	42	No stated direction	Single	Local	Clients, employees	Output, effectiveness	Archival	Longitudinal	U.S.A.
Langbein & Stazyk (2013)	1	57	No stated direction	Single	Federal	Employees	Multiple	Perceptual	Cross-section	U.S.A.
Lee & Whitford (2013)	2	36	Larger = higher P	Single	Federal	Budget, employees	Effectiveness	Perceptual	Longitudinal	U.S.A.
Meier & O'Toole (2013)	6	526	Larger = higher P, smaller = higher P	Single	Local	Budget, clients	Output, effectiveness	Perceptual	Cross-section	U.S.A.
Melkers & Willoughby (2005)	3	194	No stated direction	Multiple	Local	Population	Efficiency, responsiveness, multiple	Perceptual	Cross-section	U.S.A.
Nicholson-Crotty & O'Toole (2004)	2	544	Null, larger = higher P	Single	Local	Population, employees	Effectiveness	Archival	Longitudinal	U.S.A.
Nicholson-Crotty et al (2006)	6	711	Null, no stated direction	Single	Local	Budget, clients	Effectiveness	Archival	Longitudinal	U.S.A.
Nicholson-Crotty et al (2012)	1	121	No stated direction	Single	Local	Clients	Effectiveness	Archival	Longitudinal	U.S.A.
Resh & Pitts (2013)	1	344	No stated direction	Single	Local	Clients	Effectiveness	Archival	Cross-section	U.S.A.
Salge & Vera (2012)	1	153	Larger = higher P	Single	Local	Employees	Output	Archival	Longitudinal	Europe

Smith (2003)	1	179	No stated direction	Single	Local	Population	Effectiveness	Archival	Cross-section	U.S.A.
Theobald & Nicholson-Crotty (2005)	3	500	Null	Single	Local	Clients	Effectiveness	Archival	Longitudinal	U.S.A.
Walker & Brewer (2009)	2	135	Larger = higher P	Multiple	Local	Population change	Multiple	Archival, perceptual	Cross-section	Europe
Walker et al (2013)	2	123	Larger = higher P, no stated direction	Multiple	Local	Population, budget	Multiple	Archival	Cross-section	Europe

Table 2. Descriptive information on moderators

Moderator	Percentage in sample of effect sizes
<i>Hypothesized direction</i>	
Null	14%
Larger = higher P	28%
Smaller = higher P	5%
No stated direction	53%
<i>Organizational goals</i>	
Single	52%
Multiple	48%
<i>Level of government</i>	
Federal	9%
Local	91%
<i>Size measure</i>	
Population	27%
Budget	32%
Clients	19%
Employees	13%
Population change	9%
<i>Performance dimension</i>	
Output	15%
Efficiency	8%
Effectiveness	36%
Responsiveness	9%
Equity	6%
Multiple	26%
<i>Performance measurement</i>	
Perceptual	26%
Archival	74%
<i>Research design</i>	
Cross-section	41%
Longitudinal	59%
<i>Country</i>	
U.S.A.	54%
Europe	40%
Other	6%

Table 3. Meta-analysis of the size-performance association

Effect sizes	Population effect size	CI 95	Cochran's Q	I²	Tau²	z
122 (45 studies)	.037	[.010, .063]	550.74***	78.0%	.0144	2.711**
45 (45 studies)	.018	[-.025, .060]	187.29***	76.5%	.0125	.816

*p < .05, **p < .01, ***p < .001

Table 4. Results of the subgroup analysis.

Moderator	Effect sizes (studies)	Population effect size	CI 95	z
Hypothesized direction				
Null	17 (7)	-.013	[-.047, .021]	-.763
Larger = higher P	34 (17)	.043	[-.001, .086]	1.923
Smaller = higher P	6 (3)	-.003	[-.035, .030]	-.152
No stated direction	65 (27)	.053	[.006, .100]	2.209*
Organizational goals				
Single	63 (26)	.020	[-.016, .056]	1.106
Multiple	59 (19)	.054	[.016, .091]	2.821**
Level of government				
Federal	11 (5)	-.014	[-.105, .077]	-.296
Local	111 (40)	.040	[.012, .067]	2.827**
Size measure				
Population	33 (18)	.042	[-.020, .103]	1.332
Budget	39 (23)	.078	[.023, .133]	2.759**
Clients	23 (12)	-.041	[-.078, -.003]	-2.141*
Employees	16 (10)	.006	[-.062, .073]	.173
Population change	11 (6)	.088	[.034, .142]	3.200**
Performance dimension				
Output	18 (7)	.149	[.059, .240]	3.223**
Efficiency	10 (5)	.057	[-.049, .163]	1.060
Effectiveness	44 (22)	-.011	[-.045, .022]	-.662
Responsiveness	11 (7)	-.046	[-.123, .032]	-1.158
Equity	7 (4)	-.003	[-.076, .070]	-.088
Multiple	32 (18)	.082	[.030, .133]	3.101**
Performance measurement				
Perceptual	32 (11)	.101	[.040, .161]	3.232**
Archival	90 (37)	.014	[-.014, .043]	.980
Research design				
Cross-section	50 (18)	.056	[.002, .110]	2.022*
Longitudinal	72 (27)	.022	[-.006, .050]	1.556
Country				
U.S.A.	66 (25)	.019	[-.016, .054]	1.069
Europe	49 (18)	.057	[.015, .099]	2.640**
Other	7 (2)	.078	[-.005, .161]	1.851

*p < .05, **p < .01, ***p < .001

Table 5. Results of meta-regression analysis.

Moderators	Coefficient	Robust standard error	CI 95
Constant	.047	.117	[-.189, .282]
<i>Hypothesized direction (Null[^])</i>			
Larger = higher P	-.032	.049	[-.130, .066]
Smaller = higher P	-.037	.085	[-.209, .135]
No stated direction	.016	.040	[-.065, .096]
<i>Organizational goals (Single[^])</i>			
Multiple	-.048	.078	[-.205, .109]
<i>Level of government (Federal[^])</i>			
Local	.064	.073	[-.083, .211]
<i>Size measure (Population[^])</i>			
Budget	.001	.068	[-.136, .137]
Clients	-.071	.065	[-.203, .061]
Employees	-.041	.075	[-.193, .111]
Population change	.065	.058	[-.053, .182]
<i>Performance dimension (Output[^])</i>			
Efficiency	-.104	.057	[-.218, .010]
Effectiveness	-.116*	.051	[-.219, .014]
Responsiveness	-.175*	.054	[-.284, -.065]
Equity	-.163*	.056	[-.275, -.051]
Multiple	-.098	.067	[-.232, .037]
<i>Performance measurement (Perceptual[^])</i>			
Archival	.096	.067	[-.039, .231]
<i>Research design (Cross-section[^])</i>			
Longitudinal	.023	.039	[-.056, .102]
<i>Country (U.S.A.[^])</i>			
Europe	.081	.072	[-.065, .226]
Other	.052	.044	[-.038, .141]
Number of observations		122	
F-value		6.64***	
R ²		.2875	
Root MSE		.99308	

*p < .05, **p < .01, ***p < .001

Notes:

CI = confidence interval,

[^] reference category,

Standard errors are clustered at the study level (N = 45)